Shuo-Yan Chou
Amy Trappey
Jerzy Pokojski
Shana Smith
Editors

Global Perspective for Competitive Enterprise, Economy and Ecology

Proceedings of the 16th ISPE International Conference on Concurrent Engineering

WITH CD-ROM

Springer

Global Perspective for Competitive Enterprise, Economy and Ecology

Shuo-Yan Chou • Amy Trappey
Jerzy Pokojski • Shana Smith
Editors

Global Perspective for Competitive Enterprise, Economy and Ecology

Proceedings of the 16th ISPE International Conference on Concurrent Engineering

Springer

Editors
Shuo-Yan Chou, PhD
Department of Industrial Management
National Taiwan University of Science and Technology (NTUST)
43 Keelung Road, Section 4
Taipei 106
Taiwan R.O.C.
sychou@im.ntust.edu.tw

Amy Trappey, PhD
Department of Industrial Engineering and Management
National Taipei University of Technology
No.1, Section 3, Chung-Hsiao E. Road
Taipei 106
Taiwan R.O.C

Jerzy Pokojski, PhD
Inst. Machine Design Fundamentals
Warsaw University of Technology
Narbutta 84
02-524 Warszawa
Poland

Shana Smith, PhD
Department of Mechanical Engineering
National Taiwan University
1 Roosevelt Road, Section 4
Taipei 106
Taiwan R.O.C.

ISBN 978-1-84882-761-5 e-ISBN 978-1-84882-762-2
DOI 10.1007/978-1-84882-762-2
Springer Dordrecht Heidelberg London New York

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2009929381

Cover design: eStudioCalamar, Figueres/Berlin

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The proceedings contain papers accepted for the 16th ISPE International Conference on **Concurrent Engineering**, held in the vibrant city of Taipei, Taiwan, from July 20 to 24, 2009. The conference is a sequel of the conference series launched by the International Society for Productivity Enhancement and has constituted an important forum for international scientific exchange on Concurrent Engineering (CE).

CE appeared in the 80's as a concept of parallel performing engineering design activities and integrating all related processes. This concept is based on general assumption that different components of product life cycle should be considered together and relatively early in the development process. The main goal of CE is to make processes more efficient and more resistant to errors. Substantive advantages can be achieved by adopting CE strategies and methodologies.

The last twenty years brought many changes in organization of product design and manufacturing. Engineers' professions received narrower specializations. Engineers became present on global market. Sometimes firms create alliances. Engineers work in firms from suppliers to final producers. Engineers cooperate and collaborate cross the border of countries. They need to use different methods and tools supporting their engineering and development activities. As a result, CE has been further expanded to support many aspects of product development. Meantime the whole CE approach has got different forms and names and has become omnipresent. Industrial presence of CE differs in particular cases, from well-established corporation implementations to little small firm applications.

From the beginning the role of information systems in Concurrent Engineering was treated as an indispensible facilitation technology. First methodologies and tools were concentrated on offering possibility to contact people and processes, to make available right information and knowledge at the right time. The presence of computer tools in CE is treated as a standard and covers a spectrum of activities. The functions are no longer limited to passively managing data and processes; they have become intelligent and proactively assisting design and development activities.

If one looks now at what is going on and at how many different issues are important in design, manufacturing, supply, distribution, etc., it is evident why Concurrent Engineering context is so rich and so complicated, why we have so many CE specializations and why CE2009 Conference's main topic is the following: **Global Perspective for Competitive Enterprise, Economy and Ecology.**

The plurality of CE specializations mentioned above was transformed on the following plurality of CE2009 Conference tracks: Systems Engineering, Advanced Manufacture, Product Design, Design for Sustainability, Knowledge Engineering, SCM, Collaborative Engineering, Web Technologies, Service Solutions. Apart of the enumerated tracks the conference has also seven special sessions: Special Session in RFID, in Collaborative Product Development, in Multi-disciplinary Design and Optimization, in Design Knowledge Utilization, in Competitive Supply Chain Performance, in Value Engineering, in Competitive Design.

The proceedings contain 84 papers by authors from 14 countries. If we concluded that they belong to different tracks and special sessions then we see how multi-perspective the content of this volume is. There are papers which are theoretic, conceptual and papers which have very strong industrial roots. There are also papers very detailed, made from a narrow view and very close to specific industrial case studies. We can also find papers which are based on real processes but which operate on abstractive models and which offer a bridge between an industrial reality and an academic research. This heterogeneous nature of Concurrent Engineering brings together diverse and significant contribution to product design and development, which is also what the proceedings intend to offer.

Concurrent Engineering doesn't develop equally in each direction. The way of development depends on many factors. We think that the content of this volume reflects what is actual, noticeable and the issues' variety in the present stage of CE methods and phenomena. As a consequence of this fact careful readers can build their own view of present problems and methods of CE.

Shuo-Yan Chou
General Chair, CE 2009
National Taiwan University of
Science and Technology, Taiwan

Amy Trappey
General Co-Chair, CE 2009
National Taipei University of
Technology, Taiwan

Jerzy Pokojski
Program Chair, CE 2009
Warsaw University of Technology, Poland

Shana Smith
Local Chair, CE 2009
National Taiwan University, Taiwan

Program Committee

General Chair:
Shuo-Yan Chou, National Taiwan University of Science and Technology, Taiwan

General Co-chair:
Amy Trappey, National Taipei University of Technology, Taiwan

Program Chair:
Jerzy Pokojski, Warsaw University of Technology, Poland

Local Chair:
Shana Smith, National Taiwan University, Taiwan

General Secretariat:
Yi-Kuei Lin, National Taiwan University of Science and Technology, Taiwan

Scientific Chair:
Chih-Hsing Chu, National Tsinghua University, Taiwan

Financial Chair:
Kung-Jeng Wang, National Taiwan University of Science and Technology, Taiwan

Publicity Chair:
Shih-Wei Lin, Chang Gung University, Taiwan

Logistics Chair:
Chih-Hsien Chen, Fo Guang University, Taiwan

ISPE Technical Committee

Ahmed Al-Ashaab, Cranfield University, UK

Amy Trappey, National Taipei University of Technology, Taiwan

Geilson Loureiro, International Society for Productivity Enhancement, Brazil

Jerzy Pokojski, Warsaw University of Technology, Poland

John Cha, Beijing Jiaotong University, China

M.J.L. van Tooren, Delft University of Technology, Netherlands

Mike Sobolewski, Texas Tech University, Texas, USA

P.M. Wognum, Wageningen University, Netherlands

Parisa Ghodous, University of Lyon, France

Rajkumar Roy, Cranfield University, UK

Richardo Goncalves, Instituto de Desenvolvimento de Novas Tecnologias, Portugal

Ricky Curran, Delft University of Technology, Netherlands

Shuichi Fukuda, Stanford University, USA

International Program Committee (IPC)

Sung-Hoon Ahn, Seoul National University, Korea

Ahmed Al-Ashaab, Cranfield University, UK

John Cha, Beijing Jiaotong University, China

Kuei-Yuan Chan, National Cheng Kung University, Taiwan

Wei-Lun Chang, Tamkang University, Taiwan

Chih-Hsien Chen, Fo Guang University, Taiwan

Jahau Lewis Chen, National Cheng Kung University, Taiwan

Kai-Ying Chen, National Taipei University of Technology, Taiwan

Shuo-Yan Chou, National Taiwan University of Science and Technology, Taiwan

Chih-Hsing Chu, National Tsinghua University, Taiwan

Richard Curran, Delft University of Technology, Netherlands

Jerry Y.H. Fuh, National University of Singapore, Singapore
Shuichi Fukuda, Stanford University, USA
Liang Gao, Huazhong University of Science and Technology, China
Parisa Ghodous, University of Lyon, France
Richardo Goncalves, Instituto de Desenvolvimento de Novas Tecnologias, Portugal
Kazuo Hiekata, The University of Tokyo, Japan
George Huang, The University of Hong Kong, Hong Kong
Roger J. Jiao, Georgia Institute of Technology, Atlanta, USA
Ajay Joneja, Hongkong University of Science and Technology, Hong Kong
Weidong Li, Coventry University, UK
Yi-Kuei Lin, National Taiwan University of Science and Technology, Taiwan
Shih-Wei Lin, Chang Gung University, Taiwan
Geilson Loureiro, International Society for Productivity Enhancement, Brazil
Masato Inoue, The University of Electro-Communications, Japan
Takashi Niwa, The University of Tokyo, Japan
Jerzy Pokojski, Warsaw University of Technology, Poland
Rajkumar Roy, Cranfield University, UK
Shana Smith, National Taiwan University, Taiwan
Mike Sobolewski, Texas Tech University, USA
Qiang Su, Shanghai Jiaotong University
Chika Sugimoto, The University of Tokyo, Japan
Kenji Tanaka, The University of Tokyo, Japan
Tien-Lung Sun, Yuan Ze University, Taiwan
Kai Tang, Hong Kong University of Science and Technology, Hong Kong
Dunbing Tang, Nanjing University of Aerospace and Aeronautics, China
M.J.L. van Tooren, Delft University of Technology, Netherlands
Amy Trappey, National Taipei University of Technology, Taiwan
Kung-Jeng Wang, National Taiwan University of Science and Technology, Taiwan

Charlie, C.L. Wang, The Chinese University of Hong Kong, Hong Kong

Gary Wang, Simon Fraser University, Canada

Yan Wang, University of Central Florida, USA

Nel Wognum, Wageningen University, Netherlands

Yong Zeng, Concordia University, Canada

Sponsors

ISPE : International Society for Productivity Enhancement

National Science Council, Taiwan

National Taiwan University of Science and Technology, Taiwan

National Taipei University of Technology, Taiwan

Chinese Institute of Industrial Engineers

SME : Society of Manufacturing Engineers
Taipei Chapter

Contents

Advanced Manufacture

Collaborative Engineering

Collaborative Product Development

Competitive Design

Design for Sustainability

Design Knowledge Utilization

Knowledge Engineering

Multi-Disciplinary Design and Optimization

Product Design

Radio Frequency Identification (RFID)

Competitive Supply Chain Performance

Value Engineering

Web Technologies

Advance Manufacture

To Calculate the Sequence-Dependent Setup Time for a Single-Machine Problem with Uncertain Job Arriving Time

Ming-Hsien Yang[a,1], Shu-Hsing Chung[b], and Ching-Kuei Kao[b]

[a]Department of Business Management, National United University, Miao-Li, Taiwan, ROC.
[b]Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu, Taiwan, ROC.

Abstract: Consider a finite-capacity single machine responsible for processing several product types of jobs, in which the inter-arrival time of jobs of specific type is independently and exponentially distributed to reflect the uncertainties in market demand. Since the uncertainties of the time and the type of customer demand make the machine setup decision very complicated, the development of analytic models to calculate the expected sequence-dependent setup time under FIFO and FSR rules contributes to the quick evaluation of capacity waste due to changing machine setting. Analytic model also shows the saving of machine utilization rate by changing the considered dispatching rule from FIFO to FSR. ([1]E-mail: ymh@nuu.edu.tw (M. H. Yang))

Keywords: analytic models, expected sequence dependent setup time, Poisson process jobs arriving, capacity reduction, change the setting of machine

1 Introduction

We consider a manufacturing system designed for producing several types of products according to customer order, machine setup to switch from the current setting to a different one usually is necessary and can not be regarded as a part of job processing time. When a job of specific product type, indicating a customer order, arrives at the system, it will enter the queue line and waits a period of time according to the dispatching or scheduling rule, and then starts the job processing on machine. In this paper, the first-in first-out (FIFO) rule and the family-based scheduling rule (FSR) are considered, in which FSR [1, 2, 3] can reduce the setups, comparing with FIFO. Before starting the processing of a new job, the machine should pause and proceeds setup if the previously completed job is different from the new job. Therefore, not only job processing but also machine setup consumes

[1] Corresponding Author : Department of Business Management, National United University, Miao-Li, Taiwan, ROC; Tel: 886-(37)-381591; Fax: 886-(37)-332396; E-mail: ymh@nuu.edu.tw

machine capacity, which indicates that total setup time is related to the efficient utilizing of machine capacity.

There are two types of setups, (1) sequence independent and (2) sequence dependent, the second is a generality of the first and is considered in this paper. Besides, there are two levels of setup time problem [4], including the lower level concerning the role of setup in scheduling jobs [5, 6, 7], and the higher level concerning that in decision making at the production planning and control [8]. Missbauer [8] and Vieira *et al.* [6, 7] considered the dynamic behavior of job arriving in their researches. Missbauer [8] defined the probability of a setup is necessary for an arriving job complying with Poisson process underlying the FIFO and the FSR dispatching rules, however, did not address the setup probability for an arriving job of specific product type. Vieira *et al.* [6 ,7] assumed an independent Poisson arrival of a job with product type j and arrival rate λ_j and defined the probability of requiring a setup in a time interval. However, their equations simplified the setup probability to constantly be $(1-1/J)$ without considering the effect of the product type of an arriving job.

To analyze the accuracy of proposed models, an experimental design with various arriving conditions among several types of product, which corresponds to various resource utilization rates, is conducted.

2 Development of the probabilistic model for setups

The number of setups and the sequence dependent setup time are observed for a period of time RT, which begins initially at time 0. Let $N_j(t)$ be the number of arriving jobs of product type j by time t. That is, $N_j(t)$ is the random variable to occur in the fixed interval $(0, RT]$ and is a Poisson process with arrival rate λ_j, and there would be n_j independent jobs in the time interval $(0, RT]$ for the product type j, where $n_j=\lambda_j \times RT$, $j = 1, 2, \ldots, J$, and J is the number of product types. The total arrival rate λ is equal to the sum of λ_j, $\sum_{j=1}^{J} \lambda_j$.

2.1 Development of the probabilistic model of setups under FIFO

A setup is necessary given the condition that if a job arrives in the time interval $(0, RT]$ encountering a previously completed job of different product type, in which there are no jobs or at lease one job in the system. Considering the i^{th} arriving job of product type j, then the waiting time for this job (T_{ij}) is a gamma variable with parameters i and λ_j. Thus, the probability for this job arriving in the time interval $(0, RT]$ is shown as Equation (1), and the probability for this job arriving in time interval $(0, RT]$ with a setup under FIFO ($P_{s,ij,FIFO}$) is shown as Equation (2).

$$\Pr\left[T_{ij} \le RT\right] = \int_0^{RT} \frac{\lambda_j^i}{\Gamma(i)} t_{ij}^{i-1} e^{-\lambda_j t_{ij}} dt_{ij} \tag{1}$$

$$P_{s,ij,FIFO} = \Pr\left[T_{ij} \le RT\right]\left[p_{0,FIFO}\left(1-\frac{\lambda_j}{\lambda}\right)+\sum_{n=1}^{\infty} p_{n,FIFO}\left(1-\frac{\lambda_j}{\lambda}\right)\right] \tag{2}$$

where the probability of an arriving product type j job requiring a setup is $(1-\lambda_j/\lambda)$, $p_{0,FIFO}$ and $p_{n,FIFO}$ represent the probabilities that there are no jobs and there are $n \geqq 1$ jobs in the system under FIFO respectively.

According to the probability $P_{s,ij,FIFO}$, this arriving job faces two possible cases, in which setup occurs with the probability $P_{s,ij,FIFO}$ and no setup occurs with the probability $(1-P_{s,ij,FIFO})$. Thus, the expected number of setups for this job is $E[NS_{ij,FIFO}]=P_{s,ij,FIFO}$. If there would arrive n_j independent product type j jobs in the time interval (0, RT], then the expected number of setups for product type j is calculated as $E[NS_{j,FIFO}]=\sum_{i=1}^{n_j} E[NS_{ij,FIFO}]$, where j=1, 2, …, J.

Equation (3) shows the expected setup time under FIFO. Let s_{jr} be the setup time for the job of product type j, in which the previously processed job belongs to product type r. The first part indicates this job does not arrive in the time interval (0, RT], then the setup time should be zero with the probability $\Pr[T_{ij}>RT]$. The second part indicates this job arrived in the time interval (0, RT] and in the same product type j as previously processed job on machine, the setup time would be $s_{jj}=0$ and the probability would be $\Pr[T_{ij}\leqq RT](\lambda_j/\lambda)$. The third part indicates a different type r job arrived in the time interval (0, RT] and the setup time would be s_{jr}, then the probability would be $P_{s,ij,FIFO}(\lambda_r/\lambda^c)$, where $\lambda^c=\sum_{r=1,r\neq j}^{J}\lambda_r$.

$$
\begin{aligned}
E\left[S_{ij,FIFO}\right] &= \Pr\left[T_{ij} > RT\right]\times 0 + \Pr\left[T_{ij}\leq RT\right]\left(\lambda_j/\lambda\right)s_{jj} + P_{s,ij,FIFO}\sum_{\substack{r=1\\ r\neq j}}^{J}\frac{\lambda_r}{\lambda^c}s_{jr} \\
&= P_{s,ij,FIFO}\sum_{\substack{r=1\\ r\neq j}}^{J}\frac{\lambda_r}{\lambda^c}s_{jr}
\end{aligned}
\tag{3}
$$

Suppose that there are n_j independent product type j jobs arrive in the time interval (0, RT], then the expected mean setup time for product type j jobs arriving in the time interval (0, RT] is defined as $E[S_{j,FIFO}]=E[\sum_{i=1}^{n_j} S_{ij,FIFO}/n_j]$, where j=1, 2, …, J. Finally, the expected mean setup time of jobs arriving in the time interval (0, RT] is derived as $E[S_{FIFO}]=E[\sum_{j=1}^{J}\sum_{i=1}^{n_j} S_{ij,FIFO}/\sum_{j=1}^{J} n_j]$.

2.2 Development of the probabilistic model of setups under the FSR

Comparing with FIFO, the difference in the probability of setups under FSR occurring at the condition there are $n \geqq 1$ jobs in the system. Suppose that there are $n \geqq 1$ jobs in the system, a setup is needed given the condition if a job of type j arrives in the time interval (0, RT] encountering none job in queue belonging to type j then the probability would be $(1-\lambda_j/\lambda)^n$. Therefore, the probability that a setup is necessary for the i^{th} job of product type j under FSR is shown as Equation (4).

$$
P_{s,ij,FSR} = \Pr\left[T_{ij}\leq RT\right]\left[p_{0,FSR}\left(1-\frac{\lambda_j}{\lambda}\right)+\sum_{n=1}^{\infty}p_{n,FSR}\left(1-\frac{\lambda_j}{\lambda}\right)^n\right] \tag{4}
$$

where $p_{0,FSR}$ and $p_{n,FSR}$ are the probabilities that there are no job and $n \geqq 1$ jobs in the system under FSR respectively.

Then the expected number of setups for the i^{th} job of product type j is shown as $E[NS_{ij,FSR}]=P_{s,ij,FSR}$. Suppose that there would arrive n_j independent product type j jobs in the time interval (0, RT], which indicates the expected number of setups for product type j is computed as $E[NS_{j,FSR}]=\sum_{i=1}^{n_j}E[NS_{ij,FSR}]$, where j=1, 2, …, J.

Furthermore, the expected setup time under FSR is computed according to the same idea of Equation (3), to replace $P_{s,ij,FIFO}$ with $P_{s,ij,FSR}$, then the calculation of the expected setup time for the i^{th} job of product type j in the time interval (0, RT] under FSR is shown as $E[S_{ij,FSR}]=P_{s,ij,FSR}\sum_{r=1,r\neq j}^{J}(\lambda_r/\lambda^c)s_{jr}$. Then the expected mean setup time for product type j jobs arriving in the time interval (0, RT] is computed as $E[S_{j,FSR}]=E[\sum_{i=1}^{n_j}S_{ij,FSR}/n_j]$, where j=1, 2, …, J. The expected mean setup time of jobs arriving in the time interval (0, RT] is derived as $E[S_{FSR}]=E[\sum_{j=1}^{J}\sum_{i=1}^{n_j}S_{ij,FSR}/\sum_{j=1}^{J}n_j]$.

3 Accuracy analysis of the probabilistic model of setups under FIFO and FSR

Suppose that there is a sequence of jobs that has been dispatched by FIFO and FSR respectively. We are interest to the achieved magnitude of reduction in capacity utilization rate by replacing the FIFO with FSR. For single machine system, the machine utilization rates under FIFO and FSR are shown as $\rho_{FIFO}=\lambda E[ST_{FIFO}]$ and $\rho_{FSR}=\lambda E[ST_{FSR}]$, where $E[ST_{FIFO}]$ and $E[ST_{FSR}]$ are the expected mean service time of jobs under FIFO and FSR.

Suppose that the service time of jobs equals its processing time plus its setup time, and then the expected mean service time of jobs under FIFO is calculated according to Equation (3). First, the service time of the i^{th} product type j job would be zero with the probability $\Pr[T_{ij}>RT]$. Second, the service time of the i^{th} product type j job would be equal to its processing time with the probability $\Pr[T_{ij}\leq RT](\lambda_j/\lambda)$. Third, the service time of the i^{th} product type j job would be equal to its processing time plus its setup time with $P_{s,ij,FIFO}(\lambda_r/\lambda^c)$, where r= 1, 2, …, J and $r\neq j$. Let $ST_{ij,FIFO}$ be the random variable of the service time of the i^{th} job of product type j under FIFO. Then, the expected mean service time of jobs under FIFO is derived as $E[S_{FIFO}]=E[\sum_{j=1}^{J}\sum_{i=1}^{n_j}S_{ij,FIFO}/\sum_{j=1}^{J}n_j]$ and is given as Equation (5). To replace $E[S_{FIFO}]$ with $E[S_{FSR}]$, then the expected mean service time of jobs under FSR ($E[ST_{FSR}]$) is shown as Equation (6).

$$E\left[ST_{FIFO}\right]=\left(\sum_{j=1}^{J}n_j\right)^{-1}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\Pr\left[T_{ij}\leq RT\right]pt_j+E\left[S_{FIFO}\right] \quad (5)$$

$$E\left[ST_{FSR}\right]=\left(\sum_{j=1}^{J}n_j\right)^{-1}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\Pr\left[T_{ij}\leq RT\right]pt_j+E\left[S_{FSR}\right] \quad (6)$$

Substituting Equation (5) into Equation (6), then $E[ST_{FSR}]$ is rewritten as $E[ST_{FSR}]=E[ST_{FIFO}]-E[S_{FIFO}]+E[S_{FSR}]$.

Comparing the effects of the setup time and the machine utilization between FIFO and FSR

Because the probability $p_{0,FSR}$ and $p_{n,FSR}$ are equal approximately to $(1-\rho_{FSR})$ and $(1-\rho_{FSR})(\rho_{FSR})^n$. Then, $E[S_{FSR}]$ is reformulated as Equation (7). The part of Equation (7) is shown as Equation (8). It is seen that $(1+(\rho_{FSR}/(1-\rho_{FSR}))(\lambda_j/\lambda))\geqq 1$ with $0\leqq\rho_{FSR}<1$ and $\lambda_j>0$ for all j, which implies that $E[S_{FIFO}]\geqq E[S_{FSR}]$, where the setups is reduced by FSR if the probability of the number of jobs in queue occurs. It is noticed that $E[ST_{FSR}]=E[ST_{FIFO}]-E[S_{FIFO}]+E[S_{FSR}]\leqq E[ST_{FIFO}]$. Thus it is seen that $\rho_{FSR}=\lambda E[ST_{FSR}]\leqq\rho_{FIFO}=\lambda E[ST_{FIFO}]$, which implies that the machine utilization under FSR is small or equal to the machine utilization under FIFO.

$$\begin{aligned} E\left[S_{FSR}\right] &= \left(\sum_{j=1}^{J} n_j\right)^{-1} \sum_{j=1}^{J}\sum_{i=1}^{n_j} P_{s,ij,FIFO} \sum_{\substack{r=1\\ r\neq j}}^{J} \frac{\lambda_r}{\lambda^c} s_{jr} \left\{1-\rho_{FSR}\left[1-\left(1+\frac{\rho_{FSR}}{1-\rho_{FSR}}\frac{\lambda_j}{\lambda}\right)^{-1}\right]\right\} \\ &\leq \left(\sum_{j=1}^{J} n_j\right)^{-1} \sum_{j=1}^{J}\sum_{i=1}^{n_j} P_{s,ij,FIFO} \sum_{\substack{r=1\\ r\neq j}}^{J} \frac{\lambda_r}{\lambda^c} s_{jr} = E\left[S_{FIFO}\right] \end{aligned} \tag{7}$$

$$1+\frac{\rho_{FSR}}{1-\rho_{FSR}}\frac{\lambda_j}{\lambda} \begin{cases} =1, \text{ if } \rho_{FSR}=\sum_{n=1}^{\infty} p_{n,FSR}=0 \text{ with } \lambda_j>0,\ \forall j \\ >1, \text{ if } 0<\rho_{FSR}=\sum_{n=1}^{\infty} p_{n,FSR}<1 \text{ with } \lambda_j>0,\ \forall j \end{cases} \tag{8}$$

Relationship of the machine utilization rate between FIFO and FSR

The reduced range of the machine utilization rate by changing FIFO into the FSR is defined as $\Delta\rho=\rho_{FIFO}-\rho_{FSR}=\lambda E[ST_{FIFO}]-\lambda E[ST_{FSR}]$. Thus, the machine utilization by changing FIFO into FSR is defined as $\rho_{FIFO\to FSR}=\rho_{FIFO}-\Delta\rho$, where $\Delta\rho$ is obtained by solving the simultaneous equations $E[ST_{FSR}]=E[ST_{FIFO}]-E[S_{FIFO}]+E[S_{FSR}]$ and $E[ST_{FSR}]=\rho_{FSR}/\lambda$. Substituting $E[ST_{FSR}]=\rho_{FSR}/\lambda$ into the equation $E[ST_{FSR}]=E[ST_{FIFO}]-E[S_{FIFO}]+E[S_{FSR}]$, then this equation is rewritten as Equation (9).

$$\begin{aligned} f\left(\rho_{FSR}\right) &= E\left[ST_{FSR}\right]-\frac{\rho_{FSR}}{\lambda} \\ &= \left(\sum_{j=1}^{J} n_j\right)^{-1} \sum_{j=1}^{J}\sum_{i=1}^{n_j} \Pr\left[T_{ij}\leq RT\right] pt_j + E\left[S_{FSR}\right]-\frac{\rho_{FSR}}{\lambda} \end{aligned} \tag{9}$$

where $E[S_{FSR}]$ is given by Equation (7). Because $f(\rho_{FSR})$ is differentiable, the Newton's method is used to solve the nonlinear equation $f(\rho_{FSR})=0$.

According to $f(\rho_{FSR})$ and its derivative with respect to ρ_{FSR}, we begin with a first guess $0<\rho_{FSR}^0\leq 1$. A approximate solution ρ_{FSR}^1 can be obtained by calculating $\rho_{FSR}^0-f(\rho_{FSR}^0)/f'(\rho_{FSR}^0)$, in which should be an even better approximation to the solution of $f(\rho_{FSR})=0$. Once we have ρ_{FSR}^1, we can repeat the process to obtain ρ_{FSR}^2. After n steps, if we have an approximate solution ρ_{FSR}^n, then the next step is to calculate $\rho_{FSR}^{n+1} = \rho_{FSR}^{n} - f(\rho_{FSR}^{n}) / f'(\rho_{FSR}^{n})$. Notice that if the value for ρ_{FSR}^{n} become closer and closer to ρ_{FSR}^{n+1}. This means that we have found the approximate solution of $f(\rho_{FSR}) = 0$ after n steps.

Given a specific total arrival rate (λ^*) and the vector of job processing time (**PT**) in Figure 1, a slope and a intercept are given by $1/\lambda^*$ and zero for a line $E[ST] = \rho/\lambda^*$, where x-axis represents the machine utilization rate and y-axis represents the expected service time. Thus, the expected service time under FIFO is calculated by Equation (5) ($E[ST_{FIFO}]\big|_{\lambda^*,\mathbf{PT}}$) given λ^* and **PT**, and then the machine utilization rate under FIFO ($\rho_{FIFO}\big|_{\lambda^*,\mathbf{PT}}$) equals $\lambda^* E[ST_{FIFO}]\big|_{\lambda^*,\mathbf{PT}}$. Next, the expected service time under FSR ($E[ST_{FSR}]$) and the function $f(\rho_{FSR}) = 0$ for various ρ_{FSR} are depicted in Figure 1 according to Equation (5) and Equation (9). It is seen that the function $f(\rho_{FSR}) = 0$ is $E[ST_{FSR}]$ to shift down and the shift quantum is ρ_{FSR}/λ. Then, a root of $f(\rho_{FSR}) = 0$ (ρ_{FSR}^*) is found by using the Newton's method. Substituting ρ_{FSR}^* into $E[ST_{FSR}]$, thus $E[ST_{FSR}]\big|_{\lambda^*,\mathbf{PT},\rho_{FSR}^*}$ is obtained given $\rho_{FSR} = \rho_{FSR}^*$. Thus, the machine utilization by changing FIFO into FSR is given by Equation (10).

$$\rho_{FIFO\to FSR}^{*} = \rho_{FIFO}\Big|_{\lambda^*,\mathbf{PT}} - \Delta\rho\Big|_{\lambda^*,\mathbf{PT}} = \lambda^* E\big[ST_{FSR}\big]\Big|_{\lambda^*,\mathbf{PT},\rho_{FSR}^*} \tag{10}$$

Then, the total arrival rate is changed from λ^* to λ^{**} and with the same vector of job processing time (**PT**), where $\lambda^{**}<\lambda^*$. Thus, the slope of a line $E[ST] = \rho/\lambda^{**}$ is equal to $1/\lambda^{**}$, where $1/\lambda^{**} > 1/\lambda^*$. Repeating the above step, the machine utilization by changing FIFO into FSR is derived by Equation (11) given the condition that $\lambda = \lambda^{**}$.

$$\rho_{FIFO\to FSR}^{**} = \rho_{FIFO}\Big|_{\lambda^{**},\mathbf{PT}} - \Delta\rho\Big|_{\lambda^{**},\mathbf{PT}} = \lambda^{**} E\big[ST_{FSR}\big]\Big|_{\lambda^{**},\mathbf{PT},\rho_{FSR}^{**}} \tag{11}$$

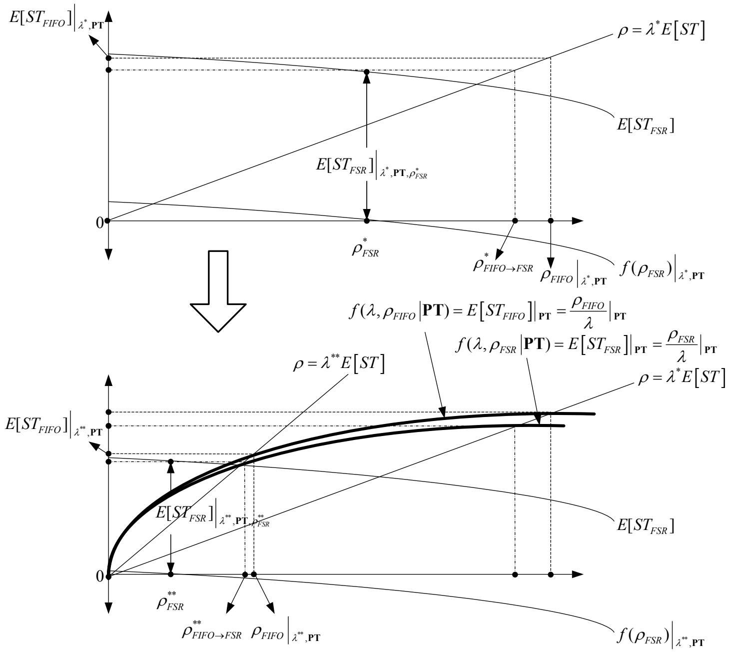


Figure 1. Relationship of the machine utilization rate between FIFO and FSR

Thus, the bold lines in Figure 1 are depicted, which are the relationships between the machine utilization rate and the expected service time for various total arrival rates under FIFO and FSR.

Moreover, $\Delta\rho$ is rewritten as Equation (12) according to Equation (7).

$$\begin{aligned}\Delta\rho &= \lambda\left(E\left[ST_{FIFO}\right]-E\left[ST_{FSR}\right]\right) = \lambda\left(E\left[S_{FIFO}\right]-E\left[S_{FSR}\right]\right)\\ &= \rho_{FSR}\left[1-\left(1+\frac{\rho_{FSR}}{1-\rho_{FSR}}\frac{\lambda_j}{\lambda}\right)^{-1}\right]\left[\lambda\left(\sum_{j=1}^{J}n_j\right)^{-1}\sum_{j=1}^{J}\sum_{i=1}^{n_j}P_{s,ij,FIFO}\sum_{\substack{r=1\\ r\neq j}}^{J}\frac{\lambda_r}{\lambda^c}s_{jr}\right]\end{aligned} \tag{12}$$

Thus, the derivative of $\Delta\rho$ with respect to ρ_{FSR} is written Equation (13).

$$\begin{aligned}\frac{d\Delta\rho}{d\rho_{FSR}} &= \left[\lambda\left(\sum_{j=1}^{J}n_j\right)^{-1}\sum_{j=1}^{J}\sum_{i=1}^{n_j}P_{s,ij,FIFO}\sum_{\substack{r=1\\ r\neq j}}^{J}\frac{\lambda_r}{\lambda^c}s_{jr}\right]\times\left[1-\left(1+\frac{\rho_{FSR}}{1-\rho_{FSR}}\frac{\lambda_j}{\lambda}\right)^{-1}+\right.\\ &\quad\left.\frac{\lambda_j}{\lambda}\frac{\rho_{FSR}}{\left(1-\rho_{FSR}\right)^2}\left(1+\frac{\rho_{FSR}}{1-\rho_{FSR}}\frac{\lambda_j}{\lambda}\right)^{-2}\right]\geq 0\end{aligned} \tag{13}$$

where $d\Delta\rho/d\rho_{FSR} \geqq 0$ with $0 \leqq \rho_{FSR} < 1$ and $\lambda_j > 0$. Hence, if $\rho_{FSR1} \geqq \rho_{FSR2}$, then $\Delta\rho(\rho_{FSR1}) \geqq \Delta\rho(\rho_{FSR2})$, where $0 \leqq \rho_{FSR1} < 1$ and $0 \leqq \rho_{FSR2} < 1$. By replacing FIFO with FSR, the reduced range of the machine utilization rate is increasing when the utilization rate of machine is rasing, which implies that the machine utilization rate is saved by FSR especially at the condition of high workload on machine.

3.1 Experimental design

To evaluate the accuracy of the probabilistic model on calculating expected setups and setup time, a simulation model is built to simulate the inter-arrival time of jobs with exponential distribution, in which the next job selected from queue to be processed on machine follows FIFO or FSR rule. If the selected next job requires changing machine setup, the sequence-dependent setups is included.

The simulation is implemented according to the experimental design by varying three control parameters: run time (*RT*), total arrival rate (λ), and the coefficient of variation of the jobs arrival rate (*CV*). Three levels of run time (*RT*) are considered: 8, 16, and 24 hours. The vector of job arrival rate among eight product types is shown as $\boldsymbol{\lambda}=[\lambda_1\ \lambda_2\ \lambda_3\ \lambda_4\ \lambda_5\ \lambda_6\ \lambda_7\ \lambda_8]$, and then six levels of the total arrival rate (λ) are considered as $a\boldsymbol{\lambda}$ ($\lambda = a\sum_{j=1}^{8}\lambda_j$) jobs in 60 seconds, where a = 1.00, 0.95, 0.90, 0.85, 0.80, and 0.75.

The coefficient of variation of the jobs arrival rate is considered, where the coefficient of variation is reported as a percentage and calculated from the mean and the standard deviation of the arrival rate of jobs. Table 1 shows three vectors of the arrival rate among eight product types and their corresponding *CV*, in which the total arrival rate of these three vectors are all equal to 0.01. Thus, the coefficient of variation of these three vectors of the arrival rate is calculated as 0, 27.9753, and 53.7234, which implies that the dispersion of the jobs arrival rate is greater when the coefficient of variation is increasing.

Table 1. Three vectors of arrival rate among eight product types and their corresponding *CV*

Arrival rate vector	Product type								*CV* (%)
	1	2	3	4	5	6	7	8	
1	0.001250	0.001250	0.001250	0.001250	0.001250	0.001250	0.001250	0.001250	0
2	0.001342	0.001577	0.001804	0.000917	0.001145	0.001443	0.000817	0.000955	27.9753
3	0.001821	0.001255	0.000297	0.001721	0.000946	0.000363	0.001467	0.002130	53.7234

For each combination, the simulation model contains the vector of job processing time among eight product types, **PT** = [15 75 85 45 55 10 80 125], and the matrix of sequence-dependent setup time (**ST**), shown below, for switching the setting of machine in changing product types. “Second” is the unit of processing time and setup time. Note that the simulation results are collected for each combination in 8 hours after 10,000 independent simulation runs.

$$\mathbf{ST} = \begin{bmatrix} 0 & 15 & 30 & 45 & 60 & 45 & 60 & 15 \\ 90 & 0 & 60 & 75 & 75 & 30 & 45 & 30 \\ 60 & 75 & 0 & 90 & 45 & 30 & 60 & 15 \\ 15 & 30 & 45 & 0 & 45 & 30 & 15 & 30 \\ 15 & 45 & 90 & 45 & 0 & 75 & 45 & 60 \\ 30 & 75 & 90 & 30 & 45 & 0 & 15 & 30 \\ 45 & 90 & 75 & 60 & 75 & 60 & 0 & 45 \\ 30 & 45 & 60 & 45 & 15 & 75 & 45 & 0 \end{bmatrix}$$

3.2 Sensitivity analysis of the probabilistic model

Figure 2 shows the expected mean setup time per job ($E[S_{FIFO}]$ and $E[S_{FSR}]$) for varying the *CV* of the job arrival rate, the total arrival rate, and the run time under FIFO and FSR in the probabilistic model. Figure 3 shows the mean setup time per job by the simulation model. Comparing with the data in Figure 2 and Figure 3, the mean setup time per customer by the simulation model seems to be a bound, $E[S_{FIFO}]$ and $E[S_{FSR}]$ are close to this bound as the total arrival rate and the run time increase. Moreover, in Figure 2 and Figure 3, for CV=53.7234%, the setup time under FIFO and FSR by the probabilistic model and the simulation model have the lowest level of value. As the increasing of the coefficient of variation of job arrival rates, the arrival of job tends to concentrate on fewer product types, which are the types having possible reduction of setups.

3.3 Accuracy of the probabilistic model

To test whether there is significant differences of setup time between FIFO and FSR (△ST), the statistical paired *t*-test is used. This test is to validate or falsify the null hypothesis that the difference of the setup time between FIFO and FSR is small or equal to zero and the alternative hypothesis that this difference is greater than zero. Table 2 presents the results of the paired *t*-test. With the P-value list in Table 2, we conclude that the difference of the setup time between FIFO and FSR is larger than zero, which implies that FSR can be used to save the setup time.

Figure 4 plots the mean absolute error of expected setups under FIFO and FSR, in comparison with the data from simulation model, by varying the run time (*RT*), the coefficient of variation (*CV)* of the arrival rate of jobs, and the total arrival rate. The mean absolute error of setups under FIFO is smaller than that of FSR, which implies that the probabilistic model under FIFO can calculate the setups more accurately than that under FSR. The overall mean absolute errors under FIFO and FSR are 3.3632 times and 5.0404 times respectively. Moreover, for *CV* equal to 27.9753% having lowest absolute error of expected setups. When the total arrival rate is increasing, the absolute error increases correspondingly, that is, the larger absolute error occurring at higher level of the utilization rate of machine.

Figure 5 displays the mean absolute error percentage of expected setup time under FIFO and FSR, in comparison with the data from simulation model, by varying the run time (*RT*), the coefficient of variation (*CV*) of the arrival rate of jobs, and the total arrival rate. The mean absolute error percentage of setup time for the model under FIFO is also smaller than that under FSR. The overall mean error percentage of setup time under FIFO is 6.4080% and that under FSR is 9.2299%. Comparing with FSR, the probabilistic model under FIFO can calculate the setup time more accurately. Furthermore, when the *CV* equals 27.9753%, there has the lowest absolute error percentage of expected setup time. While the total arrival rate is increasing, the absolute error percentage increases respectively, that is, the larger absolute error percentage occurring at higher level of the machine utilization rate.

Figure 6 shows the differences of the number of setups and the setup time by varying the *CV* of the job arrival rate, the total arrival rate and run time under FIFO and FSR by the probabilistic model and the simulation model. For the probabilistic model, the reduced ranges of the number of setups and the setup time, by replacing FIFO with FSR, increases as the raising of the total arrival rate, which are shown in Figure 6(b) and 6(d). However, for the simulation model, the reduced ranges of the number of setups and the setup time, by replacing FIFO with FSR, increases firstly and then decreases as the raising of the total arrival rate, which are shown in Figure 6(a) and 6(c). Furthermore, we can see that the reduced rage of the setup time by replacing FIFO with FSR has larger magnitude at *CV*=53.7234%. This means that the reduced rage of the setup time due to the replacing is increasing as the enlarging of the coefficient of variation of the job arrival rates.

4 Conclusions

This paper provides the probabilistic models to calculate the number of setups and the sequence-dependent setup time for a single-machine system facing uncertain job arrivals, that is, the inter-arrival time of jobs being exponentially distributed. Since the uncertainties of the time and the type of customer demand make the machine setup decision very complicated, the development of analytic models to calculate the expected sequence-dependent setup time under FIFO and FSR rules contributes to the quick evaluation of capacity waste which is due to the changing of machine setting among several product types.

The computational results show that the overall mean absolute errors of setups under FIFO and FSR are 3.3632 times and 5.0404 times respectively, and the overall mean absolute error percentages of setup time under FIFO and FSR are 6.4080% and 9.2299% respectively, which both are related to the coefficient of variation (*CV*) of the arrival rate of jobs and the total arrival rate. In general, the probabilistic models under FIFO rule can calculate the number of setups and the sequence-dependent setup time more accurately than those under FSR. However, comparing with FIFO, FSR can reduce setups and then leads to the saving of utilization rate of machine, especially at the condition of higher total arrival rate. Therefore, our models can, in some extent, accurately calculate the total setup time and efficiently evaluate the capacity waste coming from changing machine setting to respond to the uncertainties of job arrivals.

Table 2. The results of the paired *t*-test

	Null hypothesis	Alternative hypothesis	Sample size	Mean	Std. deviation	Std. error	Test statistic	P-value
Probabilistic model	△ST≦0	△ST＞0	432	1.9503	0.7782	0.0374	52.1471	<0.0001
Simulation model	△ST≦0	△ST＞0	432	2.9413	1.2898	0.0621	47.3639	<0.0001

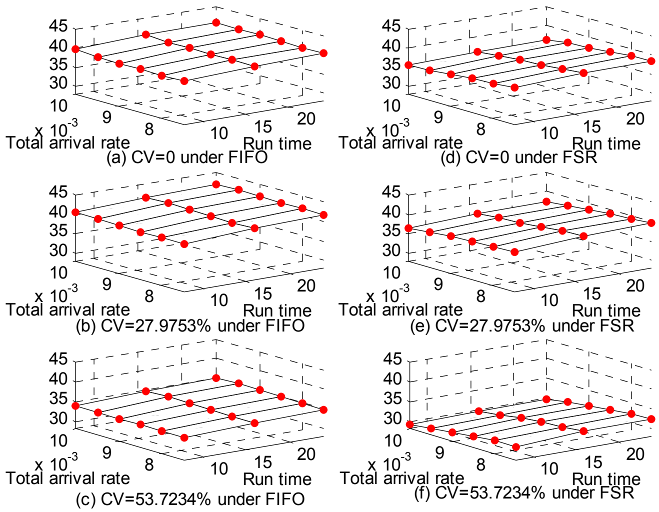


Figure 2. Expected mean setup times of jobs for various the *CV* of the job arrival rate, the total arrival rate, and the run time under FIFO and FSR in the probabilistic model

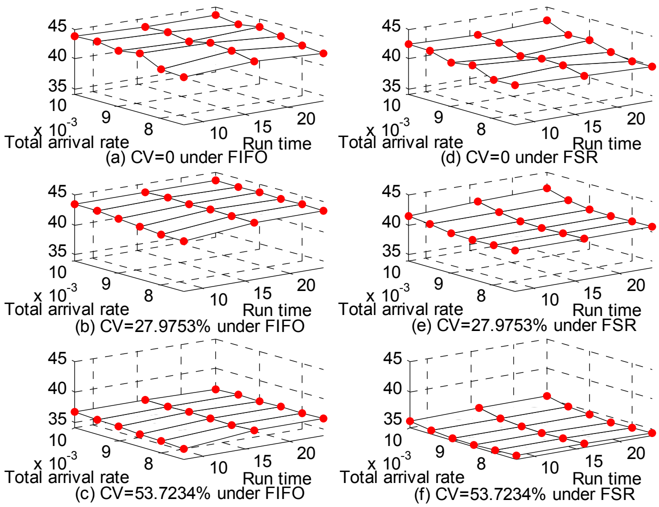


Figure 3. Mean setup times per job for various the *CV* of the job arrival rate, the total arrival rate, and the run time under FIFO and FSR in the simulation model

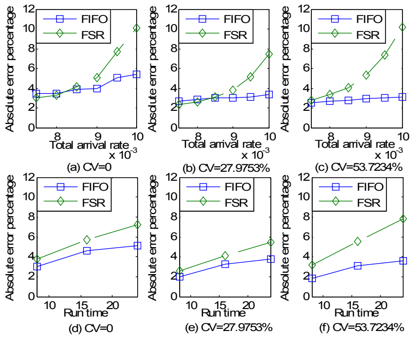


Figure 4. Mean absolute error of setups under FIFO and FSR by varying the *CV* of the job arrival rate, the total arrival rate, and the run time

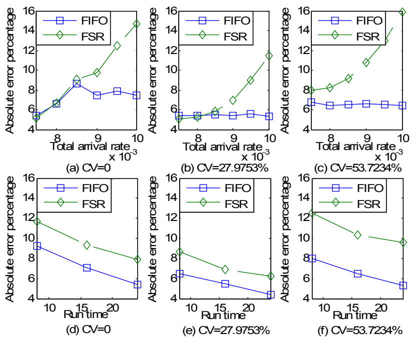


Figure 5. Mean absolute error percentage of setup time under FIFO and FSR by varying the *CV* of the job arrival rate, the total arrival rate, and the run time

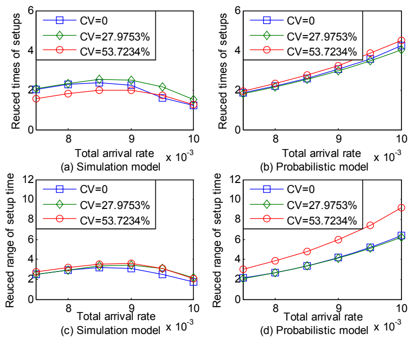


Figure 6. Differences of the number of setups and the setup time by varying the *CV* of the job arrival rate, the total arrival rate, and the run time under FIFO and FSR by the probabilistic model and the simulation model

5 Acknowledgments

This paper was supported by the National Science Council, Taiwan, ROC, contract NSC 96-2221-E-239 -004.

6 References

[1] Jensen JB, Malhotra MK. Philipoom PR, Family-based scheduling of shops with functional layouts. International Journal of Production Research 1998;36:2687-2700.
[2] Kannan VR, Lyman SB. Impact of family-based scheduling on transfer batches in a job shop manufacturing cell. International Journal of Production Research 1994;32:2777-2794.
[3] Nomden G, Van Der Zee DJ, Slomp J. Family-based dispatching: anticipacting future jobs. International Journal of Production Research 2008;46:73-97.
[4] Allahverdi A, Gupta JND, Aldowaisan T. A review of scheduling research involving setup considerations. Omega 1999;27:219-239.
[5] Chern CC, Liu YL. Family-based scheduling rules of a sequence-dependent wafer fabrication system. IEEE Transactions on Semiconductor Manufacturing 2003;16:15-25.
[6] Vieira GME, Herrmann JW, Lin E. Analytical models to predict the performance of a single machine system under periodic and event-driven rescheduling strategies. International Journal of Production Research 2000;38:1899-1915.
[7] Vieira GME, Herrmann JW, Lin E. Predicting the performance of rescheduling strategies for parallel machine systems. Journal of Manufacturing Systems 2000;19:256-266.
[8] Missbauer H. Order release and sequence-dependent setup times. International Journal of Production Economics 1997;49:131-143.

A Two-Level Genetic Algorithm for Scheduling in Assembly Islands with Fixed-Position Layouts

Wei Qin[a,1] and George Q. Huang[b]

[a] Department of Industrial and Manufacturing Systems Engineering, the University of Hong Kong.

Abstract. This paper focuses on the scheduling problem in assembly islands environment with fixed layouts. A fixed-position assembly line is always used when products (e.g., ships and planes) are too fragile, large or heavy to move. In such configuration, products normally remain in one location for its entire manufacturing period while machines, materials and workers are moved to an assembly site called an assembly island. Such layouts can afford necessary flexibility and competitive operational efficiency for products of modest variety and production volumes. However, the high dynamics of material, equipment and manpower flows in assembly islands make the production scheduling quite difficult. The authors give the definition and mathematical model for the scheduling problem. A two-level genetic algorithm is used to obtain a near optimal solution to minimize the makespan. Experimental results show that this algorithm is more effective in airline or ship industrial manufactures than in other machine or tool final assembly companies. It also can be found that some function of the number of jobs and the number of islands is the most important factor to the time of scheduling.

Keywords. Scheduling, two-level GA, assembly islands.

1 Introduction

Fixed-position layouts are normally used in One-of-a-Kind Production (OKP) industry, when products are too fragile, large, bulky, or heavy to move, such as in some aircraft and ship manufacturing factories, container manufacturing companies and etc (Trostmann et al. [1]). Among 80,000+ Hong Kong manufacturers in Pearl River District (PRD), many lifts and moulds manufacturers are also of this type. In such configurations, machines, material, and workers are moved to an assembly site (often called an assembly island) while products normally remain in one location for its entire manufacturing (assembly) period. Advantages of fixed-position layout include reduced movement of work items; minimized damage or cost of movement, and more continuity of the assigned work force since the item does not go from one place to another (Huang et al. [2]).

[1] Corresponding Author E-mail: qoolooq@yahoo.com

The authors in this paper focus on this scheduling problem in assembly islands. It’s a new form of assembly line and the mechanism is quite different from conventional job shop or flow shop problem which make the scheduling problem in such environment much more complex than in common assembly lines. Here are some distinct characteristics as follows.

Firstly, the islands are always far away from each other, so the affection of distance is quite important here. The time spent on moving and delivering can not be omitted as in conventional assembly lines. When calculating the setup times between jobs and operations in assembly islands environment, we should take account of the distance affection.

Secondly, in assembly islands layouts, the schedule is more complex than in common job shop or flow shop problems. The schedule here should contain a part that which island the job is to be operated on. And there is also a buffer restriction in assembly islands, which makes the scheduling problem like a blocking or no-buffer scheduling question with some differences.

The remainder of this paper is organized as follows. In section 2, we present a brief literature review of scheduling in assembly islands. In section 3, we give the description and definition of the problem. Section 4 discusses the proposed genetic algorithm scheduling problem. In section 5, we report experiments and computational results. And finally is the summary.

2 Literature Review

The scheduling problem considered in this paper can be described against the literature along following directions.

2.1 Research of Assembly Islands

A typical conventional assembly line has one (or more) stationary worker who stays at a given workstation to perform an assembly operation and, when this operation is accomplished, the part assembled is transferred to another worker at the next workstation for the next assembly operation.

Linear walking-worker assembly lines are a novel form of flexible assembly system where fitters travel along the line carrying out each assembly task at each workstation. They attempt to combine the flexibility of the workbench system with the efficiency of the conventional fixed-worker assembly line.

Wang, Owen and Mileham [3] investigated a local manufacturing company in Bath, UK, where the walking-worker assembly line has been implemented. They compared the system's performance based on the same production line operated with fixed workers or with walking workers. The results have shown that by using multiple-skilled walking workers this novel system has several advantages over the conventional fixed-worker assembly line under similar conditions.

Unfortunately, quite limited reference can be found for modeling and solving the scheduling problem in the configuration of assembly islands.

2.2 Scheduling with Setup Times

In the classical scheduling problem, the job's setup times are either included in the processing times or ignored. However, in many real-life industrial problems, the setup is not a part of processing and the required time is sequence-dependent. As surveyed in Cheng *et al*. [5], operations of a job can be presented in three phases: separable setup, processing and separable removal times. Furthermore, the flow shop problems with setup time can be put into four categories according to the sequence dependence and family setup time. Considering the parallel machines with sequence-dependent setup time, Lee and Pinedo [6] proposed a three phase heuristic for minimizing the sum tardiness. A specific industrial case of parallel machines has been studied by Luh *et al*. [7]. A near-optimal solution was obtained through an integrated solution methodology based on a combined Lagrangian relaxation, dynamic programming and heuristics.

Based on the above brief literature review, we could notice that setuptimes have been well investigated in job shop and flow shop problems, but no references can be found in assembly islands scheduling problem.

2.3 Summary

Based on the above brief literature review, several observations can be made. Firstly, on the aspect of assembly islands, quite limited reference can be found. And most of them have not yet come to the scheduling problem. Secondly, in terms of scheduling with setup times and no-buffer, there are some references. However, little of them take account of the affection of distance which is important in assembly islands configuration and jobs sequences into assembly islands.

3 Problem Definition and Formulation

3.1 Problem Definition

We start with a formal definition for the scheduling problem. In a typical assembly islands environment with fixed-position layout, there are k identical assembly sites $I_1, I_2, ..., I_k$, m types of assembly operators $A_1, A_2, ..., A_m$, l logistics workers $L_1, L_2, ..., L_l$ and n jobs $J_1, J_2, ..., J_n$. Each type of operators may have several alternates which could be a group of workers or a single operator. For example, type A_j has m_j alternatives $\left\{A_{1,j}, A_{2,j}, ..., A_{m_j,j}\right\}$. But here for simplicity, we assume that there is only one for each type of operators, e.g., $m_j = 1$ for any A_j , $j = 1, 2, ..., m$. A job $J_j, j = 1, 2, ..., n,$ consists of a chain of sets of sequential operations. For example, a job J_j contains n_j operations

$\left\{O_{1,j}, O_{2,j}, \ldots, O_{n_j,j}\right\}$, which means it should be processed by operators $A_1, A_2, \ldots, A_{n_j}$ sequentially. The operator of type A_i finished the operation $O_{i,j}$ of job J_j in assembly island $I(j)$ and then go to assembly island $I(k)$ to process the operation $O_{i,k}$ of job J_k.

Each job J_j has its specified characteristics such as its operations, due date, weight and so on. And usually the total order size is relatively small. In general case ($n > k$), when the order gets started, the first batch of jobs are sent to the assembly islands $I_1, I_2, \ldots, I_k$ and the operators, machines and materials according to the requirement of the first operation of these jobs are also moved to the corresponding assembly island. Here we don't take account of the case of recirculation, that is, one type of operators never processes the same job more than one time.

Each operation $O_{i,j}$ has the processing time $p_{i,j}$. The release time of operation $O_{i,j}$ is denoted by $r_{i,j}$, the start time is denoted by $s_{i,j}$ and the completion time is denoted by $c_{i,j}$. We have precedence constraints of the form $O_{i,j} \rightarrow O_{i+1,j}$. Each job can only be processed only by one machine at a time and each machine can only process one job at a time. We denote the ready (release) time of job J_i by R_i, the start time of job J_j by S_j and the completion time of job J_j by C_j.

When job J_i is finished, Job J_j is moved in the same assembly island with the setup time ST_{ij}. The setup time between operations $O_{i,j}$ and $O_{k,j}$ is denoted by st_{ikj}.

Here to establish a reference frame, the inventory warehouse is set as the origin of coordinate. The assembly islands are set a line with coordinates denoted by $CI_1, CI_2, \ldots, CI_k$. The coordinates of assembly operators are denoted by $CA_1, CA_2, \ldots, CA_m$. The coordinates of logistic operators are denoted by $CL_1, CL_2, \ldots, CL_l$.

3.2 Problem Formulation

In our analysis we assume that all jobs are simultaneously available at time zero and that all processing times are positive. On each assembly island, only one job can be processed at a time, which means once a job starts, it cannot be moved out from the island until all its operations are completed. And because of the space restriction, there is no buffer in each island. We also assume that there is no preemption, i.e., once started, an operation cannot be interrupted before completion.

We also require the assembly operators to move to the next operation as soon as they finished the previous one. Suppose the assembly operator A_k has just finished the operation $O_{k,q}$, and we have the equation $t_A(k,j) = c_{k,q}$. Then we can determine that

$$s_{k,j} = \max\left\{ c_{i,j} + st_{ikj}, \frac{CI_{I(q)} - CI_{I(j)}}{\upsilon_A} + c_{k,q} \right\}$$

Another important assumption is that logistic operators are enough. In this manufacturing company, the assembly working operation takes much more time than the delivering operation does. And also because the logistic operators are paid much less than assembly operators, the company can hire relatively more workers to do the delivering jobs. This means that the logistic operators can always finish the logistic task in time to match the completion time of the previous job or operation.

Actually this is one case of 4 different situations shown in Figure 1 (Qin, Huang and Chen [4]).

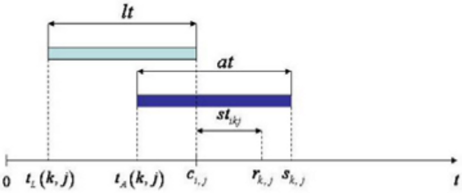


Figure 1. The logistic operators are enough.

Here the assumption indicates that

$$lt\left(CL_{L(k,j)}\left(t_L(k,j)\right), CI_{I(j)}\right) + \delta_L(k,j) = 0$$

So we get a simple equation

$$r_{k,j} = c_{i,j} + st_{ikj}$$

We consider the criterion for optimality is the makespan $C_{\max}$.

Minimize $C_{\max} = \max\{C_1, C_2, ..., C_n\}$ (1)

Subject to

$s_{k,j} - s_{i,j} \geq p_{i,j}$ for all $O_{i,j} \rightarrow O_{k,j}$

$C_{\max} - s_{i,j} \geq p_{i,j}$ for all $O_{i,j}$

$s_{i,j} - s_{i,l} \geq p_{i,j}$ or $s_{i,l} - s_{i,j} \geq p_{i,j}$ for all $O_{i,j}$, $O_{i,l}$

$s_{i,j} \geq 0$ for all $O_{i,j}$

In this formulation the first set of constraints ensures that operation $O_{k,j}$ cannot start before operation $O_{i,j}$ is completed. The third set of constraints ensures that some ordering exists among operations of different jobs that have to be processed by the same assembly operator.

4 Proposed Two-Level Genetic Algorithm

Genetic algorithms (GA) are motivated by the theory of evolution. They date back to the early work described in Rechenberg [8], Holland [9] and Schwefel [10], also Goldberg [11] and Michalewicz [12]. They have been designed as general search strategies and optimization methods working on populations of feasible solutions.

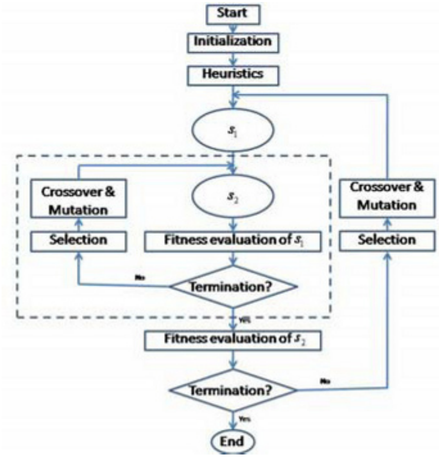


Figure 2. The flowchart of proposed GA.

In this study, Figure 2 shows the flowchart of the proposed GA approach to the scheduling problem in assembly islands. It includes two sub-problems: jobs sequences optimization and operations sequences optimization. To improve the effectiveness of the algorithm, the GA for operations sequences optimization is not executed until the optimal solution for the corresponding jobs sequence is obtained, but only executed for some generations. So in this GA algorithm, a max generation of operations sequences is pre-defined.

4.1 Solution Space

In a general scheduling problem, a schedule should include all jobs. But in our research, the situation is somewhat different. Let S be a schedule for this

$A_{nkm} \,||\, C_{\max}$ problem. Because of the restriction of assembly islands, the schedule S should contain three parts denoted by S_1 and S_2.

S_1 is the permutation of jobs moved into the assembly islands which has the form:

$$\left\{\left(J_{I_1,1},J_{I_1,2},\ldots\right),\left(J_{I_2,1},J_{I_2,2},\ldots\right),\ldots,\left(J_{I_k,1},J_{I_k,2},\ldots\right)\right\}$$

S_2 is the arrangement of operations being processed by corresponding operators, which has the form:

$$\left\{\left(O_{A_1,1},O_{A_1,2},\ldots\right),\left(O_{A_2,1},O_{A_2,2},\ldots\right),\ldots,\left(O_{A_m,1},O_{A_m,2},\ldots\right)\right\}$$

Actually, here $\left(J_{I_i,1},J_{I_i,2},\ldots\right)$, $i=1,2,\ldots,k$ is the list of jobs to be sequentially operated on assembly island I_i. $\left(O_{A_j,1},O_{A_j,2},\ldots\right)$, $j=1,2,\ldots,m$ is the list of operations to be processed by assembly operator A_j. Encoding and decoding of this solution space would be introduced in section 4.2.

4.2 Encoding and Decoding

The design of chromosomal representation to encode the solutions has significant impact on the efficiency of algorithm (Ho *et al.* [13]). Actually, encoding is the most important part and also bottleneck in GA.

Here in this problem, normally, there are n jobs waiting for processing at time zero and only k assembly islands ($k<n$). So there are at most only k jobs can be processed simultaneously. We cannot use only one chromosome to describe the solution.

For an n-job k-island and m-operator scheduling problem, as mentioned in above section, we propose to use two chromosomes s_1 and s_2. s_1 tells the information about sequences of jobs and which assembly island they are processed on. s_2 indicates the sequences of operations. As for s_1, we adopt preference list-based representation to encode the part solution S_1. A chromosome s_1 is formed of k subchromosomes, each for one island. Each subchromosome is a string of symbols and each symbol identifies a job that has to be processed on relevant island. Once s_1 is confirmed, this problem becomes much the same as a classical job-shop scheduling problem with the differences of setup times and recirculation. We use operation-based representation to encode the part solution S_2. We named all operations for a job with the same symbol and then interpreted it according to the order of occurrence in the given chromosome. It is easy to see that any permutation of the chromosome s_2 is feasible.

Table 1. Summary of a Simple Example

Jobs	Operations Sequence	Processing times
J_1	1, 2, 3	4, 2, 2
J_2	2, 3, 1	6, 5, 6
J_3	1, 3, 2	1, 2, 7
J_4	3, 1, 2	8, 3, 5

Consider the 4-job 2-island and 3-operator scheduling problem shown in Table 1. Suppose a chromosome s_1 is given as [(1 3) (2 4)]. This indicates that J_1 and J_3 are processed on assembly island I_1 sequentially, while J_2 and J_4 are processed on assembly island I_2 sequentially. Now the problem is simplified to a 2-job 3-machine job-shop scheduling problem shown in Figure 3. We combine J_1 and J_3, J_2 and J_4, with J_{I_1} and J_{I_2} to denote them separately.

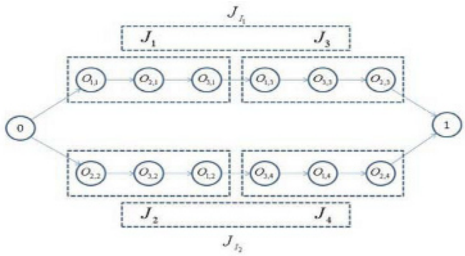


Figure 3. The disjunctive graph.

Now suppose a chromosome s_2 has the form [2 1 1 2 2 2 1 1 1 2 1 2]. Because each job has three operations, it occurs exactly six times in the chromosome. Each gene uniquely indicates an operation and can be determined according to its order of occurrence in the sequence. Let $O_{i,j}$ denote the ith operation of job J_j. The chromosome can be translated into a unique list of ordered operations of [$O_{2,2}$ $O_{1,1}$ $O_{2,1}$ $O_{3,2}$ $O_{1,2}$ $O_{3,4}$ $O_{3,1}$ $O_{1,3}$ $O_{3,3}$ $O_{1,4}$ $O_{2,3}$ $O_{2,4}$]. Operation $O_{2,2}$ has the highest priority and is scheduled first, then $O_{1,1}$, and so on. The resulting active schedule is shown in Figure 4 and Figure 5.

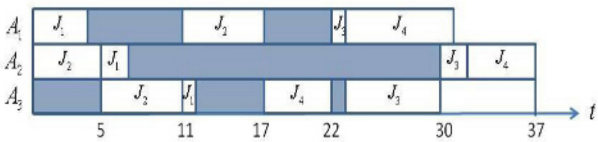


Figure 4. The active schedule.

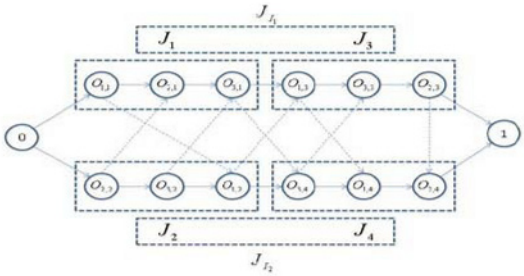


Figure 5. The disjunctive graph with the active schedule.

4.3 Fitness

The fitness is used to measure the performance of each individual. It depends on the objective. In this study, the fitness can be presented as the deviation of individual's makespan from the maximal one among all individuals.

$$fitness(x) = \max_{x \in pop}(C_{\max}(x)) + 1 - C_{\max}(x)$$

4.4 Selection

The aim of selection is to keep good individuals and eliminate bad ones. This selection is based on the fitness value of individuals. The fittest individual of current population will be chose by using the elitism selection. And then, the roulette wheel method will be used to select rest $pop-1$ individuals. All of the pop selected individuals will form a mating pool for the crossover and mutation. We perform the roulette wheel selection in this paper.

4.5 Crossover

In general, the crossover operator is regarded as a main genetic operator and the performance of the genetic algorithms depends, to a great extent, on the performance of the crossover operator used. Conceptually, the crossover operates on two chromosomes at a time and generates offspring by combining features of both chromosomes.

In this study, we adopt the linear order crossover (LOX) operator. The crossover operator can preserve both the relative positions between genes and the absolute positions relative to the extremities of parents as much as possible. The extremities correspond to the high and low priority operations.

4.6 Mutation

It is relatively easy to make some mutation operators for permutation representation. Here we adopt shift mutation. We first choose a sublist (a job) randomly and then shifts it to a random position of right or left from the sublist's position.

5 Computational Results

This section presents computational experiments using the above algorithm. Three research questions will be discussed: (1) How does the proposed algorithm perform when solving problems of different case? (2) How do some factors such as the number of assembly islands influence the algorithm performance, (3) What are the optimal evolutionary parameters for problems of different case and complexity? In this paper, the first two questions are successfully answered and the last one will be the future work. In the following sub-section, some test-problems are firstly presented, based on which all of the computational experiments are conducted.

5.1 Test-Problems and Experimental Results

Since in assumption the logistic operators are enough, so the solution space of scheduling problem in assembly islands grows only in proportion to (1) the number of jobs n, (2) the number of operations m (equal to operators here), (3) the number of assembly islands k. The first and the third factors determine the search space of jobs sequences. The second and the third factors determine the search space of operations sequences. The test-problems should reflect the independent influence of each factor.

There are two main types of industrial manufacturers using fixed-position assembly islands. One is the airline and ship assembly, in which there are few assembly islands but quite a lot of complex operations, i.e. $m >> k$. Another type is the machine and tool final assembly, in which there are many fixed-position assembly islands but the operations of each job are quite easy, i.e. $k >> m$. In both these two cases, n is always larger than k, but in the second case the quantity of ordered products, i.e. the number of jobs n is always very large. The test-problems will discuss these two typical and different cases.

The proposed genetic algorithm has been implemented in JAVA on a Pentium IV, 2.8GHz and 512RAM. For the first proposed question, the test-problems and results are shown in Table 2. For the second question, the test-problems and results are introduced in Table 3.

One thing need to be noticed is that during the experiments, the operations and their sequence of each job are randomly chosen first.

Table 2. Algorithm performance in two cases

	Case 1						Case 2					
	n	k	m	n	k	m	n	k	m	n	k	m
	10	**4**	**16**	**10**	**4**	**20**	**40**	**4**	**1**	**80**	**8**	**2**
CPU time (s)	**32**			**43**			**463**			**2054**		

Table 3. Influence of different factors to algorithm performance

	m =4, k =4					n =20, m =4					n =20, k =4				
n / k / m	4	8	12	16	20	1	2	4	8	16	1	2	4	8	16
CPU time	5	10	26	81	387	13	182	387	365	105	106	299	387	606	1532

5.2 Algorithm Performance Analysis

First in Table 2, the computational results show that the performance of algorithm effectiveness is better in case 1 than in case 2. The reason can be explained from Table 3. Figure 6, 7 and 8 tell that the relationships between the algorithm effectiveness and the three factors.

When the number of jobs n and the number of operations m increase, the used time increases sharply. But if these two parameters are fixed, the used time of GA is not an increasing function of the number of islands k . As mentioned above, n and k together determine the search space of solution S_1 . Although with the increase of k , the search space of solution S_2 also increases, but the most influenced factors are n and k . When they reach a certain relation, there is a maximum of used time. This is why in case 1, the performance of proposed GA is much better. Because in case 2, the number of jobs n is quite large, which makes the search space of solution S_1 grows fast, although the limited number of operations decrease the search space of solution S_2 .

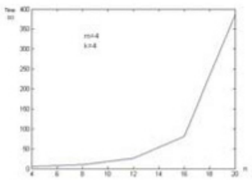


Figure 6. The relation between algorithm effectiveness and the number of jobs n.

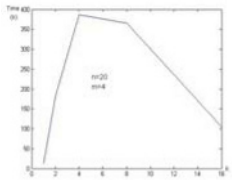


Figure 7. The relation between algorithm effectiveness and the number of islands k.

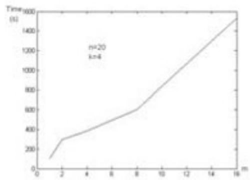


Figure 8. The relation between algorithm effectiveness and the numbers of operations m.

6 Conclusion and Future Work

This paper proposes a genetic algorithm for the scheduling problem in assembly islands with fixed-position layouts. The algorithm consists of two evolutionary processes of jobs sequences and operations sequences. They are interwoven with each other to improve the effectiveness. Experimental results show that this algorithm is more effective in airline or ship industrial manufactures than in other machine or tool final assembly companies. It also can be found that some function of the number of jobs n and the number of islands k is the most important factor to the time of scheduling. When this function gets a certain value, the used time of algorithm will reach a maximum. The future work is to do more computationary experiments to determine this function and to find the maximum point. Besides, to take account less logistic operators into modeling is also another area for further investigation.

7 References

[1] E. Trostmann, F. Conrad, H. Holm and O. Madsen, "Cybernetic modeling and control in integrated production systems – A project review", Proceedings of the Eighth IPS Research Seminar, Denmark, pp. 213-225, 22-24 March 1993.
[2] G. Q. Huang, Y. F. Zhang, and P. Y. Jiang, "RFID-based wireless manufacturing for walking-worker assembly islands with fixed-position layouts", in Robotics and Computer-Integrated Manufacturing, vol. 23, pp. 469-477, 2007.
[3] Q. Wang, G W Owen, and A R Mileham, "Comparison between fixed-and walking-worker assembly lines", in Proc IMechE Part B: J.Engineering Manufacturing, vol. 219, pp. 845–848, 2005.
[4] P. E. Chin, G. Q. Huang and Xin Chen, "Scheduling for Assembly Islands with Fixed-Position Layouts", DET 2008, Nantes, France.
[5] T.C.Cheng, J.N.D.Gupta and G.Q.Wang, "A review of flowshop scheduling research with setup times" in Operations Management, vol. 9, no. 3, pp. 262–282, 2000.
[6] Y. H. Lee and M. Pinedo, "Scheduling jobs on parallel machines with sequence-dependent setup times", in European Journal of Operational Research, Vol. 100, pp. 464-474, 1997.
[7] P. B. Luh, L. Gou and Y. Zhang, "Job shop scheduling with group-dependent setups, finite buffers, and long time horizon", in Annals of Operations Research, 76, pp.233-259, 1998.
[8] I. Rechberg, "Opimierung technicher Systeme nach Prinzipien der biologischen Evolution", Problemate, Frommann-Holzboog, 1973.
[9] J. H. Holland, "Adaptation in Natural and Artificial Systems", The University of Michigan Press, Ann Arbor, 1975.
[10] H. P. Schwefel, "Numerische Optimierung von Computer-Modellen mittels del Evolutionsstrategie",Birkhauser, Basel, 1977.
[11] D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, Reading, Mass., 1989.
[12] Z. Michalewicz, "Genetic Algorithms + Data Structures = Evolution Programs", Springer, Berlin, 1997.
[13] N. B. Ho, J. C. Tay and E. M. K. Lai, "An effective architecture for learning and evolving flexible job-shop schedules", European Journal of Operational Research, 179, pp. 316-333, 2007.

Exact Controllability for Dependent Siphons in S^3PMR

Yu-Ying Shih[a,b], Te-Chung Liu[c], Chui-Yu Chiu[d] and D. Y. Chao[e,1]

[a]Department of Information Management,
National Taipei College of Business –Taipei, R.O.C.
[b]Department of Graduate Institute of Industrial and Business Management,
National Taipei University of Technology –Taipei, R.O.C.
[c] Management Collage, National Defense University –Taipei, R.O.C.
[d]Department of Industrial Engineering and Management
National Taipei University of Technology –Taipei, R.O.C.
[e]Department of MIS, National Cheng Chi University –Taipei, R.O.C.

Abstract. This papers reports controllability results of compound siphons in an S^3PMR— a generalization of S^3PR by allowing a job place to use more than one resource and when entering the next operation place, it may not release the resource. We categorize dependent siphons into Types I and II. The former (resp. latter) shares resources between the same (resp. different) set of processes. There are two types of Type I dependent siphons: strong and weak. In an earlier paper, we showed that the dependent condition may be relaxed so that an elementary siphon (actually a compound siphon), while requiring a monitor previously, may be controlled after some elementary siphons get controlled. However, the result may not be optimal so that further linear integer programming (as in Li & Zhou's approach to reduce the total number of monitors from exponential to linear) is needed to determine whether to adjust control depth variables. We derive the exact (both sufficient and necessary) controllability for a Type I strongly 2-dependent siphon and show that any n-dependent (n>3) needs no monitor. Thus, the total time complexity to verify controllability for all Type I strongly relaxed dependent siphons is reduced from exponential to linear. Furthermore, we derive the time complexity for controllability verification of Type II dependent siphons to be $|\Pi_E||\Pi_D|$ (product of total number of elementary siphons and total number of dependent siphons).

Keywords. Petri nets, siphons, controllability, FMS, S^3PMR.

[1] Corresponding Author E-mail: yuhyaw@gmail.com

1 Introduction

Deadlocks in an FMS occur when all resources (modeled by tokens in Petri nets) leak out from a siphon *S* (a structure object in Petri nets) into its complementary set, denoted by [*S*]. *S* plus [*S*] is the support of a so-called P-invariant since these tokens stay either in *S* or [*S*]. Once unmarked (i.e., empty of tokens), *S* remains unmarked forever and output transitions of *S* can never be fired.

To prevent a siphon *S* unmarked, control place V_S and some control arcs are added so that [*S*] plus V_S form part of the support of a new P-invariant. The initial number of tokens (denoted by $M_0(V_S)$) in V_S restricts the maximal of tokens leaking from *S* into [*S*]. *S* is said to be invariant-controlled by the new P-invariant.

However, the number of siphons grows in the worst case grows exponentially fast with respect to the Petri net size leading too many control places and arcs as well as a much more structurally complex liveness-enforcing Petri net supervisor than the plant net model that is originally built.

Unlike other techniques, Li & Zhou [1, 2] add control nodes and arcs to only elementary siphons greatly reducing the number of control nodes and arcs. For the rest, called dependent siphons, no monitors need to be added by adjusting control depth variables [if satisfying neither a marking linear inequality (MLI) nor a subsequent linear integer programming (LIP, Theorem 4 in [1]) test] of elementary siphons. A characteristic T-vector is assigned to each problematic siphon. A maximal set of independent T-vectors (corresponding to the set of elementary siphons) can be computed to form a basis of the vector space comprising all T-vectors. Li & Zhou [1] said, "It should be noted that the proposed elementary siphon and related concepts are applicable to general Petri nets."

We [2] discovered that elementary (resp. dependent, called compound), called basic siphons in an S^3PR (systems of simple sequential processes with resources) may be synthesized from elementary (resp. compound) resource circuits. This has the advantage of avoiding the above computation. This paper revises the definition of elementary siphons so that basic and compound siphons in an S^3PGR^2 (systems of simple sequential processes with general resource requirements) remain to be elementary and compound siphons, respectively. In addition, we will derive the exact controllability (both sufficient and necessary) so that the subsequent LIP can be eliminated.

We need not do the MLI test for n>2 dependent siphons. The number of dependent siphons with n=2 is $O(n^2)$; a simple algebraic test is both sufficient and necessary to determine whether control depth variables needs to be adjusted, eliminating the time-consuming LIP test completely.

2 Preliminaries

A marked Petri Net (PN) is defined by a quadruple N= (P, T, F, M_0), where P is the set of places, T is the set of transitions, F: (P x T) ∪ (T x P) → Z^+ (the set of nonnegative integers) is the flow relation, and M_0: P→ Z^+ is the net initial marking assigned to each place p ∈ P, M_0(p) tokens. In the special case that the flow relation F maps onto {0, 1}; the Petri net is said to be ordinary (otherwise, general). the incidence matrix of N is a matrix [N]: P×T→Z (the set of integers) indexed by P and T such that [N](p, t)= F(t, p)-F(p, t) where F(p, t) is the weight of the arc from place p to its output transition t, and F(t, p) is the weight of the arc from transition t to its output place p.

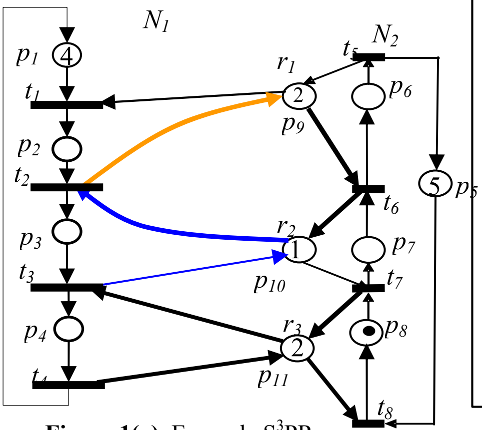


Figure 1(a). Example S³PR.

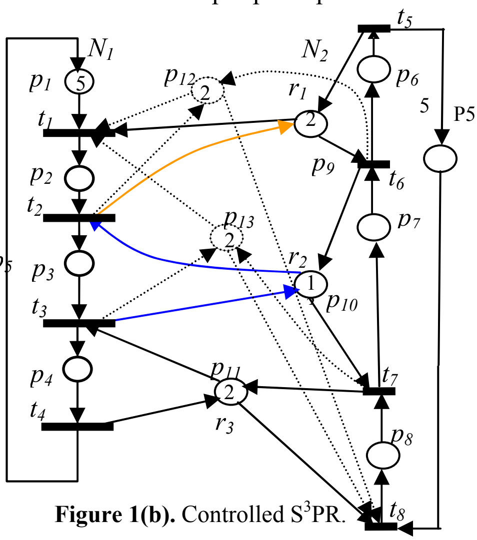


Figure 1(b). Controlled S³PR.

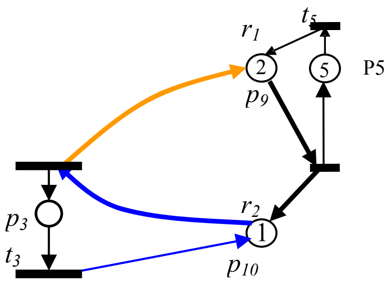


Figure 1(c). I-subnet of S_1.

The set of input (output) transitions of a place p is denoted by $^\bullet p$ (resp. $p^\bullet$). Similarly, the set of input (output) places of a transition t is denoted by $^\bullet t$ (resp. $t^\bullet$). Finally, an ordinary PN such that (s.t.) t∈T, $|t^\bullet|=|^\bullet t|=1$, is called a State Machine (SM).

Given a marking M, a transition t is enabled $p\in{}^{\bullet}t$, $M(p)\geq F(p, t)$, and this is denoted by M[t>. $t\in T$ is said to be disabled by $p\in{}^{\bullet}t$ at M iff M(p) < F(p, t). Firing an enabled transition t results in a new marking M_1, which is obtained by removing F(p, t) tokens from each place $p\in{}^{\bullet}t$, and placing F(t, p') tokens in each place $p'\in t^{\bullet}$ moving the system state from M_0 to M_1. Repeating this process, it reaches M' by firing a sequence $\sigma=t_1t_2\ldots t_k$ of transitions. M' is said to be reachable from M_0; i.e., $M_0[\sigma>M'$.

A transition $t\in T$ is live under M_0 iff $\forall M\in R(N, M_0)$, $\exists M'\in R(N, M)$, t is firable under M'. A PN is live under M_0 iff $\forall t\in T$, t is live under M_0. A Petri net is said to be deadlockfree, if at least one transition is enabled at every reachable marking.

A P-vector is a column vector L: $P\rightarrow Z$ indexed by P and a T-vector is a column vector J: $T\rightarrow Z$ indexed by T, where Z is the set of integers. For economy of space, we use $\sum L(p)p$ (resp. $\sum J(t)t$) to denote a P (resp. T)-vector.

For a Petri net (N, M_0), a non-empty subset S (resp. τ) of places is called a siphon (resp. trap) if ${}^{\bullet}S\subseteq S^{\bullet}$(resp. $\tau^{\bullet}\subseteq{}^{\bullet}\tau$), i.e., every transition having an output (resp. input) place in S has an input (resp. output) place in S (resp. τ). If $M_0(S)=\sum_{p\in S} M_0(p)=0$, S is called an empty siphon at M_0. A minimal siphon does not contain a siphon as a proper subset. It is called a strict minimal siphon (SMS), denoted S, if it does not contain a trap. A siphon is said to be controlled if it is marked under all reachable markings. I_S is the I–subnet [the subnet derived from (S, ${}^{\bullet}S$); see Fig. 1(c)] of an SMS S. Note that $S=P(I_S)$; S is the set of places in I_S.

Y is a P-invariant (place invariant) if and only if $Y\neq 0$ and $Y^T\bullet[N]=0^T$ hold. $||Y||=\{p\in P|Y(p)\neq 0\}$ is the support of Y. A minimal P-invariant does not contain another P-invariant as a proper subset. A siphon S is a set of places where tokens can leak out into another set of places called complementary set [S] of the siphon. Thus, these tokens stay either in S or [S]. S and [S] together form the support of a so-called P-invariant Y; i.e., $S\cup[S]=||Y||$. Thus, $[S]=||Y||\backslash S$. The total number of tokens is conservative in ||Y||.

Definition 1: An elementary resource circuit is called a basic circuit, denoted by c_b. *The siphon constructed from* c_b *is called a basic siphon. An n-compound circuit* $c_0=c_1 \circ c_2 \circ\ldots\circ c_n$ *denotes a circuit consisting of multiply interconnected elementary circuits* $c_1, c_2, \ldots, c_n$ *such that* $c_i\cap c_{i+1}=\{r_i\}$, $r_i\in R$ *(i.e.,* c_i *and* c_{i+1} *intersects at a resource place* r_i*). The SMS* $S_0=S_1 \circ S_2 \circ\ldots\circ S_n$ *synthesized from n-compound circuit* c_0 *is called an n-compound siphon, where* S_i, $i=0,1,2,\ldots,n$, *is the siphon synthesized from* c_i.

Definition 2 [8]: An S^3PMR *net N is a net that results from adding a set R of initially marked places (resource places) to a set of independent process nets. (1) Each resource place r is associated with a set of operation places, H(r). This implies that these operation places require r. (2) For each input transition t of some* $p\in H(r)$, *there exists an arc from r to t if* ${}^{\bullet}t\cap H(r)=\varnothing$ *(empty set, i.e., input places of t but not operation places that use r) . (3) For each output transition t of some* $p\in H(r)$, *there* *exists an arc from t to r if* $t^{\bullet} \cap H(r) = \varnothing$.

Note that we have used H(r) instead of OP(r) in [1]. H(r) is also called the set of holder places that use r. Define H(A)={H(r)| r∈A} as the set of holder places that use resources in A.

Define ρ(r)={r}∪H(r) to be the support of a minimal P-invariant that contains r. Each operation place in an S^3PR (an S^3PMR) uses a unique resource place, and two consecutive operation places use different resource places. S^3PMR [1] is a generalization of S^3PR by allowing an operation place to use more than one resource. Also, when entering the next operation place, it may not release the resource. An example of S^3PMR (resp. S^3PR) and its controlled model are shown in Figs. 2(a) &2(b),respectively.

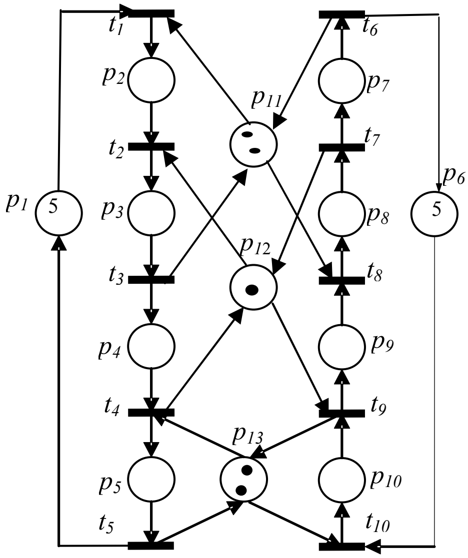


Figure 2(a). Example S^3PMR. $\eta_3 \neq \eta_1 + \eta_2$.

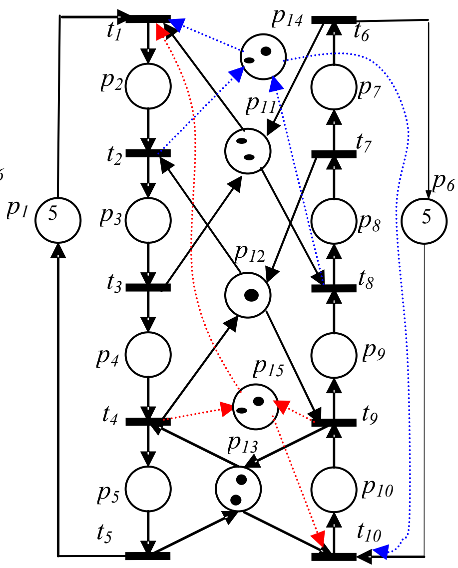


Figure 2(b). Controlled S^3PMR. $M_0(p_{12})$

A control policy involves three factors: for each monitor place VS, (1) its input and output arcs, (2) its initial marking, and (3) the siphon it controls.

In Fig.2(b), we add a control place $V_{S1}=p_{14}$ and the associated arcs for S_1 so that we can consider V_{S1} as another shared resource since the structure involved is similar. V_{S1} plus its holder set of places forms the support $\{V_{S1}, p_2, p_9, p_{10}\}=\{V_{S1}\}\cup H(V_{S1})$ of a new P-invariant. Comparing with the support $S_1\cup[S_1]$ of a P-invariant, we define the controller region $[V_{S1}]=H(V_{S1})$.

If $[S_1]\subseteq[V_{S1}]$, then max $M([S_1])\leq$max $M([V_{S1}])=M_0(V_{S1})$ [since $M_0(V_{Si}) =M(V_{Si}) + M([V_{S1}])$]. If $M_0(V_{Si1})$ is set to $M_0(S_1)$ —*1*, then $M_0(S_1)$ —$1\geq$ max $M[S_1])$ or $1< M(S_1)$ [since $M([S_1]) + M(S_1) = M_0(S_1)$] and S_1 is controlled. On the other hand, the same conclusion cannot be held if $[S_1] \supset [V_{S1}]$. To achieve more states, set $[S_1]=[V_{S1}]$ since the controller region $[V_{S1}]$ is disturbed by the control arcs and V_{S1}. The larger the $[V_{S1}]$, the more the disturbance is.

$[S_1] \subseteq [V_{S1}]$ is achieved in [1] by setting $V_S^{\bullet} \subseteq P_0$ with the added functionality of preventing new siphons from being generated—occurred since V_{S1} is another kind of resource.

Definition 3: Let $M_0(V_S)= M_0(S)- \xi_S$ where $\xi_S \geq 1$ is called the control depth variable. S is said to reach its limit state when $M(S)=1$ or $M([S])=M_0(S)-1$; it is limit-controlled iff it is able to reach its limit state but not able to reach unmarked state; i.e., $\xi_S=1$ or min $M(S)=1$ or max $M([S])=M_0(S)-1$.

Tokens in a siphon S can either leak out to the complementary set $[S]$ of S or stay in S. Thus the sum of tokens in $S \cup [S]$ is a constant $M_0(S)$ and $S \cup [S]$ forms the support of a minimal P-invariant. In Fig. 2(a), $[S_1]=\{p_2, p_9\}$, $[S_2]=\{p_3, p_4, p_{10}\}$, and $[S_3]=\{p_2, p_4, p_9, p_{10}\}$. Note that $\{S_3] \subseteq [S_1] \cup [S_2]$, $([S_3] \setminus ([S_1] \cup [S_2])=\{p_3\} \subseteq S_3)$, while for a dependent siphon $S_3= S_1 o S_2$ in an S^{3}PR, $[S_3]=[S_1] \cup [S_2]$. Yet,

$$\max M([S_3])=\max M([S_1])+\max M([S_2]) \qquad (1)$$

for S_3 in an S^{3}PMR and an S^{3}PR [2].

This is because $[S_1] \cup [S_2] =[S_4] \cup \{p_8\}$, p_8 being in S_4, and $M(p_8)=0$ to empty S_4. To make $M(p_8)=0$, t_7 and t_8 must fire the same number of times in a firing sequence leading the system from M_0 to M.

Definition 4: An elementary resource circuit is called a basic circuit, denoted by c_b. The siphon constructed from c_b is called a basic siphon. A n-compound circuit c is a circuit consisting of multiply interconnected elementary circuits $c_1, c_2, \ldots, c_n$ such that $c_i \cap c_{i+1}=\{r_i\}$, $r_i \in R$ (i.e., c_i and c_{i+1} intersects at a resource place r_i). The SMS synthesized from n-compound circuit using the Handle-Construction Procedure is called an n-compound siphon.

3 Motivation

After the failure of two tests, Li & Zhou adjust control depth variables ξ_S of elementary siphons associated with a dependent siphon to satisfy a Marking Linear Inequality (MLI) as follows:

Theorem 1: Let S be an SMS in an S^3PMR, I_S be its I-subnet, and υ the core subnet by removing all resourceless handles from I_S. υ is strongly connected.

Proof: Assume contrary. Then $\exists t$ in υ such that $\bullet t \cap T(\upsilon)=\varnothing$ or $t\bullet \cap T(\upsilon)=\varnothing$. The latter is impossible since by definition, t is an input transition of some place p in S. The arc (t p) plus a resourceless handle H from t to another place in υ forms another resourceless handle H'. This violates the definition of υ since all resourceless handles have been removed and H' is not among the removed ones. Thus, neither $\bullet t \cap T(\upsilon)=\varnothing$ nor $t\bullet \cap T(\upsilon)=\varnothing$., and υ is strongly connected.

Theorem 2(Theorem 1 in [1]): Let (N_0, M_0) be a net system and S_0, S_1, S_2,... , and S_n be its SMS. Assume that S_0 is a strict dependent SMS w.r.t. elementary siphons S_1, S_2,... , and S_n where $\eta_0=\sum^n_{i=1}(a_i\eta_i)$. S_0 is controlled if 1) N_0 is extended by n additional control places V_{S0}, V_{S1}, V_{S2}, ..., V_{Sn} such that S_1, S_2,..., and S_n are controlled and 2) if $M_0(S_0) > \sum^n_{i=1}(a_iM_0(S_i) - a_i\xi_{Si})$ where ξ_{Si} is the control depth variable for S_i.

We will develop a better (sufficient and necessary) test than that in Theorem 2 so that one adjusts control depth variables if and only if the new test fails. This avoids the time-consuming integer programming test completely whether the new test fails or not.

Set $b=b_1+b_2$, where $b_1=M(p_3)$ and $b_2=M(p_7)$. First we explore the condition under which S_3 is emptied. In order to empty S_3, all tokens in p_9 and p_{11} must go to p_2 and p_8 respectively, and all tokens in p_{10} must distribute to p_3 and p_7. Thus,

$M(p_2)=a$, $M(p_8)=c$, $M(V_{S1})= M(V_{S2})=0$,

$M_0(V_{S1})=c+b_2+a=a+b-\xi_{S1} \Rightarrow b_2=b-\xi_{S1}-c \Rightarrow b_1=c+\xi_{S1}$

$M_0(V_{S2})=c+b_1+a=b+c-\xi_{S2} \Rightarrow b_1=b-\xi_{S2}-a \Rightarrow b_2=a+\xi_{S2}$

Adding the two equations, we have $M_0(r_2)=b=c+a+\xi_{S1}+\xi_{S2}\geq c+a+2$ since $\xi_{S1}\geq 1$ and $\xi_{S2}\geq 1$. $b=c+a+2$ is the condition to empty S_3 when $\xi_{S1}=\xi_{S2}=1$, and the condition for S_3 to be marked (or controlled) is

$b<c+a+2$. (2)

Note that to empty S_3, $b=c+a+2$, $b_1=c+1$ and $b_2=a+1$ when $[V_{Si}]\supset[S_i]$ (i=1,2), versus $b=2$, $b_1= 1$ and $b_2= 1$ when $[V_{Si}]= [S_i]$ (i=1,2). b is increased by

$c+a=M_0(R([V_{S1}]\backslash[S_1])) + M_0(R([V_{S2}]\backslash[S_2]))$.

Physically, $[V_{Si}]$, i=1,2, covers more places than $[S_i]$ due to the movement of output nodes of control arcs to output, called source, transitions of idle places. As a result, when all tokens in $[V_{Si}]$ (i=1,2) are used to trap tokens, some of $[V_{Si}]$ are in $[V_{Si}]\backslash[S_i]$ (i=1,2) and fewer tokens are in $[S_3]$ than the case when $[V_{Si}]=[S_i]$.

This reduces the number of tokens in V_{Si} to trap the tokens in S_3. To compensate for this, we increase $M_0(V_{Si})$ via increasing b by Δb_i (called compensation factor); $\Delta b=\Delta b_1+\Delta b_2=\Delta V_{S1} +\Delta V_{S2} =c+a$. Thus b is increased to c+a+2 (from 2) to empty S_3. And neither S_1 nor S_2 can be limit-controlled since $M(S_1)\geq M(p_3)=b_1=c+1>1$ and $M(S_2)\geq M(p_7)=b_2=a+1>1$. Thus, it seems that the MLI can now be modified to $M_0(S_3) > (M_0(S_1) -(\xi_{S1}+c)) + (M_0(S_2) -(\xi_{S2}+a))$.

To extend to more general cases, we should consider $M_0(R(([V_{Si}]\cap[S_3])\backslash[S_i]))$ rather than $M_0(R([V_{Si}]\backslash[S_i]))$. This is because to empty S_3, tokens in V_{Si} may not need to be trapped in $[V_{Si}]\backslash[S_3]$

4 Theory

We first propose the basic theory below to decide whether a siphon is dependent.

Definition 5: An n-dependent siphon is a dependent siphon depending on n elementary siphons.

Lemma 1: Let (N, M_0) be an ordinary Petri net (PN) system, S an SMS, Y_S the minimal P-invariant associated with S where $\forall p\in||Y_S||$, $Y_S(p)=1$; $\forall p\in P\backslash||Y_S||$, $Y_S(p)=0$, and $M\in R(N, M_0)$. 1) $M([S])+ M(S)=M_0(S)$. 2) $M([V_S])+ M(V_S)=M_0(V_S)$. 3) $M([S])\leqq M_0(S)$-1. 4) $M([V_S])\leqq M_0(V_S)$.

Proof: 1) $S\cup[S]$ is the support of a P-invariant Y_S. By the assumption, $\forall p\in|| Y_S ||$, $Y_S(p)=1$; $\forall p\in P\backslash|| Y_S ||$, $Y_S(p)=0$. Based on Property 1, $Y_S{}^T\bullet M =M_0(S) \Rightarrow M([S])+M(S)=M_0(S)$. 2) The proof is similar to that for (1). 3) S cannot be empty; hence, $M(S) \geq 1 \Rightarrow M([S])\leqq M_0(S)$-1. 4) $M([V_S]) =M_0(V_S) - M(V_S)$ from 2) $\Rightarrow M([V_S])\leqq M_0(V_S)$, since $M(V_S) \geq 0$.

Lemma 2: If siphon S is never empty, $[S]\subseteq[V_S]$.

Proof: Assume $[S]\supset[V_S] \Rightarrow M([S]) >M([V_S])\Rightarrow \max M([S]) > \max M([V_S])= M_0(V_S)= M_0(S)$-1. It is possible that $M([S]\backslash[V_S])=1$ when $M([S])= \max M([S])$. Thus, $\max M([S]) = M([S]\backslash[V_S])+ \max M([V_S])= M_0(S) \Rightarrow M(S)=0$ (empty siphon) — contradiction.

To further explore the controllability for an n-dependent siphon, n>2, specific cases of n=3 will be presented. From which, a general theorem is proposed to conclude that any n-dependent siphon, n>2 is already controlled and needs no monitor if every elementary siphon is limit-controlled.

Definition 6: Let $M_0(V_S)= M_0(S)- \xi_S$ where $\xi_S\geq 1$ is called the control depth variable. S is said to reach its limit state when M(S)=1 or $M([S)=M_0(S)$-1; it is limit-controlled iff it is able to reach its limit state but not able to reach unmarked state; i.e., $\xi_S=1$ or min M(S)=1 or $\max M([S])=M_0(S)$-1.

Based on Lemma 2, we should set $[V_S]$ = [S] to keep the disturbed region as small as possible. To do so, we have $V_S{}^\bullet={}^\bullet[S]\backslash[S]^\bullet$ and ${}^\bullet V_S = [S]^\bullet\backslash{}^\bullet[S]$ where ${}^\bullet[S] =\{{}^\bullet x| x\in[S]\}$ and $[S]^\bullet =\{x^\bullet| x\in[S]\}$.

Theorem 3: Let $S_0= S_1oS_2o\ldots oS_n$.

1) $M([S_0]) \leq M([S_1])+ M([S_2])+\ldots + M([S_n])$.

2) $M([S_i]))= M_0(S_i) - \xi_{Si} - \mu_{Si}$, where i=0,1, ...,n, $\mu_{Si}=M(([V_{S1}]\cap[S_0])\backslash[S_i])$ (maximum tokens of $[S_i]$ equals initial marking of each S_i minus the sum of control depth variable and compensation factor.

3) If $M_0(S_0)> \sum^n_{i=1} (M_0(S_i) -\xi_{Si} -\mu_{Si})$, then S_0 is controlled.

Proof:

1) By Observation 1, $[S_0] \subseteq[S_1]\cup[S_2] \cup\ldots\cup[S_n]$.
Hence, $M([S_0]) \leq M([S_1])+ M([S_2])+\ldots + M([S_n])$.

2) Based on Lemma 1.1, $M([S_i])+ M(S_i) = M_0(S_i)$.
Since $[V_{Si}] \supseteq[S_i]$, the controller region $[V_{Si}]$ can be separated into two: $[S_i]$ and $[V_{Si}]\backslash[S_i]$. The latter can be further divided into $[V_{Si}]\backslash[S_0]$ and $([V_{Si}]\cap[S_0])\backslash[S_i]$; i.e.,$[V_{Si}]= [V_{Si}]\backslash[S_0] +([V_{Si}]\cap[S_0])\backslash[S_i] + [S_i]$.

Similarly, the marking of the controller region $[V_{Si}]$ is the sum of that of the above three subregions:
$M([V_{Si}])= M([V_{Si}]\backslash[S_0]) +M(([V_{Si}]\cap[S_0])\backslash[S_i])+ M([S_i])$
Rearranging the terms, we have
$M([S_i]) = M([V_{Si}])- M([V_{Si}]\backslash[S_0]) - M(([V_{Si}]\cap[S_0])\backslash[S_i]))$
max $M([S_i])$ occurs when $M([V_{Si}]\backslash[S_0])=\min M([V_{Si}]\backslash[S_0])=0$ and $M([V_{Si}])=\max M([V_{Si}])=M_0(V_{Si})$. Thus, $\max M([S_i]) = M_0(V_{Si}) -\mu_{Si}$
where we have set $\mu_{Si} = M(([V_{Si}]\cap[S_0])\backslash[S_i])$ so that the compensation effect occurs the most when $M([V_{Si}]) = M_0(V_{Si})$ and $M(R(S_0))=0$.
By Def. 6, $M_0(V_{Si})=M_0(S_i) - \xi_{Si}$; thus, $\max M([S_i]) = M_0(V_{Si}) -\mu_{Si}= M_0(S_i) - \xi_{Si}-\mu_{Si}$
1) By Observation 1, $\{[S_1]\cup[S_2]\cup...\cup[S_n]\}\backslash S_0\subseteq S_0$. max $M([S_0])$ occurs when $M(\{[S_1]\cup[S_2]\cup...\cup[S_n]\}\backslash S_0)=0$; hence,
$M(\{[S_1]\cup[S_2]\cup...\cup[S_n]\})= M([S_0])$.
Taking the maximum on both sides of this equation, we have
$\max M([S_0]) = \max M([S_1])+ \max M([S_2])+... + \max M([S_n])$. (3)
By Theorem 3.2, $\max M([S_i])= M_0(S_i) -\xi_{Si} -\mu_{Si}$; substituting this equation into Eq. (3), we have $\max M([S_0]))= \sum^n_{i=1\ i}(M_0(S_i) - \xi_{Si} -\mu_{Si})$.
S_0 is controlled if $\max M([S_0]) = \sum^n_{i=1\ i}(M_0(S_i) - \xi_{Si} -\mu_{Si}) < M(S_0)$ or $M_0(S_0)> \sum^n_{i=1} (M_0(S_i) -\xi_{Si} - \mu_{Si})$. (4)
This proves the theorem.

Based on this lemma, we will show that any *n>2* dependent siphon is already controlled and needs no monitors. For a 2-dependent siphon, we now derive the condition for it to be controlled based on Eq. (4) when each elementary (or basic) siphon is limit-controlled.

Example 1: Consider $S_4=S_1oS_2$. $[S_1]=\{p_2, p_8, p_9\}$, $[S_2]=\{p_3, p_7\}$, $[S_4]=\{p_2, p_3, p_7, p_9\}$, and $[S_4] \subseteq [S_1]\cup[S_2]$, satisfying the condition in the above theorem. Let $M_0(p_{19})=a$, $M_0(p_{20})=b$, and $M_0(p_{21})=c$. To empty S_4, set $M(p_2)=a$, $M(p_7)=c$, $M(V_{S1})= M(V_{S2})=0$, $b=b_1+b_2$, where $b_1=M(p_3)$ and $b_2=M(p_9)$ indicating that all tokens in p_{19} and p_{21} must go to p_2 and p_7, respectively, and all tokens in p_{20} must distribute between p_3 and p_9 since no tokens are left in any resource place $r\in S_4$.

$M_0(V_{S1})=c+b_2+a\leq M_0(V_{S1})=a+b-\xi_{S1}$ ($M(V_{S1})\geq 0$ & $M(V_{S1})+ M([V_{S1}])=M_0(V_{S1})$) $\Rightarrow b_2\leq b -\xi_{S1}-c \Rightarrow b_1\geq c+\xi_{S1}$

$M_0(V_{S2})=c+b_1+a\leq M_0(V_{S2})=b+c -\xi_{S2}$ ($M(V_{S2}) \geq 0$ & $M(V_{S2})+ M([V_{S2}])=M_0(V_{S2})$) $\Rightarrow b_1\leq b -\xi_{S2}-a \Rightarrow b_2\geq a +\xi_{S2}$

Adding the two inequalities, we have $M_0(p_{20})=b=c+a +\xi_{S1}+\xi_{S2}\geq c+a+2$ since $\xi_{S1}\geq 1$ and $\xi_{S2}\geq 1$. $b\geq c+a+3$ is the condition to make S_4 non-max-marked when $\xi_{S1}=\xi_{S2}=1$, and the condition for S_3 max-controlled is $b<c+a+2$

Note that to empty S_3 when $[V_{S1}]=[S_1]$ ($i=1,2$), $b\geq 2$, but $b\geq c+a+2$, when $[V_{S1}]\supset[S_1]$ ($i=1,2$). b is increased by $c+a=M_0(R([V_{S1}]\backslash[S_1])) + M_0(R([V_{S2}]\backslash[S_2]))$.

Physically, $[V_{Si}]$, $i=1,2,$ covers more places than $[S_i]$ due to the movement of output nodes of control arcs to output, called source, transitions of idle places. As a result, when all tokens in $[V_{Si}]$ $(i=1,2)$ are used to trap tokens, some of $[V_{Si}]$ are in $[V_{Si}]\backslash[S_i]$ $(i=1,2)$ and fewer tokens are in $[S_3]$ than the case when $[V_{Si}]=[S_i]$.

This reduces the number of tokens in V_{Si} to trap the tokens in S_3. To compensate for this, we increase $M_0(V_{Si})$ via increasing b by Δb_i (called *compensation factor*); $\Delta b=\Delta b_1+\Delta b_2=\Delta V_{S1}+\Delta V_{S2}=c+a$. Thus b is increased to $c+a+2$ (from 2) to empty S_3. And neither S_1 nor S_2 can be limit-controlled since $M(S_1)\geq M(p_3)=b_1=c+1>1$ and $M(S_2)\geq M(p_7)=b_2=a+1>1$. Thus, the MLI can now be modified to $M_0(S_4) > (M_0(S_1) -(\xi_{S1}+c)) + (M_0(S_2) -(\xi_{S2}+a))$, $a=\mu_{S1}$, and $c=\mu_{S1}$ —consistent with the above theorem.

In the sequel, we assume that resources are shared between N_1 and N_2 and the net is a Type I S^3PR to simplify the presentation. The above theorem can be extended to show that n>2 dependent siphon is always controlled. The following theorem helps in this aspect.

Theorem 4: Let $S_0= S_1 o S_2 o \ldots o S_n$. Then

$$\mu_{S1}+\mu_{S2}+\ldots+\mu_{Sn}=M_0(V_{S2})+M_0(V_{S3})+\ldots+M_0(V_{Sn})+(M_0(r_1)+M_0(r_{n+1})). \quad (5)$$

This theorem expresses the sum of compensation factors in terms of known quantities (initial markings). Note that this theorem holds only if $a_1=a_2=\ldots=a_n$, or when the dependent siphon is a compound one.

Corollary 1: Let $S_0= S_1 o S_2 o \ldots o S_n$ and each S_i $(i=2,3,\ldots,n)$ is limit-controlled. Then S_0 is controlled if $n>2$.

Example 2: In Fig. 3, for $S_6=S_1 o S_2 o S_3$, (see Tables 1&2) applying the same method as Example 1. Let $M_0(p_{19})=a$, $M_0(p_{20})=b$, $M_0(p_{21})=c$, and $M_0(p_{22})=d$. To empty S_6, set $M(p_2)=a$, $M(p_{16})=d$, (all tokens in p_{19} and p_{22} must go to p_2 and p_7, respectively) $M(V_{S1})= M(V_{S2})= M(V_{S3})=0$, $b=b_1+b_2$, where $b_1=M(p_3)$ and $b_2=M(p_9)$; i.e., all tokens in p_{20} must distribute between p_3 and p_9 since no tokens are left in any resource place $r\in S_4$. Similarly, $c=c_1+c_2$, where $c_1=M(p_5)$ and $c_2=M(p_7)$; i.e., all tokens in p_{21} must distribute between p_5 and p_7.

Thus,

$\mu_{S1}=d+c_2=M(p_{16})+M(p_7)$,

$\mu_{S2}=d+a= M(p_{16})+M(p_5)$,

$\mu_{S3}=a+b_1=M(p_2)+M(p_3)$, where $\mu_{Si} = M(([V_{Si}]\cap[S_6])\backslash[S_i])$, $M_0(V_{S2})= a+d+b_1+c_2$

Summing the above three equations, we have

$\mu_{S1}+\mu_{S2}+\mu_{S3}=(a+d+b_1+c_2)+(d+a)= M_0(V_{S2})+ (d+a)$ —consistent with Eq. (5) and hence S_6 is controlled.

Remarks: Note that 1) Theorem 4 for n-dependent siphons n>3, holds only when all $a_i=1$ and all $b_j=1$. When $\mu_{Si}=0$ $\forall i\in[1,2,\ldots n]$ and $S_0= S_1 o S_2 o \ldots o S_n$, there is no need for the time-consuming integer programming test after we perform the MLI` test by adjusting control depth variables in an S^3PR. However, if some $\mu_{Si}\neq 0$, then Theorem 4 shows the dependent siphon is already controlled.

Total Time Complexity: Case 1) Only Type I siphons exist. The worst total time complexity is $O(n^2)$ since only 2-dependent siphon needs to check the new MLI based on Theorem 3.3 and there are at worst $O(n^2)$ 2-dependent siphons. In practice, Type I strongly 2-dependent occurs between adjacent resource places shared between two processes and there are linear number of 2-dependent siphons as shown in [7]. As a result, the total time complexity to check controllability of all strongly dependent siphons is reduced from exponential to linear. Case 2) Only Type II siphons exist. The time complexity to verify the MLI of a dependent siphon is $|\Pi_E|$. They are $|\Pi_D|$ depend siphons. As a result, the total time complexity is $|\Pi_E||\Pi_D|$, where $|\Pi_E|$ is total number of elementary siphons and $|\Pi_D|$ total number of dependent siphons).

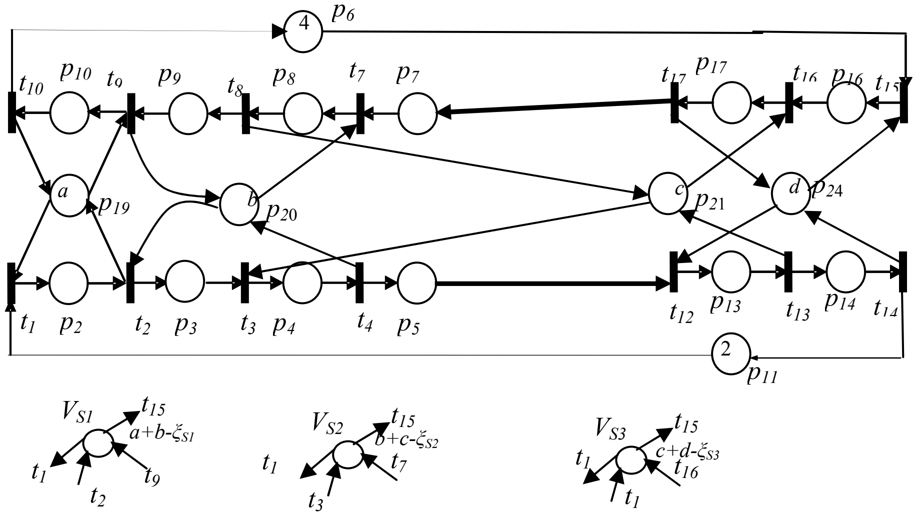


Figure 3. Another example S^{3}PMR

Table 1. Core circuits and synthesized siphons of the net in Figure 3.

Basic siphons	places	c_b	[S]
S_1	$p_3, p_4, p_{10}, p_{19}, p_{20}$	$c_1 = [p_{19}\ t_9\ p_{20}\ t_2\ p_{19}]$	p_2, p_8, p_9
S_8	$p_4, p_5, p_8, p_9, p_{13}, p_{20}, p_{21}$	$c_2 = [p_{20}\ t_7\ p_8\ t_8\ p_{21}\ t_3\ p_4\ t_4\ p_{20}]$	p_3, p_7, p_{17}
S_9	$p_7, p_8, p_{13}, p_{17}, p_{21}, p_{24}$	$c_3 = [p_{21}\ t_{16}\ p_{17}\ t_{17}\ p_{24}\ t_{12}\ p_{13}\ t_{13}\ p_{21}]$	p_4, p_5, p_{16}

Table 2. Compound core circuits and synthesized siphons of the net Figure 3.

Compound siphons	places	c_p	[S]
S_4	$p_4, p_5, p_8, p_{10}, p_{13}, p_{19}, p_{20}, p_{21}$	c_1 o c_2	$p_2, p_3, p_7, p_9, p_{17}$
S_5	$p_4, p_8, p_9, p_{13}, p_{14}, p_{17}, p_{20}, p_{21}, p_{24}$	c_2 o c_3	p_3, p_5, p_7, p_{16}
S_6	$p_4, p_8, p_{10}, p_{13}, p_{14}, p_{17}, p_{19}, p_{20}, p_{21}, p_{24}$	c_1 o c_2 o c_3	$p_2, p_3, p_5, p_7, p_9, p_{16}$

5 Conclusion

Our contributions are as follows: 1) Improve the sufficiency test to avoid the need for the ensuing time-consuming linear integer programming test. 2) 1) Develop a new MLI test— both sufficient and necessary, much better than the only sufficient MLI or LPP (linear programming problem) in [1-5]. The Marking Linear Inequality (MLI) needs to be modified by adding a constant μ_{Si} to each control depth variable ξ_{Si}. 3) Discover type II dependent siphons where μ_{Si}=0 and the MLI is the same as that in [1-2]. 4) Prove that any n-compound (n>3) needs no monitor. 5) Infer that controllability verification is needed only for n=2 compound siphons. Therefore, the total time complexity is linear compared with the exponential one in [1] if all relaxed dependent siphons are compound ones.

6 References

[1] Z. Li and M. Zhou, "Elementary Siphons of Petri Nets and Their Application to Deadlock Prevention in Flexible Manufacturing Systems," IEEE Trans. Syst., Man, Cybern. A., 34(1), pp. 38-51, Jan. 2004.

[2] Z. W. Li and M. C. Zhou, "Clarifications on the Definitions of Elementary Siphons in Petri Nets," IEEE Trans. Syst., Man, Cybern., A: Systems and Humans, 36(6), pp. 1227-1229, November, 2006.

[3] Li, Z. and M. C. Zhou, "Control of Elementary and Dependent Siphons in Petri Nets and their Application," IEEE Trans. on Systems, Man, and Cybernetics: Part A, 38(1), pp. 133-148, Jan. 2008.

[4] Z. W. Li and M. Zhao, "On Controllability of Dependent Siphons for Deadlock Prevention in Generalized Petri Nets," IEEE Transactions on Systems, Man, and Cybernetics, Part A, 38(2), pp.369-384, 2008.

[5] Z. W. Li and M. C. Zhou, "On Siphon Computation for Deadlock Control in a Class of Petri Nets," IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Human, vol.38(3), pp. 667-679, 2008.

[6] Chao, D. Y.: 'Incremental Approach to Computation of Elementary Siphons for Arbitrary S3PR', Control Theory & Applications, IET 2(2), pp.168-179, 2008, doi: 10.1049/iet-cta:20070130.

[7] D. Y. Chao, "An incremental approach to Extract Minimal Bad Siphons," Journal of Information Science and Engineering, 23(1), pp. 203-214, 2007.

[8] Y. S. Huang, M.D. Jeng, X. L. Xie and D. H. Chung, "Siphon-based deadlock prevention policy for flexible manufacturing systems," IEEE Transactions on System, Man, and Cybernetics—PART A: SYSTEMS AND HUMANS, 36(6), pp. 1248-1256, November, 2006

A New MIP Test for S^3PGR^2

Yu-Ying Shih[a,b], D. Y. Chao[c,1]and Chui-Yu Chiu[d]

[a] Department of Information Management,
National Taipei College of Business –Taipei, R.O.C.
[b]Department of Graduate Institute of Industrial and Business Management,
National Taipei University of Technology –Taipei, R.O.C.
[c]Department of MIS, National Cheng Chi University –Taipei, R.O.C.
[d]Department of Industrial Engineering and Management
National Taipei University of Technolog–Taipei, R.O.C.

Abstract. For general Petri nets such as S^3PGR^2 (systems of simple sequential processes with general resources requirement), the mixed integer programming (MIP) test by Chu *et al*. is no longer valid. Park et al. modified MIP test and claimed that it can determine the net is live if there are no feasible solutions. However, the net can have livelocks even though it is deadlock-free. The set of places with dead input transitions may not form a siphon and cannot be detected by the above modified MIP test which detects siphons and not their subsets. We show one counter example to confirm this and propose a revised MIP test to fix the problem.

Keywords. Petri nets, siphons, deadlocks, MIP.

1 Introduction

Chu *et al.* [1] proposed the mixed integer programming (MIP) method to detect deadlocks (tied to unmarked siphons) quickly. It needs no explicit enumeration of siphons which grows exponentially with the size of the net and hence is useful for checking deadlock-freeness of large systems. Its computational efficiency is relatively insensitive [1] to the initial marking and more efficient than classical state enumeration methods. The MIP method is able to find an unmarked maximal siphon in an ordinary Petri net (all arcs unit weighted).

It [2] has been applied to design liveness-enforcing supervisors such as S^3PR (systems of simple sequential processes with resources) [3] for flexible manufacturing systems (FMS) to avoid complete siphon enumeration. Hence, it has high

[1] Corresponding Author E-mail: yuhyaw@gmail.com

computational efficiency compared with the existing ones in the literature. The deadlock detection method in [4] and [5] were then based on MIP.

For general Petri nets (GPN) such as S^3PGR^2 (systems of simple sequential processes with general resources requirement), the MIP test in [1] is no longer valid. This is because a deadly marked siphon in a GPN may not be unmarked. Park et al. [6] modified the MIP test and claimed that it can determine the net is live if there are no feasible solutions.

However, the net can have livelocks even though it is deadlock-free. The set of places with dead input transitions may not form a siphon [7] and cannot be detected by the above modified MIP test which detects siphons and not their subsets. We will cite one counter example in Section 4 to confirm this and propose a revised MIP test to fix the problem.

2 The GPN Challenge

An S^3PR is live if no siphons ever become empty [3]. However, this is no longer true for S^3PGR^2 due to the fact an S^3PGR^2 is a general Petri net (GPN). Hence, we tackle the problem first for an arbitrary GPN. One can no longer ensure liveness by making all siphons never empty. The new condition is called max-controlled [8]. To understand this, we explore the condition under which a GPN behaves like an OPN (ordinary PN). It occurs if $\forall M \in R(N, M_0)$, $\exists\, p \in P$, $M(p) \geq W(p, t)$ (the arc weight from p to t), $\forall t \in p^{\bullet}$. That is for every reachable marking, there is a max-marked *p* (see Def. 1); i.e., the amount of its tokens is greater than the weight of any outgoing arc.

In general, when all siphons are max-controlled, then it behaves like an OPN. Thus, for an S^3PGR^2, it is live as long as all siphons are max-controlled. We will show that this condition can be relaxed in the next section. For the moment, we present some basic theories of max-controlled siphons.

Definition 1: *p_i is called max-marked under M, if $\forall t \in p_i^{\bullet}$, $M(p_i) \geq W(p_i, t)$ (the arc weight from p_i to t). Let $D=\{p_1, p_2, \ldots, p_K\}$. D is called max-marked under M, if $\exists\, p \in D$, p is max-marked. If D is a siphon, it is said to be max-controlled iff D is max-marked at any reachable marking. Let y be an S-invariant with components y_i and $M_D=[a_x(p_1)-1\ a_x(p_2)-1\ \ldots\ a_x(p_K)-1\ 0\ 0\ \ldots\ 0]^T$ be a marking where $a_x(p)=max_{t \in p^{\bullet}} W(p, t)$ is the maximal weight of all outgoing arcs from p. That is, $\forall p \in D$, $M(p)=a_x(p)-1$; $\forall p \in P \backslash D$, $M(p)=0$, where $P \backslash D=\{x | x \in P, x \notin D\}$. The weighted sum of tokens under M_D is*

$$W(M_D)= W_D= M^T_D \bullet y=\sum_k (a_x(p_k)-1)y_k.$$

The support of an S-invariant is also a set of places where the weighted total number of tokens is conserved. A minimal siphon is in one such support. Hence all the unloaded tokens remain in the support of the invariant. If they are also in that of another invariant (ν), the minimal siphon is said to be *invariant-controlled* [8]. By controlling the number of tokens in ν, we may prevent the minimal siphon from being completely unloaded.

*Definition 2***:** *Let y be an S-invariant with components y_i, $\forall p_i \in P$, and $D \subseteq P$ a siphon of N. The siphon is called max-controlled by the S-invariant or max-controlled under M_0 iff the weighted sum of tokens $W(M_0)= y^T \bullet M_0 > W_D$, $\forall p_k \in D$, $\mathbf{y}_k=\mathbf{y}(p_k)>0$ and $\forall p_i \in P \backslash D$: $\mathbf{y}_i \leq 0$.*

Lemma 1: *Let (N, M_0) be a net-system and let $D \subseteq P$ be a siphon of N. If D is max-controlled by the S-invariant under M_0, then $\forall M \in R(N, M_0)$: D is max-marked under M.*

By this lemma, if it is initially max-marked, it remains so for all reachable markings.

*Lemma 2***:** *For a dead Petri net (N, M_0), there exists a non-max-controlled siphon at M_0.*

Lemma 3: *N is deadlock-free under M_0 if every siphon D is max-controlled by an S-invariant under M_0.*

For S^3PGR^2, the condition of max-controlled siphons may be overly constrained [7] as shown in the next section.

3 A Better Liveness Condition

The relaxation of the condition is shown in Fig. 1(a) where $M_0(p_1)=1$, $M_0(p'_1)=1$, $M_0(r_1)=6$, $M_0(r_2)=3$. There is only one SMS S={r_1, r_2, p_3, p'_2}. We push as many tokens out from *S* as possible to make some transitions dead. t_1 and t_2 can never fire and the rest are live; hence it is deadlock free. Note that t_2 is neither live nor potentially firable even under M_0; hence it is not quasi-live. After adding a token to r_2, it becomes live and also quasi-live since all transitions are potentially firable (i.e., quasi-live) under M_0. However, neither r_1 nor r_2 is max-marked; hence *S* is not max-controlled. This motivates us to relax the liveness condition.

In Fig. 1(b), all transitions in circuit [r_2 t''_2 p''_2 t''_1 r_2] of WP_3 are live, while all transitions in WP_1 and WP_2 are dead. The net is said to be in a livelock state.

Note that it is not live, yet siphon *S* is not deadly marked. Unlike S^3PR, the absence of deadly marked siphons (*DMS*) does not imply the liveness of an S^3PGR^2. Simple extension of *DMS* is insufficient for the liveness analysis of S^3PGR^2.

Thus, the constraint of max-controlled may be relaxed as follows.

*Definition 3***:** *Let $a_y(p)=\max_{t \in p\bullet} W(p, t)$, $t \in I_D$, (for all output t which are in I_D) and t is not on a PP-circuit (denoted by H^{PP}_c and called a PP'-circuit) of the form [r $t_2 p_1 ... p_k$ t_k r], $r \in P_R$, $p_i \in H(r)$, i=1,2,...,k. p is called max'-marked under M in a SMS, if $a_y(p) \leq M(p)$. Let D={p_1, p_2, ..., p_K}. D is called max'-marked under M, if $\exists p \in D$, p is max'-marked. If D is a siphon, it is said to be max'-controlled iff D is max'-marked at any reachable marking. Let* **y** *be an S-invariant with components y_i and $M'_D=[a_y(p_1)-1\ a_y(p_2)-1\ ...\ a_y(p_K)-1\ 0\ 0\ ...\ 0]^T$. That is, $\forall p \in D$, $M(p)=a_y(p)-1$; $\forall p \in P \backslash D$, $M(p)=0$. The weighted sum of tokens under M_D is $W(M_D)= W'_D=M'^T_D \bullet y=\sum_k (a_y(p_k)-1)y_k$*

In Fig. 1(a), $t'_2 \in r_1\bullet$, $t'_2 \in I_D$, and $M_0(r_1)=2> W(r_1, t'_2)=1$. Hence, r_1 is max'-marked. Note that $t_1 \in r_1\bullet$, $t_1 \in I_D$ and $M_0(r_1)=2<W(r_1, t_1)=4$. We do not consider H^{PP}_c (a circuit with only one resource place r; the rest belong to $H(r)$) above. An example of H^{PP}_c ($[r_2\ t''_2 p''_2 t''_1 r_2]$) is shown in Fig. 1(b).

Let B denote the set of all H^{PP}_c in I_D and $I'_D= I_D \backslash B$. D is max'-marked if it is max-marked in I'_D. If I'_D is separated from N to become a whole (not sub-) net, then, by Lemma 3, it is deadlock-free if D is max-marked in I'_D. When the net N is not live, then there is a permanently dead I'_D, an input place p' (that is not in I'_D.) of transition t' in I'_D are marked to enable the corresponding input arcs ($p'\ t'$). This implies that if every I'_D is max'-controlled in N, then some transition in I'_D will become enabled and the net is live as will be proved in Theorem 1.

Definition 4: *Let y be an S-invariant with components* y_i, $\forall p_i \in P$, *and* $D \sqsubseteq P$ *a siphon of N. The siphon is called max'-controlled by the S-invariant under* M_0 *iff the weighted sum of tokens* $W(M_0)=y^T\bullet M_0> W'_D$ *and* $\forall p_i \in P \backslash D$: $y_i=y(p_i)\leq 0$.

Lemma 4: *Let* (N, M_0) *be a net-system and let* $D \sqsubseteq P$ *be a siphon of N. If D is max'-controlled by the S-invariant under* M_0, *then* $\forall M \in R(N, M_0)$: *D is max'-marked under M.*

By this lemma, if it is initially max'-marked, it remains so for all reachable markings. Note that the condition $W(M_0)=Y^T\bullet M_0> W'_D$ in Def. 2 is only sufficient (not necessary) for D to be max'-marked under M.

Proposition 1 [6]: *Let* $N=(P\cup P^0\cup P_R,T,W)$ *be an* S^3PGR^2 *and* $D(\neq 0)$ *a siphon so that it does not contain the support of any S-invariant. Then* $|D\cap P_R|>1$.

Lemma 5: *Let* (N, M_0) *be an marked* S^3PGR^2, $M \in R(N,M_0)$ *and* $t \in T$ *a dead transition under M. Then* $M_0 \notin R(N, M)$.

Corollary 1: *Let* (N, M_0) *be an marked* S^3PGR^2, $M \in R(N, M_0)$ *and* $t \in T$ *a dead transition under M. Then there exist* $M' \in R(N, M)$ *and two subsets* $J \subset I_N$ *and* $H \subset I_N$ *such that* $I_N = J\cup H$, $I_N=\{1, 2, ...,k\}$, $J\cap H=\emptyset$, $J\neq\emptyset$ *and: 1)* $\forall h \in H$, $M'(p_h^0)= M_0(p_h^0)$; *2)* $\forall j \in J$, $M'(p_j^0)< M_0(p_j^0)$ *and* $\{p\bullet |\ p \in P, \text{and } M'(p)>0\}$ *is a set of dead transitions.*

The above lemma and corollary are taken from [6] and their proofs are similar and omitted here. Lemma *5* states that if there exists a dead transition *t*, then M_0 is not reachable. Otherwise, *t* is potentially firable (since N is assumed to have well-marked as in Def. 2) and not dead — contradiction. Corollary 1 states that if there exists a dead transition *t*, then some WPs are blocked and cannot proceed to complete operations. If they can return to the initial idle state, then *t* is potentially friable and not dead — contradiction.

Based on Corollary 1, we now prove the main property about siphon: If there is a dead transition *t* under M, then there exists a non-max'-marked siphon under $M' \in R(N, M)$. The basic idea behind the proof is the construction of a non-max'-marked siphon. It consists of two sets of places: 1. resource places *r* such that one of its output arcs is disabled. 2. unmarked holders of these *r*. The proof is similar to that for S^3PR with some differences due to the condition of non-max'-marked instead of empty siphons.

Theorem 1: Let (N, M_0) be an marked S^3PGR^2, $M\in R(N, M_0)$ and $t\in T$ a dead transition under M. Then $\exists M'\in R(N, M)$, $\exists D$ a siphon so that D is nonempty and non-max'-marked.

4 MIP for OPN

This section introduces the MIP basics for OPN (helpful to understand MIP for GNP) via Subsection 4.1 on the constraints in an MIP and Subsection 4.2 on the objective function.

4.1 Constraints

Definition 5: Let S be a siphon. (1) $\mathbf{v}_p$ is a binary variable for p, and $\mathbf{v}_p=1$ if $p\notin S$. (2) $\mathbf{z}_t$ is a binary variable for t, and $\mathbf{z}_t=1$ if $t\notin S^\bullet$.
Clearly, any p with $\mathbf{v}_p=1$ and any t with $\mathbf{z}_t=1$ will be removed if the classical algorithm is used. Note that $S=\{p|\mathbf{v}_p=0\}$.

Lemma 6: Let S be a siphon and $p\in S$. Then $\forall t\in {}^\bullet p$, (1) $\mathbf{z}_t=0$. (2) $\mathbf{v}_p \geq \mathbf{z}_t$.

Thus, if p is in S, then all input transitions of p are in I_{SS}. If one of input transition t is not in I_S, then after infinite firings of t, p becomes unbounded in contradiction to the notion of siphon whose tokens can only decrease. Based on this lemma, we have

Theorem 2: $\forall p\in P$, $\forall t\in {}^\bullet p$, $\mathbf{v}_p \geq \mathbf{z}_t$.

The inequality $v_p \geq z_t$ is a constraint used in the MIP test in [1].

Theorem 3: *Let $M\in R(N, M_0)$ and siphon S be empty of tokens under M, and $M(p)>0$, then (1) $\mathbf{v}_p=1$. (2) $\forall t\in T$, if $\forall p\in {}^\bullet t$, $M(p)>0$, then (a) $\mathbf{z}_t=1$. (b) $\mathbf{z}_t\geq\sum_{p\in t^\bullet} v_p-|{}^\bullet t|+1$.*

The above inequality $z_t \geq \sum_{p\in t^\bullet} v_p -|{}^\bullet t|+1$ in Theorem 3.2.b is another constraint used in the MIP test in [1]. Theorem 3.1 implies that if a places is marked, then $v_p=1$. Similarly, if a transition is enabled, then $z_t=1$. Further, if all input places are not in an S, then t cannot be in I_S since places in an unmarked siphon are permanently empty of tokens; any input transition of these places cannot be live.

Note that even if t is not enabled, but all of its input places carry $v_p=1$, $z_t=1$ by Theorem 3.2.b. We say that transition t is pseudo-enabled.

Corollary 2: *$\mathbf{z}_t=1$ for any enabled transition.*

Theorem 4: $\forall p\in t^\bullet$, $\mathbf{v}_p=1$, if $\mathbf{z}_t=1$.

This theorem helps to propagate binary variables in a backward fashion. Another constraint comes from the state equation; i.e., $M=M_0+\Theta\bullet x$ where $M\in R(N, M_0)$, Θ is the incidence matrix of net N, and $\boldsymbol{x}$ a firing vector. The constraints so far obtained are summarized as follows:

$$z_t \geq \sum_{p\in t^\bullet} v_p -|{}^\bullet t|+1 \quad (1)$$
$$v_p \geq z_t,\ \forall (t\, p)\in W \quad (2)$$
$$v_p, z_t \in \{0, 1\} \quad (3)$$

$v_p = 1$, if $M(p)>0$ (4)

$M=M_0+\Theta \bullet x$, $M\geq 0$, $x\geq 0$ (5)

where Θ is the incidence matrix of net N.

Note that Eq. (4) is nonlinear; it can be linearized as follows:

$v_p \geq M(p)/SB(p)$, $\forall p \in P$ (4')

for a structurally bounded net, where the structural bound (SB) is defined as $SB(p)=\max\{M(p)|\ M=M_0+\ \Theta\bullet x,\ M\geq 0,\ x\geq 0\}$.When $M(p)>0$, $v_p=1\geq M(p)/SB(p)>0$. When $M(p)=0=M(p)/SB(p)$, $v_p=0$ or 1 and again $v_p \geq M(p)/SB(p)$.

4.2 Objective Functions

There are only two possible objective functions: (1) min-objective function: $G(M)=\min \sum_{p\in P} v_p$, and (2) max-objective function: $G(M)=\max \sum_{p\in P} v_p$. The following theorems and lemmas help to find the correct objective function to find empty or unmarked siphons.

Lemma 7: If t is potentially friable from M_0, then $v_p=1$ $\forall p \in t^\bullet$, if $z_t=1$.

Proof: If $v_p=0$, then $v_p< z_t$ against the fact that $v_p \geq z_t$ in Theorem 2.

The following theorem confirms that $\mathbf{z}_t=1$ and hence $\mathbf{v}_p=1$ in the above lemma.

Theorem 5: (1) if $t\in T$ is potentially friable under M_0, then $z_t=1$. (2) If $p\in P$ is potentially marked, then $v_p=1$.

Proof: Prove by induction w. r. t. $|\sigma|$, the length of firing sequence σ. First prove for $|\sigma|=1$, $\sigma=t$. t is enabled under M_0. By Theorem 3.1, $\forall p \in {}^\bullet t$, $M(p)>0$, and $v_p=1$. Thus, $\sum_{p\in t\bullet} v_p=|{}^\bullet t| \Rightarrow \sum_{p\in t\bullet} v_p - |{}^\bullet t|+1=1= z_t$. After firing t, $\forall p' \in t^\bullet$, $M(p')>0$, and hence $v_{p'}=1$. Assume the theorem holds for any σ' such that $|\sigma'|<|\sigma|=m$; now prove it also holds for σ, where transition t' is enabled; i.e., potentially firable under M_0. $\forall p\in {}^\bullet t'$, let σ'' be the firing sequence such that p is potentially marked. Thus, $v_p=1$ by the assumption and $\sum_{p\in t\bullet} v_p=|{}^\bullet t| \Rightarrow \sum_{p\in t\bullet} v_p - |{}^\bullet t|+1=1= z_t$. This proves (1). By Lemma 7, $\forall p' \in t'^\bullet$, $v_{p'}=1$. This proves (2).

Corollary 3: $\forall t\in T$ in a strongly connected (SC) net, if t is potentially friable under M_0, then $\sum_{p\in P} v_p=|P|$.

Proof: By Theorem 5, $z_t=1$. $\forall p\in t^\bullet$, $v_p=1$. Since the net is SC, $\forall p\in P$, $\exists t\in T$, such that $p\in t^\bullet$. Thus, $v_p=1$ for every place p in the net and $\sum_{p\in P} v_p=|P|$.

Thus, if all transitions are potentially friable from initial marking, then all $v_p=1$ and $S=\{p|v_p=0\}=\emptyset$. One cannot find any unmarked siphon under M_0. This remains so for any reachable marking M as shown by the next theorem.

Theorem 6: If every t is potentially friable from an $M\in R(N, M_0)$, then $\max \sum_{p\in P} v_p=|P|$.

Proof: From Corollary 3, $\sum_{p\in P} v_p=|P|$, which is the maximal of $\sum_{p\in P} v_p$ corresponding to an empty siphon since $S=\{p|v_p=0\}$.

This theorem immediately leads to the following:

Corollary 4: (1) If there is an $M \in R(N, M_0)$ such that every t is potentially friable from M, then the MIP method with max-objective function produces an incorrect solution of empty siphon. (2) The MIP method is able to produce an empty siphon only if $\forall M \in R(N, M_0)$, there is a transition that is not potentially firable.

Proof: (1) By Theorem 6, max $\sum_{p \in P} v_p = |P|$ and $S=\{p | v_p=0\}=\varnothing$, which may not hold always. (2) Assume contrarily that all transitions are potentially friable, and then by (1), only null siphon can be obtained —contradiction.

Therefore, if the objective function G(M)= max $\sum_{p \in P} v_p$ is chosen, one would not be able to obtain unmarked siphons if all transitions are potentially firable under the initial marking (true for a well-designed FMS). On the other hand, the maximal siphon unmarked at a given marking can be determined by the following MIP problem, and there exist siphons unmarked at M iff G(M)<|P|: $G(M)=\min \sum_{p \in P} v_p$ under constraints (1)–(4) and $M=M_0+ \Theta \bullet x$, $M \geq 0$, $x \geq 0$ where Θ is the incidence matrix of net N.

Note that Eq. (4) implies that $v_p=1$ if M(p)>0 and $v_p=1$ or 0 if M(p)=0. To minimize $\sum_{p \in P} v_p$, $v_p=0$ if M(p)=0. Also from (1), $z_t \geq 0$ when $0 \geq \sum_{p \in t\bullet} v_p - |{}^\bullet t|+1$; $z_t=0$ or 1. We pick $z_t = 0$ to have more $v_p=0$ since $v_p \geq z_t=0$, $\forall (t\ p) \in W$ by (2). Based on above results, we can conclude that there is no unmarked siphon in a net system N=(P, T, F, M) if G(M)=|P| is true. In order to find maximal unmarked siphons, the method (in the next subsection) to trace a net to find all v_p is helpful to find G(M).

5 MIP Test for S^3PGR^2

Definition 6: Given a well-marked S^3PGR^2 net (N, M_0) and $M \in R(N, M_0)$, the modified marking $\overline{M}$ is defined by

$$\overline{M}(p) = \begin{cases} M(p) & if \quad p \notin P_0 \\ 0 & otherwise \end{cases} \quad (6)$$

Furthermore, the set of all modified markings induced by the reachable markings is defined by $\overline{R(N, M_0)} = \{\overline{M} \mid M \in R(N, M_0)\}$.

They proceed to the development of these results by first presenting an IP formulation that, given a modified marking $\overline{M} \in \overline{R(N, M_0)}$, of a CS^3PGR^2 (Controlled System of Simple Sequential Processes with General Resource Requirements) net, it computes the maximal deadly marked siphon S, such that (i) $S \cap (P_R \cup P_W) \neq \varnothing$ (P_R, P_W are the set of resource and control places respectively) ; and (ii) every place in $S \cap (P_R \cup P_W)$ is a disabling place.

They propose a sufficient condition for the non-existence of deadly marked siphons S, such that $S \cap (P_R \cup P_W) \neq \varnothing$ and all places in $S \cap (P_R \cup P_W)$ are disabling places, in the entire space $R(N, M_0)$ of a given CS^3PGR^2 net N. Furthermore, in light of Theorem 7, this condition constitutes a sufficient condition for the liveness of CS^3PGR^2 nets.

Theorem 7 (Theorem 4 in [6]): Let $N=(P, T, W, M_0)$ be a well marked CS^3PGR^2 net. Then, given a marking $\overline{M} \in \overline{R(N, M_0)}$, the maximal deadly marked siphon S such that (i) $S\cap(P_R\cup P_W) \neq \emptyset$; and (ii) every place in $S\cap(P_R\cup P_W)$ is a disabling place at M, is determined by:

$$S=\{p \mid p\in P,\ v_p=0\} \quad (7)$$

where v_p is obtained through the following IP formulation:

$$G(M)=\min \textstyle\sum_{p\in P} v_p \quad (8)$$

s.t.

$$f_{pt} \geq (\overline{M}(p)- W(p\ t)+1)/SB(p),\ \forall W(p\ t)>0 \quad (9)$$

$$f_{pt} \geq v_p,\ \forall W(p\ t)>0 \quad (10)$$

$$z_t \geq \textstyle\sum_{p\in {}^\bullet t} f_{pt} - |{}^\bullet t|+1 \quad (11)$$

$$v_p \geq z_t,\ \forall W(t\ p)>0 \quad (12)$$

$$\textstyle\sum_{r\in PR\cup PW} v_r \leq |P_R\cup P_W|-1 \quad (13)$$

$$z_t \geq \textstyle\sum_{p\in \Box t} f_{pt} - |{}^\bullet t|+1 \quad (14)$$

$$v_p, z_t, f_{pt} \in \{0, 1\},\ \forall p\in P,\ t\in T \quad (15)$$

where the structural bound (SB) is defined as $SB(p)=\max\{ M(p) \mid M=M_0+ \Theta\bullet x, M\geq 0, x\geq 0\}$. Eq. (14) expresses the fact that if all output arcs of resource place r are enabled, then r must be max-marked and $v_r=1$. Otherwise, r is not max-marked and $v_r=1$ or 0. Thus, Theorem 7 finds a maximal non-max-controlled siphon. However, according to [6], Theorem 7 tests the presence of DMS under a given marking M. In order to find deadly marked siphons (DMS), it is necessary to apply Theorem 7 to all reachable M using the following state equation:

$$M=M_0+\Theta\bullet x,\ M\geq 0,\ x\geq 0 \quad (16)$$

Corollary 5 [6]: Given a reachable marking M of a CS^3PGR^2 net $N=(P, T, W, M_0)$, M contains no deadly marked siphon S such that $S\cap(P_R\cup P_W) \neq \emptyset$ and all the places in $S\cap(P_R\cup P_W)$ are disabling places iff the integer program of Theorem 7 is infeasible.

Corollary 6 [6]: Let $N=(P, T, W, M_0)$ be a well marked CS^3PGR^2 net. Then, if the mixed integer program defined by Eqs. (7)-(16) and Eq. (6) is infeasible, N is live.

6 Counter Example

A counter example is shown in Fig. 1(b) where the net *N* is a well-marked S^3PGR^2 net; all transitions in N_3 are live, while all transitions in N_1 and N_2 are dead. Thus, *N* is not live, yet the only problematic siphon $S=\{r_1, r_2, p_3, p_5, p_8\}$ is not deadly marked. Thus, unlike S^3PR, the absence of deadly marked siphons (DMS) does not imply the liveness of an S^3PGR^2.

However, the net is declared live since there is no feasible solution as shown below using the MIP test based on Corollary 5.

Under the marking M shown in Fig. 1(b), $f_{pt}=1$ for arcs (r_2 t_8), (p_2 t_2), and (p_6 t_5) by Eq. (9). z_{t8} =1 by Eq. (11)⇒ $\boldsymbol{v}_{p8}$=1 by Eq. (12)⇒ z_{t7} =1 by Eq. (11)⇒ v_{r2}=1 by Eq. (12) ⇒ $\boldsymbol{z}_{t2}$ =1 by Eq. (11)⇒ $\boldsymbol{v}_{p3}$=1 by Eq. (12) and $\boldsymbol{v}_{r1}$=1 by Eq. (12). Now $\sum_{r\in PR\cup PW} \boldsymbol{v}_r=\boldsymbol{v}_{r1}+\boldsymbol{v}_{r2}=2>|P_R\cup P_W|-1=1$ violating Eq. (13) and hence the solution is infeasible under *M*. For all other reachable markings *M'*, every transition is potentially firable and hence $\boldsymbol{v}_{r1}=\boldsymbol{v}_{r2}=1$; again the solution is infeasible.

Another example is shown in Fig. 1(a) where all places in $S=\{r_1, r_2, p_3, p_5\}$ are non-max-marked and will be reported using the MIP test. Yet, *S* is not deadly marked. Based on Eq. (14), r_1 is not max-marked; hence $\boldsymbol{v}_{r1}$ = 0. Similarly, $\boldsymbol{v}_{r2}$=0. Now $\forall t\in T$, *t* is pseudo-disabled and $\boldsymbol{z}_t$=0, which leads to $\boldsymbol{v}_p=0$, if $M(p)=0, \forall p\in P$. Thus, it reports $S=\{r_1, r_2, p_3, p_5\}$ as a deadly marked siphon (DMS) and the net not live, which is incorrect.

On the other hand, r_1 is max'-marked, $\boldsymbol{v}_{r1}$ =1 if Eq. (14) is replaced by one that for a max'-marked place. By Eq. (11), $\boldsymbol{z}_{t'2}$ is pseudo-enabled; hence $\boldsymbol{z}_{t'2}$=1. Now $\boldsymbol{v}_{r1}\geq \boldsymbol{z}_{t'2}$=1; hence $\boldsymbol{v}_{r1}$=1. Similarly, $\boldsymbol{v}_{p'2}$=1. Repeating this process, $\boldsymbol{v}_p$=1 $\forall p\in P$ and $G(M)=|P|$; this remains true for all other *M* that satisfies the state equation in Eq. (16). Thus, the net is live.

It remains a challenging task to develop a new MIP test to find siphons responsible for livelocks.

7 Revised MIP TEST

This section introduces two approaches. The first one does not work, while the second is correct with an example. The I-subnet I_S in Fig. 1(b) is separated into two parts: A dead (blue colored) and B live (red colored). The two intersects at r_2. The value of the binary variable $\mathbf{v}_p$ is ambiguous. If $\mathbf{v}_p$ =1 (resp. 0), then p is in B (resp. A). Thus, $\mathbf{v}_p$ is no longer effective to identify permanently dead subnets and will be removed from the improved MIP test. Specifically, when place p is in a siphon S, $\mathbf{v}_p$ =o, all input transitions of p are in the I-subnet of S. However, the set of places in a permanently dead subnet is no longer, but a subset of, a siphon. As a result, $\mathbf{v}_p$ is useless to portray an empty siphon and will not be considered in the new MIP test. Constraint $\mathbf{v}_p \geq \mathbf{z}_t$ will be deleted.

Instead, a new binary variable *f* is assigned for each arc in the net. In a dead (resp. live) subnet, all arcs are disabled (resp. enabled) whose *f*=0 (resp. 1). For *f*, unlike v_p, there is no ambiguity. Similarly, there is no ambiguity for z_t, the binary variable for transition *t*. In a dead (resp. live) subnet, all transitions are disabled (resp. enabled) whose z_t =0 (resp. 1). Eq. (1) expressess the fact that if $v_p =1$ for each input place of *t* in an ordinary net, then z_t =1. Recall that $v_p =1$ if *p* is marked or arc (*p t*) is enabled. Thus, z_t =1 if *t* is enabled. When $v_p =1$ but *p* is marked, z_t remains at 1, *t* is said to be pseudo-enabled. For general Petri nets, arc (*p t*) may not be enabled even though *p* is marked

(v_p=1). Hence variable v_p in Eq. (1) should be replaced by f. This results in the following constraint. Similarly Constraint $v_p \geq z_t$ should be replaced by $f_{tp} \geq z_t$ by substituting f_{tp} for v_p. Physically, when a transition t is enabled, so is each output place p since after firing t, tokens can flow along arc (t p). This is consistent with $f_{tp} \geq z_t$ where f_{tp} =1 if z_t =1. On the other hand, if t is an output, rather than input, transition of p, f_{tp}=1 does not imply z_t =1; f_{tp}=0 implies z_t =0 since t is disabled if any input arc is disabled.

$$G(M)=min\sum_{(p\ t)\in W} f_{pt} \quad (17)$$

s.t.

$$f_{pt} \geq (M(p)- W(p\ t)+1)/SB(p), \forall W(p\ t)>0 \quad (18)$$

$$f_{tp} \geq z_t \ \forall W(t\ p)>0 \quad (19)$$

$$z_t \geq \sum_{p\in \bullet t} f_{pt} -|{}^{\bullet}t|+1 \quad (20)$$

$$M=M_0+ \Theta \bullet x, M\geq 0, x\geq 0 \quad (21)$$

$$z_t, f_{pt} \in \{0, 1\}, \forall p\in P, t\in T \quad (22)$$

Eq. (18) expresses the fact if the number of tokens at place p is no less that the weight of arc (p t), then arc (p t) is enabled and f_{pt}=1. Eq. (19) expresses the fact if z_t =1, then arc (t p) is enabled and f_{tp}=1 for each output place of transition t. Eq. (20) expresses the fact if each input arc of a transition is enabled, then transition t is enabled and z_t =1.

However, experimental results do not show correct solution, where M(p'$_2$)>0 and I_D is max-marked. Hence the resulting siphon is not non-max'-marked. This is due to the removal of Eq. (2).

Eq. (14) expresses the fact that if all output arcs of resource place r are enabled, then r must be max-marked and v_r=1. Otherwise, r is not max-marked and v_r =1 or 0. However, as mentioned earlier, one should consider max'-controlled rather than max-controlled siphons. Thus, Eq. (20) should be revised as follows:

$$v_r \geq \sum_{t\in r\bullet \cap Tu} f_{rt} -|r^{\bullet}\cap T_u|+1 \ \forall r\in P_R$$

where f_{pt} is the binary variable for arc (p t) that is enabled if M(p) is larger than W(p t) and T_u [={t_2, t'$_2$} in Fig. 1(b)] is the set of transitions in N_u, the minimal strongly connected resource subnet that contains all resource places in an S^3PGR^2. Eq. (18) for S^3PGR^2 thus remains valid.

Recall that the I-subnet I_S in Fig. 1(b) is separated into two parts: A dead (blue colored) and B live (red colored). Using the manual tracing rules in Section 4, v_p=z_t=1 (resp. v_p=z_t=0) in live (resp. dead) subnets. Thus, it seems that the set of places with v_p =0 is no longer a siphon and the MIP test breaks down. This indicates that z_t should not be judged by the enabling of input arcs of t. Rather, one should reemploy Eqs. (1) and (2) since they hold for any siphon. By Eq. (2), if all input places of t are not in siphon S (i.e., potentially marked), neither is t (potentially firable). On the other hand, if all input arcs are enabled, t is not necessarily outside S. Note that the set of non-max'-marked places in Fig. 1(b) forms a siphon.

$$G(M)=min\sum_{p\in P} v_p \quad (23)$$

s.t.

$$f_{pt} \geq (M(p)- W(p\ t)+1)/SB(p),\ \forall W(p\ t)>0 \qquad (24)$$
$$z_t \geq \sum_{p\in{}^\bullet t} v_p - |{}^\bullet t|+1 \qquad (25)$$
$$v_p \geq z_t,\ \forall W(t\ p)>0 \qquad (26)$$
$$M=M_0+ \Theta\bullet x,\ M\geq 0,\ x\geq 0 \qquad (27)$$
$$v_r \geq \sum_{t\in r^\bullet\cap Tu} f_{rt} - |r^\bullet \cap T_u|+1\ \ \forall r\in P_R \qquad (28)$$
$$v_p,\ z_t,\ f_{pt}\in \{0, 1\},\ \ \forall p\in P,\ t\in T \qquad (29)$$

where the structural bound (SB) is defined as SB(p)=max{ M(p)| $M=M_0+ \Theta\bullet x$, M≥0, x≥0}.

Comparing with the MIP test in [6], among all f binary variables, only some f on N_u appears in MIP, thus reducing the size of the MIP.

LINDO is a linear programming tool to test the above MIP test. The input file to LINDO for the net in Fig. 1 has no binary variables f at all (reducing the size of the problem) as explained below.

For r_1, $|r_1^\bullet \cap T_u|=1$, i.e., there is only one output arc (r_1, t'$_2$) from r_1 that is on N_u. Thus, $v_{r1}=1$ if $f_{r1t'2}=1$ (arc (r_1, t'$_2$) is enabled). On the other hand, if arc (r_1, t'$_2$) is disabled, then $f_{r1t'2}=1$ or 0, which implies that $v_{r1}=0$ or 1. Thus, for r_1 with only one output arc in N_u, Eqs. (24) and (28) can be combined into a single constraint: $v_{r1} \geq (M(r_1) - W(r_1, t'_2)+1)/SB(r_1)$ eliminating the binary variable $f_{r1t'2}$. Similar conclusion applies to r_2: $v_{r2} \geq (M(r_2) - W(r_2, t'_3)+1)/SB(r_1)$.

Note that Eq. (28) only applies to resource place r since for an operation place p, the weight of every output arc is one. Hence, p is max-marked if and only if M(p)>0 and there is no need for Eq. (28) which is used to indicate that if a resource place is max'-marked, then v_r =1.

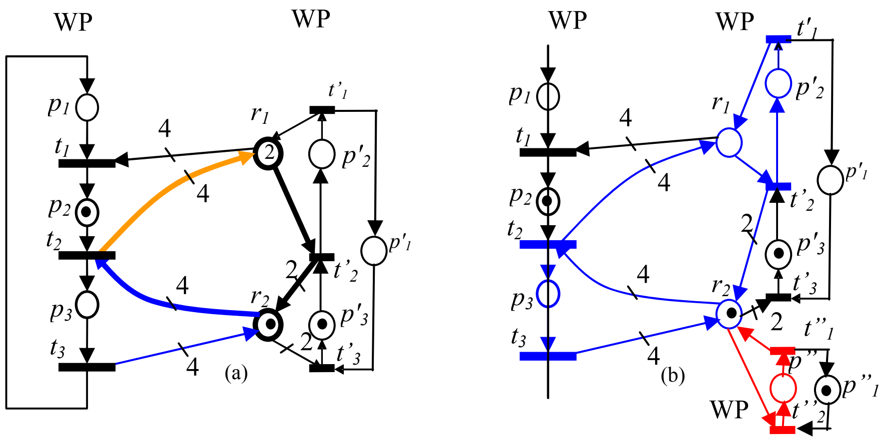


Figure 1. a) An example of weakly live S3PGR2 (M0(p1)=1, M0(p'1)=1, M0(r1)=6, M0(r2)=3). t1, t2 and t3 can never fire. The rest are live. After adding a token to r2, it becomes live. It is live if all transitions are potentially firable from M0 or WP1 and WP2 are quasi-live. b) An example of HPPc = [r2 t''2 p''2 t'1 r2]. S={r1, r2, p3, p5, p8} is not deadly marked (t7 and t8 are live; all the rest are dead).

8 Conclusion

Contributions of this paper are summarized below:

1) Indicated that livelocks may emerge using the MIP test by [6].
2) Demonstrated a counter example.
3) Proposed a revised MIP test to resolve the problem.

The obtained partial siphon using the proposed is, similar to that for ordinary Petri nets, a maximal one.

An extra step is needed to extract minimal partial siphon. In an earlier paper, we proposed a direct MIP test to compute minimal siphons without the extra step. Future efforts should be directed to extending this direct method for ordinary PN to general PN on top of the revised MIP test presented in Section 5.

9 References

[1] F. Chu and X. L. Xie (1997) Deadlock analysis of Petri nets using siphons and mathematical programming. IEEE Trans. Robot. Automat., 13, pp. 793–804.

[2] Huang, Y. S., Jeng, M. D., Xie, X. L., Chung, S. L. (2001) Deadlock prevention policy based on Petri nets and siphons. International Journal of Production Research, 39(2), pp.283-305.

[3] J. Ezpeleta, J. M. Colom, and J. Martinez (1995) A Petri net based deadlock prevention policy for flexible manufacturing systems. IEEE Trans. Robot. Automat., 11(2), pp. 173–184.

[4] Li, Z. W., Hu, H. S., Wang, A. R. (2007) Design of liveness-enforcing super-visors for flexible manufacturing systems using Petri nets. IEEE Transactions on Systems, Man, and Cybernetics, Part C, 37(4), pp.517-526.

[5] Li, Z. W., Zhou, M. C., Uzam, M. (2007) Deadlock control policy for a class of Petri nets without complete siphon enumeration. IET Control Theory and Applications, 1(6), pp.1594-1605.

[6] J. Park and S. A. Reveliotis (2001) Deadlock avoidance in sequential resource allocation systems with multiple resource acquisitions and flexible routings. IEEE Trans. Automat. Contr., 46(10), pp. 1572–1583.

[7] Daniel Yuh Chao (2007) Max'-Controlled Siphons for Liveness of S^3PGR^2. IET Control Theory and Applications, 1(4), pp. 933-936.

[8] K. Barkaoui and J. Pradat-Peyre (1996) On liveness and controlled siphons in Petri nets. In Reisig, (Ed.), "Petri Nets, Theory and Application", 1091 in LNCS. Springer-Verlag, pp. 57-72.

Simplifying Abrasive Waterjet Cutting Process for Rapid Manufacturing

Nguyen Van Ut, Pisut Koomsap [1], and Viboon Tangwarodomnukun

Industrial and Manufacturing Engineering, Asian Institute of Technology, Thailand.

Abstract. Abrasive waterjet machining (AWJ) is a nontraditional machining process offering a good quality cut with low heat affected zone, but similar to other CNC machines, AWJ machine is not flexible and not user friendly. Many lines of G-code commands are required for generating tool path to perform any contour cut. The more complex the contour is, the longer and more complicated the tool path program will be. Consequently, working on a new contour, unavoidably, consumes time, is prone to error, and requires experience. This research focuses on improving man-machine interface that allows inexperience users to complete cutting any complicated contours in very short period of time without writing a single G-code. A two-axis servo table with a user-friendly LabVIEW-based program that allows a user to input commands by dragging a mouse to draw a desired shape has been developed and tested by cutting ceramic tiles. The cut quality is comparable, but tool path generation is effortless. This research makes it possible to achieve several quality cuts of different contours in minutes rather than in days.

Keywords. Abrasive waterjet, Man-Machine interface, LabVIEW.

1 Introduction

Abrasive waterjet machining (AWJ) is a nontraditional machining process where its main activity is in cutting process for a board spectrum of materials ranging from thin flexible fabric to thick hard granite. AWJ delivers a good quality cut with low heat affected zone comparing to other contact and non-contact machining processes; therefore, it is suitable for applications that are sensitive to thermal deformation, and for new material with low machinability.

This non-contact process removes material through erosion process. Trains of high velocity abrasive particles impinge onto the surface of workpiece, and remove material through shear deformation process. By control the movement of the

[1] Industrial and Manufacturing Engineering, School of Engineering and Technology, Asian of Institute of Technology, Pathumthani 12120, Thailand, Tel: +66-2-5245678: Fax: +66-2-5245697; Email: pisut@ait.ac.th; http://www.ait.ac.th

abrasive waterjet, drilling, cutting, turning or even milling can be performed, and majority of researches on AWJ machining has been focusing on determining the effect of process parameters on material removal rate, penetration rate and cut-wall surface finish [1-5].

Similar to other machining processes, AWJ machining process is performed by controlling relative motion between nozzle (cutting tool) and workpiece. Tool paths are generated by curve tracing method. Linear interpolation is applied for segmenting curve where inherent conflict between accuracy and efficiency need to be balanced. The results are converted to be an NC program, an ordered sequence of G (preparatory function) and M (miscellaneous) codes, to control machine operations. G-code, a 50-year-old international standard, provides one-way information flow that focuses on programming the path of the cutting tool with respect to the machine axes, rather than the machining tasks with respect to the part [7]. Standard code numbers can be used to perform simple tasks but may be difficult for complex task; therefore, additional codes have been created for specific features [8] but they may vary from vendor to vendor.

Even though hardware-based CNC systems are well accepted for machine control, but its operating and maintenance costs are high. Spare parts may not always be available for old systems, and part interchangeability between manufactures is a concern. Furthermore, CNC code generation is not trivial. These reasons open an opportunity for personal computer to invade into this field. Rapid advancement of computing technology in recent years makes personal computer more powerful, more reliable, and more robust even in harsh environment, like on factory floor. Its open architecture removes manufacturer proprietary that leads to low system and maintenance costs, and allows third party to support any specific applications. PC-based CNC systems have been developed to perform various diverse tasks needed for control many machines (e.g., PC-based NC milling machine [8]. Hace and Jezernik [9] developed a PC-based CNC system for waterjet cutting machine by incorporating the basic waterjet control routines in the CNC package that run on QNX operating system. Two-processor architecture was implemented because CAD/CAM software used to generate NC program runs on Windows OS.

This paper presents a development of a user-friendly man-machine interface that allows inexperience operators to complete cutting any complicated contours in very short period of time without writing a single G-code.

2 PC-Based Two-Axis Servo Table

Currently, for operating the waterjet machine in CIM laboratory at Asian Institute of Technology, a user is required to write several lines of G-codes to direct abrasive waterjet to cut a planar contour on a workpiece that is put stationary on a machine table, the assembly of consumable grid slabs. The more complex the contour is, the longer and more complicated the tool path program will be. Consequently, working on a new contour, unavoidably, consumes time, is prone to error, and requires experience. Therefore, the main purpose of this research is to improve the usage of the waterjet machine by making it easier to be used so that inexperience users can operate and complete cutting any complicated contours with no difficulty. Furthermore, this attempt should also be applicable on other waterjet machines with minimum modification.

After considering all these requirements, a PC-based two-axis servo table has been developed as a separate module with its own controller. The architecture of this module is illustrated in Figure 1. With this architecture, the waterjet machine and the two-axis servo table are controlled independently, but their communication can be made for synchronization. Consequently, this design is not restricted by any particular model or manufacture of a waterjet machine.

The designed unit composes of two major components: a two-axis servo table put on the waterjet table, and a G-code free control program. Except for cutting profile, and feed rate that are set on the control unit of the servo table, other parameters (e.g., pressure and stand-off distance) are still set on the waterjet controller. A LabVIEW-based control program has been developed to control the movement of the XY table. The program has two parts: front panel for human-machine interface and diagram for the table controller. Graphical representation on the front panel allows the user to input cutting profile, feed rate, and a starting point. Depended up on a desired part, the starting point will either be inside the contour for obtaining the outside part or be outside the contour for the inner part. Five options are available for generating the cutting profile: line, polyline, circle

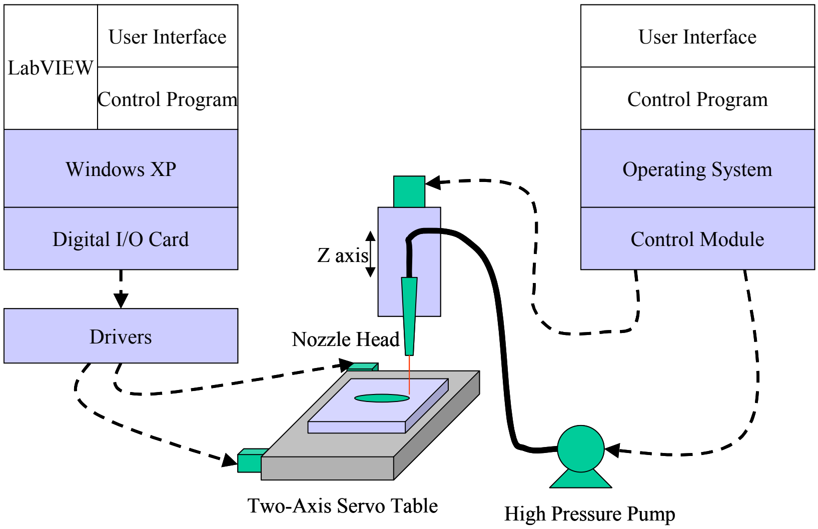


Figure 1. Architecture of the proposed PC-based waterjet machining system

with known radius, circle with points on contour, and arbitrary profile. For arbitrary profile, for example, the user can input a profile by dragging a mouse to draw a desired shape (Figure 2). The LabVIEW has been programmed to record the current positions of the cursor. The profile input is sent to a path planner to generate setpoints and sequence them in a list for point-to-point execution. Linear interpolation is used for 2D movement to calculate pulse ratio between the axes. Pulse trains are generated and sent through digital I/O board to drive the XY table immediately after the user confirms the input.

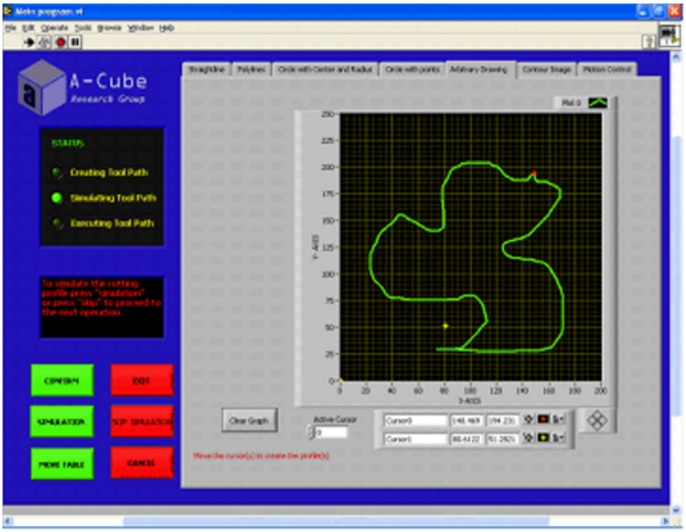


Figure 2. Front panel for arbitrary profile

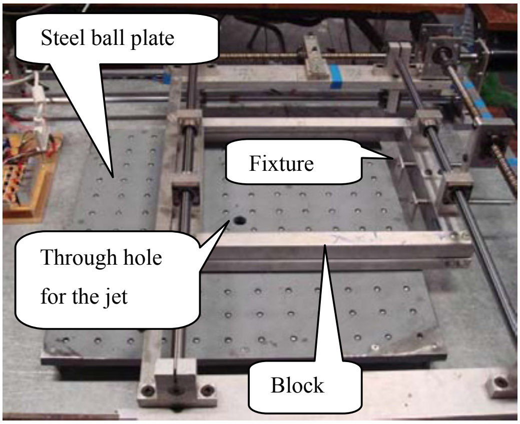


Figure 3. A two-axis servo table

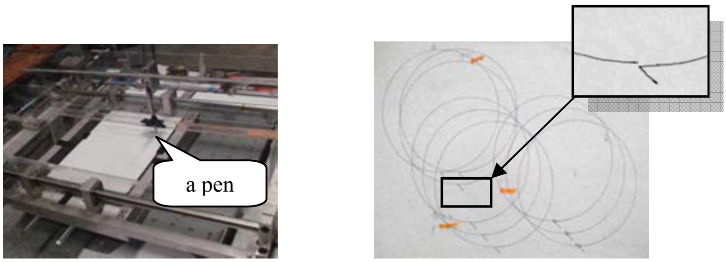
The table consists of a block, fixtures, steel ball plate, base plate, ball screws, linear rods, and stepping motors and their drivers. As illustrated in Figure 3, the table was designed in the way that a workpiece is sit on the steel ball plate inside the block and held in place with the simple screw-type fixtures. The block is driven by two stepping motors through the ball screws, and its movement is confined on the XY plane by the four linear rods. These ball screws and rods are mounted on the base plate. The steel ball plate is inserted between the block and the base plate to support the workpiece. Since the workpiece is moved relative to a stationary waterjet nozzle, grid slab consumption is reduced. The table can also be used to move a tool as illustrated in the next section.

3 Experiments and Results

A prototype of the PC-based two-axis servo table was constructed, and put in test for proof of concept. Experiments were conducted in two sets to evaluate the performance of this prototype. Dry runs were performed first to check the accuracy of the table before experimenting with abrasive waterjet. The experiments and results are presented in the following subsections.

3.1 Dry Run Experiments

The main purpose of this set of experiments was to check the table itself before integrating with the waterjet machine. As shown in Figure 4a, a pen, representing the waterjet nozzle, was fixed at the zero point on the block, and touched on a paper below. By moving the block; a line was drawn on a paper. Since the pen was moved on a stationary paper, the drawings obtained were the mirror of what should appear on parts cut with waterjet. Various shapes such as circles, lines, zigzag lines and other arbitrary shapes were tested with 10 runs each to meet 95% confidence interval aspect. Some of the results are shown in Table 1. Inherent backlash contributed error to straight-line movement. Number of points visited and rounding off of calculated pulse ratio caused error on 2D movement to be higher than straight line. The difference in pitch length between the two ball screws contributed to curve distortion that can be noticeable in the form of an offset between starting and end point as shown in Figure 4b.



a. Experimental setup b. Experimental results

Figure 4. Dry run experiment

Table 1. Results from dry run experiments

Shape (boundary)	Setting value (mm)	Avg. Measured value (mm)	Error (mm)	%Error
Straight Line	100	99.6	-0.4	0.40
Circle	∅ 120	∅ 119	-1	0.83
	∅ 100	∅ 98.68	-1.32	1.32
	∅ 80	∅ 79.4	-0.6	0.75
	∅ 40	∅ 39.4	-0.6	1.5

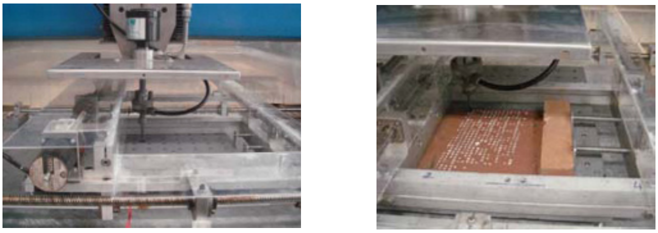
3.2 Experiments with Abrasive Waterjet

In this step, the prototype was integrated with the waterjet machine to cut ceramic tiles. During the setup period, the nozzle was moved to align with the axis of the through hole on the table. A tile was then loaded and mounted in the block where the bottom of the tile sits on the steel ball plate. The nozzle was lowered to the stand-off distance of 3mm, and a tile was moved to align a starting point with the nozzle tip. For all runs, the pressure was fixed at 800 bars, and the feed rate was set at 100 mm/min. The jet diameter was about 0.6 mm. Similar to dry run experiments, various shapes were input and tested. The results, presented in Table 2, show that there is small variation between two approaches, and from visual inspection, cut-wall surfaces were also comparable.

Figure 6 shows the comparison between a contour cut by abrasive waterjet and a contour drawn in dry run experiment. It can be seen that all measured diameters of the cut contour are smaller than those diameters of a drawn contour at the same positions because the jet diameter was not offset. In addition, small errors, occurring in dry run, were amplified by the jet diameter as can be seen from the offset between starting point and end point that enlarge by 40%.

Table 2. Abrasive waterjet cutting results from two control approaches

Shape Profile	Set Value (mm)	Measured Value (mm)	
		Waterjet Controller	Two-Axis table
Line	100	99	98.6
Square	100x100	99x99	98.4x98.6
Circle	∅100	∅99	∅98.2



(a) Setting up the two-axis table on the waterjet machine

(b) Load ceramic tile on the table

Figure 5. Experimental setup with waterjet machine

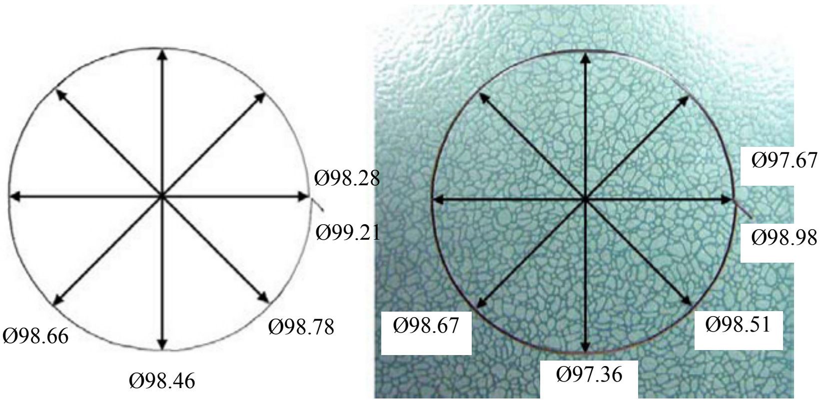
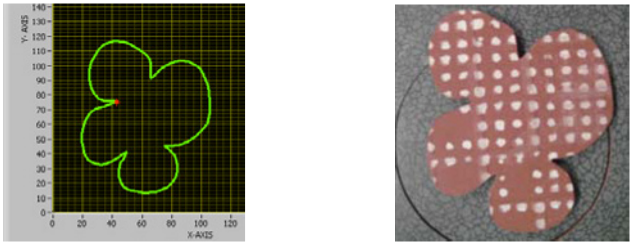


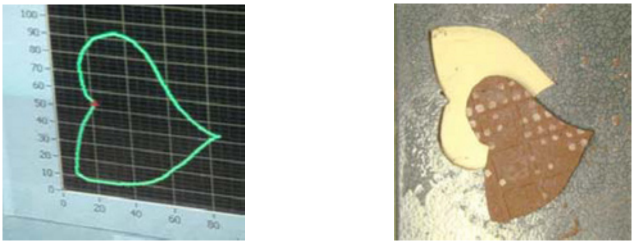
Figure 6. Comparing the results of dry run and AWJ cutting

4 Conclusions

The PC-based two-axis servo table has been researched with the objective to make the operation of a waterjet cutting process simple, and its first prototype has been constructed and tested with the waterjet machine. Its cut results were evaluated and compared with the similar parts obtained by generating G-codes for operating the machine. Despite of larger errors, comparing to running G-codes, the results obtained from the prototype shows promising future. User friendly interface not only shorten production time but also allow inexperience user to complete task with no difficulty. It is expected that by minimizing mechanical errors and implementing offset function, the table will produce much more accurate results. To make it more useful, channels for input cutting profile will be extended in the near future to accept a contour image and 2D drawing from commercial software.



(a) Arbitrary profile 1 (b) Part obtained for arbitrary profile 1



(c) Arbitrary profile 2 (d) Part obtained for arbitrary profile 2

Figure 7. Experiments on arbitrary profiles

5 References

[1] Arola D, Ramulu M. Material removal in abrasive waterjet machining of metals surface integrity and texture. Wear 1997; 210: 50-58.

[2] Chen FL, Siores E. The effect of cutting jet variation on striation formation in abrasive water jet cutting. International Journal of Machine Tools and Manufacture 2001; 41: 1479-1486.

[3] Hashish M. A model study of metal cutting with abrasive waterjets. ASME Journal of Engineering Materials and Technology 1984; 106: 88-100.

[4] Lemma E, Chen L, Siores E, Wang J. Optimising the AWJ Cutting Process of Ductile Materials Using Nozzle Oscillation Technique. International Journal of Machine Tools and Manufacture 2002; 42(7): 781-789.

[5] Chen FL, Siores E, Patel K. Improving the cut surface qualities using different controlled nozzle oscillation techniques. International Journal of Machine Tools and Manufacture 2002; 42: 717-722.

[6] Reynolds TD, Richards P. Unit Operations and Processes in Environmental Engineering. PWS Publishing Company, MA, 1996.

[7] Xu XW, Newman ST. Making CNC machine tools more open, interoperable and intelligent – a review of the technologies. Computers in Industry 2006; 57: 141-152.

[8] Omirou SL, Barouni AK. Integration of new programming capabilities into a CNC milling system. Robotics and Computer-Integrated Manufacturing 2005; 21: 518-527.

[9] Hace A, Jezernik K. Control system for the waterjet cutting machine. IEEE/ASME Transactions on Mechatronics 2004; 9: 627-635.

Hybrid System Supporting Flexible Design of Flat Rolling Production Processes in Collaborative Environment

Lukasz Rauch[a,1], Michal Front[a], Marek Bigaj[a], Lukasz Madej[a]

[a]Department of Applied Computer Science and Modelling, AGH – University of Science and Technology, Krakow, Poland

Abstract: The paper is devoted to advanced production processes design, based on numerical simulations of material behaviour under complex loading conditions. The computer system proposed in this work facilitates creation of the sophisticated flat rolling facilities composed of different subesequent stages e.g. heating, roughing and finishing mills, cooling, cutting, descaling. Each stage is treated as a separated module with its own features and methods that implement its functionality. However, the most demanding part of proposed system lies in reliable simulation of connection between these separated modules. To deal with this the highly fexible numerical solutions are required. Creation of this appro ach is the main goal of the work and is described in details including examples obtained results. Disscusion on accurate material models taking into account dynamic recrystallization or grain growth as well as on application of the optimization procedures inorder to obtain desired final properties is also presented in the paper.

Keywords. digital manufacturing, production design, flat rolling, optimization

1 Introduction

The ever-changing global market is facing modern steel manufacturers into a state of continuous adaptation to meet new customer expectations. In steel industry the problem of obtaining required exploitation properties of products is crucial. Such parameters as i.e. fatigue resistance, wear resistance and thermal resistance, but also environmental and recycling problems, are crucial for both increasing the safety by extending the life cycle of product and for meeting the environment protection requirements. On the other hand, the life cycle of products becomes shorter due to rise and fall of demand and to increasing demand for customized products. To satisfy these demands new steel grades are being developed in laboratory conditions and then new technology have to be transferred to the manufacturers of the raw steel material for further processing i.e. foundries or steel

[1]Corresponding Author : AGH – University of Science and Technology, Krakow, Poland, Department of Applied Computer Science and Modelling; Tel: +48 12 6172921; Fax: +48 12 6172921; Email: lrauch@agh.edu.pl; www.isim.agh.edu.pl

mills. The steel industry in order to meet such market pressures and fast progress in new steel grades development requires efficient numerical tool for planning and organization of the production system, which is flexible, reconfigurable and cost efficient [13]. Progress in the work towards development of such computer system is presented in the following sections.

2 State of the art

The system proposed in this work can be classified into a group of decision support systems, based on numerical calculations and artificial intelligence [17,18]. They can be applied to Production Planning Systems (PPS) and play important role in everyday business life [2]. These systems manage also the supply chains, optimize work of employees, maximize incomes or plan and rearrange the warehouses space. Commercial versions of such systems were developed in early 90's of the last century and are based on the task schedulers or the management of Gantt charts [11]. This simple functionality was constantly improved by implementation of the following ideas [3]: mass production, flexible manufacturing, computer integrated manufacturing, lean manufacturing and Material Resource Planning (MRP). These steps were necessary to achieve the milestone in the lifecycle of PPS, i.e. conversion from MRP to Enterprise Resource Planning (ERP). Further evolution of the functionality of the PPS systems focused on implementation of the methods supporting concurrent engineering and, finally, to agile manufacturing. The latter systems were created to manage production processes held in unstable environment and to satisfy individual fast changing customer needs.

Following the PPS development, the systems were modified and equipped with the algorithms based on the artificial intelligence and soft computing, to create the Intelligent Manufacturing Systems (IMS). One of the first examples, proposed by Giachetti in 1998 [5], was based on the formal multi-attribute decision model and the relational database. This system was helpful in selection of materials and manufacturing processes, however its functionality became out of date very fast. Nowadays, such approaches often use the expert systems [10] or the knowledge bases [6]. The first proposed a framework to create the customize rule based system, using semantic net structure, where calculation of semantic hulls allows to obtain solution and to determine the optimal decision. The second suggested to create the knowledge-based “road-map”, which facilitates decision making in production planning by introducing the flexibility and dynamics of the manufacturing process.

As presented the lack of the complex computer systems, which employ the numerical simulations, optimization methods and knowledge base at the same time, can be noticed. This fact and the needs of metal forming branch inspired Authors to create the system based on the numerical simulations of industrial processes combined with the recent scientific knowledge [16]. This system is able to predict final properties of products and optimize the manufacturing process. To predict mentioned properties, the system uses production chain modelling approach, combined with the optimization algorithms. However, in Authors earlier attempts the system was dedicated to one specific flat rolling process without possibility of easy reconfiguration to other rolling facilities. On the other hand, there exists

HSMM system offered by INTEG Process Group Inc. [20]. This system very accurately addresses needs of flexible production process design, however material behaviour is mainly based on the rheological models and microstructural evolution models are still too simplified to satisfy requirements of metallurgists and material scientists. Overcoming this limitation became inspiration to combine the functionality of rolling system created by the Authors [16] and HSMM software [20]. Thus, the main objective of the work is to create computer system facilitating the work of technologists by functionality of flexible design of production processes supported by numerical simulations as well as quantitative optimization methods. Similar solutions presented in scientific literature e.g. [12] also possess huge functionality, which is additionally extended with application of multi objective optimization methods allowing consideration of qualitative factors. However, most of such solutions can not be probably implemented directly for industrial purposes and put into practice without crucial modifications. Such approaches are usually limited to multi-software architecture or composed of many separated implementations of various algorithm. Such solutions force users to switch the data, stored in external input-output files, between different computer programmes, which highly impedes the usage of the software. In this paper the compact computer system is proposed, where all the models, input data and numerical algorithms are composed into one application. This facilitates the work with data improving final efficiency of technologists and a company.

3 Main idea description

The presented literature review revealed lack of software able to simulate flexible multistage production processes in metal forming branches. Thus, the main idea of this work is focused on creation of computer system, which would be able to simulate production processes designed on the basis of collections of separated elements.

Each of designed elements is related to various production devices e.g. rolling table, finishing mill, descaler, coiler, etc. (Figure 1). However, establishing a connection and common influence between two or more elements is the main problem of such an approach. It is caused by highly sophisticated relations, which occur especially during rolling of long products e.g.:

- long products (up to 600m long) can be formed in different places at the same time – while the head of material is inside a coiler with $v=6.5m/s$, the tail still moves ($v=0.6m/s$) under encopanel (thermo tunnel dedicated to preserve material temperature). In fact, each material point has to be simulated within different boundary conditions in rapidly changing temperature,
- in the case of n mills composed into one stage of rolling e.g. finishing rolling, the material is rolled in each of them at the same time. Therefore all devices' velocities depend on the velocity established at last rolling device. Thus, the processes designed flexibly should be equipped with numerical mechanisms able to cope with such constraints,

- additionally introduced parameters like acceleration (constant or variable) drastically increase calculation complexity of numerical algorithms, influencing distribution of temperature.

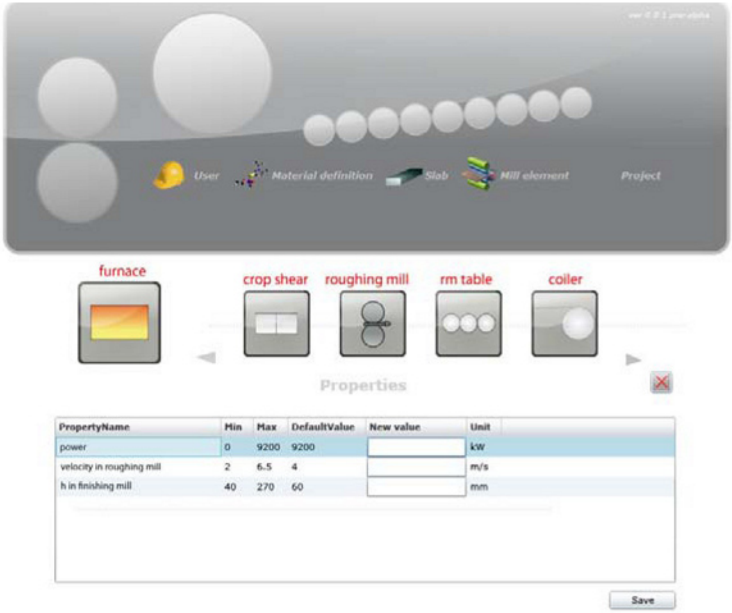


Figure 1. Example of screen presenting properties of roughing mill.

Most of the mentioned problems are solved in proposed system by the object oriented representation of production devices, which are composed of characteristic properties and methods describing their functionality. All the devices have a set of common parameters, which determine behaviour of the subsequent stages of manufacturing chain. Due to such approach final user obtains flexible tool to design production process built of smaller pieces, then he can describe their properties and simulate sophisticated rolling schedules. Additionally, optimization procedures can be used for the devices groups to obtain rolling schedule that will provide desired properties of the final product. More specific description of these possibilities is presented in the next section.

4 Implementation details

4.1. Design

The system is designed on the basis of layered architecture (Figure 2). The top-most tier is User Interface part, containing all administration logic and panels to manage rolling mill elements, their properties, material groups, chemical compositions of strips, system users etc.

User Interface contains also the Visual Rolling Mill Designer, which facilitates design of a whole production process using drag and drop method. User can drag available mill elements, drop them in the proper place of the rolling facility or rearrange them in desired way. Afterwards, selected properties for each separated mill element or groups of devices can be established. User Interface layer cooperates closely with the finite element (FE) Library, passing all the details of rolling mill model together with the required properties of used material. Due to received input data, FE library simulates the production process by using FE method. During last decade FE was broadly used in industrial environment to simulate various forming operations (e.g. cold forging, hot forging, rolling, extrusion, stamping) and can be treated as most popular and well known numerical approach. The FE simulations provide information regarding changes in shape, stress, strain and temperature distribution, and rheological models of materials. These models are used for calculation purposes and describe relations between the flow stress and some external variables like strain ε, strain rate $\dot{\varepsilon}$ or temperature T. All calculated results are returned is structured form to visualization module in the upper tier.

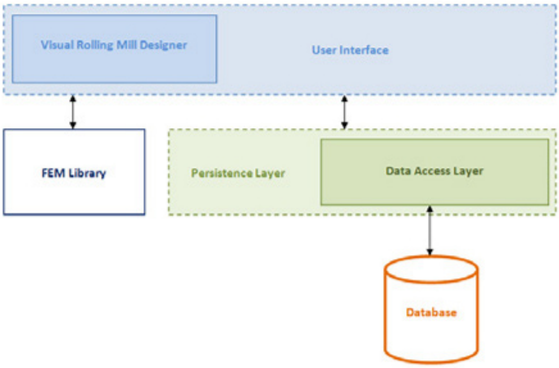


Figure 2. Architecture of proposed system

To load or save project data the User Interface layer communicates with Persistence Layer. This layer was created using NHibernate framework, which allows implementation of the Object-Relational mapped and persistent classes. It uses data structure descriptions written in XML, thus functional objects inside the system are converted directly to the database tables. The conversion is performed in Data Access Layer. Additionally, the state of loaded object is stored during the whole lifetime of that object and used in further connections to database. Such solution provides higher efficiency by optimizing database queries and minimising amount of loaded data.

The database stores all project and additional data i.e. material groups with constraints of chemical composition, materials with specific chemical composition, differences in chemical composition between default material and material defined for project purposes, all rolling mill elements' definitions, types, properties, designed rolling mill models and knowledge in form of topics. The database was designed to be as flexible as possible to facilitate the development of the system's functionality by adding new rolling devices. New elements can be added only on the basis of other already existing devices, inheriting all its characteristic features and methods (including FE methods). This approach assures the possibility of using newly created elements in reliable numerical simulations.

The software is dedicated to collaborative environment, thus it was designed to work in Client-Server architecture with "thick" Server side responsible for systems' logic and FE calculations, and Client side based on Internet browser. System was developed using Microsoft .NET 3.5 in C# language, while Visual Rolling Mill Designer was created in Silverlight 2.0 framework dedicated for so called Rich Internet Applications. It is one of the newest Microsoft activities supporting JavaScript/Ajax programming model as well as implementation of the applications without necessity of using client-side code. The database was designed in Microsoft SQL Server 2005, however it can be easily transferred to other database engines such as PostgreSQL or Oracle. Numerical simulations module is implemented in unmanaged C++ and compiled into dll library, which is required to obtain high efficiency of calculations.

4.2 Used models and procedures

A number of mathematical models are used for the calculation of temperatures, forces, microstructure evolution, phase transformations and mechanical properties of material. These models can generally be classified into two interconnected, areas:

- Thermo-mechanical, including: temperature evolution in the work piece roll bite parameters (flow stress, strain, strain rate, rolling force), motor torques, powers and load ratios, production rates, shape,
- Microstructure/properties, including: recrystallization (dynamic, meta-dynamic, static), grain growth after recrystallization, precipitation, phase transformations, ferrite grain size, yield strength, tensile strength, hardness, elongation.

For the FE calculations, available in literature flow stress models and conventional microstructure evolution models are used [7,8,14,15,19]. However, in the future also an advanced alternative multi scale approach can be developed based on the hybrid CAFE method [1,4,9]. The multi scale models extend the computational time, however they provide very accurate results with graphical illustration of the microstructure morphology in the selected locations.

Additional advantage is the possibility of application of optimization procedures for groups of elements. Two procedures were implemented for the purposes of the system i.e. Simplex and Rosenbrock. They are used to establish constant acceleration in the group of rolling devices to obtained stable temperature along the material length. This functionality is crucial to receive stable material properties. It is still in the phase of development, however obtained results proves its powerful capabilities.

5 Results

Several different configurations of the rolling mill were tested in the developed system. Obtained results were validated with the measurements of real production processes in the ArcelorMittal company. Examples of the system capabilities are presented in Figure 3 and 4.

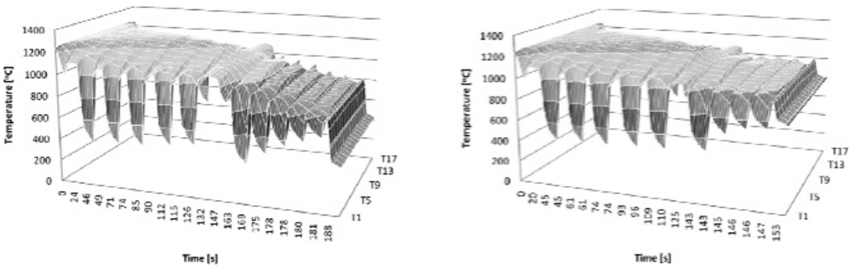


Figure 3. Distibution of temperature as an effect of tracing of material head slice in case of two schemes: (a) 5RE5F with cooling and (b) 5R4F without cooling.

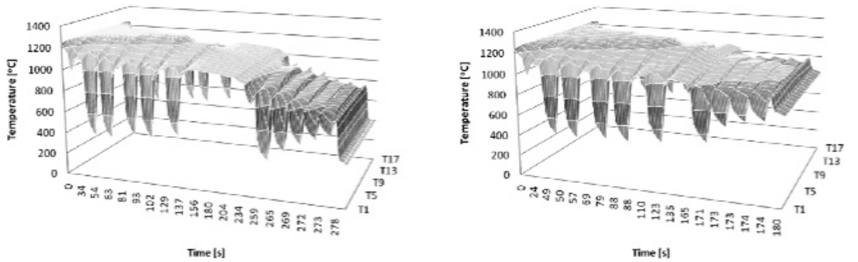


Figure 4. Distibution of temperature as an effect of tracing of material tail slice in case of two schemes: (a) 5RE5F with cooling and (b) 5R4F schemes without cooling.

This section presents only two selected cases regarding 5RE5F and 5R4F configurations (5RE4F means 5 passes in roughing mill equipped with edger followed by 5 finishing mills, 5R4F means5 passes in roughing mill followed by 4 finishing mills). The analyzed cases are also equipped with furnaces, descalers and encopanels. Figures 3 and 4 show 3D plots of temperature distribution obtained for mentioned configurations for head and tail slices of material respectively. Clearly visible differences between the results prove high capabilities of created models as well as their universality in flexible design of manufacturing processes.

6 Conclusions and future plans

The paper describes sophisticated computer systems able to simulate flexibly designed flat rolling production processes. The system offers rich functionality allowing users to manage manufacturing process built of separated elements, which are related to specific production devices. Large number of problems with simulation of different configuration of elements and their common interactions were solved by the object-oriented implementation and usage of FE calculations. Obtained results proved high efficiency and universality of the software, encouraging to develop another models of materials and devices. That way due to flexible 'drag and drop' interface, the proposed approach will not only be limited to simulation and development of the rolling manufacturing chain. When numerical models of other manufacturing devices are created (e.g. forging, stamping, drawing), this system can be used to design and simulate other forming processes. These will be the main objective of the future plans. Moreover, the emphasis will be put on the development of optimization procedures in connection with the advanced material models and multi scale simulations.

Acknowledgements. The financial support of the Polish Ministry of Science and Higher Education, project no. 10.10.110.853 is acknowledged.

7 References

[1] ALLIX O. Multiscale strategy for solving industrial problems. Comp. Meth. Appl. Sci. 2006;6;107-126.
[2] BRUCCOLERI M, LO NIGRO G, PERRONE G, RENNA P, NOTO LA DIEGA S. Production planning in reconfigurable enterprises and reconfigurable manufacturing systems. Annals of the CIRP 2005;54;433-436.
[3] CHENG K, HARRISON DK, PAN PY. Implementation of agile manufacturing – AI and Internet based approach. J of Mat Proc Tech, 1998;76:96-101.
[4] GAWAD J, PIETRZYK M. Application of CAFE multiscale model to description of microstructure development during dynamic recrystallisation. Arch. Metall. Mater. 2007;52;257-266.
[5] GIACHETTI RE. A decision support system for material and manufacturing process selection. J of Intell Manuf 1998;9:265-276.
[6] HALEVI G, WANG K. Knowledge based manufacturing system (KBMS). J of Intell Manuf 2007;18:467-474.
[7] HODGSON PD, BROWNE KM, COLLINSON DC, PHAM TT, GIBBS RK. A mathematical model to simulate the thermomechanical processing of steel, Quenching and Carburizing, Melbourne, 1991, 139-159.
[8] LENARD JG., PIETRZYK M, CSER L. Mathematical and physical simulation of the properties of hot rolled products, Elsevier, Amsterdam, 1999.
[9] MADEJ L, HODGSON PD, PIETRZYK M. Multi-scale rheological model for discontinuous phenomena in materials under deformation conditions. Computational Materials Science 2007;38;685-691.
[10] MAHL A, KRIKLER R. Approach for a rule based system for capturing and usage of knowledge in the manufacturing industry. J of Intell Manuf 2007;18:519-526.

[11] MCKAY KN, BLACK GW. The evolution of a production planning system: A 10-year case study. Computers in Industry 2007;58:756-771.
[12] ODUGUWA V, ROY R, FARRUGIA D, Development of a soft computing-based framework for engineering design optimisation with quantitative and qualitative search spaces, Applied Soft Computing, 2007;7:166–188.
[13] PEREIRA J, PAULRE B. Flexibility in manufacturing systems: a relational and a dynamic approach. European J. Operational Research 2001;130:70-82.
[14] PIETRZYK M Finite Element Based Model of Structure Development in the Hot Rolling Process, Steel Res., 61, 1990, 603-607.
[15] PIETRZYK M, KUZIAK R. Coupling the Thermal-Mechanical Finite-Element Approach with Phase Transformation Model for Low Carbon Steels, 2nd ESAFORM Conf. on Material Forming, ed., Covas J., Guimaraes, 1999;525-528.
[16] RAUCH Ł, MADEJ Ł, MATUSZYK PJ. Knowledge based optimization of the manufacturing process supported by numerical simulations of production chain, Collaborative Product and Service Life Cycle Management for a Sustainable World, 15th ISPE International Conference on Concurrent Engineering: 2008, Bangor, Springer, 435-442, 2008.
[17] RAUCH L, MADEJ L, PIETRZYK M. Hybrid system for modeling and optimization of production chain in metal forming. J. Machine Eng. 2008; 8: 14-22.
[18] RAUCH L, MADEJ L, WEGLARCZYK S, PIETRZYK M. System for design of the manufacturing process of connecting parts for automotive industry. Archives of Civil and Mechanical Engineering, 2008; 8: (in press).
[19] SELLARS CM. Modelling Microstructural Development During Hot Rolling, Mat. Sci. Techn., 6, 1990, 1072-1081.
[20] WWW site of INTEG Process Group Inc., HSMM software, http://www.integpg.com/Products/HSMM_Functionality.asp, 2009.I.

Collaborative Engineering

Organization and Interoperation of Engineering Design Services in Service-Oriented Architecture

Nan Li[a,1], Jianzhong Cha[a] and Wensheng Xu[a]

[a]School of Mechanical, Electronic and control Engineering, Beijing Jiaotong University – BEIJING, China.

Abstract. Current CAD/CAE systems are not ideal for engineering design in distributed knowledge and intelligent resources. This paper investigates the approach of distributed collaborative design in service-oriented architecture. Service-ORiented Computing EnviRonment (SORCER) is employed to form a federation of distributed services that provide engineering data, applications and tools on a network. In order to implement interoperability between design services, standardized operation interface and design information transfer protocol should be developed for target domain. This paper introduces a standardized engineering layout design service interface to demonstrate the mechanism. We develop a flexible software architecture to organize services, to manage design process and to monitor design status in real time. This system aims to enable asynchronous distributed collaborative design with ease of alternative design services, reduced design cycles, and improved design resolution quality.

Keywords. Design service organization, design service interoperation, layout, service-oriented architecture.

1 Introduction

With the recent occurrence of collaborative complex product design among designers, manufacturers, suppliers and vendors is one of the keys for designers to improve product design quality, reduce cost, and shorten design cycle in today's global competition.

A critical success factor for engineering design in SOA (Service-Oriented Architecture) is the ability to reuse existing design assets in the architecture. Due to that, SOA-based design becomes a useful approach to solve the problem above mentioned.

One of the first SOA application in engineering design was Federated Intelligent Product Environment (FIPER) [1], in which engineering tools like CAD, CAE, PDM, optimization, cost modeling, etc., act as federating service

[1]Corresponding Author : Ph.D. Candidate, School of Mechanical, Electronic and control Engineering, Beijing Jiaotong University, No. 3 of Shangyuan Residence Haidian District in Beijing, China; Tel: 8610-51685335; Email: 06116257@bjtu.edu.cn; http://www.bjtu.edu.cn/en

providers and service requestors. The Service-Oriented Computing Environment (SORCER) [2-4] builds on top of FIPER to introduce a metacomputing operating system.

Compared to traditional stand-alone CAD/CAE system, there are new issues that need to be resolved in distributed design based on service-oriented architecture (SOA). For example,

(1) Design service provider and service requestor: design resources, design knowledge and design tools should be wrapped as service providers or service requestors, so that they can work in distributed computing environment based on SOA.

(2) Service registry, service lookup and service proxy.

(3) Service organization and management.

(4) System security.

(5) Design Interface and information transfer protocol.

Due to SORCER can deal with most of issues abovementioned, we investigates the approach of design services organization and interoperation on top of SORCER platform.

2 Service-Oriented Design Environment

SORCER is a federated service-to-service metacomputing environment that treats service providers as network objects with well-defined semantics of a federated service object-oriented architecture [2].

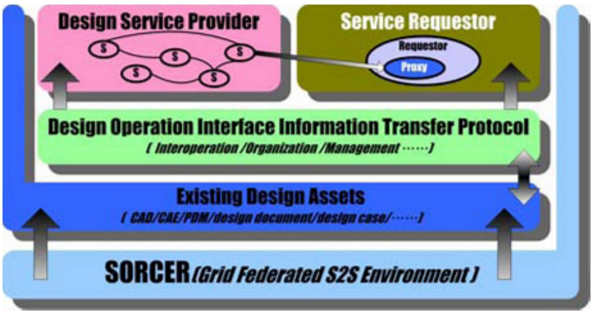


Figure 1. Engineering design based on SORCER.

Figure 1 illustrates the framework of engineering design environment based on SORCER. The design requestors should be wrapped as services so that they can join in the environment. A design proxy—net objects implementing the same design operation interface and information transfer protocol as its service provider—always ready for calling by service requestors. In our environment, the technology detail of service registry, service lookup and service employ will be hid by SORCER. The designer only needs to deal with design services building, services management and design process control.

The interface and protocol are fixed and known beforehand by the provider and requestor. Using our mechanism, a requestor can use this fixed protocol and a service description obtained from a service registry to create a proxy for binding to the design service provider and for remote communication over the fixed protocol.

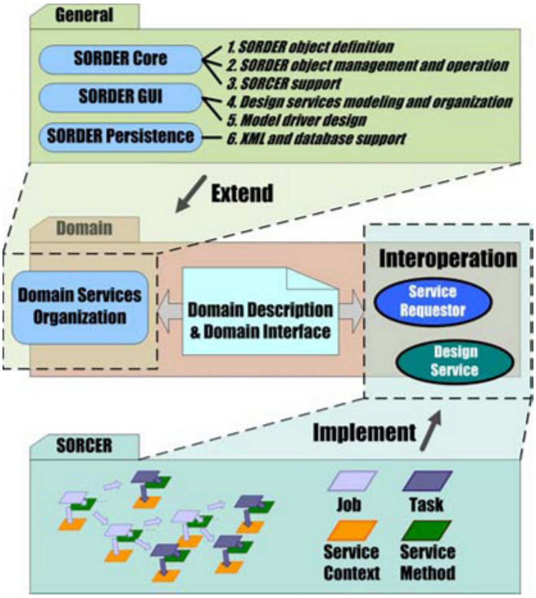


Figure 2. Three parts of SORDER structure.

A whole Service-ORiented Design EnviRonment (SORDER) includes three parts at least, shown in figure 2:

(1) floor-layer SOA technology: SORCER plays an important role in our environment to support floor-layer SOA technology. As shown in figure 2, In the SORCER, two types of basic exertions are defined: tasks and jobs. An engineering design problem can be decomposed to different tasks or jobs (a job can include many tasks) which is the atomic exertion. In our environment, the service context works as design data carrier, and the service method plays a service executer role. A service is an act of requesting a service (exertion) operation from a service provider. An exertion is accepted if its method matches one of provider's interface and one of the interface's methods.

(2) General design environment: domain-independent objects and functions should be defined to support general design process. As figure 2 shown, the SORDER Core can create general design object, like design project object, design services object, design services organization object and so on. The project object include mostly domain-independent design information as project name, user affiliation, date information, project description etc. SORDER Core also need to support SORCER object and SORCER function. In other word, SORDER builds on top of SORCER to support engineering design problem solving. In order to

allow end-user or designer to manager design process and to organize design services, SORDER supplies a general GUI framework for domain application. Due to most of design problems need to process a mass of data and information, SORDER supplies persistence function to support using database and XML.

(3) Domain application: SORCER and SORDER can be used as infrastructure to implement engineering design in a specific domain. With extension of SORDER GUI and SORDER core, domain design expert can use SORDER and SORCER to build domain design environment in SOA. In this structure, domain knowledge can be abstracted to domain interface to implement domain design services. And the detail of domain interface will be discussed in section 3.

A domain service must be a SORCER service first, and two ways are used to build a SORCER service: to wrap general software as SORCER services or to build SORCER applications directly.

3 Service Interoperation based on Standardized Design Interface

Interface and protocol play an importment role in SOA-based design environment. The interoperation of design services depend on standardized design interface and information transfer protocol.

Each service should find match service provider according to the interface. Both design data interoperability and design information commnication need implemention of this fixed interface and protocol. In order to get same kind of service, A service requestor can employ different design service providers, which implement same interface.

In order to explain the function of design interface and protocol, we introduce a standardized engineering layout design service interface to demonstrate the mechanism.

Szykman and Cagan[5] and Kolli et al.[6] defined the general 3-D component layout problem as：

Given a set of three-dimensional objects of arbitrary geometry and an available space (possibly the space of a container), find a placement for the objects within the space that achieves the design objectives, such that none of the objects interferes (i.e. occupy the same space), while satisfying optional spatial and performance constraints on the objects.

According the definition of layout problem, we build the engineering layout design interface and protocol.

The mainly content of this interface and protocol include:

(1) Layout components and containers model format: a standard representation for 3-D layout components and containers modeling. If implemented this interface protocol, a general CAD system can supply layout components modeling service for FLDS as a service provider.

(2) Design parameter: a design service requestor can implement this interface protocol to submit user needs to FLDS.

(3) Layout state description: every service which wants to use layout resolution should implement this protocol. This interface protocol describes all information of layout result.

(4) Layout constrains model format: a standard representation for 3-D layout constrains modeling. A constrain modeling tool should implement this interface protocol to supply constrains modeling service to the FLDS.

(5) Algorithm interface: some algorithms which implement this interface can supply layout optimization service for the FLDS.

(6) Evaluation parameter structure: evaluation service should implement this interface protocol to supply evaluation service for FLDS.

(7) Human-computer interaction command: the command is used to operate some services with GUI in batch mode, for example: modeling command stream is used to build components model automatically on components modeling service.

(8) Multimedia report interface protocol: this protocol support to build a multimedia layout result report. A report service should to implement this interface for custom-built report generation.

As shown in the list, the design interface includes a mass of domain design knowledge. Thus, more information and rules will be added into the protocol structure in the future.

Due to powerful description ability of domain data, SORCER Context [3] and XML are good carrier for the interface and protocol. SORCER Context is used as runtime communication carrier, and XML is employed to be data store format.

4 Design Service Organization

Although the end-users and designers don't need to worry about the detailed running process of each design service which is invoked in whole design process, we also need a mechanism to driver, to monitor, and to organize useful design services.

Figure 3 illustrates the three levels of service invoking. First level is direct service invoking. This level includes important domain design services which implement the domain design interface strictly. These services can supply mainly useful functions in whole design process, and the service requestor can organize them according the design intention. Designer can employ the service organization tool (SORDER GUI supplies this tool) to build service organization model which can drive the design process.

Services in level 2 are still domain design services, but they are invoked by services in level 1. In this case, services in level 1 also work as service requestors. Due to that, services in level 2 are invoked indirectly by the main service requestor, so that they can't be controlled or be driven by the designer. But, the designer still can monitor the run process of services in level 2, that due to the main service requestor also can be used as a special service— monitor service. The monitor service can be invoked by every service in level 1 and level 2 to show the employer service running status.

Services in level 3 don't belong to given domain, but they can be invoked by domain services for some reasons. The unknown services don't follow the domain service interface, thus they can't be monitored by the monitor service we just mentioned. In other words, they are invisible for the designer.

Every service provider can become a service requestor whenever they want to. The multi-role of services imply that to be a service requestor, they can invoke other services (include domain services and non-domain services), and monitor the service running process, and to be a service provider, they can supply service for the other service requestors. It's need to point out that some popular services, like optimization algorithm service, can cover different domain, while necessary domain interfaces are implemented.

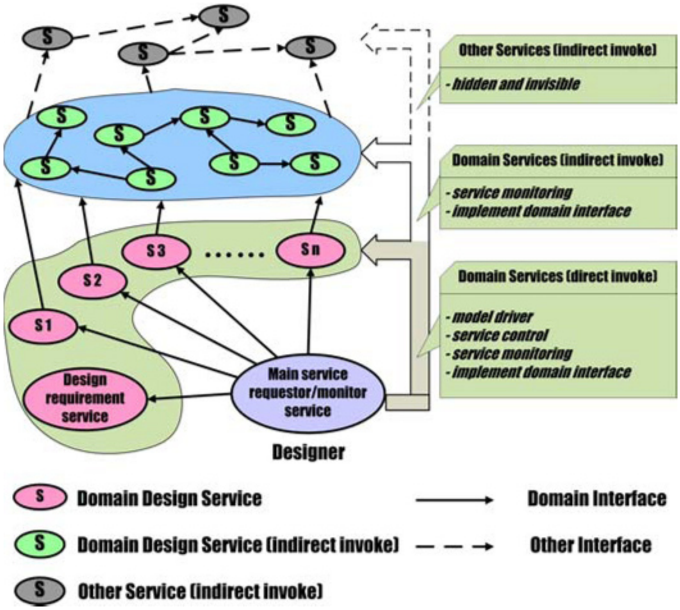


Figure 3. Three levels of service invoking.

5 Engineering Application

We still use engineering layout design to demonstrate the architecture. Whatever knowledge resources and intelligent resources can build their own service according to the layout domain interface, and launch the service to design environment as a component. It is allowable that more than one service can implement same function, and that's lead to services competition. The users or the service employer can choose the best service form all services with same function in the environment. The "best" means best quality, best efficiency, or least cost etc. Figure 5 illustrates that the hierachical services structure of layout design service. Users should play two roles in our system: service provider and service requestor. As service provider, a user should supply layout design requirement to other services in this domain. To be a service requestor, a user can monitor the whole design process and get final layout design result. Every service is autonomous and can call other service in the environment. The employer service don't need to care about what happend in employee services, even though the employee service calls other services either.

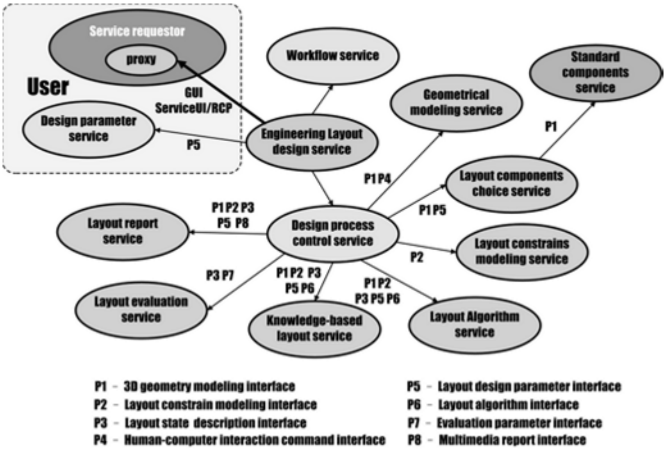


Figure 4. Hierarchical services structure of layout design

As shown in figure 4, the engineering layout design service is an integrated service to supply whole layout design function. The user only needs to call engineering layout design service singly to start layout design process. Every call between services should follow matchable interface protocol.

Some pivotal design tools of layout design were developed and the system was demonstrated through some real engineering application, as vehicle engine compartment layout design [7](As shown in figure 5).

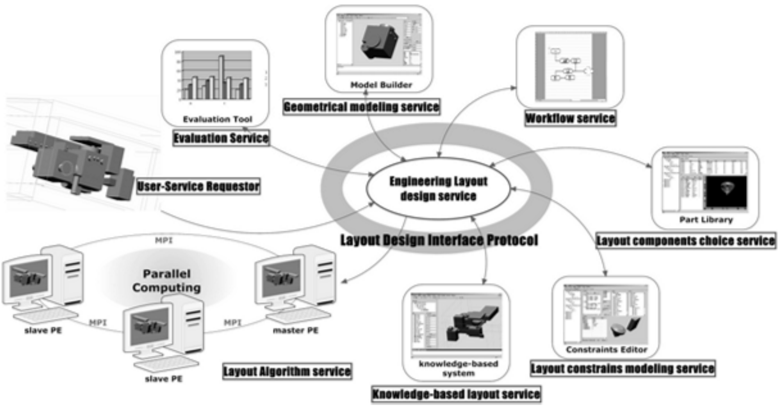


Figure 5. Pivotal services of engineering layout design.

Not all of the services shown in figure 3 are necessary for an engineering layout design task. The users or designers can organize the services and add or delete them according their desire.

We extend SORDER GUI to build a service organization tool (shown in figure 6) for engineer layout design service (shown in figure 4), a designer can create a layout project and build a service organization model to drive the other domain services.

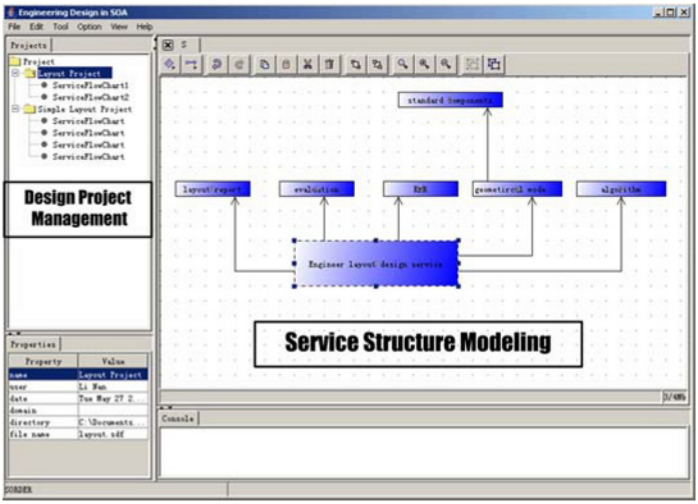


Figure 6. Service organization tool.

A service provider can supply Java-based GUI to service requestor for human-computer interaction. Some simple interaction can be wrapped as SORCER GUI, which can be loaded by Jini [8] service browser—IncaX [9]. In this case, service users only need to run IncaX to call service GUI to get human-computer interaction. In contrast, complicated interaction application should be wrapped to Rich Client Program (RCP) (As knowledge-based layout service shown in figure 5), so that users must run integrated RCP to call the services what they want.

6 Conclusion

This paper presents an approach for distributed collaborative engineering design in SOA to enhance automatic design ability. As a computing and metacomputing grid environment, SORCER was employed as bottom platform to build SORDER—a highly flexible software system. SORDER can be extended easily to match different domain, and design interface and information transfer protocol play important role in SORDER application. Using SORDER and layout design interface, engineer can arbitrary organize and manage the domain design services. The SORDER enable asynchronous distributed collaborative design with ease of alternative design services, reduced design cycles, and improved layout resolution quality.

Due to including a mass of domain knowledge and expert experience, design interface building is a challenging work for engineer in different design domain. A successful SOA-based design environment deployment needs collaboration between SOA technology (SORCER), general SOA design environment (SORDER) and domain knowledge (design interface), and our future work will focus on the three points.

7 References

[1] M. Sobolewski, "FIPER: The Federated S2S Environment", JavaOne, Sun's 2002 Worldwide Java Developer Conference, 2002. Available at: http://sorcer.cs.ttu.edu/ publications/papers/ 2420.pdf. Accessed on: March 15, 2007.

[2] M. Sobolewski, "SORCER: computing and metacomputing intergrid". Retrieved Mar. 23, 2008; from: http://sorcer.cs.ttu.edu/publications/papers/2008/iceis-intergrid-08.pdf.

[3] M. Sobolewski, "Service-oriented programming". SORCER Technical Report SL-TR-13. Retrieved Mar. 23, 2008; from: <http://sorcer.cs.ttu.edu/publications/papers/ 2008/SLTR-13.pdf>.

[4] SORCER Lab. Retrieved March 23, 2008, from: <http://sorcer.cs.ttu.edu/>.

[5] S. Szykman, and J. Cagan, "Constrained three dimensional component layout using simulated annealing". ASME Journal of Mechanical Design Vol.119, no.1, pp. 28–35, 1997.

[6] A. Kolli, J. Cagan, and R. A. Rutenbar, "Packing of Generic, Three Dimensional Bomponents based on Multi-resolution Modeling", in Proceedings of the 1996 ASME Design Engineering Technical Conferences and Computers in Engineering Conference: Design Automation Conference, 96-DETC/DAC-1479, 18–22, Irvine, CA, August, 1996.

[7] Li N, Cha J Z, Lu Y P, el a1. "A simulated annealing algorithm based on parallel cluster for engineering layout design". In: Complex Systems Concurrent Engineering: Proceedings of 14th ISPE International Conference on Concurrent Engineering, 83–89, Sao Paulo, 2007.

[8] Jini architecture specification, Version 1.2. Available at: <http://www.sun.com/software/jini/specs/jini1.2html/jini-title.html>. Accessed on: Mar. 23th, 2008.

[9] Inca X Service Browser for Jini Technology. Available at: <http://www.incax.com/index.htm?http://www.incax.com/servicebrowser.htm>. Accessed on: Mar. 23th, 2008.

Taxonomy and Attributes of Business Collaborative Models: an Exploratory Study

JrJung Lyu [a,1] and Ping-Shun Chen[b]

[a] Department of Industrial and Information Management, National Cheng Kung University, Taiwan, ROC.
[b] Department of Industrial and Systems Engineering, Chung Yuan Christian University, Taiwan, ROC.

Abstract. As the global market becomes prevalent, enterprises need to cooperate with partners in the supply chain. After a collaborative alliance is formed, companies share mutual information or resources in order to increase their core competence. Their business collaborative model (BCM) and strategies of an alliance belong to collaborative commerce which requires collaborative management (CM) disciplines. Therefore, the aims of this paper are to study BCMs and to propose the taxonomy and attributes of these BCMs. Enablers of the CM projects are also analyzed and discussed through a field study approach for future reference.

Keywords. Collaborative commerce, collaborative management, business collaborative models.

1 Introduction

The exponential growth of internet-based e-commerce demonstrated the value of the new economics. To meet the full potential of e-commerce, a business relationship needed to be transformed into a more sophisticated interaction [16]. In the later 1990s, due to the progress of information technology and the emergence of internet, enterprises invested in Internet infrastructure, networks and websites for e-business. In the early 2000s, e-business changed not only the trade model with suppliers and customers, but also the models of collaboration. Moreover, electronic data interchange (EDI) and enterprise resource planning (ERP) with partners in the supply chain data could benefit all supply chain members by reducing the inventory cost, shortening the response time or streamlining the manufacturing process by sharing forecasting and planning data [6].

[1] Corresponding Author : Professor, Department of Industrial and Information Management, National Cheng Kung University, Tainan 701, Taiwan, ROC. Tel: (886) 62080576; Fax: (886) 62501734; E-mail: jlyu@mail.ncku.edu.tw

The most important element of collaboration between organizations is the establishment of mutual trust. If trust exists, companies can create strategic vertical or horizontal alliances. In addition, collaborative commerce, the "second generation" of e-commerce, has recently attracted intense attention from e-business professionals [1]. Many world-class enterprises, such as Wal-Mart and Procter & Gamble, have achieved business excellence through the integration of business process reengineering, the implementation of collaborative planning, forecasting and replenishment (CPFR) programs with their supply-chain partners and customer relationship management (CRM) [12]. Nevertheless, the business collaborative models (BCMs) and associated cross-functional cooperation processes in collaborative commerce are still unclear and worthwhile to do further research [2, 3, 9, 15].

To the best of our knowledge, few models for developing collaborative management (CM) and their enablers have been studied. A field study can provide valuable insights for CM practitioners aiming to adopt the right BCM. Furthermore, the development of taxonomy of BCMs is a basis for designing and implementing a CM initiative. Therefore, the goal of this paper is to identify the disciplines and enablers of BCMs. Since the authors have supervised and advised the Corporate Synergy Development Center in the execution of nearly 60 CM projects, the classification scheme and critical attributes are summarized on the basis of the field study approach. The findings from recent CM projects can offer a reference for further study of CM area.

2 Taxonomy of BCMs

Ulieru, et al. developed a multi-resolution collaborative architecture to integrate the global manufacturing [13]. From the functional perspective, their architecture could be classified into two categories: a vertical (agent) hierarchy and a horizontal (communication platform) hierarchy. Agents specialized in dealing with different tasks and could be able to communicate to other agents easily through the communication platform. Therefore, extending from their concepts, this research reformulates the BCMs into the four quadrants, which are based on the levels of function integration and organization relationships (Figure 1). The taxonomy is used to identify the BCM of a collaborative initiative. Details of each category are described below.

2.1 Supply Chain

Many industries, e.g., the automobile industry or PC industry, have formed a number of supply chains. A manufacturer forms a relationship with its suppliers and retailers. Here, the supply chain represents a traditional vertical integration. For the automobile industry, manufacturing a car requires thousands of parts provided by suppliers or outsourcing partners. Hence, establishing a good communication channel between a manufacturer and its suppliers is important in order to solve the problems of engineering design change or raw material shortage.

However, the old concept of a supply chain is just a loose relationship among partners without sturdy sanctions; communications between partners may lack efficiency. The new concept of a supply chain is, therefore, reformed, transforming a loose relationship into an intimate collaboration. Members in a supply chain become an integral organization, sharing information with partners and producing the goods for customers in order to make profits. Moreover, there are many ways to collaborate with partners and the most common method of collaboration in a supply chain is the functional collaborative approach. For instance, many automobile manufacturers have constructed a strong cooperation relationship with their first- and even second-tier suppliers to execute a either collaborative-procurement or collaborative-manufacturing model [11]. Thus, the supply chain is a vertical and functional collaborative model.

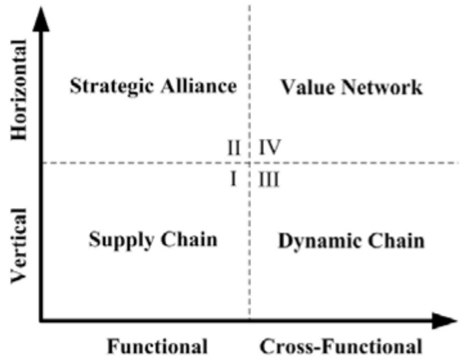


Figure 1. Categories of business collaborative models

2.2 Strategic Alliance

For organizations on the same tier in the supply chain, forming an intra-organizational collaborative strategic alliance can entail the pooling of skills and resources by partners in order to achieve mutual objectives of a specific function [5]. Due to the intra-organizational collaboration, the strategic alliance belongs to a horizontal collaborative model which is an ordinary approach in many Eastern-style companies [14]. In addition, some hospitals in the same geographies have adopted a collaborative operational model to form a strategic alliance [7]. After forming such an alliance, allied hospitals share medical resources or instruments and have more power to negotiate with insurance or drug companies. As a result, such an alliance can reduce operational costs of hospitals and increase hospitals' profits.

In sum, the strategic alliance is a horizontal and functional collaborative model. Typically, small- and medium-sized enterprises prefer strategic alliances which can complement the company's core competence with allied companies. However, it is challenging for enterprises to maintain their valuable processes and to assure the interoperability of systems in the intra-organizations [4].

2.3 Dynamic Chain

To apply an integrated collaborative model in a supply chain, the value network is constructed by the delivery of values to the end-customers [8]. The dynamic chain evolves from the supply chain. The dynamic chain not only inherits the vertical feature of the supply chain, but also extends the functional feature to the cross-functional feature. Hence, the dynamic chain is a vertical and cross-functional collaborative model and CPFR is a famous application.

2.4 Value Network

The most complicated potential category is the horizontal and cross-functional collaborative model, value network. A value network encompasses a diverse range of companies and each company interacts with its chosen customers and suppliers depending on its business model. The UK telecommunication industry has moved in this direction [10].

Overall, the proposed taxonomy can serve as the basis of a better BCM. For each model, in-depth analysis can be applied after the completion of CM projects. Prior to an empirical analysis, the backgrounds of the 34 CM projects in this research are depicted in the next section.

3 The Field Study

The collaboraitve commerce has evloved from the vertical/horizontal integration between the organizations to supply chain integration, and finally to the business collaboration for the developed countries. After the vertical/horizontal integration in the domestic industries matured, the Industrial Development Bureau (IDB) begun to advocate the CM projects in 2003 in order to increase the company's competitive advantage. Hence, many companies wrote their CM proposals to the IDB during the past several years. While CM programs were initialized, the IDB organized a steering committee to select good CM proposals. The criteria of funding the CM proposals were based on their feasibility, potential benefits/risks and the supporting resources. All CM projects had been reviewed in three phases - proposal presentation, mid-term report and final report. The performance of each CM project was justified where the project owner had to provide a written report for review. In order to verify the deliverables of CM projects, members of the steering committee audited the practices of enterprises, either through a field investigation or document reviews during the project period. It was, therefore, able to access the field information of CM projects with the assistance of Corporate Synergy Development Center.

During the interviews of CM project owners and the participation of panel sessions held by the Corporate Synergy Development Center, the difficulties incurred in the implementation process were widely discussed. There were also several workshops and seminars held in which the knowledge of collaborative approach practices was shared. For instance, the president of Voluntary Interindustry Commerce Solutions Association was invited to report on the CPFR methodology and its applicability in industry practices. What is more, a CM survey was conducted in order to measure CM benefits and related impacts. The attributes of CM projects will be elaborated in the following sections.

3.1 Background of CM Projects

Since few reference models of CM can be found in the literature, there is a need for the government to be more proactive in advocating CM from academy to industry. With the help of non-profit organizations or consultancies, approximately 100 companies applied for different CM projects in the hope of obtaining partial funding from the Industrial Development Bureau. Through this mechanism, 57 CM projects were executed and 34 questionnaires were received. CM project owners were in the textile (4), ceramics (1), food (2), chemistry (1), computer or IC chip (7), automobile or bicycle parts (7), metal manufacturing (4), machinery (3), electronics (3) and other industry (2). The data collected from these CM projects were the subject of our field study.

3.2 CM Project Portfolios

34 CM projects are classified into the 5 collaborative models (4 collaborative-design models, 2 collaborative-procurement models, 17 collaborative-manufacturing models, 2 collaborative-marketing models and 2 collaborative-logistics models), 3 CPFR models and 4 integrated collaborative models. Half of the cases adopt the collaborative-manufacturing models. A primary reason is that enterprises which submitted CM proposals may face an intense global competition and they firmly desire to improve their manufacturing capability. Hence, the collaborative-manufacturing model is a reasonable way to reach the targets of good product quality and on-time delivery in such a growing complicated supply chain.

CM projects last for several months to one year. Most of the CM projects observed in this study are from the manufacturing side. The CM project owners are either vice presidents or higher and each project involves from 3 to 17 participants. More than 70% of CM projects with 3 to 5 collaborative participants seem to be small-scale initiatives of CM projects and can be extended to more business partners in the future.

3.3 Information Technology Infrastructure of CM Companies

In order to recognize information technology (IT) infrastructure of CM companies, the questionnaire investigates the existing IT infrastructure of companies and IT annual investment cost. The basic IT systems are composed of enterprise resource planning (ERP), product data management (PDM), supply chain management (SCM) and other systems. (Note that a company may have more than one system.) The data displays that 61.8% of companies have implemented ERP systems; 29.4% have implemented PDM systems; 23.5% have implemented SCM systems; and 14.7% have developed electronic commerce platforms.

In IT investment, comparing the company's IT investment cost with total revenue, the result shows that 73.5% of CM companies invest less than or equal to 1.5 % of total revenue and only 23.5% of CM companies invest more than 3% of total revenue. Therefore, CM companies may invest more money in their IT infrastructure.

3.4 Performance of CM Projects

The 4 important key performance indicators (cost reduction, process improvement, total revenue increment and customer satisfaction) in the CM projects are summarized in Table 1. 11 of 34 CM projects are excluded because of insufficient data. For the remaining 23 projects, more than half of the CM cases have great or excellent performance in all of these criteria.

Table 1. Key performance indicators of CM projects

Key performance indicator	Not significant	Fair	Great	Excellent
Cost reduction	1	10	8	4
Process improvement	2	0	10	11
Total revenue increment	3	8	8	4
Customer satisfaction	1	0	22	0

By analyzing the CM projects (Figure 2), 23 CM projects are located in the 'Supply Chain' category. 6 companies have the cross-function and vertical integration and therefore belong to the 'Dynamic Chain' category. Only 5 companies are in the 'Strategic Alliance' category (horizontal and functional integration) and it is still a small number compared to that of vertical integration. The 'Value Network' category requires the more mature collaborative experience to implement this kind of CM projects and the value network will be a future trend.

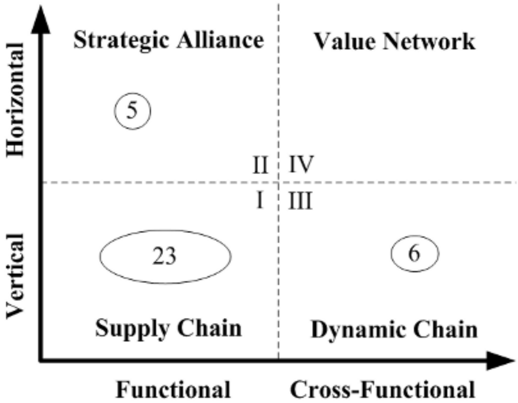


Figure 2. The location of CM projects in the BCMs

4 Attributes of BCMs

The payoffs of each CM project are significant in terms of more efficient operations and lower costs. Although each CM project is challenging to be successful, there is the potential of delivering a strategic value to a community. From the field information of several industries, the attributes of BCMs can be summarized as follows.

4.1 BCM Matters

The first step to implement a CM project is to decide which BCM is suitable for an organization. Considering a collaborative model between an auto-part manufacturer and its major dealer, it is a vertical model of collaborative-marketing initiative and it belongs to the 'Supply Chain' category. The value increased from the customer perspective through the CM initiative is a key performance indicator of CM project. In another example is that a food company applies an integrated collaborative model to streamline its supply chain. The response time from end to end (for a five-tier supply chain) dramatically declines in order to maintain the freshness of products. This case is in the 'Dynamic Chain' category and the customer requirement decision is crucial. The third case, the steel company, has built a community for collaborative-marketing with a major customer. The synergy generated from this community has recently attracted several large orders from world-class buyers. It is the 'Strategic Alliance' category and the game-rule formulation is the key to be successful.

While other collaborative-design models have been applied in the textile and other industries, the point is that adopting an appropriate BCM based on taxonomy in Section 2 is the top priority before proposing a CM project. Each case may involve a different scenario, but the value proposition is clear: to increase the total value of the entire community through collaborative efforts.

4.2 Timing Is Critical

Communication among parties is vital in order to form a cross-organizational project team. It is, however, very time-consuming to gain trust through communications. The expected deliverables and infused resources in a CM project are also difficult to determine and evaluate. During the execution of a CM project, a good consultancy is hard to find. Without appropriate guidelines from outside parties, arguments among enterprises can never be settled easily. Thus, timing is critical in executing a CM project since a CM project manager may squander resources trying to communicate and negotiate with different parties and thereby lose the momentum to finish the CM project on time and to deliver the expected outputs.

5 Conclusion and Future Work

A CM project involves the inter- and intra-organizational business collaborative process reengineering and applications of new information-sharing mechanism. This paper presents the four basic BCMs and their classification. It is obvious that a CM project is of strategic value to a supply chain and a successful CM project requires the full involvement of top managers.

In this research, a taxonomy is proposed and its key attributes are discussed. The findings in this study are based on the field study, consisting of an in-depth survey of 34 CM projects. Innovation of the BCMs and the disciplines in executing the collaborative approach are both important attributes. The field information and the summary can provide a reference for practitioners to execute a CM project more cost-effectively and successfully. Furthermore, while the importance of collaboration is becoming prevalent for every industry, the future research direction will be how to explore the in-depth relationship between the practical business model and collaborative commerce.

6 References

[1] Centola C, Myer CJ, Raisinghani MS, Virgil D. Collaborative commerce for non-technical industries: is it worth the investment? Int J Inf Manage 2004;24(5):433-440.

[2] Chen M, Zhang DS, Zhou L. Empowering collaborative commerce with Web services enabled business process management systems. Decis Support Syst 2007;43(2):530-546.

[3] Chiang TA, Trappey AJC. Development of value chain collaborative model for product lifecycle management and its LCD industry adoption. Int J Prod Econ 2007;109(1-2):90-104.

[4] Datta S. Unified theory of relativistic identification of information: convergence of unique identification with syntax and semantics through internet protocol version 6 (IPv6). Engineering Systems Division, MIT. IOT 2008, Lecture Notes in Computer Science, Springer-Verlag (in press). 2007.
[5] Gebrekidan DA, Awuah GB. Interorganizational cooperation: a new view of strategic alliances - The case of Swedish firms in the international market. Ind Mark Manage 2002;31(8):679-693.
[6] Hammer M. The superefficient company. Harv Bus Rev 2001;79(8):82-91.
[7] Huckman RS. Hospital integration and vertical consolidation: An analysis of acquisitions in New York State. J Health Econ 2006;25(1):58-80.
[8] Kothandaraman P, Wilson DT. The future of competition - Value-creating networks. Ind Mark Manage 2001;30(4):379-389.
[9] Li EY, Du TC, Wong JW. Access control in collaborative commerce. Decis Support Syst 2007;43(2):675-685.
[10] Li F, Whalley J. Deconstruction of the telecommunications industry: from value chains to value networks. Telecommun Policy 2002;26(9-10):451-472.
[11] Lyu J, Lin Y. e-Business strategies in the automobile industry - case studies approach. Int J Electron Bus Manage 2004;1(2):85-90.
[12] Seifert D. Collaborative planning, forecasting, and replenishment: how to create a supply chain advantage. New York: AMACOM, 2003.
[13] Ulieru M, Norrie D, Kremer R, Shen WM. A multi-resolution collaborative architecture for web-centric global manufacturing. Information Sciences 2000;127(1-2):3-21.
[14] Xu T, Bower DA, Smith NJ. Types of collaboration between foreign contractors and their Chinese partners. Int J Proj Manage 2005;23(1):45-53.
[15] Yang MH, Lin WS, Liu SC, Chao HY, Chen SH. Developing the partner relationship management system for franchised electronic stores. Int J Technol Manage 2008;43(1-3):176-193.
[16] Zhao J, Huang WV, Zhu Z. An empirical study of e-business implementation process in China. IEEE Trans Eng Manage 2008;55(1):134-147.

Applying Petri Net to Analyze a Multi-Agent System Feasibility - a Process Mining Approach

C. Ou-Yang[a,1] Yeh-Chun Juan[b] C.S. Li[a]
[a]Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan
[b]Department of Industrial Engineering and Management Information, Huafan University, Taipei, Taiwan

Abstract: Using agent development tools to construct an agent-based system is a popular approach. However, the developed agent system might not be feasible due to certain mutual conflicted rules stored in various agents. Currently, the development tools do not have the function to check the feasibility about the workflow of the agent system during it implementation stage. Therefore, in this research, a Petri Net (PN) based evaluation approach was developed. There were three stages in the proposed system. In the conceptual stage, the pitfall of the current agent system developing process was examined and an improvement analysis process was specified. Then, in the system design stage, an evaluation approach which extracted the process log file from a developing agent system into a PN model in terms of a process mining approach-α algorithm was proposed. This model was simulated in a PN simulation package. The agent system performance was evaluated in terms of analyzing the deadlock phenomena of the PN model. Finally, in the implementation stage, the proposed concept was implemented by using an agent developing tool JADE and a PN simulation tool CPN. An agent-based robotic assembly system was used to examine the possible deadlock of the agent system.

Key Words：Agent-based Systems, Workflow Feasibility, Process Mining, Petri Nets

1. Introduction

Recently, multi-agent systems have been applied in many fields such as manufacturing control, supply chain collaboration and e-commerce. There are many approaches developed to design a multi-agent system. Basically, the system development can be categorized as modeling stage, and system developing stage. In the system modeling stage, most of the developed modeling approaches, such as MaSE and WARP, were based on UML. Currently, most of the development tools still lack of the capability of system verification. That is once an agent system was

---

[1] Corresponding Author E-mail: ouyang@mail.ntust.edu.tw

constructed, it is necessary to assure that the system can be operate feasibly before it was formally operated. In this research, an evaluation approach was developed, which extracted the process log file from a developing agent system and then transferred into a PN model in terms of a process mining approach. The agent system feasibilities were evaluated in terms of analyzing the deadlock of the PN model.

2. Background

In the agent system development life cycle, once an agent-based system has been constructed, its performance should be verified. Most of the research works in system verification were based on simulation. For instance, a simulation approach has been used to verify that the strategies made by a cooperative agent system could reach a convergence value for a manufacturing environment [1]. Another rule-based verification schema has been developed for a cooperative traffic agents system to validate the situation such as the worst case behaviors of a traffic system [2].

In addition to the works on performance verification, very few researches were addressed on the agent system feasibility verification. The system feasibilities indicated that an agent system could be carried out smoothly in most of the conditions. PN has been used to verify the feasibility for various systems. For instance, the deadlock of a PN can be analyzed by calculating the state transformation matrix of the net. The system bounded condition can verified in terms of the place invariant properties of the net.

The idea of applying process mining in the context of workflow management system was first introduced in [3]. Most of the approach would discover the workflow process from the event logs. Today, Process mining has reached a certain level of maturity and has been used in a variety of real life cases [4] [5].

Various mining algorithms have been developed recently, such as α miner [6], heuristic miner [7], α++ miner [8], and genetic miner 9]. Among those, α miner is the most straightforward approach. A comparison about these miners based on four indexes: fitness, precision, generalization and structure, can be found from [10].

In this research, an evaluation approach which extracted the process log files from an agent system and transferred into a PN model in terms of a process mining algorithm would be proposed. This transferred model was imported into a PN simulation package CPN. The agent system feasibilities were evaluated in terms of analyzing the deadlock of the PN model.

3. The Proposed Evaluation Framework

A three stage approach has been proposed to support the feasibility analysis of an agent system (Fig-1). In the conceptual stage, the concept of using the behavior information for verifying of an agent system was discussed. Then, in the design stage, an approach of extracting the behavior information from a testing agent system and transforming into a PN model in terms of α algorithm was proposed. The feasibility of the agent system would be examined based on the simulation results of the respective PN model. Finally, in the implementation stage, the proposed approach would be implemented in terms of the agent development tool, JADE and a PN tool Colored Petri Net (CPN). A case about a robotic assembly cell would be used to examine the proposed concepts.

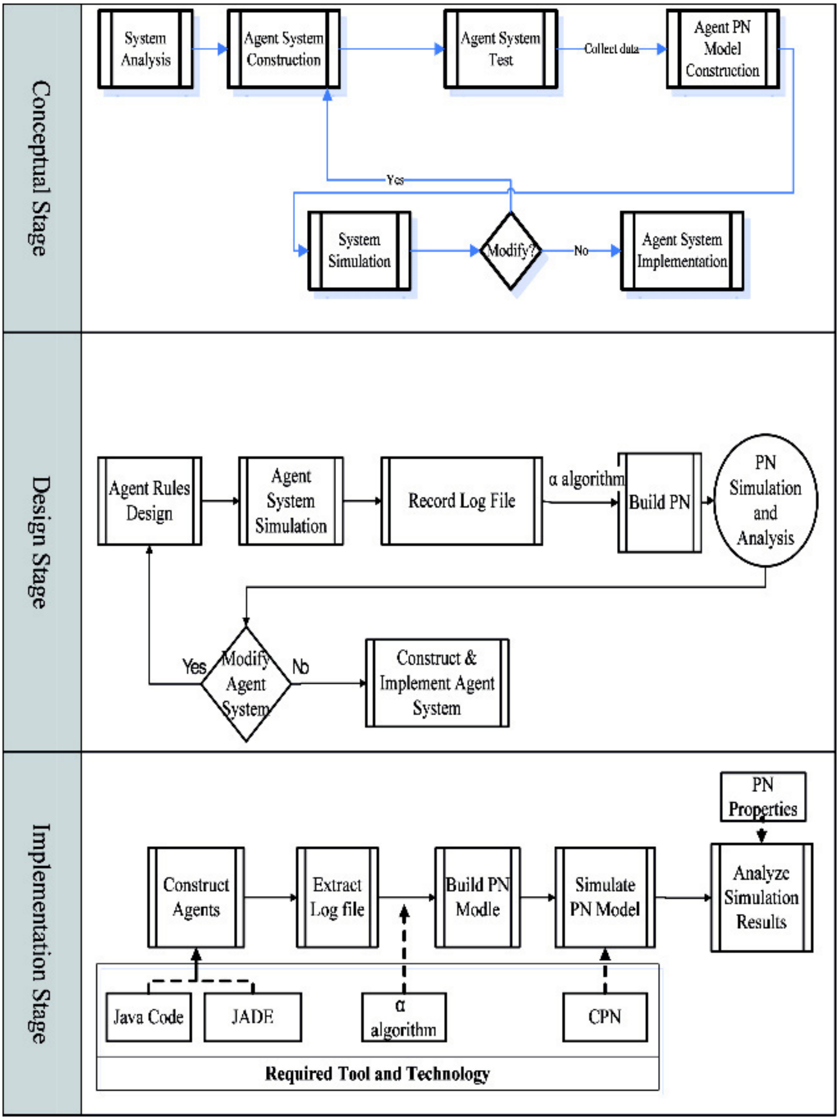


Figure 1. Proposed framework

3.1 The Conceptual Stage

In this stage, an agent system would be constructed based on UML analysis. The constructed system would be tested and the triggering sequence of the system would be recorded in a log file. Then, the contents of the log file would be transferred to a PN model through α algorithm. The transferred PN model would be simulated and certain infeasible conditions such as deadlock could be verified through the simulation data. Finally, the verified system would be implemented. The main features of this approach were that it could reflect certain dynamic condition occurred during the system test phase.

3.2 The Design Stage

The main objective of this section was to develop an approach to verify the feasibility of an agent system. The basic idea was to extract the interaction information from a multi-agent system and then transfer the behaviors of the extracted data into a PN. The main issue in this stage was how to transfer the information about the agent interactive logic into a proper PN model. In this work, a process mining approach -α algorithm was applied and modified to fulfill this goal [6].

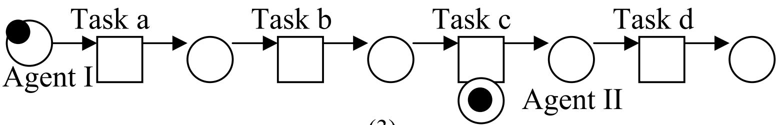
3.2.2 Multi-Agent System Application and Modified α Algorithm

Process	Operation	Notation
Case 1	Task a	Beginning Task
Case 1	Task b(Request)	Agent I send a request to Agent II
Case 1	Task c(Agree)	Agent II agree with Agent I
Case 1	Task d	Ending Task

(1)

Process	Operation	Notation
Case 2	Task b(Request)	Agent I send a request to Agent II
Case 2	Task c(Agree)	Agent II agree with Agent I

(2)



(3)

Figure 2. (1) log file for agent I, (2) log file for agent II, (3) The PN for agent I and agent II

Fig-3 showed the modified α algorithm. Basically, the main modification was to add a step (Step-6) into the original algorithm. The main role of this step was to find the succession task from the set of duplicated tasks in two agents' log files. Based on the log files shown in Fig-2-1 and Fig-2-2, the related PN can be translated according to the following procedure:

1. Find the transition set: T_{W1}={a,b,c,d}, T_{W2}={b,c}
2. Find the initial transition set: T_I={a,b}
3. Find the final transition of each set: T_O={d,c}
4. Find the set of transitions which has only one directional precedence: X_W={(a,b), (b,c), (c,d), (b,c) }
5. Remove the duplicated transition set. $Y_W = X_W$ –(b,c)={(a,b), (b,c) (c,d),}
6. For the duplicated transition set find the successor transition. T'_I={c}.
7. Develop the place set: P_W={i_{W1},i_{W2}, $p_{(A, B)}$, $p_{(B, C)}$, $p_{(C, D)}$, o_W}
8. Develop the arrow set to link the places and transitions. F_W={(i_{W1}, a), (a, $p_{(a, b)}$), ($p_{(a, b)}$, b), (b, $p_{(b, C)}$), ($p_{(b, c)}$,c), (i_{W2},c), (c, $p_{(c, d)}$), ($p_{(c, d)}$, d), (d, o_W),
9. Complete the final PN: α(W)=(P_W, T_W, F_W)

The translated PN was shown in Fig-2-3. This model can then be simulated in terms of certain commercial tools for further analysis.

4. Empirical Verification

The proposed approach has been implemented and verified in terms of a robotic assembly cell example from Zhou's book. [11]. Fig-4 showed the assembly cell. There are two robots and two assembly stations. Each station required both of the robots to carry out an assembly job. If workstation 1 required to perform the task, it would acquire robot 1 first. Once it got the authorization of robot 1, it would acquire robot 2. For workstation 2, the sequence of acquiring robots would in a reverse way. Therefore, it could be found that if both of the workstation 1 and 2 acquired robots at the same time, then, neither of them could get another robot. The assembly cell would be in a deadlock situation.

$$1.\mathrm{T}_{\mathrm{wi}} = \{\mathrm{t} \in \mathrm{T} \mid \exists_{\sigma \in \mathrm{w}} t \in \sigma\}, i \in n$$

$$2.\mathrm{T}_{\mathrm{I}} = \{\mathrm{t} \in \mathrm{T} \mid \exists_{\sigma \in \mathrm{w}} t \in first(\sigma)\}$$

$$3.\mathrm{T}_{\mathrm{O}} = \{\mathrm{t} \in \mathrm{T} \mid \exists_{\sigma \in \mathrm{w}} t \in last(\sigma)\}$$

$$4.\mathrm{X}_{\mathrm{w}} = \{(\mathrm{A},\mathrm{B}) \mid \mathrm{A} \subseteq \mathrm{T}_{\mathrm{w}} \wedge \mathrm{B} \subseteq \mathrm{T}_{\mathrm{w}} \wedge \forall_{\mathrm{a} \in \mathrm{A}} \forall_{\mathrm{b} \in \mathrm{B}} \rightarrow_w \mathrm{b}$$
$$\wedge \forall_{\mathrm{a}_1, a_2 \in \mathrm{A}} \mathrm{a}_1 \#_w \mathrm{a}_2 \wedge \forall_{\mathrm{b}_1, \mathrm{b}_2 \in \mathrm{B}} \mathrm{b}_1 \#_w \mathrm{b}_2\}$$

$$5.\mathrm{Y}_{\mathrm{w}} = \{(\mathrm{A},\mathrm{B}) \in \mathrm{X}_{\mathrm{w}} \mid \forall_{(\mathrm{A}',\mathrm{B}') \in X_w} \mathrm{A} \subseteq \mathrm{A}' \wedge \mathrm{B} \subseteq \mathrm{B}' \Rightarrow (A,\mathrm{B}) = (\mathrm{A}',\mathrm{B}')\}$$

$$6.\mathrm{T}'_I = \{t \in last(\mathrm{A},\mathrm{B}) \mid (\mathrm{A},\mathrm{B}) \subseteq \mathrm{T}_{\mathrm{w}} \wedge \mathrm{A} \rightarrow_w B \wedge (\mathrm{A},\mathrm{B}) \subseteq T'_W\}$$

$$7.P_W = \{p_{(A,B)} \mid (\mathrm{A},\mathrm{B}) \in Y_W\} \cup \{i_{W_i}, o_W\}, i \in n$$

$$8.F_W = \{(a, p_{(A,B)} \mid (\mathrm{A},\mathrm{B}) \in \mathrm{Y}_{\mathrm{W}} \wedge a \in \mathrm{A}\} \cup \{p_{(\mathrm{A},B)}, b) \mid (\mathrm{A},\mathrm{B}) \in \mathrm{Y}_{\mathrm{W}} \wedge b \in \mathrm{B}\}$$
$$\cup \{(i_w, t) \mid \mathrm{t} \in \mathrm{T}_{\mathrm{I}}\} \cup \{(i_w, t) \mid t \in T'_I\} \cup \{(t, o_w) \mid \mathrm{t} \in \mathrm{T}_{\mathrm{o}}\}$$

$$9.\alpha(W) = (P_w, T_w, F_w)$$

Figure 3. Modified α algorithm

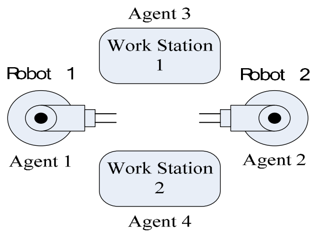


Figure 4. A robotic assembly cell

4.1 Log Files Generation

Fig-5 showed the window of the sniffer agent and its retracted information. The left side of the window showed the four developed agents (R1: Robot 1, R2: Robot 2, WS1: Workstation 1, and WS2: Workstation 2) as well as the AMD and DF. The right side of the window showed the sequence diagram for the process of WS1 requesting R1 agent and R2 agent. The communication messages among these agents were shown in the bottom of the diagram. The information about the log files was extracted from the message files.

The sequence diagram showed in the sniffer agent window indicated the sequence of WS1 requesting R1 and R2 robots. Mainly, the WS1 would send a "request" to DF for searching required robot. Then, DF would acknowledge WS1's request through "inform" message. In the third stage, W1 would "request" R1 for service, and R1 would "agree" the service in the fourth stage. Then W1 would "request" R2 in the fifth stage. The information and sequence about these tasks can be extracted from the message file of sniffer agent and converted into the log file for W1 agent.

The information about the log files for W2, R1 and R2 also can be extracted from the same message file and there interaction behaviors were showed in Fig-6. It can be found that W2 would be refused by R2 since it already assigned to W1. As for the log file for R2, it showed that it agreed the request from W1 and refused R2.

4.2 Log Files Conversion to PN Model

Once a log file has been created, the next stage was to convert the log file into a PN model. For the WS1 agent, its log file can be shown as {REQUEST, INFORM, REQUEST, AGREE, REQUEST, AGREE, CANCEL, CANCEL}and{ REQUEST, INFORM, REQUEST, AGREE, REQUEST, AGREE, CANCEL, CANCEL}. Table 1 showed the messages in the log file and their mapped symbols. It can be found that although the messages for G and H are CANCEL, G was the message sent from WS1 to R1 and H was sent from WS1 to R2.

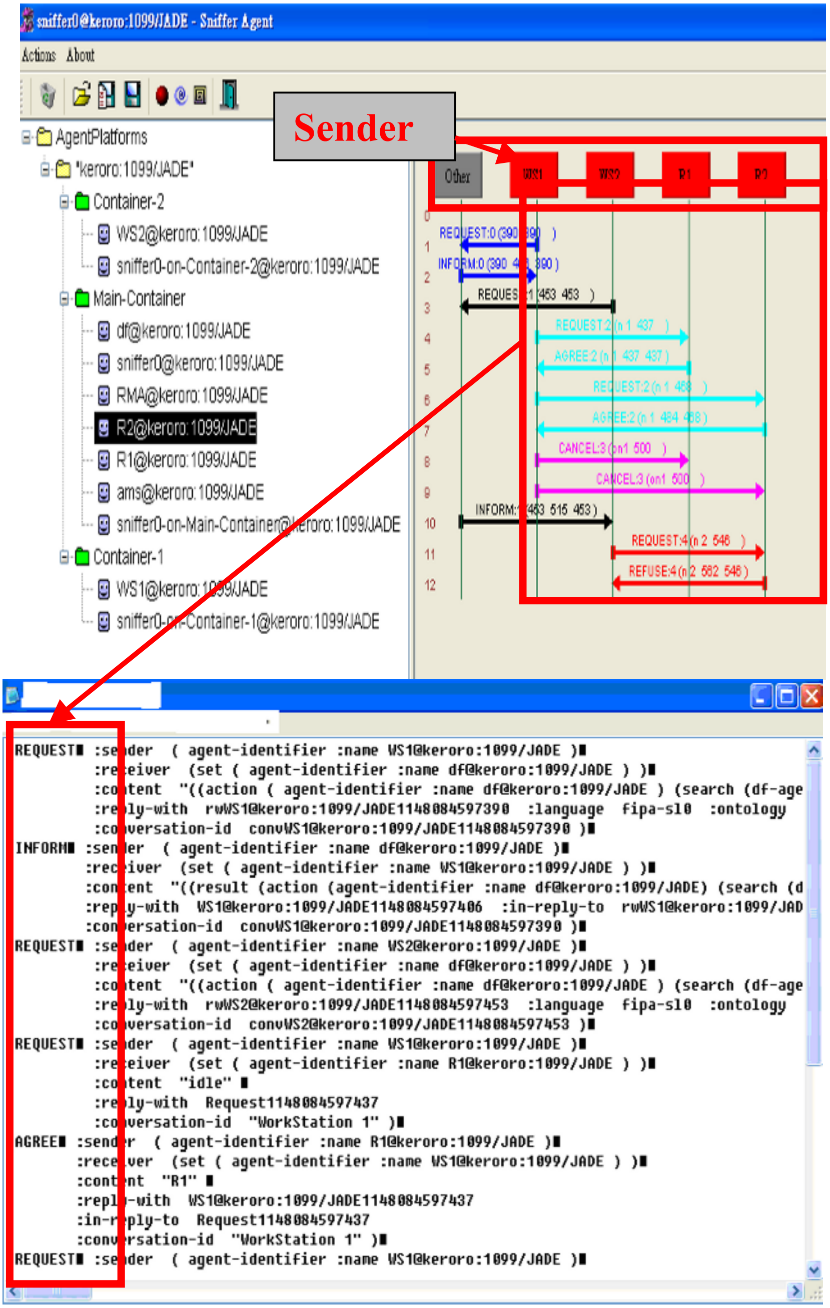


Figure 5. JADE window and the generated communication messages

According the log file, the respective PN can be transferred as the following steps:

1. *Find the transition set:* $T_W=\{A, B, C, D, E, F, G, H\}$ 。
2. *Find the initial transition set:* $T_I=\{A\}$ 。
3. *Find the Final transition set:* $T_O=\{G, H\}$ 。
4. *Find the set of transitions which has only one directional precedence:* $X_W=\{(A, B), (B,C), (C,D), (D,E), (E,F), (F,G), (F,H)\}$.
5. *Remove the duplicated transition set. There was no duplicated set in this log file. So,* $Y_W= X_W =\{(A, B), (B,C), (C,D), (D,E), (E,F), (F,G), (F,H)\}$.
6. *Develop the place set:* $P_W=\{i_W, p(A,B), p(B,C), p(C,D), p(D,E), p(E,F), p(F,G), p(F,H), o_W\}$.
7. *Develop the arrow set to link the places and transitions:* $F_W=\{(i_W, A), (A, p(A,B)), (p(A,B), B), (B, p(B,C)), (p(B,C), C), (C, p(C,D)), (p(C,D), D), (D, p(D,E)), (p(D,E), E), (E, p(E,F)), (p(E,F), F), (F, p(F,G)), (F, p(F,H)), (p(F,G), G), (p(F,H), H), (G, o_W), (H, o_W)\}$.
8. *Complete the final PN:* $\alpha(w)=(P_W, T_W, F_W)$
9.

The converted PN model was shown in Fig-7.

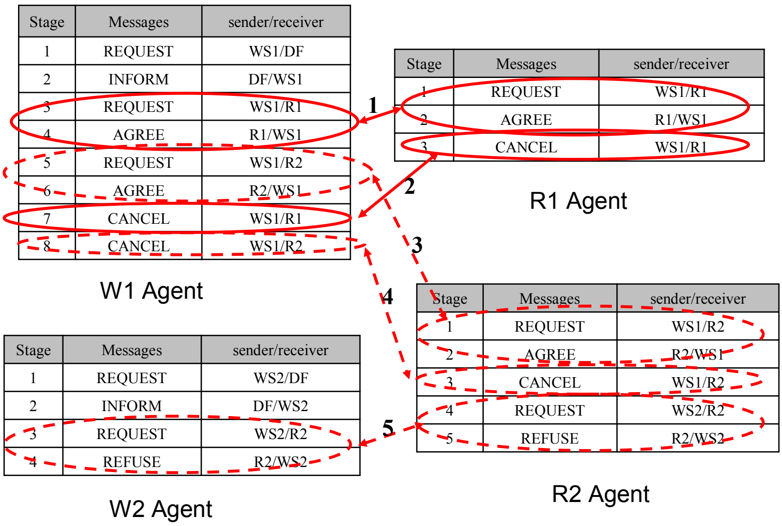


Figure 6. The interactions among the four agents

Table 1. The messages and its mapped symbols for the log file of WS1

Messages	REQUEST	INFORM	REQUEST	AGREE
Symbol	A	B	C	D
sender/receiver	WS1/DF	DF/WS1	WS1/R1	R1/WS1
Messages	REQUEST	AGREE	CANCEL	CANCEL
Symbol	E	F	G	H
sender/receiver	WS1/R2	R2/WS1	WS1/R1	WS1/R2

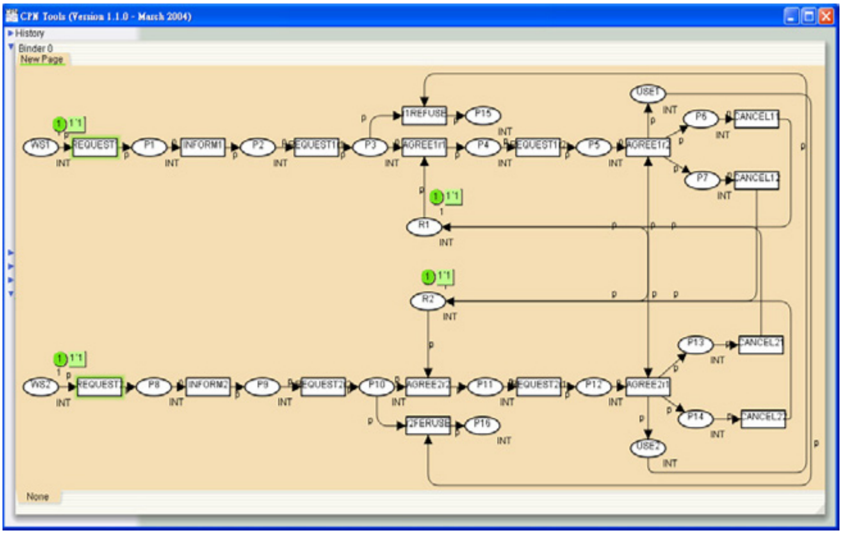
As mentioned before, the empirical assembly cell would encounter a deadlock if W1 issued a request to R1 and W2 sent a request to R2 at the same time. This situation can be simulated through the PN model by issuing two tokens at place "WS1" and WS2" respectively. The CPN report was shown that two tokens were halted in place P5 and P12 respectively. From the PN model, P5 indicated that W1 was waiting the response from R2, and P12 specified that W2 was waiting the response from R1. Since both of the assembly workstations occupied a robot and waited another robot, a deadlock would happen.

5. Discussion and Conclusions

This research tried to bridge the gap between a multi-agent systems and a PN tool, such that the developed PN model can be used to simulate and to analyze the feasibility situation, such as deadlock, of the multi-agent system. The major contributions of this work were:

1. The concept of verifying the feasibility of a multi-agent system during its developing stage can be accomplished.
2. A process mining approach was applied to transfer the log files of a multi-agent system into its respective PN model.

One possible issue is that in this work, α algorithm was modified such that the log files for two interactive agents can be transformed to a PN model. The main assumption of the modification was based on that two agents should have "direct succession" relation. That was two agents would send a "request" and "response" messages between each other in the beginning of the interaction. However, in the real environment, the phenomena for agent interaction might be more complicated. For instance, if both of the two agents having reactivity property, they might have interaction due to the trigger from outside environment. One possible example was that if the inventory level was lower than the safety level, an inventory control agent might send a replenishment order to the supplier agent without sending the "request" message first. Since these agent might not have a common "request", "response" messages in their log files, the modified α algorithm could not link these agents in a PN model.



(3)

Figure 7. The PN model

6. Acknowledgement

The authors would like to acknowledge the financial support by National Science Council, Taiwan, through project no. NSC-95-2221-E-011-015.

7. References

[1] S. M. Deen *and* R. Jayousi, "A Preference Processing Model for Cooperative Agents", Journal of Intelligence Information System, Vol.26, 2006, p115-p147.

[2] W. Damm, H. Hungar and E.R. Olderog, „Verification of cooperating traffic agents", International Journal of Control, Volume 79, Number 05, 2006, p395-p421.

[3] R. Agrawal, D. Gunopulos, and F. Leymann, Mining Process from Workflow logs, Sixth International Conference on Extending Database Technology, 1998, p. 469-p483.

[4]W.M.P. van der Aalst, H.A. Reijers, A.J.M.M. Weijters, B.F. van Dongen, A.K. Alves de Medeiros, M. Song, and H.M.V. Verbeek, Business Process Mining: An Industrial Application, Information Systems, Vol.32, No. 5, 2007, p713-p732.

[5] A. Rozinat, I.S.M. de jong, C.W. Gunther, and W.M.P. van der Aalst, Process Mining of Test Processes: A Case Study, BETA Working Paper Series, Eindhoven University of Technology, Eindhoven, 2007.

[6] W.M.P. van der Aalst, A.J.M.M. Weijters, and L. Maruster, "Workflow Mining: Discovering Process Models From Event Logs", IEEE Transactions on Knowledge and Data Engineering, Vol. 16, No. 9, Sep 2004.

[7] A.J.M.M. Weijters, W.M.P. van der Aalst, "Rediscovering Workflow Models from Event-Based Data", Engineering of Intelligent Systems, 2003, p65-p65,

[8] L. Wen, J. Wang, J.G. Sun. "Detecting Implicit Dependencies Between Tasks from Event Logs", Lecture Notes in Computer Science, Vol.3841, pp.591- 603, 2006

[9]W.M.P. van der Aalst, V. Rubin, B.F. van Dongen, E. Kindler, C.W. Gunther, “Process Mining: A Two-Step Approach using Transition Systems and Regions”, BPM Center Report BPM-06-30, BPMcenter.org, 2006

[10] A. Rozinat, A.K. Alves de Mediros, C.W. Gunther, A.J.M.M. Weijters, W.M.P. van der Aalst, “Towards an Evaluation Framework for Process Mining Algorithms”, Lecture Notes in Computer

[11] M. Zhou, Petri Net Synthesis for Discrete Event Control of Manufacturing Systems, Kluwer Academic Publishers, 1993.

The Impact on Global Logistics Integration System to Concurrent Collaborative Process

Yao Chin Lin [a,1,] and Ping Heng Tsai

[a]Associate Professor of Information Management Department and Director of Innovation and Incubation Center in Yuan-Ze University, Chung-Li, TW (R.O.C)

Abstract. This study was focused on front of five metal stamping industry of the world named JF Machine Industrial who is the biggest maker of the mechanical power presses in Taiwan. JF Machine Industrial builds Global Logistics Integration System including four main systems. The goal of these four systems is strengthen interaction between their clients and engineers, bring some innovation service process to their customers and integration the inside process. The value of this study has constructed in three views that is Information Systems, Interaction Processes, and Business Benefits. After the discussion, JF Company makes some Concurrent Collaborative Working to increase interaction between clients and company rising up the customer satisfaction to keep their customers. Finally, this paper may help global companies who want to organize the enterprise resources concerned with the development, support and operation of IT systems to look beyond the systems of their collaborative working to consider the effect of supporting localization service in their systems.

Keywords. Global Logistics Management, Information Technology, Collaborative Working, Competitive Advantage, Concurrent Engineering

1 Introduction

A Review of competitive characteristics for Taiwan in which Manufacturing industry is a fundamental, such as, produces, heavy export processing even The Original Equipment Manufacturing/Original Design Manufacture (OEM/ODM) and earlier stage inexpensive labourer, demonstrates an inseparable, rigidly precise correlation between Taiwan’s competency and the manufacturing industry.
Particularly, Engineering Industry usually indicated to machine tools, Industry Machinery, General Machinery and part of Mechanical, and Machine Tool separated into two categories one is Metal Cutting, other is Metal Forming.
The Engineering Industry in Taiwan in 2007 has had 6274 hundred million NT dollars output value, and Metal Cutting Machine Tool was occupied 771 hundred million NT dollars in Engineering Industry; Metal Forming Machine Tool was occupied 223 hundred million NT dollars in Engineering Industry, and total of

[1] Corresponding Author : Associate Professor of Information Management Department and Director of Innovation and Incubation Center in Yuan-Ze University, Chung-Li, TW (R.O.C); Tel:03-4638708/ Fax: 03-4638497; Email: imyclin@saturn.yzu.edu.tw

Machine Tools occupied 994 hundred million NT dollars. Therefore, Machine Tool Industry have important affected in economic layer of Taiwan (ITIS, 2007).
The start point of this study is to discuss the function in Global Logistics which is dynamic, not only in what is done but the way is done, has a main goal is supporting good service to customers and rapidly fulfilling customers' demand to obtain the competitive advantages. Thus, in this study draws into conception of Concurrent Collaboration Working to discuss the execution outcome and the creation value to obtain the competitive advantage on market.
This research is going to be near a step to discuss case company who provides customer real-time business model and Concurrent Collaborative Working system. Through the case company regard with the global customers in highly adaptable model of Local Market demand that can help they promoting their delivery efficiency and increasing customer satisfaction, than reducing the location installation cost and working suspension. In the other side, if the company can use existing information system as foundation not only to consider building Global Logistics Management but also to use information system into supply chain and supply source management, can raise the brand advantage and focus on different service driven in market.

2 Literature Review

IT capability is viewed as a catalyst to avoid supply chain management failure. IT can improve services and reduce costs simultaneously, and significantly influences overall logistics competency. Thus, leading edge firms strongly invest in state-of-the-art IT [3],[6].
Information sharing, which refers to a firm's willingness to share key information that is timely, accurate, responsive and useful, with supply chain partners, is fundamental to supply chain coordination and critical to supply chain partnerships' success [7],[10].

2.1 Global Logistics Management

Global Logistics Management refers to the overall management system of a corporation's undertaking of worldwide market distribution, product design, customer satisfaction, production, procurement, logistics, suppliers and inventory. A Global Logistics Management system aims to address and accomplish real-time refection of market changes, decision to business and inventory risks, reduction of distributor cost, mitigation of supply chain and increase of market reaction agility.
The logistics approach can be regarded as a key strategic resource or capability for acquiring sustained competitive advantage [7],[9] when capabilities (e.g. information-based, benchmarking, and flexibility) are valuable, scarce and difficult to imitate. An information-based capability comprises information technology and information sharing.
Information Technology is a capability that can improve distribution performance, facilitate logistics integration, and contribute to supply chain success [16],[14],[4].
Information sharing, which refers to a firm's willingness to share key information that is timely, accurate, responsive and useful, with supply chain partners, is fun-

damental to supply chain coordination and critical to supply chain partnerships' success [5],[8].

2.2 Concurrent Engineering

Concurrent Engineering is a systematic approach [2] to the integrated, concurrent design of products and their related processes, including manufacturing and support defined by Carter and Baker. This approach is intended to cause the developers, from the outset, to consider all elements of the product life cycle from concept through disposal, including quality, cost, schedule, and user requirements.
CE represents an organization's ability to carry out product development as a series of overlapping phases, which delivers product on time, to provide customer satisfaction at the right price [12]. The key challenge to Concurrent Engineering in this wider context is complexity. The multiple facets that have to be managed in a large-scale project demand the support of IT systems.
The availability of information in near real time reduces the time and space of decision making and demanding changes in the way of work. These new ways of work are further enhanced by the use of better IT systems. To achieve this requires a 'right-first-time' approach by applying numerous tools and techniques during the project management of product development, to enable effective decision-making [1][13].
This accelerated cycle in tightening the physical and information flow is going to impact our way of life in a Large-scale engineering is an industry area that deals with complex products, large project teams made-up of companies from different technical and commercial disciplines and even nationalities.

2.3 Collaborative Working

Following the National Council for Voluntary Organizations' (NCVO) Collaborative Working Unit defines Collaborative Working as partnership between voluntary or community organizations. An organization may work with one or two others or may belong to a wider consortium. Organizations can work together in a spectrum of ways, from informal networks to joint delivery of projects, and for a range of purposes. Collaborative Working can last for a fixed time or be permanent.
Through the Collaborative Working integrated the enterprise resource is more important things. Especially the Machine tool industry, which has the highly output value and the rate of turnover in Taiwan, uses the resource integration in supply chain activity to do collaborative production planning and collaborative design development in business activity linking to whole supply chain, increasing the difference on products, and promoting the competitive in industry.

3 Research Design

This study has been running for two years. The research methods have included the several times meetings, conferences, and practices. In this period, it was set many inspect spots on their time schedule, such building up the consensus in organization, business process reengineering, system architecture analysis, and popularization.

In each times meeting, this study has invited some core people who join the discussion and help to answer our finding in JF Company, such vice president, MIS managers, expert professors, customers and IT providers. Each two or three months set a spot review, ask the some questions about their working and observe the salutations. The case company hopes integration their whole suppliers, group factories, global business footholds, and sales partnership.

Therefore, JF Company introduces Global Logistics Management into their company process that has included five key component factors, such building domestic and foreign suppliers' cooperation system in thirty-five countries, Union Quotation Center and Product Segmentation, establishing the technology transfer and knowledge transfer in factories, setting forty sales patterns cooperation system, including seven self footholds and thirty-three dealers. Establishing Global Intelligence Center supported location installation and localization design.

This study discusses the efficiency on Collaborative Working into machinery industry that helps bringing the some synergys into machine tools industrial who wants to get the competitive advantage. Finally, the study is focused on how to integrate the Collaborative Working into IT, such as, Global Logistics System, and Enterprise Resource Planning, and discussing the difference in Collaborative Working between before and after. The research model likes Figure 1.

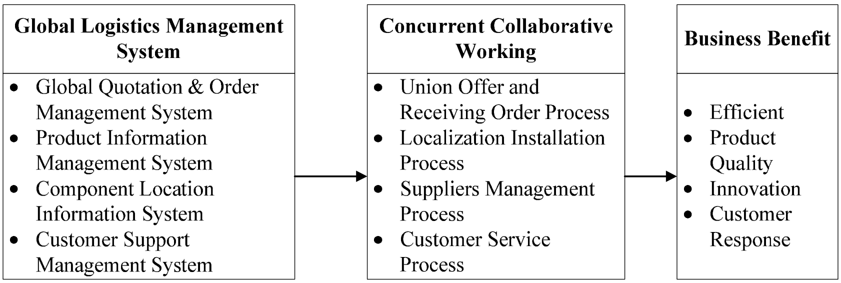


Figure 1. Research Framework

3.1 Global Logistics Management System (GLMS)

The Global Logistics Management has starting from the customers' vision, and through integration the global resources into company's strategy planning that help company to get some business benefit. Therefore, this research thinks the GLM as an information system, via some interaction process between clients and company to earn the best business benefit and create the more value between each other. GLMS has three main objects. First is Quick response to customers' need, second is service differentiation, and solving the problems to build up the inside knowledge. Here are list four sub systems in GLMS as we talk about:

1. Global Quotation and Order Management System: It takes care about the enquiry, order, procedure rate of progress, design rate of progress, and installation function management.
2. Product Information Management System: supporting the product knowledge base, information management of competitive product, and function of product FAQ management.
3. Component Location Management System: supporting supply resources and components management.

4. Customer Support Management System: supporting service resources, maintain service, customer order management, and customer suit function.

Concurrent Collaborative Working

The Concurrent Enterprise is a powerful platform for assuring a comprehensive overall perspective on topics mainly related to Collaborative Working and Information Systems. It is also a powerful instrument for the progressive integration of all participants in the value chain and more specifically, it provides a road-map for Small and Medium Enterprises' (SMEs') participation.

Therefore, this study proposed four relating process to improve the process in each information systems.

1. Union offer and receiving process: making a register and union Quotation platform to avoid snatches the list and let the profit loss. Hoping the salesperson who can focus on undeveloped customer helps company expanding the benefit and market share.
2. Localization Installation Process: Building the global location installation system will control the time schedule, supply source and parts information to provide localization service and save the maintain procedure.
3. Suppliers Management Process: Through the supply source management coordinate with the service process provides maintenance and schedule information for supplier and outside contractor as well as cutting the installation cost, improving the delivery rate of product and controlling the parts quality.
4. Customer Service Process: Through customer order and suit process management and customer satisfaction procedure provides the relational function and message system for their customers to help product system reformation and improve the company's image and quickly response for their customer.

Business Benefit

The major competitive advantage in a company is made by creating value that was proposed by Porter in 1985. There are three parts of components, cost leadership, differentiation and focus. Hence, this study based on these components to discuss the business benefits, such as, efficient, technology capability, industry service, cost down, curtail delivery, quick response, and product quality. If company can follow these seven basic rules, that may success to win the competitive advantage.

Whatever, a good customer response system needs company supporting a lot of service more than opponents. Therefore, this study used the list dimensions, such as customization, reducing response time, and supporting other service, to discuss whether company has competitive advantage on customer response.

1. Efficient: In generally, company always analyzed employees base on their product and their performance outcome, but in this article we joined another outcome variable are including development production, sales production and information production. We are trying to use these variables to be our analysis units.
2. Product Quality: A high quality products can bring both two sides affect for company's competitive advantage. One is standing on value of product by customers imagine, because product which win the customer interesting has high quality that will bring the connection between efficient. However, this study through process of information system on customer interesting and efficient to

discuss whether company can get the competitive advantage in quality side.

3. Innovation: Innovation always separated into two dimensions, one is production innovation and other is procedure innovation. These two forms of innovation can help company bringing some advantages, such as, raising the customers' perspective on company's value, and cutting down the production cost. Therefore, in this study tried to use these indicates to evaluate whether company has competitive advantage in innovation.
4. Customer Response: The Company quickly responds in time for customer's demands and keeps eye on their customer is very important. Because if company can obtain reliance on their customer on positive side, that will affect surviving problem in company. Whatever, a good customer response system needs company supporting a lot of service more than opponents. Therefore, we used these dimensions, such as customization, reducing response time, and supporting other service, to discuss whether company has competitive advantage on customer response.

4 Case Description

This study was focused on front of five metal stamping industry of the world named JF Machine Industrial who is the biggest maker of the mechanical power presses in Taiwan. They want to satisfy their customer's demand at each location, and to confront emerging markets releasing extendable customer's demand on market share. Hence, JF start from service, through Global Headquarter Center in Taipei to provide whole supply chain system and combined with information system. Finally, bring the opposing competitive for JF. Consequently, we adapted to single case design to present the different and difference in JF.

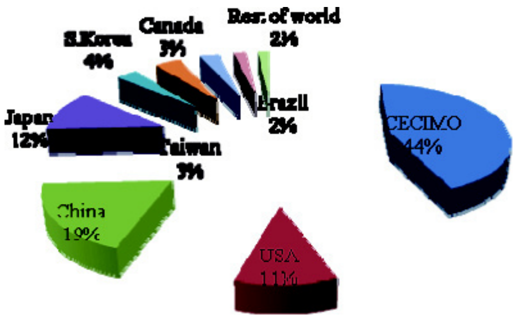


Figure 2. Global machine tool requires of the rate in some main nations or locations in 2003. Resource: Modifying from the result of information project in Ministry of Economic Affairs, R.O.C. (2004)

JF Company who is trying marketing to the world has their own Private Brand named "JF" and "STAMTEC". Therefore, JF Company has national brand is a machine tool provider. Developing countries suddenly appear on the horizon as a result of huge market, and against the feature in machine tool industry, such as

Capital-Intensive, Technology-Intensive, customer around the world and higher uniqueness product.
JF Company must ponder on how to think outside the box of the traditional imagine and fundamental expanding the market share in the world. For this reason, they must preserve the original productivity in standard types, and must aim to customers demand to develop more customization products to conform the customers demand. That will bring more difference in product, and create more brand satisfaction and loyalty.

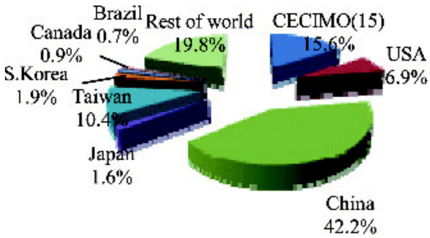


Figure 3. Machine tool exported from Taiwan to other countries in 2003
Resource: Modifying from the result of information project in Ministry of Economic Affairs, R.O.C. (2004)

5 Research Finding

In this study found that each strategy established following company demand. The goal is all for acquiring the market share in machine tool industry, and promoting the customer satisfaction. Hence we have been choosing improvement interaction to be a function to solve the problem between clients and company. Interaction relied on Collaborative Working process. Therefore this research regarded the whole machine industry and hoped that has the following influence.
First is the related innovation business model. In previously, the most of business model is the product manufacturer role; taking the high quality product as the goal. The competition market has the impact regarding the mechanical industry. This research in which expectation can provide Taiwan mechanical industrial from manufacture design driven turn into global localization service changes the service as the guidance.
Second, from the internationalization industry, in past time the machine which produces in Taiwan is a standard in manufacture industry. Following customer will increase from the many countries; manufacture must consider the customer location components standard, and the whole supply chain bring implement system integration as a goal to expect the machine tool industry elaborate their synergy.
Third focused on accelerating efficiency on R&D and reducing technology and commercialized time interval. In previously, R&D data cannot integrate and separate in each R&D engineers' computer, therefore the blueprint searching makes the waste on human cost, and R&D engineers unable to induce the standardized specification product. The R&D center collects the components and blueprint informa-

tion supporting R&D department searching and inducting customization products to develop the standard products.

Last is Off-Site integration source service, building after Global Logistics. Off-site service will expand the outsourcing component suppliers cooperate with maintain service team using outsourcing resource. Strategy must have an applicable tool to help decision maker.

Thus Global Logistics System exists in JF Company which not only through the information to reform in-outside problem but also helps company brings up a difference service, such as collaborative with their partners, to obtain the business value from their clients. However JF Company utilized the information system that obtain the added-value and application to make difference, promote their position and get the competency within the same industry.

From JF's implementation process affects between each Information System (Seeing figure 4). The global logistics system has been considering based on the collaborative, changing process and improving technique; here are two features in the JF's Information Systems. First, they are setting standard type and customization type in the different place of production, making sure the lower cost on standard type and inputting the more interactive process to response their customer in real-time.

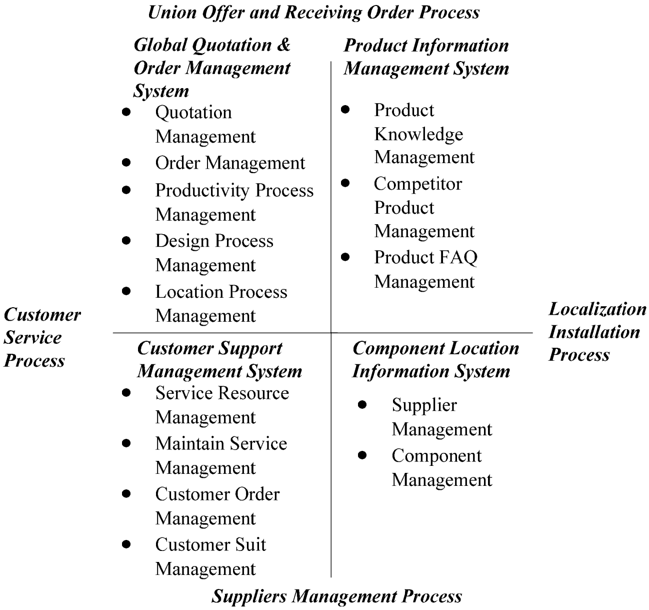


Figure 4. The IS Systems with Each Collaborative Working. (This study)

Second, the transparency on supply chain operation will promote the service quality and precise the delivery date. Heavy Industry technology that always direct to customization type develop capability is a more important thing in this industry. It

represents the high profit, but heavy machine tool which manufacture made must care about on response customer demand in time.
Quickly achieve the machine production, installation, and testing to avoid the stop working lost. Therefore, upgrading the global customization capability, focusing on localization service and reducing the delivery date those are JF Company's vision. It focused on private brand going to do global logistics service, combining the Overseas Sales office, coordinating operation process with partners, engaging the information and management and sharing the information to form a whole interdependent supply chain network. Through coordination commerce integrates upstream and downstream and understands the customer demand strengthening the capability on development technique.
After this application in strategy system, we are aware of the promotion and efficient in JF Company and let JF Company can win the business benefit. Through Concurrent Collaborative Working coordinated with Global Logistics Management, helping JF Company changing in their business model and operation processing. Because of the main systems are connecting to the theoretical practice and general conceptions, but subsystems are discussing deep into the process, useful data filtering and information integration.

6 Discussion

The value of this study has created in three dimensions. First is strategy. JF Company, who has the different thinking, aims to the global industrial service integration. In previously, when company worked on the product research and development always from the passive catalog sales into customer demand, through the customer demand changes the products type. JF Company who has 5% growth rates in the automobile industry without booming situation.
Second is organization process. In the beginning, service department belongs to sales department. After this study that we can figure out, through the BPR based on the service customers and turn service unit into service department that can display the benefit. In the process side, we found JF Company making the Register and Union Quotation Platform to avoid snatching the list mutually and letting the profit loss. Hoping the salesperson who can focus on undeveloped customer helps company expanding the benefit and market share.
Finally, this study bases on the service side to find that introducing the customer service module making location component standards can share with the inside suppliers, promoting the company competency. In the other side, building the supply chain source can improve the location installation quality. Company's service quality can obtain the customer the approval. In the other hand, JF Company started from new service alignment with Global Logistics Management not only promoting business performance and obtaining competency but changing JF's position in market and aiming to customer demand making response.
JF Company will be a leader that is Number one from Manufacture driven turn into Global Industry Service driven in Taiwan, and through the Concurrent Collaborative Working combined with Global Logistics Management to integrate whole supply chain increasing the interaction between suppliers and company. Finally,

Machine tool will be elaborated synergy from which JF Company will get some helps in supply chain.

7 References

[1] Blanchard, D., (2008) “Global logistics is on the move”, Industry Week, 257 (4), pp. 66-67.

[2] Boothroyd G.(1999). Development of DFMA and its impact on US industry. In: Proceedings of the 6th ISPE International Conference on Concurrent Engineering—Research and Applications (ISPE), Bath, UK.

[3] Burgess, R., (1998). “Avoiding supply chain management failure: lessons from business process re-engineering.” International Journal of Logistics Management vol. 9 (1), pp. 15–23.

[4] Carter DE, Baker BS., (1991). Concurrent engineering, product development for the 90’s. London: Addison Wesley Publishing Company.

[5] Chopra, S., Meindl, P., (2001). “Supply Chain Management: Strategy, Planning, and Operation.” Upper Saddle River. Prentice-Hall, NJ.

[6] Closs, D.J., Goldsby, T.J., Clinton, S.R., (1997). “Information technology influences on world class logistics capability.” International Journal of Physical Distribution and Logistics vol. 27(1), pp.4–17.

[7] Cooper, M.C. and J.T. Gardner,(1993). “Building good business relationships - more than just partnering or strategic alliances?”, International Journal of Physical Distribution and Logistics Management, 23, 6: 14-26.

[8] Hanshaw, S., Carter, L., (2008), “Using information technology for strategic growth from single-mission transportation company to multi-faceted global logistics corporation”Journal of Cases on Information Technology, 10 (3), pp. 10-20.

[9] Hoskisson, R.E., Hitt, M.A., Wan, W.P., Yu, D., (1999). “Theory and research in strategic management, swings of a pendulum.”, Journal of Management vol. 25(3), pp. 417–456.

[10] Lambert, D.M., Emmelhainz, M.A., Gardner, J.T., (1996). “Developing and implementing supply chain partnerships.”, The International Journal of Logistics Management vol. 7(2), pp. 1–17.

[11] Oliver, C., (1997). “Sustainable competitive advantage: combining institutional and resource-based view.”, Strategic Management Journal vol. 18(9), pp. 697–713.

[12] Porter, M. E., (1985). *Competitive strategy*, New York: The Free Press.

[13] Porter, M. E., (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*, New York: The Free Press.

[14] Porter, M. E., (1996). What is Strategy, Harvard Business Review.

[15] Prasad B., (1996). Concurrent Engineering Fundamentals, Integrated Product and Process Organization. London: Prentice Hall.

Collaborative Product Development

Using DEA and GA Algorithm for Finding an Optimal Design Chain Partner Combination

Chih-Ling Chuang[a], Tzu-An Chiang[b, 1], Z. H. Che[b] and H. S. Wang[b]

[a] Department of Industrial Engineering & Management, National Taipei University of Technology, Taipei (106), Taiwan, R.O.C.
[b] Department of Commerce Automation and Management, National Pingtung Institute of Commerce, Pingtung (900), Taiwan, R.O.C.

Abstract. To assist enterprises in building the optimal design chain partner combination, this research focuses on the development of a weight restricted DEA model, in which appropriate design chain partners are evaluated and selected according to different partner roles, and appropriate partner sets are formed correspondingly. As product development time and costs depend closely on coordination efficiency among different design chain partner members, this study takes this factor into account by developing a multi-objective design chain partner combination evaluation and selection model. This study uses multi-objective genetic algorithm (MOGA) to search for the optimal design chain partner combination in order to achieve the objective of minimized new product development costs and time in conjunction with maximized product reliability.

Keywords. Design chain partner combination, DEA, multi-objective genetic algorithm, product development.

1 Introduction

Nowadays, a single enterprise is unlikely to complete the product development tasks on its own with limited R&D resources in a short period of time, given the significant increase in product design complexity. Therefore, enterprises can now make full use of the R&D and innovation capacity of the design chain members to enhance the efficiency of collaborative development of low-cost, high-quality, and customer-oriented new products. It is obvious that the collaborative design chain formed by members of different organizations spread across geographical barriers will definitely create far-reaching impacts on new products' market competitiveness and enterprises' profitability in the future. Hence, this study aims to develop a specific and feasible integrated decision-making model to help enterprises in building design chain network and formulating the optimal strategy for design chain partner combinations.

[1] Corresponding Author E-mail: phdchiang@gmail.com

2 Literature Review

Cao and Wang [1] integrate Monte Carlo simulation method with genetic algorithm to evaluate and select suppliers. Wang et al. [2] bring fuzzy theory into genetic algorithms and take the risk and duration of project failure into account for finding an optimal partner combination. Ip et al. [3] use genetic algorithms to solve the cooperation partner selection problem, which takes into consideration the issues of project start and completion time. Chan et al. [4] successfully solve the decision-making problem of supplier selection by FHA using cost, services, quality, and risk as criteria. Mikhailov [5] selects partners by FHA using financial stability, price, quality, customer services as criteria.Wang [6] uses fuzzy multi-criteria xzoutranking method to first obtain performance evaluation for design chain partners by pair-wise comparison. Then, genetic algorithms can be applied to find the optimal combination under time and cost constraints. However, this design chain partner selection model fails to directly represent its quantitative criteria evaluation by the raw measured values, which are always distorted after objective pair-wise comparison. Hence, some researches [7] [8] [9] [10] apply Data Envelopment Analysis (DEA) for evaluation and the selection process. Talluri and Baker [11] present a multi-phase mathematical programming approach for effective selection by: (1) evaluation of supplier, manufacturer and distribution candidates, (2) identification of the optimal number of suppliers, and (3) identification of optimal routing decisions. Wu et al. [12] propose a DEA-based integrated model that combines the previous DEA achievements into a model such as specific factors, unrealistic weight elimination and full ranking technique. Besides, it helps to solve two problems: one is to determine which indices should be considered in the partner selection; the other is to decide which partner is the best one to select. But, DEA loses objectivity in the evaluation and selection of design chain partners as it uses criteria weights most favorable to the decision making unit (DMU).

3 Methodology for Optimal Design Chain Partner Combination

For finding the most appropriate design chain partners, the partner roles are categorized into three types. Based on each role's task types, this study develops corresponding partner evaluation and selection criteria whose relative importance can be expressed through Saaty's nine-scale values [13] to represent the relative triangular fuzzy weights between criteria i and j, namely $\tilde{a}_{ij} = (\alpha_{ij}, \beta_{ij}, \delta_{ij})$, from which the fuzzy positive reciprocal matrix is derived. By applying Equation 1, criteria's weight can be obtained. Due to subjectivity of pair-wise comparison between criteria weights, it is necessary to maintain consistency for the fuzzy positive reciprocal matrix. According to Buckley [14], if a matrix [β_{ij}] has consistence index (CI) less than 0.1, the consistency is maintained for its fuzzy positive reciprocal matrix [$\tilde{a}_{ij}$].

$$w_i = [\tilde{\alpha}_{i1} \otimes \cdots \otimes \tilde{\alpha}_{in}]^{1/n}, \forall i \quad (1)$$

3.1 Selection Criteria System for Integration Partner And Criteria Assessment Model

Fig. 1 shows the hierarchy for system integration partner evaluation and selection in three layers. For qualitative criteria, we use Likert scale [16] to express performance. The quantitative criteria for performance evaluation model are explained as follows.

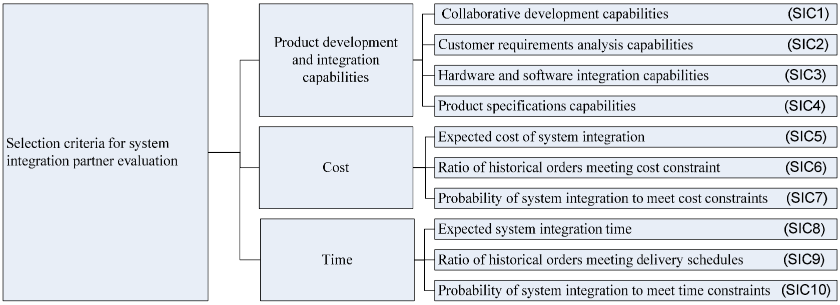


Figure 1. The hierarchy for system integration partner evaluation and selection criteria

˙Expected cost of system integration：This study assumes beta distribution for the cost variable of product development activities. A design chain partner s provides its optimistic cost (oc), most likely cost (mc), and pessimistic cost (pc),to calculate the expected cost of system integration by using into Equation 2.

$$ec_s = \frac{oc_s + 4mc_s + pc_s}{6} \tag{2}$$

˙Ratio of historical orders meeting cost constraints: It is the ratio of the system integration partner's past transactions that meet cost constraints.

˙Probability of system integration to meet cost constraints: Assuming beta distribution for the cost variable of product development activities, two parameters α and β need to be obtained first. Regnier recommends to use $\alpha = 2$ and to use Equation 3 for obtaining parameter β [15]. Finally, the above information can be used to estimate the probability of meeting cost constraints for partner s.

$$\frac{pc_s - oc_s}{mc_s - oc_s} \tag{3}$$

˙Expected system integration time: This study assumes the time variable of product development activities corresponds to beta distribution as well. A design chain partner s provides its optimistic time (ot_s), most likely time (mt_s), and pessimistic time (pt_s) to calculate the expected time for partner s to complete system integration by using Equation 4.

$$et_s = \frac{ot_s + 4mt_s + pt_s}{6} \tag{4}$$

˙Ratio of historical orders meeting delivery schedules: It is the ratio according to the system integration partner's past transactions that meet delivery schedules.
˙Probability of system integration to meet time constraints: Assuming beta distribution for the time variable of product development activities, parameter $\alpha = 2$ and parameter β is obtained by using Equation 5. With these parameters, the probability can be estimated for partner s to meet time constraints.

$$\frac{pt_s - ot_s}{mt_s - ot_s} \tag{5}$$

3.2 Selection Criteria for Functional Module Development Partner And Criteria Assessment Model

Fig. 2 shows the hierarchy for functional module development partner evaluation and selection. The criteria for performance evaluation model are explained as follows.

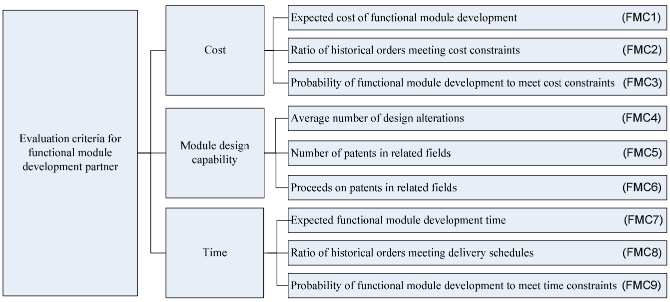


Figure 2. Hierarchy for functional module development partner evaluation criteria

• Number of patents in related fields: It is the number of patents owned by the design chain partner that are related to functional module f .

• Expected cost of functional module development: It is the expected cost of development of functional module f calculated from Equation 4.

• Ratio of historical orders meeting cost constraints: Based on past transactions of the design partner of functional module f , it is the ratio of transactions meeting cost constraints.

• Probability of functional module development to meet cost constraints: The probability is estimated using the same method as for the system integration partner.

• Average number of design alterations: Based on past record of the design chain partner for development of functional module f , it is the average number of design alterations.

• Proceeds on patents in related fields: The net cash proceeds generated by patents that the functional module f development partner owns and licenses to other enterprises, as shown by Equation 6.

$$\sum_{t=1}^{n}\frac{cf_t}{(1+r)^t}-cf_0 \qquad (6)$$

cf_0 : Patent application fee (an initial cost) ；

cf_t : Proceeds acquired from the patent in period t ；

r : Discount rate ； n : Total number of periods for the valid patent 。

• Expected functional module development time: The estimation model for design chain partner is development of functional module f , as shown in Equation 6.

• Ratio of historical orders meeting delivery schedules: It is the ratio of transactions meeting delivery schedules.

• Probability of functional module development to meet time constraints: The probability is estimated using the same method as for the system integration partner.

3.3 Selection Criteria for Software And Component Supplier Partner And Criteria Assessment Model

Fig. 3 shows the hierarchy for supplier evaluation and selection. The criteria for performance evaluation model are explained as follows.

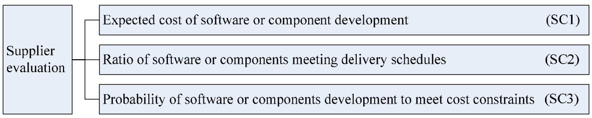


Figure 3. Hierarchy for software and components supplier evaluation and selection

• Expected cost of software or component development: Equation 4 shows the estimation model for a design chain partner's expected cost of software or components development.

• Ratio of software or components meeting delivery schedules: Based on past transactions of the design partner of software or components, it is the ratio of transactions meeting delivery schedules.

• Probability of software or components development to meet cost constraints: The estimated probability for design chain partner to meet cost constraints in software or components development. The probability is estimated using the same method as for the system integration partner in meeting cost constraints.

3.4 Candidate Performance Evaluation Model for Design Chain Partner

By using Equation 7, the composite performance evaluation can be calculated for system integration partners, functional module development partners, and software and components suppliers, respectively. Then, this paper screens out those partners with efficiency value being 1 to form appropriate partner sets. Finally, a design chain partner combination is formed by selecting a partner from each set.

$$\begin{aligned}
&MAX \quad E_s = \sum_{k=1}^{K} b_k Y_{ks} - c_r \\
&s.t. \quad \sum_{k=1}^{K} b_k Y_{ks} - \sum_{i=1}^{I} a_i X_{is} - c_r \le 0, \quad \mathrm{s} = 1 \cdots n \\
&\qquad \sum_{i=1}^{I} a_i X_{is} = 1 \\
&A_i^- \le a_i \le A_i^+, a_i \ge 0 \\
&B_k^- \le b_k \le B_k^+, b_k \ge 0
\end{aligned} \tag{7}$$

E_s : Composite performance for design chain partner s；

c_r : Returns to scale index；

a_{i} , b_{k} : Weights for a design chain partner's input criterion i and output criterion j;

A_{i}^- 、 A_{i}^+ 、 B_{k}^- 、 B_{k}^+ : Upper and lower bounds for the weights.

3.5 Overall Performance Evaluation Model for Design Chain Partner Combination

This study develops a overall performance evaluation model for multi-objective design chain partner combination, as shown in Equation 8. The model's performance evaluation includes such dimensions as total expected cost (TEC) and total expected time (TET) for product development, as well as product reliability (PR). This study incorporates the coordination efficiency factor into the design chain partner combination overall performance evaluation model to adjust for impact of different design chain combinations on time and cost.

$$\begin{aligned}
&Min \quad TEC = \sum_{s=1}^{S} (sic_s * sp_s) + \sum_{f=1}^{F} \sum_{j=1}^{J} (fc_{fj} * fp_{fj} * fce_{fjs}) + \sum_{f=1}^{F} \sum_{p=1}^{P_f} \sum_{k=1}^{K_p} (pc_{fpk} * pp_{fpk} * pce_{fpkj}) \\
&Min \quad TET = \sum_{s=1}^{S} (sit_s * sp_s) + \sum_{f=1}^{F} \sum_{j=1}^{J} (ft_{fj} * fp_{fj} * fce_{fjs}) + \sum_{f=1}^{F} \sum_{p=1}^{P_f} \sum_{k=1}^{K_p} (pt_{fpk} * pp_{fpk} * pce_{fpkj}) \\
&Max \quad PR = \prod_{f=1}^{F} \prod_{p=1}^{P_f} \left(\sum_{k=1}^{K_p} (pr_{fpk} * pp_{fpk}) \right) \\
&subject\ to: \\
&\sum_{s=1}^{S} sp_s = 1; \\
&\sum_{j=1}^{J} fp_{jf} = 1 \qquad \forall f \in F; \\
&\sum_{k=1}^{K_p} pp_{fpk} = 1 \quad \forall f \in F; \quad \forall p \in P
\end{aligned} \tag{8}$$

sic_s ：Expected cost for system integration by partner s；

$sp_s: \begin{cases} 1 & \text{System integration partner s being selected as a design chain combinations member} \\ 0 & \text{otherwise} \end{cases}$；

fc_{fj} ：Expected cost for functional module f development by functional module development partner j ；

$fp_{fj}: \begin{cases} 1 & \text{Functional module development partner j participating in development of functional module f} \\ 0 & \text{otherwise} \end{cases}$；

fce_{fjs} ：Coefficient of cooperation between development partner j of functional module f and system integration partner s；

pc_{fpk} ：Expected cost for development of software or component p in functional module f by candidate supplier k ；

$pp_{fpk}: \begin{cases} 1 & \text{Candidate supplier k develops software or component p in functional module f} \\ 0 & \text{otherwise} \end{cases}$；

pce_{fpkj} ：Coefficient of cooperation between candidate supplier k of component p used in functional module f and functional development partner j；

sit_s ：Expected time for system integration by partner；

ft_{fj} ：Expected time for functional module f development by functional module development partner j ；

pt_{fpk} ：Expected time for development of software or component p in functional module f by candidate supplier k ；

pr_{fpk} : Reliability for development of software or component p in functional module f by candidate supplier k 。

In order to find the optimal design chain partner combination for minimizing total expected cost and time in product development, and for maximizing product reliability, this study uses the MOGA to solve for the design chain partner combination overall performance evaluation model, with steps outlined as follows:

- Step 1: Chromosomes encoding. Each chromosome represents a partner combination. Each chromosome contains N variables. Fig. 4 shows the chromosome structure of a design chain partner combination, in which S represents a system integration partner; F_i represents a developer of functional module i ; p_{ij} represents a supplier participating in development of software or component j used in functional module i .

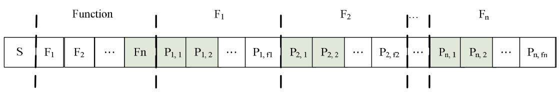


Figure 4. Chromosome structure of a design chain partner combination

● Step 2: Initial population generation. The initial design chain partner combination is formed the individual by randomly selecting a partner from each appropriate partner set. *N* individuals are selected in this method to generate an initial population.

● Step 3: Calculation of objective function values.

● Step 4: Sorting individuals. In this study, chromosomes are sorted according to a method proposed by Deb [17]. First, identify all non-dominated solutions and label corresponding individuals as level 1 before removing them temporarily from the population. Then, select from remaining population those individuals with non-dominated solutions and label them as level 2, followed by temporary removal from the population. Finally, sort the chromosomes according to the non-dominated level.

● Step 5: Calculating fitness function values. A fitness function value is calculate for each individual according to its non-dominated level, as shown in Equation 9 [18], where $F(X_i)$ is the fitness function value for individual i, and $rank_j$ is the non-dominated level for individual j.

$$F(X_i) = 1 - rank_i \Big/ \sum_{j=1}^{N} rank_j \tag{9}$$

● Step 6: Niche. This study uses a shared function to generate penalty on individuals in a relatively dense area with feasible solutions, so as to reduce the fitness function value. Equation 10 shows the shared function for that X and Y, where d(X, Y) is the distance between individual X and individual Y, and σ the specified distance parameter.

After obtaining the shared function, Equation 11 is used to calculate the revised fitness function value $F'(X)$ for individual X.

$$\delta(X,Y) = \begin{cases} 0 & d(X,Y) \geq \sigma \\ 1 - d(X,Y)/\sigma & d(X,Y) < \sigma \end{cases} \tag{10}$$

$$F'(X) = \begin{cases} F(X) & \sum_{Y \in P} \delta(X,Y) \leq 1 \\ \dfrac{F(X)}{\sum_{Y \in P} \delta(X,Y)} & \sum_{Y \in P} \delta(X,Y) > 1 \end{cases} \tag{11}$$

● Step 7: Reproduction. Individuals with higher fitness function values are more likely to be reproduced to mating pool [19].

● Step 8: Crossover. Two individuals are randomly selected from the mating pool for two-point mating, which is done by swapping between two individuals the genetic codes contained in the segment between two specified points.

● Step 9: Mutation. Randomly select from an individual the position for genetic mutation and replace the design chain partner with a newly selected one.

● Step 10: Generating new population. As this study operates on elite selection strategy, individuals will be sorted according to revised fitness function values, and only the top 50% of the population is retained, while the bottom 50% is eliminated.

● Step 11: Repeat until stopping condition is reached

4 Case Study

4.1 Design chain partner selection

Fig. 5 shows the design chain of the network structure for TV box. Before evaluation and selection of design chain partners, the ranges of weights for criteria need to be determined. This study uses FHA method to compile experts' opinions for obtaining fuzzy weights for criteria and, through the consistency test, transitivity of weights is confirmed. Finally, the upper and lower bounds for the weights of criteria can be obtained. Table 1 show the weight ranges of criteria for systems integration partners.

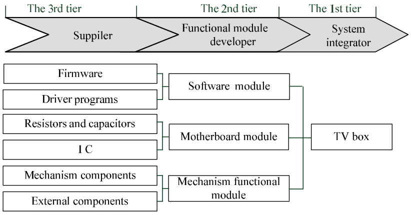


Figure 5. Design chain of the network structure for TV box

Table 1. Weight range for selection criteria of the system integration partner

Surface	Index	lower	upper
Product development and integration capabilities	Collaborative development capabilities	0.707	2.266
	Customer requirements analysis capabilities	1.206	2.975
	Hardware and software integration capabilities	1.737	4.261
	Product specifications capabilities	0.344	0.936
Cost	Expected cost of system integration	0.663	1.781
	Ratio of historical orders meeting cost constraint	0.846	2.012
	Probability of system integration to meet cost constraints	1.080	3.279
Time	Expected system integration time	0.189	0.487
	Ratio of historical orders meeting delivery schedules	0.270	0.794
	Probability of system integration to meet time constraints	0.600	1.668

Equation 9 can be applied to obtain composite performance evaluation for each candidate. Based on the analysis, an efficiency value of 1 represents an efficient partner. Therefore, 10 appropriate partner sets can be formed by screening out appropriate candidates. Finally, this study uses multi-objective genetic algorithm to select from each of the 10 partner sets a partner to form the optimal design chain partner combination. This study begins by setting the size of initial population to be 50, number of generations to be 2500, crossover rate to be 0.7, and mutation rate to be 0.1 Through the setting of above parameters, the optimal design chain partner combination's total expected cost, total expected time, and product reliability are obtained as NT$3,043,400, 307 days, and 0.911 averagely. Table 2 shows the optimal design chain partner combination.

Table 2. Optimal design chain combination

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
candidate in combination #1	8	3	2	7	1	2	4	4	1	4
candidate in combination #2	8	3	1	7	1	2	4	4	1	2

5 Conclusion

Given the significant increase in product design complexity, a modern day enterprise is unlikely to complete the product development tasks on its own with limited product development resources. Hence, an outsourcer often has to set up different evaluation and selection criteria corresponding to various tasks for the selection of appropriate design chain partners. Furthermore, different design chain partner combination will generate different collaborative synergy, which will directly affect the product development time and costs. Faced with this critical decision-making problem, an enterprise needs an objective scientific method to assist a company executives in building the optimal design chain partner combination. Therefore, this study builds a framework for appropriate evaluation and selection criteria based on respective roles of design chain partners, and develops weight restricted DEA model to assist enterprises in finding appropriate design chain members. To build the optimal design chain partner combination, this research develops a multi-objective design chain partner combination performance evaluation model for identifying the design chain partners combination with lowest expected product development costs and time, as well as the optimal product reliability. Finally, this paper uses a case of TV Box product development project to illustrate the practical value and significant benefits of this methodology.

6 Acnowledgments

This research was partially financially supported by the National Science Council under project No. NSC 97-2221-E-131 -021. We also wish to thank the Editor-in-Chief and the two anonymous referees for their valuable comments.

7 References

[1] CAO, H. Y.; WANG, D. W. A simulation based genetic algorithm for risk-based partner selection in new product development. International Journal of Industrial Engineering, vol. 10, pp. 16-25, 2003.

[2] WANG, D. W.; YANG, K. L.; IP, W. H. A heuristic genetic algorithm for subcontractor selection in a global manufacturing environment. Systems, Man and

Cybernetics, Part C: Applications and Reviews. IEEE Transactions on, vol. 31, no. 2, pp. 189-198, 2001.
[3] IP, W. H.; HUANGB, M.; YUNGA, K. L.; WANG, D. W. Genetic algorithm solution for a risk-based partner selection problem in a virtual enterprise. Computers and Operations Research, vol. 30, pp. 213–231, 2003.
[4] CHAN, F. T. S. Interactive selection model for supplier selection process: an analytical hierarchy process approach. International Journal of Production Research, vol. 41, pp. 3549–3579, 2003.
[5] MIKHAILOV, L. Fuzzy analytical approach to partnership selection in formation of virtual enterprises. OMEGA – International Journal of Management Science, vol. 30, no. 5, pp. 393–401, 2002.
[6] WANG, J.; LIN, H. Y. A fuzzy hybrid decision-aid model for selecting partners in the design chain. International Journal of Production Research, vol. 44, no. 10, pp. 2047–2069, 2006.
[7] RAM, N.; SRINIVAS, T.; DAVID, M. Supplier Evaluation and Rationalization via Data Envelopment Analysis: An Empirical Examination. The Journal of Supply Chain Management, vol. 37, no. 3, p28–37, 2001.
[8] WEBER, C.; CURRENT, J.; DESAI, A. Non-cooperative negotiation strategies for vendor selection. European Journal of Operations Research, vol. 108, pp. 208–223, 1998.
[9] CELEBI, D.; BAYRAKTAR, D. An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information. Expert Systems with Applications, vol. 35, pp. 1698–1710, 2008.
[10] WEBER, C.; CURRENT, J.; DESAI, A. Non-cooperative negotiation strategies for vendor selection. European Journal of Operations Research, vol. 108, pp. 208–223, 1998.
[11] TALLURI, S.; BAKER, R. A multi-phase mathematical programming approach for effective supply chain design. European Journal of Operations Research, vol. 141, pp. 544–558, 2002.
[12] WU, J.; LIANG, L.; YANG, F. Partner Selection in Virtual Enterprise: A DEA Approach. Asia Pacific Management Review, vol.12, pp. 223-231, 2007.

[13] SAATY, T. L. Multi-criteria Decision Making: The analytic hierarchy process, McGraw-Hill, New York, 1980.
[14] BUCKLEY, J. J. Fuzzy Hierarchical Analysis. Fuzzy Sets and Systems, vol. 17, p. 233-247, 1985.
[15] REGNIER, E. Activity Completion Times in PERT and Scheduling Network Simulation, DRMI Newsletter, vol. 12, 2005.
[16] LIKERT, R. A Technique for the Measurement of Attitudes. New York: Archives of Psychology, vol. 140, 1932.
[17] DEB K, Multi-objective Genetic Algorithms: Problem Difficulties and Construction of Test Problems, Department of Mechanical Engineering, 1999.
[18] HIRANI, H.; SUH, N. P. Journal bearing design using multi-objective genetic algorithm and axiomatic design approaches. Tribology International, Vol. 38, pp. 481-491, 2005.
[19] GOLDBERG, D. E. Genetic Algorithm in search. Optimization and Machine Learning, Addison Wesley, 1989.

Adaptive Architecture for Collaborative Environments

Youssef Roummieh[1] and Parisa Ghodous[1]

LIRIS Laboratory – SOC Team – Lyon 1 University, France.

Abstract. This paper is an answer to the question: how to make a collaborative architecture adaptive. Firstly, we propose to represent an application or a project by a unique document, the project document, which is a multi-actor document. Secondly, we propose an algorithm that allows generating, for an actor, workspaces from the project document.

Keywords. Collaborative environment, multi-actor document, multi-agent, XML.

1 Introduction

Nowadays, achieving a project has become a complex activity. Many team, distributed geographically, composed of actors come from several disciplines, cooper to achieve a common goal. For example, a lot of actors – belong to different teams, and come from different disciplines such as : computer science, medical, microbiology, etc. – collaborate to follow-up the nosocomial infection [1]. Indeed, each actor hold a part of the whole information and they must collaborate to get the necessery information distributed through the actors.

Much architecture is proposed to realize the collaboration between actors. In this paper, we are based on the multi-agent architecture proposed in [2]. Figure 1 show this architecture which the actors are organized around a blackboard system. It uses this blackboard to communicate its results to other actors. The blackboard represents the collaborative workspace.

This paper does not aim to propose a new collaborative approach, but to extend the current collaborative approach presented in [2] and make it adaptive. Indeed, the proposed architecture is not intended to manage the access to information. Each actor can be procured all information published in the blackboard. However, some information is shared by some actors, but not all. For example, to follow-up the nosocomial infection, patient's information represents a sensitive data and does not access by any actor. Doctors are the only actors that can be access to this

[1] Corresponding Author : Lyon Research Center for Images and Intelligent Information Systems (LIRIS), Service Oriented Computing Team, Bâtiment Nautibus, 43, bd. du 11 novembre 1918, 69622 Villeurbanne Cedex, France; Tel: +33(0)4.72.44.58.84; Fax: +33(0)4.72.43.13.12; Email: {youssef.roummieh, parisa.ghoudous}@liris.cnrs.fr

information. So, each actor must get the relevant information according to his profile. For example, a technician is allowed to retrieve information about the water system, but not about patient. Therefore, any collaborative architecture must be adaptive, i.e. each actor can access only to the information relates to his profile.

This paper presents an algorithm that answer to the question: how to generate workspaces corresponding to the profile's actor? This algorithm takes a multi-actor document and an actor A, and rendered workspaces according to the profile of the actor A. We means by multi-actor document a multi-structure document which each structure correspond to one actor (each actor has a workspace structure different from others).

This paper is organized as follow. Section 2 presents the related work. Section 3 provides some basic principles of our approach. Section 4 provides our proposition, and finally we conclude in section 5.

2 Related Works

Blackboard and Multi-agent systems are used to achieve effective collaboration among a group of independent software entities [3]. The communication in the blackboard system is realized by making data and information accessible to all participants. It is subdivided into spaces and levels; each one corresponding to a particular kind of information. This approach has adopted the central setting. We quote some examples of projects that used this system: DICE (Distributed and Integrated Collaborative Engineering Design) project [4], DESIGN-KIT project [5], GUIDE and FELEX GUIDES (Globally-Unified Interactive Design and Evaluation) [6], ICM (Interdisciplinary Communication Medium) [7] and ANAXAGORE [8].

Otherwise, in the multi-agent system, the agents collaborate concurrently in a distributed environment. Some Various approaches are used to define collaborative design agent-based systems [9]. Among the most important works that have adopted the agent paradigm to define systems, the PACT project [10] can be quoted as one of the first multi-agent systems for collaborative design.

This paper does not aim to present a new collaborative system, but to provide a new technique that makes the collaborative approach adaptive. This work is based on the collaborative approach proposed by [2].

3 Background

Collaborative workspace is a representation of the common knowledge and each actor, during the development process, can retrieve information from this shared knowledge space, and make requests on it.

[2] proposed a multi-agent architecture (figure 1) in which each participant is represented by an agency. An agency is a multi-agent system which each agent represents the activity related to a particular discipline of an application. In this paper, an agent means an actor. Figure 1 shows the agencies organised around a

blackboard system. Every agency use the blackboard to communicates its results to other agencies.

In each agency, as in terms of architecture, the actors are organised around a blackboard system. Therefore, the communication can be achieved in hybrid manner: directly by an exchange of messages, telephone, etc., or by the blackboard.

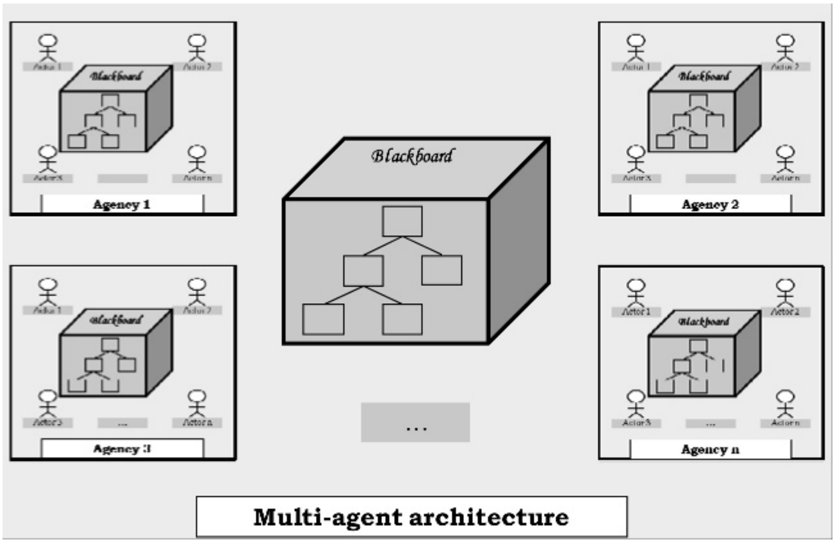


Figure 1. A collaborative architecture.

Each actor has his own profile that is different from another. Therefore, an actor cannot retrieve that information's relates to his profile. A collaborative architecture must be adaptive, i.e. it must be able to extract for each actor his own workspaces.

4 Adaptive Collaborative Environment

In this section, we answer to the question: how to make a collaborative architecture adaptive? Specifically, how we can generate workspaces for an actor from a unique document? Firstly, we proposed to represent the application or the project by an unique document. We called this document a project document which has a hierarchy structure (XML document). Project document is multi-actor document, i.e. it represent the all actor's workspaces. Secondly, we proposed an algorithm that allow us to generate, for each actor, workspaces from the project doucment.

4.1 Project Document

Project document is an XML document. It is a multi-actor document. We means by multi-actor document a multi-structure document which each structure correspond to one actor. Each actor has structure different from others.

Figure 2 shows the XML Schema for the project document. A project is devised into activities. An activity can be taken by one or more actors. Consequently, many actors can collaborate to accomplish an activity. The attribute visibility in activity element indicates the set of actors which can access to an activity. For example, if visibility = “A1” in activity act1, then A1 is the only actor which can be access to act1. Moreover, the visibility = “public” in activity act2, means that all actors can be access to act2.

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
    <xs:element name="activity">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="actor" maxOccurs="unbounded" />
            </xs:sequence>
            <xs:attribute name="visibility" type="xs:string" use="required" />
            <xs:attribute name="id" type="xs:ID" use="required" />
        </xs:complexType>
    </xs:element>
    <xs:element name="actor">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="data" maxOccurs="unbounded" />
            </xs:sequence>
            <xs:attribute name="id" type="xs:ID" use="required" />
        </xs:complexType>
    </xs:element>
    <xs:element name="data">
        <xs:complexType mixed="true">
            <xs:attribute name="id" type="xs:ID" use="required" />
        </xs:complexType>
    </xs:element>
    <xs:element name="project">
        <xs:complexType>
            <xs:sequence>
                <xs:element ref="activity" maxOccurs="unbounded" />
            </xs:sequence>
            <xs:attribute name="id" type="xs:ID" use="required" />
        </xs:complexType>
    </xs:element>
</xs:schema>
```

Figure 2. The XML Schema for the Project document.

4.2 Filter Algorithm

In the collaborative environment, actors belong to several disciplines and expertises such as computer science, medical, microbiology, etc. therefore, each actor has workspaces different to another according to his own profile. To extract workspaces for a given actor we propose a filter algorithm.

Filter algorithm takes as input the projected document d and an actor A, filters out all markup elements in d that are not related to A, and produces as output two documents: D1 corresponds to the shared workspace and D2 corresponds to the private workspace.

```
Algorithm : filter
input   : XML document d
          actor A
output : XML document D1 //shared docuemnt
         XML document D2 //private docuemnt

start parsing document d;
token = nextToken();//first token : project
fputs(token, D1);
fputs(token, D2);
while (more tokens)
    token = nextToken()
    if(token is SE)
        if(tag(token) = activity)
            fputs(token, D1);
            tok = token;
            if(visibility is public or A included in visibility)
                token = nextToken(); // this token is actor
                while(tag(token) != activity)// activity is spread among actors
                    _private = false;
                    if(actor = A)
                        _private = true;
                        fputs(tok, D2); // tok is the token that correspond to model
                    token = nextToken();
                    while(tag(token) != actor)
                        fputs(token, D1);
                        if(_private is true)
                            fputs(token, D2);
                        token = nextToken();
                    token = nextToken();
            else // actor can not access to model
                token = nextToken();
                while(tag(token) != activity)
                    token = nextToken();
    else // token is EE
        fputs(token, D1);
    token = nextToken();
```

Figure 3. The pseudo-code for filter algorithm.

The algorithm parsing the document d. for this purpose, we use the Simple API for XML (SAX) [11]. SAX allows to parse an XML document and to construct your own data structures using your call-back event handlers. The most important events are the start and end of the document, the start and end of elements, and character data. In this paper, we are interesting by: start elements (SE), end elements (EE).

5 Conclusion

In this paper, we present an adaptive collaborative environment. In this environment, each actor has worspaces different from another according to his own profile. We are proposed to create one document that represents all actors : a multi-actor document. This document contains all actor's documents and we generate the document that represents the workspace of an actor from the multi-actor document. The proposed technic allows the dynamic modelisation. Indeed, the representation of all worksapces by one document implies that the project can be updated to reflect any changes. Moreover, the information remain consistent, which means that the coherence between the different workspaces is guaranteed.

6 References

[1] Roummieh Y, Ghodous P, Vanehms P, Verdier C, Dutra M. Collaborative Environment for the Detection and the Follow-up of Legionella. In: Methods and Design models for Health Information Systems, workshops in conjunction with IEEE ICDIM'2007, Lyon, France, 2007; 609-614.

[2] Slimani K. Système d'échange et de partage de connaissances pour l'aide à la Conception Collaborative [in French]. Ph.D. thesis, University of Claude Bernard Lyon 1, Lyon, France, 2006.

[3] Daniel D Corkill. Collaborating Software: Blackboard and Multi-Agent Systems & the Future. In Proceedings of the International Lisp Conference, New York, 2003.

[4] Sriram RD. Distributed and Integrated Collaborative Engineering Design. Savren, 2002. ISBN 0-9725064-0-3.

[5] Stephanopoulos G, Johnston J, Kriticos T, Lakshmanan R, Mavrovouniotis M, Siletti C. DESIGN-KIT: An object-oriented environment for process engineering. Computer & chemical engineering. Vol. 11, No 6, 1987; pp. 655-674.

[6] Morse DV and Hendrickson C. A Communications Model to Aid Knowledge-Based Design Systems. Artificial Intelligence for Engineering Design, Analysis and Manufacturing. Vol. 4, No. 2, 1990; pp. 99-115.

[7] Reiner KA, Toye G, Fruchter R and Leifer LJ. Collaborative mechatronic system design. In Concurrent Engineering: Research and Applications. Vol. 4, No 4, 1996; pp. 401-412.

[8] Trousse B. Cooperation between knowledge-based systems and CAD tools: the ANAXAGORE environment. In Third International Symposium on AI: Applications of engineering design, Manufacturing & Management in Industrialized and developing Countries, ITESM, Mexico, 1990; pp. 83-89.

[9] Pahng GDF, Bae S, and Wallace D. A web-based collaborative design modeling environment. In Proceedings of the IEEE Workshop on Enabling Technologies Infrastructure for Collaborative Enterprises (WET ICE'98), 1998; pp. 161-167.

[10] Cutkosky MR, Engelmore RS, Fikes RE, Genesereth MR, Gruber T, Mark WS, Tenenbaum JM, and Weber JC. PACT: An experiment in integrating concurrent engineering systems. IEEE Computer. Vol. 26, No 1, 1993; pp. 28-37.

[11] Simple API for XML (SAX). Available at: <http://www.saxproject.org/>.

Conceptual Modelling of Design Chain Management towards Product Lifecycle Management

Wei Liu[a,b], Yong Zeng [b,1]

[a] Department of Management – Northwestern Polytechnical University, China.
[b] Institure for Information Systems Engineering – Concordia University, Canada.

Abstract.. Exisitng design chain management reference models either address collaboration or process management issues. This paper aims to develop a comprehensive model for the design chain management by applying the Environment-based Design theory. Based on the the identified conflicts between design chain management environment components, a set of solutions are derived that cover the areas of the product design process, the collaboration hierarchy model and the design chain management process.

Keywords. design chain management, environment-based design, conceptual model.

1 Introduction

In today's industrial production, PLM - Product Lifecycle Management has become an essential tool for coping with the challenges of more demanding global competition, mass customization, ever-shortening product lifecycles and increasing product complexity [1, 2]. CIMData [3] defines PLM as "a strategic business approach that applies a consistent set of business solutions in support of the collaborative creation, management, dissemination, and use of product definition information across the extended enterprise from concept to end of life – integrating people, processes, business systems, and information". Design chain management is progressively addressed as one of such solutions in industrial products and research efforts (see e.g. [4] and [5]), due to its paramount importance for industries to develop innovative products within a shorter lead-time [6]. It is fairly well known that up to 80% of product cost is committed by the time the product is designed. Therefore, effective management of the design phase during the whole product lifecycle is necessary.

The term Design Chain Management was defined as the management of the participants, both internal and external to a focal firm, that contribute the capabilities (knowledge and expertise) necessary for the design and development of a product which, on completion, will enable full-scale manufacture to

[1] Corresponding Author E-mail: zeng@ciise.concordia.ca

commence [7]. Consequently, the design chain involves participants throughout the product development process, from concept, detail engineering, process engineering, prototype manufacturing, through to post-launch activities [8], including designers, suppliers, manufacturers and customers. Research in this field is mainly concerned with the early involvement of suppliers in new product development [9, 10], and the cross-industry diagnostic tool for design-chain management- Design Chain Operation Reference-model (DCOR) [11] as well as its extension [12], and the collaboration in the design chain [13, 14]. Other studies in design chain management are focused on selecting suitable design chain partners for co-development success [15-17]. To our best knowledge, not much attention has been paid to investigating a conceptual model of the design chain management (DCM) in addressing product lifecycle management considerations. Such a conceptual model is the foundation for further developing the design chain management.

The rest of this paper is organized as follows. In Section 2, we review collaborative product development and supply chain management literature. Section 3 introduces a scientific approach of design: Environment-Based Design (EBD). The conflicts existing in the environment components are analyzed to identify the requirements of design chain management and the proposed conceptual model is presented in Section 4. In the final section, we make a summary of our study, discuss the limitations, and suggest future directions.

2 Literature Review

Research studies in relation to the field of collaborative product development and supply chain management have been firstly reviewed in this section. Then, we move our attention to addressing the challenges most companies faced in the field of the design chain management.

2.1 Collaborative Product Development

The topic of collaborative product development and the technologies associated with implementing it are nowadays receiving ample attention. The increasing competition and globalization of market has led collaborative product development (CPD) to become a new yet inevitable trend for today's new product development (NPD). Organizations have to collaborate more with their suppliers, their customers, and other relevant parties in the current business environment, to respond to key customer needs, market opportunities, and technology changes, and reduce the time to market, cost and development risk [18]. Collaboration issues [18], supplier involvement [10], information technology [19, 20], and theoretical modeling [21] are main streams of research. According to the previous definition of design chain management, DCM and CPD emphasize different aspects, although OEMs and Tier 1’s are aggressively pursuing what CIMdata calls "collaborative product development"—collaborative product design and design chain management [6]. More precisely, DCM is a management process, involving plan, research, design, integrate, and amend, whereas CPD is an approach more associated with design and implementation of product.

2.2 Supply chain manamgement

The term supply chain management (SCM) was originally introduced by consultants in the early 1980's [22] and has subsequently gained tremendous attention. Lambert *et al* [23] defined SCM as the integration of business processes from end user through original suppliers that provide products, services, and information that add value for customers. The supply-chain council (SCC) proposed a supply chain operations reference model (SCOR) for benchmarking supply chain processes and designing IT solutions for SCM [24]. Cooper *et al* [25] identified seven business processes within supply chain management. Although the consensus remains that SCM is more than simply logistics, most supply chain literature examined procurement and value-adding activities, without explicitly defining product development as its part [8]. However, most of a product's competitive characteristics and costs are determined and committed by the activities of the design chain [26]. Furthermore, SCM can lead to a significant reduction in production costs, whereas it does not address other important factors that drive product competitiveness, especially innovation and time-to-market. Therefore, design chain management is becoming as important, if not more, as the logistics and production supply chain. While design chain is a subset of supply chain [14], managing the design chain is more challenging than managing the supply chain due to the inherent complexity such as diverse design processes, dynamic design environments, higher task uncertainty, more complex information (often in an incomplete form), and new buyer–supplier relationship[8, 14].

2.3 Challenges for design chain management

As previously mentioned, design chain management differs from CPD due to its feature of management process, and contrasts to SCM as it takes a design- centric point of departure, as depicted in Figure 1.

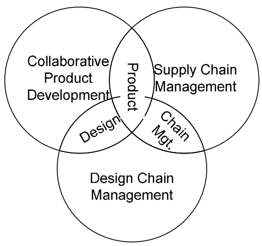


Figure 1. DCM & CPD & SCM

The ultimate objective of design chain management is to lower R&D costs, reduce time to market and support product innovation through integrating and leveraging knowledge, technologies, and resources among all participants in design chain towards effective product lifecycle management. Currently, DCM in most companies encounters a lot of difficulties, such as:

1. hard to select appropriate partners [16, 17, 27];
2. lack of effective design chain strategy [27];
3. collaboration obstacles coming from different organizational goals and strategies, diverse design processes, information sharing and protection, weak communication among collaborators, and incompatible infrastructures such as applications, facilities, network and standards [6, 12, 13, 27-29];
4. inevitable changes in the design environment [13, 28, 30].

Respective design chain management reference models have been developed to tackle partial collaboration[8, 12, 13], change management[5, 14], or design chain process[11]. However, unlike the case of supply chain management, there is as no comprehensive model to design chain management. Such a design chain management conceptual model can be regarded as a design problem by itself, which will be developed to explicitly identify the scope of design chain management issues in the following sections.

3 A brief review on Environment-Based Design (EBD)

Over the last several decades, a variety of design theories and methodologies have been proposed, such as systematic design methodology [31], Theory of Inventive Problem-Solving (TRIZ) [32], axiomatic design [33], General Design Theory [34] and Axiomatic Theory of Design Modeling [35]. Axiomatic Theory of Design Modeling is a logical tool for representing and reasoning about object structures. It states that everything in the universe is an object and there is relation between objects. A core concept is structure operator $\oplus$: $\oplus O = O \cup (O \otimes O)$, where O is an object, $\otimes$ is relation from the object to itself, and $\oplus O$ is the structure of the object O. An object may include other objects. The key in applying this theory is to identify the primitive objects and relations underlying the concerned phenomenon, based on which the phenomenon can be formally studied.

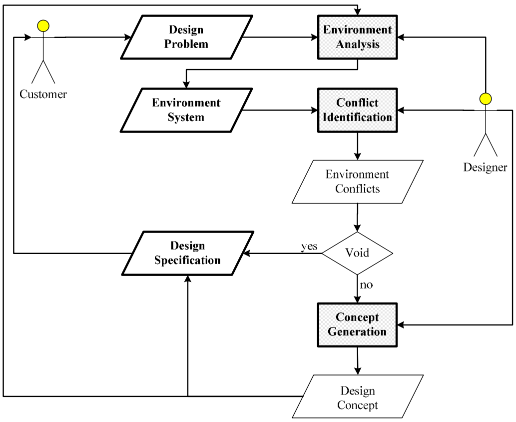


Figure 2. Environment based design: process flow [36]

Environment-based Design (EBD) [37] was derived from the Axiomatic Theory of Design Modeling in solving the design governing equation: $\oplus E_{i+1} = K_i^s(K_i^e(\oplus E_i))$. EBD conforms to a common observation that a design process starts when an unacceptable conflict is identified from an environment, which includes natural, human, and built environments. Solution concepts are then produced to resolve the environment conflict. As a descriptive model of design, the three steps included in the EBD, which are environment analysis, conflict identification, and concept generation, reflect how an effective designer designs and makes decisions. As a prescriptive model of design, the three steps included in the EBD provide an effective guideline as to how a design problem can be solved. As is illustrated in Figure 2, these three steps work together to progressively and simultaneously generate and refine the design specifications and design solutions.

The following explains the three steps included in EBD [36]:

Step 1: Environment analysis: define the current environment system $\oplus E_i$; find out the environment components, where the product exists, and the relationships between the environment components.

$$\oplus E_i = \oplus(\bigcup_{j=1}^{n_e} E_{ij}) = \bigcup_{j=1}^{n_e}(\oplus E_{ij}) \cup \bigcup_{j_1=1}^{n_e} \bigcup_{\substack{j_2=1 \\ j_2 \neq j_1}}^{n_e} (E_{ij_1} \otimes E_{ij_2}). \tag{1}$$

where n_e is the number of components included in the environment E_i at the i^{th} design state; E_{ij} is an environment component at the same design state. It should be noted that decisions on how many (n_e) and what environment components (E_{ij}) are included in E_i depend on designer's experience and other factors relevant to the concerned design problem.

Step 2: Conflict identification: identify undesired conflicts C_i between environment components by using evaluation operator K_i^e, which depends on the interested environment components.

$$C_i \subset K_i^e(\bigcup_{j_1=1}^{n_e} \bigcup_{\substack{j_2=1 \\ j_2 \neq j_1}}^{n_e} (E_{ij_1} \otimes E_{ij_2})). \tag{2}$$

Step 3: Concept generation: generate a design concept s_i by resolving a group of chosen conflicts through a synthesis operator K_i^s. The generated concept becomes a part of the new product environment for the succeeding design.

$$\exists c_{ik} \subset C_i, K_i^s : c_{ik} \rightarrow s_i, \oplus E_{i+1} = \oplus(E_i \cup s_i). \tag{3}$$

The design process above continues with new environment analysis until no more undesired conflicts exist.

4 Conceptual model of DCM

4.1 Environment Analysis

According to the environment-based design theory, a design problem is implied in a product system which is initially defined by the environment. Environment can be generally classified into natural, built, and human environments, denoted by E^n, E^b, and E^h, respectively [35].

$$\oplus E = \oplus (E^n \cup E^b \cup E^h) . \quad (4)$$

Table 2 shows the structure of the environment of design chain management. In the environment of DCM, the natural environment includes objects such as time, space, and natural resource.

The built environment includes objects such as products, organizations, standards, design data, design knowledge, methodologies, IT tools, design processes and design strategy.

Table 1 Environment components for design chain management

Environment	Components	Sub-components
E^h Human environment	E_1: design chain participants	E_{11}: developers E_{15}: suppliers E_{12}: manufacturers E_{16}: transporters E_{13}: distributors E_{17}: customers E_{14}: maintainers E_{18}: recyclers
	E_2: product design goals	E_{21}: innovative product E_{22}: fast time to market E_{23}: low cost E_{24}: high profit E_{25}: expanded market demand E_{26}: mass customization E_{27}: easy, safe to transport E_{28}: easy to use E_{29}: easy, safe to recycle E_{210}: easy to maintain E_{211}: less change …
E^b Built environment	E_3: product design processes	E_{31}: requirements E_{32}:conceptual design E_{33}: detail engineering E_{34}: process engineering E_{35}: prototype manufacturing E_{36}: testing E_{37}: post-launch activities
	E_4: product design information	E_{41}: configuration E_{42}: specification E_{43}: BOM E_{44}: engineering changes …
	E5: product design standards	E_{51}: STEP E_{52}: PDML E_{53}: U3D E_{54}: XML …
	E_6: IT Tools	E_{61}: SCM E_{62}: PM E_{63}: CAD E_{64}: CAE E_{65}: PDM E_{66}: EDM E_{64}: ERP …
	E_7: products …..	E_{71}: mechanical products E_{72}: aerospace products E_{73}: electronic products E_{74}: service

		E_{75}: automatic products …
E^n Natural environment	E_8: time E_9: space…	…

Design data is generated at the beginning of product life and rarely changes during the lifetime of the product, normally including BOMs, the specification of the product, specific component identification, configuration options, operation instructions, material content, disassembly attributes (e.g. sequence and tools) and recycling information, etc [38].

IT tools can be distinguished between four different types, depending on their functionality [39]: (a) communication tools, ranging from simple telephones and fax machines, to e-mails, video conferencing, shared databases and e-Business solutions; (b) visualization tools, three dimensional CAD systems (such as CATIA, ProEngineer and IDEAS) are the most popular type of tools; (c) calculation tools, consist mainly of simulation software and mathematical prototypes/models, such as CAE tools; (d) collaboration tools, for example, Collaborative product commerce (CPC) enterprise solutions use the internet to create an online community for all development partners allowing them to collaborate in creating, developing and managing products throughout the entire product design and development process.

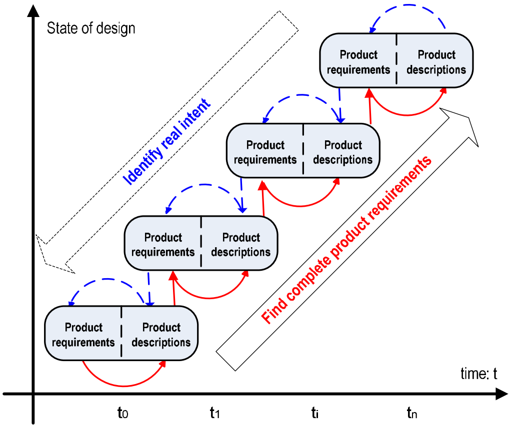


Figure 3 Evolution of the design process [40].

As previously mentioned, product development process is well known starting from concept design, detail engineering, process engineering, prototype manufacturing, and testing, through to post-launch activities. Many existing DCM models exclude the requirements management phase from design chain management [6, 8]. However, based on the logic of design [41], design is a recursive process in which a satisfying design solution must pass an evaluation defined by the design knowledge that is recursively dependent on the design solution to be evaluated. Since the design knowledge, which implies the design criteria, is part of the design problem, the generation of design solutions indeed changes the original design problem. Based on this logic, the design process is described as a series of design states defined by both product descriptions and product requirements [40], as is shown in Figure 3. Therefore, it is fundamentally impossible to distinguish design problem and design solutions. An effective DCM model must integrate design concept evolution and requirements change.

The human environment includes participants in design chain and their goals for product design. Design chains can be defined as sets of engineering and technology deployment organizations that work together to provide solutions in the form of intellectual property, such as design information, or knowledge, that are subsequently manifested as a supply chain. However, from the viewpoint of both the product lifecycle and the co-evolution design process, participants not only involves suppliers and customers but also include developers, manufacturers, transporters, distributors, maintainers, and recyclers [42].

Hence, the structure of DCM environment can be represented as the union of the structure of time, space, products, IT tools, product design data, standards, processes, participants and their goals, as well as other sub-components, and relations between them. Relations can be observed including design, use, share, inquiry, communication, etc.

4.2 Conflict Identification and Concept Generation

In the environment of DCM, various relations exist between two components or from a component to itself. Conflicts may emerge between two relations. For example, developers cannot access the information they need to design their parts of the product. Such conflict can be represented as

$$c(e_{11} \otimes_{\neg access} e_4, e_{11} \otimes_{require} e_4) \quad (5)$$

To remove this conflict, available solutions have to be designed and performed until developers can accesses the required information. Collaboration between the developers e_{11} and participants (for example, manufacturers e_{12}) who own such information would be a solution, which can be formulized as:

$$\left.\begin{matrix} c(e_{11} \otimes_{\neg access} e_4, e_{11} \otimes_{require} e_4) \\ e_{12} \otimes_{own} e_4 \end{matrix}\right\} \rightarrow s_{collaboration} \begin{pmatrix} e_{11} \otimes_{communicate} e_{12} \\ e_{12} \otimes_{share} e_4 \\ e_{11} \otimes_{access} e_4 \end{pmatrix} \quad (6)$$

Consequently, DCM can be regarded as a management process, providing a set of solutions to address various unacceptable conflicts among the relations between the environment components of DCM. A representative model of design chain management can be derived as,

$$\mathrm{S}_{\mathrm{DCM}} = \bigcup_{i=1}^{n} \left\{ \mathrm{s}_i \left| \begin{matrix} C_i \\ R_i \end{matrix} \right. \right\}, \exists C_i \subset C, \exists R_i \subset R, \quad (7)$$

where C_i is a subset of conflicts set C, and R_i is a subset of Relations set R. Providing the condition of C_i and the resource of relations R_i, solutions can be generated to remove the conflicts and change the relations. We investigate manifest conflicts existing in the environment of design chain management, and suggest corresponding solutions listed in Table 3.

Table 2. Solutions to selected conflicts

Conflict descriptions	**Suggested Solutions**
Various suppliers have different ability	Design partner selection
Incompatible goals of different participants	Goals collaboration
Various priorities, and outsourcing strategies for different participants	Priorities & strategy selection
Process interdependence	Process collaboration
consistency of engineering data	Change management
Data access problems	Secure information sharing
Communication problems in disparate systems	System integration
Various routine patterns of behavior or output	Standardization

As seen in the above analysis, diverse competitive goals, outsourcing strategies, design processes, information semantics, data format, sharing mechanism, legacy applications, and standards in the design chain make collaboration as an important feature of design chain management. Therefore, a collaboration hierarchy model is derived, as shown in Figure 4. There are six levels implied in the collaboration: technology standards, applications, information, processes, strategies, and goals. The collaboration in the design chain can be established from top to bottom, and realized from bottom to top. The solution to the collaboration of goals and strategies would be communication and negotiation among design chain participants, while the collaboration of design processes, configuration information can be supported through the functionalities such as project management, product data management, change management, configuration management, document management, and knowledge management. The applications integration and technology standards unification form the basis of information processing and exchange.

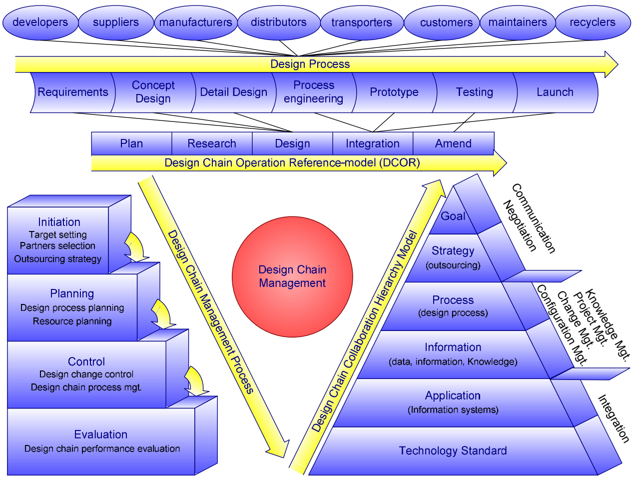


Figure 4. Conceptual model of design chain management

Furthermore, a holistic conceptual model of design chain management should enclose operation and management issues. Figure 4 presents the mapping from

DCOR to the product design process from requirements, concept design, detail design, through process engineering, prototype, testing to product launch and the involvement of all stakeholders of product lifecycle management in the DCM.

From the management perspective, design chain management involves a series of management processes including initiation, planning, control, and evaluation. In the initiation phase, OEM set competitive targets, select appropriate design partners, and chose outsourcing strategies. Change control needs to be activated in the complex design chain environment once the resource, design processes are planned. At last, design chain performance evaluation is essential to improve the cooperation ability in the design chain, thus increase effective product innovation.

In such conceptual model, diverse elements and aspects such as involved participations, design process, design chain operation, design chain management process, and design chain collaboration, are taken into account, which is significant for the effective development of potential DCM methodologies or supporting tools.

6 Conclusion and Future Work

In this paper, we summarize the challenges in the design chain management by distinguishing design chain management from supply chain management and collaborative product development. Existing design chain management reference models tend to address part of issues, such as supplier-customer collaboration, strategy collaboration, change management, and process management. This present work contributes to the design chain management by applying environment-based design theory to design a comprehensive conceptual model of DCM. The conceptual model involves a set of solutions to remove unacceptable conflicts between environment components implied in the design chain management. The proposed conceptual model reveals that the design process, the collaboration hierarchy model, and design chain management process are imperative components of the design chain management solutions.

Future work includes validation of the conceptual DCM model and more conflicts identification based on the recursive logic of design with the purpose of refining design chain management model in the context of Product Lifecycle Management.

7 Acknowledgements

This project is supported by NSERC-CRD (PJ 350114-06). We are grateful to the financial support from NSERC, CRIAQ, Pratt & Whitney Canada Corp, Bombardier Inc., CAE Inc., CMC Electronics Inc., and Rolls-Royce Canada Limited. The first author is partially supported by CSC (The China Scholarship Council).

8 References

[1] Saaksvuori, A. and A. Immonen, Product Lifecycle Management. Edition: 2, illustrated, revised ed. 2005: Birkhäuser. 247.

[2] Ameri, F. and D. Dutta, Product Lifecycle Management Needs, Concepts and Components. 2004, University of Michigan, Ann Arbor, MI. pp. p. 2.
[3] Amann, K., Product lifecycle management: empowering the future of business. 2002, CIM Data, Inc.
[4] Ming, X.G., et al., Technology solutions for collaborative product lifecycle management - Status review and future trend. Concurrent Engineering-Research and Applications, 2005. 13(4): pp. 311-319.
[5] IBM Product Lifecycle Management. [cited 2009.02.02]; Available from: ftp://ftp.software.ibm.com/software/applications/plm/resources/plm_design_chain_management.pdf.
[6] Garg, M., Supply Chain Management And Design Chain Management, Production & Industrial Engineering Department of Mechanical & Industrial Engineering, Indian Institute of Technology Roorkee.
[7] Twigg, D., Design chain management. The Blackwell Encyclopaedic Dictionary of Operations Management. 1997: in Slack, N. (Ed.), T Blackwell, Oxford.
[8] Twigg, D., Managing product development within a design chain. International Journal of Operations & Production Management, 1998. 18(5): pp. 508 - 524.
[9] Akira, T., Bridging inter- and intra-firm boundaries: management of supplier involvement in automobile product development. Strategic Management Journal, 2001. 22(5): pp. 403-433.
[10] Dowlatshahi, S., Implementing early supplier involvement: a conceptual framework. International Journal of Operations & Production Management, 1998. 18(2): pp. 143 - 167.
[11] Nyere, J., The Design-Chain Operations Reference-Model, Supply Chain Council (SCC).
[12] Wu, W.-H., S.-C. Yeh, and L.-C. Fang, The development of a collaborative design chain reference model for the motorcycle industry. The International Journal of Advanced Manufacturing Technology, 2007. 35(3-4): pp. 211-225.
[13] Choi, Y., K. Kim, and C. Kim, A design chain collaboration framework using reference models. The International Journal of Advanced Manufacturing Technology, 2005. 26(1-2).
[14] Shiau, J.-Y. and H.M. Wee, A distributed change control workflow for collaborative design network. Computers in Industry, 2008. 59(2-3): pp. 119-127.
[15] Nagarajan, R.P., et al. Performance Measures and Metrics for Collaborative Design Chain Management. in 10th International Conference on Concurrent Enterprising. 2004. Escuela Superior de Ingenieros, Seville, Spain.
[16] Lin, H.-Y., A Robust Decision Model for Selecting Partners in Design Chain Management, in *Graduate Institute of Industrial Engineering*. 2003, Feng Chia University.
[17] Wang, J. and H.Y. Lin, A fuzzy hybrid decision-aid model for selecting partners in the design chain. International Journal of Production Research, 2006. 44(10): pp. 2047 - 2069.
[18] Littler, D., F. Leverick, and M. Bruce, Factors Affecting the Process of Collaborative Product Development: A Study of UK Manufacturers of Information and Communications Technology Products. Journal of Product Innovation Management, 1995. 12(1): pp. 16-32.
[19] Huang, G.Q., J. Huang, and K.L. Mak, Agent-based workflow management in collaborative product development on the Internet. Computer-Aided Design, 2000. 32(2): pp. 133-144.
[20] Wang, L., et al., Collaborative conceptual design--state of the art and future trends. Computer-Aided Design, 2002. 34(13): pp. 981-996.

[21] Case, M.P. and S.C.Y. Lu, Discourse model for collaborative design. Computer-Aided Design, 1996. 28(5): pp. 333-345.
[22] Oliver, R.K. and M.D. Webber, Supply chain management - logistics catches up with strategy, in Logistics - Strategic Issues, M.G.E. in Christopher, Editor. 1992, Chapman and Hall: London. pp. 63-75.
[23] Lambert, D.M., M.C. Cooper, and J.D. Pagh, Supply Chain Management: Implementation Issues and Research Opportunities. The International Journal of Logistics Management, 1998. 9: pp. 1-20.
[24] Stewart, G., Supply-chain operations reference model (SCOR): the first cross-industry framework for integrated supply-chain management. Logistics Information Management, 1997. 10(2): pp. 62 - 67.
[25] Cooper, M.C., Douglas M. Lambert, and Janus D. Pagh, Supply Chain Management: More Than a New Name for Logistics. The International Journal of Logistics Management, 1997. 8(1): pp. 1-14.
[26] Designing collaboration. 2000 [cited 02/02/2009]; Available from: http://www.hinduonnet.com/businessline/2000/11/22/stories/242239ye.htm.
[27] McGrath, M.E. (2002) STRETCH YOUR R&D FUNDS.
[28] James, G. DESIGN CHAIN DISASTERS. [cited; Available from: http://www.designchain.com/featurearticle.asp?id=2&issue=fall04.
[29] Chuang, W.-C. and H.-H. Yang, Design Chain Collaboration– A Strategic View. International. Journal of Electronic Business Management, 2004. 2(2): pp. 117-121.
[30] Ouertani, M.Z., Supporting conflict management in collaborative design: An approach to assess engineering change impacts. Computers in Industry, 2008. 59(9): pp. 882-893.
[31] Hubka, V. and W. Eder, Theory of Technical Systems. 1988, New York: Springer-Verlag.
[32] Altshuller, G.S., CREATIVITY AS AN EXACT SCIENCE: The Theory of the Solution of Inventive Problems. 1984: Gordon and Breach Science Publishers.
[33] Suh, N., The principles of design. 1990, New York: Oxford University Press.
[34] Yoshikawa, H. General Design Theory and CAD System. in Man-Machine Communication in CAD/CAM, Proceedings IFIP WG 5.2 Working Conference. 1981. North Holland: Amsterdam.
[35] Zeng, Y., Axiomatic theory of design modeling. Transaction of SDPS: Journal of Integrated Design and Process Science, 2002. 6(3): pp. 1-28.
[36] Zeng, Y., Environment-based design: process model. 2004, Concordia Institute for Information Systems Engineering, Concordia University: Montreal. pp. 40.
[37] Zeng, Y., Environment-based formulation of design problem. Transaction of SDPS: Journal of Integrated Design and Process Science, 2004. 8(4): pp. 45-63.
[38] Schedit, L.G. and S. Zong. An approach to achieve reusability of electronic modules. in Procedings of the IEEE International Symposium on Electronics & the Environment. 1994. San Francisco, USA: IEEE.
[39] Yassine, A., et al., Investigating the role of IT in customized product design. Production Planning & Control, 2004. 15(4): pp. 422–434.
[40] Zeng, Y. and P. Gu, A science-based approach to product design theory Part II: Formulation of design requirements and products. Robotics and Computer Integrated Manufacturing, 1999. 14(4): pp. 341-352.
[41] Zeng, Y. and G.D. Cheng, On the logic of design. Design Studies, 1991. 12(3): pp. 137-141.
[42] Ming, X.G., et al., Mass production of tooling product families via modular feature-based design to manufacturing collaboration in PLM. Journal of Intelligent Manufacturing, 2007. 18(1): pp. 185-195.

Data Persistence in P2P Backup Systems

Rabih Naïm Tout[a,1], Parisa Ghodous[a], Aris Ouksel[b] and Mihai Tanasoiu[c]

[a] University of Lyon1 – Villeurbanne, France.
[b] University of Illinois – Chicago, USA.
[c] Alter Systems Company – Lyon, France.

Abstract Peer-to-peer (P2P) networks have been shown to be a natural and efficient paradigm for modeling most internet applications. However, data persistence remains still a major challenge, particularly in highly dynamic and unstable P2P networks. Within the framework of collaborative engineering, we propose a probabilistic approach based on peers collaboration to guarantee persistence of critical data in a system. Markov chains are used to model applications and realistically capture the behavior of practical systems. The model is first analytically investigated, and then data persistence is measured using an erasure coding redundancy scheme. The mathematical analysis allows us to determine the extent of data persistence in several important cases, and to anticipate the robustness of the large scale dynamic distributed applications.

Keywords: Peer-to-peer, Redundancy, Persistence, Erasure coding, Probabilistic guarantee.

1 Introduction

Peer-to-peer (P2P) networks have gained a lot of recognition in recent years. Internet growth resulted in a paradigm shift from the traditional client-server model to P2P systems. To ensure high scalability and prevent bottlenecks and single points of failure in the network, P2P systems do not rely on knowledge of the global state of the network, which would have required centralized information. Each peer is autonomous and the network evolves according to the peers' local connections. Peers share equally in services and responsibilities. The availability of powerful desktop computers and broadband networks led to the emergence of P2P file sharing such as CAN [1], Chord [2] and Pastry [3] where millions of autonomous users share resources in the same network. , These new structures have enabled the development of novel P2P practical applications (e.g. µTorrent, LimeWire, eMule…) and, importantly, the question of insuring resilient backup support in this environment. By exploiting free hard disk space available at each peer in the network, promising P2P backup techniques have been developed as robust as those for client-server ones.

[1] Corresponding Author E-mail: rabih.tout@liris.cnrs.fr

Some of the most notable research projects on P2P backup include pStore [4], PeerStore [5] and Pastiche [6]. These systems provide the user with the ability to securely backup and restore files in a distributed network of peers. Selected files for backup are encrypted to prevent other members of the backup network from seeing their private contents. Then, each file is broken into digitally signed blocks that are stored in different peers over the P2P network. Each peer determines other peers with whom to exchange replicas. To exchange replicas, peers must collaborate between each other. In this environment, the problem is to determine the number of block copies (replicas) to maintain in the network and the set of nodes where these replicas must be stored.

P2P networks are dynamic. Peers can dynamically join or leave the system at any time. A peer may leave the system voluntarily or as a result of a crash. This raises critical issues such as how to ensure that the availability of critical data regardless of the dynamic nature of the network. Consider the desire to have a file backed up... Our approach is to split a file into fixed-sized blocks. These blocks are first encrypted, then replicas created, and finally, blocks are distributed to a subset of nodes (core) in the network. The problem is to determine the nodes receiving the replicas such that, regardless of the dynamic scenario, there will be a minimum number of nodes sufficient to restore the original file. In addition, replication ought not to be excessive as it results in large overheads in storage and bandwidth, and certainly, update costs. In this paper, we develop probabilistic guarantees for ensuring reliability in the face of a dynamic P2P environment. The guarantees are expressed as a function of several key parameters including the size of the network, the size of the core, the number of replica and the churn rate of the system. This study is for structured P2P networks and does not address the locating scheme problem. In structured P2P, locating the resource is reliable when the resource exists.

Our model uses a Markov chain approach where four possible states are defined for each node in the system: *connected* or *disconnected*, *with* or *without critical data*. To the best of our knowledge, this paper is the first paper to model this problem in this manner. We analyze the case of using only an erasure coding scheme and generate a formula to compute the probability of restoring a file in the face of failures in the system. We then show that the use of erasure coding with replication can give the formulas more flexibility in choosing parameters values.

The remainder of the paper is organized as follows: in section 2, we discuss survey pertinent related work. In section 3, the proposed model and its analysis are presented. In section 4, simulation results are shown and their implications commented. Finally we conclude in section 5 with the review of the problem studied, the results achieved, and our perspective for future work.

2 Related Work

Storage redundancy is a critical factor in reliable P2P systems. This issue has not received extensive attention as others in the literature. In [7], the notion of a persistent set of nodes (core) is defined to guarantee the availability of critical data. Several factors are considered to achieve high reliability, including the size of the core, the frequency at which the core has re-established and the churn rate of the system. This specific model does not consider the subdivision of a file into several blocks. Whole files are instead sent across the network, and this may result in excessive overhead in the case of large files. When a file is subdivided into a sequence of blocks, each of them may be identified independently so that a block-level approach may allow downloading different parts of an object simultaneously from different peers, and thus reduce overall download time. In addition, the cost of replicating blocks may be so small that the aggregate cost of replicating the multiple blocks of one file is smaller than that of replicating a whole file.

In the same paper, an unrealistic assumption is made also: a node leaving the system does not remember its state. This means that any node that enters the system is considered as a new node. This is not the case of real systems. In real P2P networks, a node leaves the system either voluntarily or because it crashes. When a node disconnect from the system and if no problems occurs, it can return later (e.g. turning off a computer and turning it on later, or losing internet connection for a period of time). If we consider that a left node will never come back again, this means that each time a node holding critical date leaves, we lose a replica of critical data. To replace this data, we need to replicate a new copy to another node in the system. As the left node will never come back, this means that each time a node holding data leaves, new replicas should be done and so we need more replicas to maintain the persistence of the core.

Two redundancy schemes exist to achieve high durability of data. The complete replication scheme [8, 9] and parity schemes such as RAID [10]. While the first one is a simple redundancy scheme, which does not incur excessive bandwidth and storage overhead, the second one, does not provide sufficient robustness to permit survival in an environment of a high rate of failures. Besides replication schemes, Peer-to-peer DHTs have proposed another redundancy scheme based on erasure coding [11, 12]. Erasure codes, such as Reed-Solomon [13] and Tornado [14], divide an object into m blocks that are recoded into n redundant blocks (fragments) (where $n > m$). $r = \frac{m}{n} < 1$ is the rate of encoding. Using this redundancy scheme, the original object can be reconstructed from any m distinct blocks.

Comparisons [15, 16] have shown the superiority of erasure coding over replication schemes as it achieves lower storage requirements. These conclusions [17] have refined to indicate that superiority is achieved only when node availability is low. Erasure coding has been revisited in [18], where the optimization problem of determining the number n of blocks per file is studied. The theoretical analysis indicates that even higher values of n achieve higher availability, this does not seem to match realistic situation.

A hybrid replication scheme has been proposed in [19]. It runs as an erasure coding scheme, but where some of the already decoded files are kept in the system to provide a data service similar to that of the replication scheme. This increases the availability of the file and the probability of finding it without replication.

In [20], Markov chains are used to analyze the persistence problem; Formulas are derived for data persistence in structured P2P networks of three different redundancy schemes: replication, erasure coding and hybrid. The model supposes only three possible states for each node: online, offline and leave. Churn rate of the system is used to determine the probability of a node permanently leaving this system. There is ambiguity as to whether the leave is permanent or reconnection is anticipated some time down the road. Permanent leave is confirmed either from information obtained directly from the node or from an offline timer. The study focuses on the calculation of the mean time to failure for the three redundancy schemes. It does not however address the issue of determining the minimum number of replicas required to insure data availability for a period of time. Their evaluation shows that the erasure coding scheme and the hybrid scheme result in better data persistence than the replication scheme, except in situations of higher churn rate and low average node availability.

3 Proposed Model

Our model is designed to study data persistence using the erasure coding redundancy scheme. The hybrid redundancy scheme is not investigated as it relies on decoding blocks of any file. For security reasons, we want users to be able to only decode their own files. In backup context, a backed up file will be restored and decoded only by its owner.

Consider a P2P network S consisting of N nodes. It is assumed that a node may leave the system either voluntarily or because of a technical problem or a crash. A node that left the system may rejoin after a period of time and this may happen at any time.

Using erasure coding, a file is divided into m blocks that are recoded into n blocks (m < n). At the start, the system S is composed of a core Q of n nodes holding the n blocks (one block per node). The number of nodes not holding any block is $N-n$. Let $S(t)$ be the system state at time t. The system state $S(t+1)$ at time $t+1$ depends only on $S(t)$ and not on the previous states. It can be represented by a Markov chain. Each node may be in any of the four states shown in (Figure 1).

First state	Second state	Third state	Fourth state
Connected and has data	Connected and doesn’t have data	Disconnected and has data	Disconnected and doesn’t have data

Figure 1. Node states table

As the state space is finite, the transition probability distribution is represented in the following transition matrix P, where the $(i, j)th$ element of P is equal to:

$$P_{ij} = \Pr\left(X_{n+1} = j \;\middle|\; X_n = i\right) \tag{1}$$

This transition matrix is a stochastic matrix and is the transition matrix for one step. The *k*-step transition probability matrix is equal to P^k.

Consider a system $S(t)$ at time t. It is composed of 4 sets nodes (A_1, A_2, A_3, A_4) representing all the nodes in their different states. Thus, for example, A_1 and A_2 are sets of connected nodes, A_3 and A_4 are those with disconnected nodes, A_1 and A_3 are those nodes with replicas, whereas A_2 and A_4 are those without replicas. Let N_i denote the cardinality of A_i. Thus: $N_1 + N_2 + N_3 + N_4 = N$.

Let c be the percentage of nodes that leave or join the system per time unit. Thus, $p = \frac{c}{100}$ is the probability that a node change its state (leave or rejoin the system) after a unit of time, and $q = 1 - p$ is the probability that a node remains in the same state after a unit of time.

The one-step transition matrix is:

$$P = \begin{pmatrix} q & 0 & p & 0 \\ 0 & q & 0 & p \\ p & 0 & q & 0 \\ 0 & p & 0 & q \end{pmatrix} = qI + pJ \qquad (2)$$

where: $I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ and $J = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$.

For example: let $u^0 = (1, 0, 0, 0)$ represent the state of a node in set A_1 at time t=0. Then, its state at time 1 is:

$$u^1 = u^0 P = (q, 0, p, 0) \qquad (3)$$

Which means that at time t=1, this node has a probability q to remain in A_1 (to stay connected with replicas) and a probability p to be in A_3 (to disconnect with replicas);

$$u^2 = u^1 P = u^0 P^2; \ldots; u^k = u^0 P^k \qquad (4)$$

Let $X^0 = (N_1, N_2, N_3, N_4)$ denote the cardinalities of each set at time t=0. The cardinalities of the various sets at time t=1 and t=k are $X^1 = X^0 P$ and $X^k = X^0 P^k$, respectively

In our case, $X^0 = (n, N_2, 0, N_4)$ with $n + N_2 + N_4 = N$, where n is the number of connected nodes having replicas. N_2 is the number of connected nodes that don't have replicas, N_3 is equal to zero initially as there are no disconnected nodes with replicas., and N_4 is the number of disconnected nodes that do not have replicas.

This allows us to infer that:

$$X^k = X^0 P^k = (a_1, a_2, a_3, a_4) \tag{5}$$

where, after k units of time, the number of connected nodes with replicas is $\lceil a_1 \rceil = s_1$ and the number of connected nodes without replicas is $\lceil a_2 \rceil = s_2$ .All the other nodes (a_3, a_4) are offline.

Given any set of n random nodes in $S(k)$, the probability that the set has at least m unique blocks is:

$$\frac{\sum_{i=m}^{S_1} \binom{S_1}{i} \binom{S_2}{n-i}}{\binom{S_1 + S_2}{n}} \tag{6}$$

In other words, this is the probability to restore the original file (probability of success).

Parameter	Description
c	Percentage of node that join or leave the system per time unit
m	Number of blocks a file is divided into
n	Number of blocks after erasure coding (fragments)
k	Number of time units

Figure 2. Parameters table (Parameters used in simulations)

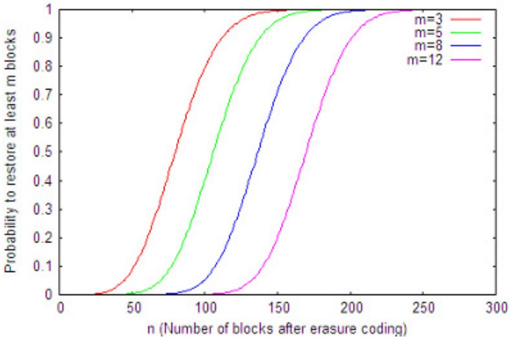


Figure 3. Probability to restore at least *m* blocks

In (Figure 3), the system is composed of 5000 nodes (2000 online and 3000 offline). At the start the offline nodes do not have any data. c=0.1, k=100. For multiple values of *m*, we compare the probability of restoring the *m* blocks (restoring the file) in relation to *n* (the number of blocks after erasure coding). As we can see, when m is bigger, we need more erasure coding blocks in order to restore these m blocks after k units of time.

A file subdivided into m blocks is first converted to a set of n blocks using the erasure coding. Following this encoding, the blocks are then distributed over the peers in the network. Given a file of *m*=3 blocks, (Figure 3) indicates that the restoration of these blocks will require an encoding into *n*=110 blocks after erasure encoding to guarantee a high probability of success. This is obviously excessive.

5 Conclusion

In this paper, we focused on data persistence in structured P2P networks. We motivated the importance of this subject in the introduction and defined the three existing redundancy schemes used for this purpose. Unlike previous methods, our approach used Markov chain modeling to define four states for peers in our system. In so doing, we bring enhance its realism. We then analyzed the problem and deduced some formulas to calculate the persistence of the size of a core set of nodes holding critical data. The formulas combine several parameters, and are usable in multiple scenarios. The values of the parameters can be determined by the requirements of the application. Future work should consider two additional states that are birth (new nodes entering the system) and death (permanent leave) of nodes in the system.

6 References

[1] Ratnasamy S, Francis P, Handley M, Karp R, Shenker S (2001) a scalabe content addressable network. In: Proc of ACM SIGCOMM, 2001, pp 161–172.

[2] Stoica I, Morris R, Karger D, Kaashoek MF, Balakrishnan H (2001) Chord: a scalable peer-to-peer lookup service for Internet applications. In: Proc of ACM SIGCOMM, 2001, pp 149–160.
[3] Rowstron A, Druschel P (2001) Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer system. In: Proc of Middleware, 2001, pp 329–350.
[4] C. Batten, K. Barr, A. Saraf, and S. Trepetin. pStore: A secure peer-to-peer backup system. Technical Memo MITLCS-TM-632, Massachusetts Institute of Technology Laboratory for Computer Science, October 2002.
[5] Landers, M., Zhang, H., and Tan, K. 2004. PeerStore: Better Performance by Relaxing in Peer-to-Peer Backup. In Proceedings of the Fourth international Conference on Peer-To-Peer Computing (P2p'04) - Volume 00 (August 25 - 27, 2004). P2P. IEEE Computer Society, Washington, DC, 72-79.
[6] Cox, L. P., Murray, C. D., and Noble, B. D. 2002. Pastiche: making backup cheap and easy. SIGOPS Oper. Syst. Rev. 36, SI (Dec. 2002), 285-298.
[7] Core Persistence in Peer-to-Peer Systems: Relating Size to Lifetime. Vincent Gramoli, Anne-Marie Kermarrec, Achour Mostéfaoui, Michel Raynal, Bruno Sericola. *Proceedings of the On The Move International Workshop on Reliability in Decentralized Distributed systems (OTM'06)*, apr 2006. Springer LNCS 4278 p.1470-1479.
[8] BOLOSKY, W., DOUCEUR, J., ELY, D., AND THEIMER, M. Feasibility of a serverless distributed file system deployed on an existing set of desktop PCs. In *Proc. of Sigmetrics* (June 2000).
[9] DRUSCHEL, P., AND ROWSTRON, A. Storage management and caching in PAST, a large-scale, persistent peer-to-peer storage utility. In *Proc. of ACM SOSP* (2001).
[10] PATTERSON, D., GIBSON, G., AND KATZ, R. The case for raid: Redundant arrays of inexpensive disks, May 1988.
[11] F. Dabek, J. Li, E. Sit, J. Robertson, F. Kaashoek, and R. Morris. Designing a DHT for low latency and high throughput. In *Proc. NSDI '04*.
[12] J. Kubiatowicz, D. Bindel, Y. Chen, S. Czerwinski, P. Eaton, D. Geels, R. Gummadi, S. Rhea, H. Weatherspoon, W. Weimer, C. Wells, and B. Zhao. Oceanstore: An architecture for globalscale persistent storage. In *Proc. ASPLOS 2000*.
[13] Reed S, Solomon G (1960) Polynomial codes over certain finite fields. J SIAM 8:300–304.
[14] Byers JW, Luby M, Mitzenmacher M, Rege A (1998) a digital fountain approach to reliable distribtuion of bulk data. In: *Proc of SIGCOMM, 1998*, pp 56–67.
[15] R. Bhagwan, K. Tati, Y. Cheng, S. Savage, and G.Voelker. Total recall: System support for automated availability management. *In Proceedings of NSDI*, 2004.
[16] H. Weatherspoon and J. Kubiatowicz. Erasure coding vs. replication: A quantitative comparison. *In Proceedings of the First International Workshop on Peer-to-Peer Systems (IPTPS 2002)*.
[17] Rodrigo Rodrigues and Barbara Liskov. High availability in DHTs: Erasure coding vs. replication. In Proceedings of IPTPS, 2005.
[18] Lin, W. K., Chiu, D. M., and Lee, Y. B. 2004. Erasure Code Replication Revisited. In Proceedings of the Fourth international Conference on Peer-To-Peer Computing (August 25 - 27, 2004). P2P. IEEE Computer Society, Washington, DC, 90-97.
[19] Chen, G., Qiu, T., and Wu, F. 2008. Insight into redundancy schemes in DHTs. J. Supercomput. 43, 2 (Feb. 2008), 183-198.
[20] LI, H. and Chen, G. 2007. Data Persistence in Structured P2P Networks with Redundancy Schemes. *In Proceedings of the Sixth international Conference on Grid and Cooperative Computing (August 16 - 18, 2007)*. GCC. IEEE Computer Society, Washington, DC, 542-549.

Competitive Design

Using DEA Approach to Develop the Evaluation and Priority Ranking Methodology for NPD Projects

Ling-Chen Hung[a], Tzu-An Chiang[b,1] , Z. H. Che[a] and H. S. Wang[a]

[a]Department of Industrial Engineering & Management – National Taipei University of Technology, Taiwan, R.O.C.
[b]Department of Commerce Automation & Management – National Pingtung Institute of Commerce, Pingtung (900), Taiwan, R.O.C.

Abstract. With the shortening of product life cycles, new product development (NPD) has become a critical factor to the sustainable growth of enterprises. This paper proposes an integrated evaluation and priority ranking methodology of NPD projects. It uses fuzzy hierarchical analysis (FHA) to set the reasonable weights of evaluation criteria, and Data Envelopment Analysis (DEA) to analyze efficiency, in order to extract NPD projects with market potential and high added value. Given limited corporate resources, priority ranking can be implemented for NPD projects with market potential, providing a basis for senior management in developing R&D strategy. Finally, this paper explains and validates the application and significant efficiency of the proposed methodology, and offers a case study of NPD projects for electronic products.

Keywords. New product development (NPD), priority ranking, FHA, DEA.

1 Introduction

According to Product Development & Management Association (PDMA) (2004), among the top 20% companies, 42% of their gross profits were contributed by sales of new products. It also indicated that, 41% of new product development (NPD) projects were unsuccessful. NPD has great significance, yet also with higher risks. Given limited corporate resources, companies need to carefully evaluate and select NPD projects before investing large NPD resources, in order to reduce NPD risk and enhance profit-making capability. This research applied Data Envelopment Analysis (DEA) to evaluate NPD projects with original data. However, DEA is based on the most favorable weights to the Decision Making Units (DMU), and from the perspective of DMU, the relative significance of NPD project evaluation criteria should be set within a reasonable range to provide the most appropriate and correct NPD evaluation results. To better assist companies

[1] Corresponding Author. E-mail: allen@npic.edu.tw

controlling the NPD process, this research aimed to develop a benchmark-based NPD project evaluation model by combining fuzzy hierarchical analysis (FHA) and DEA, and thus, identify the NPD project with potential value. Super efficiency methods need to be used for priority ranking of NPD projects. The analytical results could improve NPD project evaluation and minimize business risks.

The remainder of this paper is organized as follows: Section 2 reviews the research works related to NPD project evaluation; Section 3 describes the methodology of NPD project evaluation; Section 4 illustrates the application of NPD project evaluation cases; and Section 5 gives conclusions and suggestions based on the research findings.

2 Literature Review

NPD project evaluation plays a crucial role in the future development of companies, thus, NPD evaluation must be considered in a well-rounded manner. Faced with complex decision-making problems, many researches used Analytic Hierarchy Process (AHP) to conduct well-structured and systematic evaluation, thus helping senior management to evaluate the alternatives in a scientific and quantitative approach [1-3]. Since traditional AHP could not present the uncertainties and fuzziness of subjective judgments, FHA was gradually used to improve this problem. For instance, Sun et al. (2008) applied FHA to evaluate R&D projects, and constructed a Decision Support System (DSS), based on expert opinions [4].

In recent years, DEA has been applied to evaluate the alternative packages by using the input and output of DMU for performance evaluation. For example, Linton et al. (2002) and Eilat et al. (2006) employed a DEA model to analyze, sort, and select R&D projects [5, 6]. When calculating the relative efficiency of DMU, DEA was based on the most favorable weights. For decision-makers, the relative significance among criteria should be set within a reasonable range; otherwise, appropriate NPD projects may not be selected from alternative plans.

Based on the above literature review, it is known that traditional DEA is very suitable to evaluate the relative advantages of decision plans for DMU. Due to the characteristics of decision problems, NPD project evaluation has some blind points and unsatisfactory results. To resolve the abovementioned problems, FHA has been used to obtain the reasonable range of criteria weight, which is then set as the weight constraint condition of DEA model. This would make the analytical results of DEA more reasonable, and thus, solve NPD project evaluation problems in a practical decision environment. Given limited NPD resources, companies are often unable to implement potential NPD projects. Therefore, this research aimed to perform priority ranking for NPD projects with potential value, thus providing a valuable reference for senior management in establishing R&D strategies.

3 Evaluation methodology of NPD projects

Figure 1 shows the process of NPD project evaluation and priority ranking. Since NPD is subject to uncertain internal and external factors, such factors should be taken into account to reduce the failure rate when evaluating NPD projects. This research referenced foreign periodicals and research reports to collect key factors that affect NPD. Based on literature reviews and field interviews, 4 evaluation criteria were concluded, as listed in Figure 1. Moreover, FHA was used to calculate the weight range of evaluation criteria, and the criteria performance evaluation model was developed to obtain specific performance indicators.

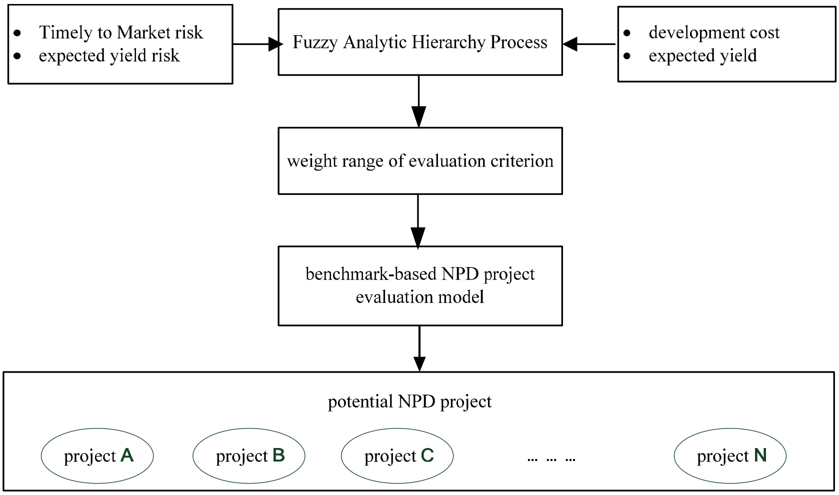


Figure 1. Process of NPD project evaluation and priority ranking

3.1 Setting of NPD evaluation model and weight range

Prior to NPD project evaluation, a performance evaluation model for NPD evaluation criteria must be decided as the basis of performance evaluation. Qualitative evaluation criteria include TTM (Time-to-Market) and expected yield risks. Experts' opinions were used to quantify risk probabilities. The quantitative model for development costs and expected yield are shown in Eqs. (1) and (2):

$$\mathrm{PTC_i} = \sum_{j=1}^{J}(\mathrm{p}t_{\mathrm{ij}} \times pc_j) + \sum_{k=1}^{K}(et_{\mathrm{ik}} \times ec_k) + \mathrm{mt_i} \times mc + \mathrm{e}tt_{\mathrm{i}} \times etc \quad (1)$$

where, $\mathrm{PTC_i}$ is the total cost of i-th project;

mt_{i} is the marketing and field interview hours of i-th project;

ett_{i} is engineering's commissioned hours of i-th project;

pt_{ij} is the R&D work-force hours of i-th project;

et_{ik} is the input hours of k-th R&D facility of i-th project;

mc is the unit time cost of marketing and investigation;

pc_j is the unit time cost of j-th R&D work-force;

ec_k is the unit time cost of k-th R&D facility;

etc is the unit time cost of engineering's commission;

$$NPV = \sum_{t=0}^{n} \frac{cf_t}{(1+k)^t} \tag{2}$$

where, cf_t is the cash earnings of new product projects in t-th phase;
K is the annual discount rate;
n is the survival years of investment project.

To calculate the weight range of NPD project evaluation criteria, this research applied a 9-point scale and triangular fuzzy numbers developed by Satty (1980), to express the relative significance of the criteria, as listed in Table 1. The Fuzzy Positive Reciprocal matrix can be obtained by pairwise comparison, as shown in Eq.(3).

Table 1. Scale of relative significance

Scale	Linguistic variable
1	Equal Importance
3	Weak Importance
5	Essential Importance
7	Very Strong Importance
9	Absolute Importance
2、4、6、8	ranging between 1.3.5.7.9

Eqs.(3) and (5) (Buckley, 1985) were employed to calculate the relative fuzzy significance weight of the evaluation criteria. Due to inconsistent expert judgment, if the Consistence Index (C.I.) of positive reciprocal matrix $[M_{ij}]$ extracted from the median value M_{ij} of triangular fuzzy number is less than 0.1, it indicates that the fuzzy positive reciprocal matrix$[\tilde{a}_{ij}]$ is consistent, as shown in Eq.(4). Finally, the relevant weights were rated by several experts together, and combined into the final rating. The opinions of experts must be combined using geometric averaging method, as shown in Eq.(5):

$$\left[\tilde{a}_{ij}\right] = \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \cdots\cdots & \tilde{a}_{1n} \\ \tilde{a}_{21} & \tilde{a}_{22} & \cdots\cdots & \tilde{a}_{2n} \\ & & & \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \cdots\cdots & \tilde{a}_{nn} \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

$$C.I. = \frac{\lambda_{\max} - n}{n-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

$$\widetilde{w}_i = (\widetilde{a}_{i1} \times \widetilde{a}_{i2} \times \widetilde{a}_{i3} \times .....\widetilde{a}_{in})^{\frac{1}{n}} \quad ....(5)$$

where, $\widetilde{a}_{ij}$ is the triangular fuzzy number of j-th column, i-th line in fuzzy positive reciprocal matrix;

$\widetilde{w}_i$ is the single fuzzy weight;

$\lambda_{\max}$ is the maximum eigenvalue;

n is the array dimension;

C.I. is the consistence index;

$\widetilde{W}_i$ is the triangular fuzzy weight of i-th criterion after combination.

3.2 NPD project evaluation model and priority ranking method

After obtaining the fuzzy weight of NPD project evaluation criterion, a DEA model was introduced to establish a mathematical model for new product evaluation, as shown in Eq. (6). This model was applied to the VRS BCC-I model of DEA. If the rating of evaluation criterion is smaller and better, it indicates that it is the input variable (e.g. cost and risk), otherwise, it is the output variable (e.g. expected yield).

$$Max \quad E_k = \sum_{r=1}^{t} u_r Y_{kr} - c_k$$

$$\text{s.t.} \quad \sum_{i=1}^{s} v_i X_{ki} = 1 \qquad (6)$$

$$\sum_{r=1}^{t} u_r Y_{kr} - c_k - \sum_{i=1}^{s} v_i X_{ki} \le 0$$

$$a \le \frac{u_r}{u_g} \le b, c \le \frac{v_i}{v_g} \le d, k = 1,...,n, r = 1,...,t, g = 1,...,t, i = 1,...,s,$$

c_k without positive/negative constraint

where, E_k is the efficiency rating of k-th case;

Y_{kr}, X_{ki} are the r-th output and i-th input of k-th project;

u_r is the relative weight of r-th output item;

v_i is the relative weight of i-th input item;

ε is the non-Archimedean quantity;

c_k is used to judge the scale returns.

Since the efficiency rating calculated in the analysis process is a relative value, many DMUs with an efficiency rating of 1 may exist in the same group, thus, making it impossible to distinguish the order of priority. Therefore, this research adopted the Super-Efficiency Method proposed by Anderson and Peterson (1993) to sort the efficient items, as shown in Eq.(7) [7]. When the super-efficiency project of the efficient items is obtained, it is possible to conduct priority ranking depending on the scores:

$$Max\ h_k = \sum_{r=1}^{t} u_r Y_{kr} - c_k$$

$$\text{s.t.} \quad \sum_{i=1}^{s} v_i X_{ki} = 1 \qquad (7)$$

$$\sum_{r=1}^{t} u_r Y_{jr} - c_k - \sum_{i=1}^{s} v_i X_{ji} \le 0$$

$$a \le \frac{u_r}{u_g} \le b\,, c \le \frac{v_i}{v_g} \le d\,, \mathrm{j} \in \mathrm{k}\,, k = 1,\dots,n\,, \mathrm{k} \notin \mathrm{j}\,, r = 1,\dots,t\,, g = 1,\dots,t\,,,$$

$i = 1,\dots,s\,, c_k$ without positive/negative constraint

where, h_k is the super-efficiency of k-th case;

Y_{kr} , X_{ki} are the r-th output and i-th input of k-th project;

u_r is the relative weight of r-th output item;

v_i is the relative weight of i-th input item;

ε is the non-Archimedean quantity;

c_k is used to judge the scale returns.

4 Case Study

This section will explain and validate the feasibility of the research method, based on 8 new products of an electronic company. After interviewing 3 experts, an FHA model was used to combine their opinions on the significance level of NPD project evaluation criteria, as listed in Table 2.

Table 2. Fuzzy performance rating of new package calculated by FHA

Evaluation criterion \ Triangular fuzzy weight		Lower limit	intermediate limit	upper limit
1	expected yield risk	1.51	2.26	2.75
2	Timely to Market risk	0.59	1.13	1.48
3	development cost	0.28	0.45	0.71
4	expected yield	0.66	0.86	2.09

With the performance-rating model of NPD project evaluation criteria, it is feasible to estimate the criteria performance of every project, as listed in Table 3. Then, the integrated efficiency of NPD project can be obtained using Eq. 6. As shown in Table 3, the efficiency rating of NPD projects A, F, and G is 1, so it is impossible to distinguish the priority level of the 3 NPD projects. Hence, priority ranking for these NPD projects is conducted using Eq. 7, and the results are G, A, and F. Thus, NPD project G has the greatest development value.

Table 3. Package-specific criterion evaluation rating

Serial number	NPD project	Expected revenue risk	Time to Market risk	NPD cost	expected revenue	efficiency rating	super-efficiency
A	ExpressCard 34 Daughter Card	0.184	0.414	43	112	1	1. 007
B	Express Card Reader to USB port	0.273	0.285	35	90	0.919	–
C	SmartCard Reader	0.367	0.142	37	84	0.805	–
D	CF I/O Card Reader	0.197	0.300	36	70	0.79	–
E	XD/SM Extender Board	0.242	0.230	29	71	0.91	–
F	XD Extender Board	0.234	0.205	34	100	1	1.004
G	SmartCard Extender Board	0.263	0.173	31	106	1	1.117
H	USB Test Fixture	0.241	0.258	29	70	0.846	–

Note: development cost, expected yield (unit: NT$ 10,000).

5 Conclusions

Discussions on operational performance have raised great concerns across different industries in recent years, and DEA is an efficient method for evaluating the performance. Due to the characteristics of NPD project evaluation problems, this research applied FHA to estimate the relevant significance of NPD project evaluation criteria, in order to obtain a reasonable weight range. The result was taken as the constraint conditions of the DEA-based project evaluation model to improve objective evaluation. NPD projects of development potential could be correctly identified by combining FHA and DEA. Given limited R&D resources, companies are unlikely to implement several NPD development projects simultaneously. Thus, NPD priority ranking can be applied to establish execution sequencing of NPD projects, providing a reference to the management of R&D departments for decision-making on the optimum product R&D strategy. This research also developed a performance evaluation model correspondingly to analyze the performance of various criteria. Finally, a case study of electronic product development projects was used to validate the practical value and contribution of this methodology.

6 References

[1] Alidi AS. Use of the analytic hierarchy process to measure the initial viability of industrial projects. International Journal of Project Management 1996;14:205-208.
[2] Liang MC. The analytic hierarchy process in project evaluation: An R&D case study in Taiwan. Benchmarking 2003;10:445-456.
[3] Ahn, BS.; Choi, S. H. ERP system selection using a simulation-based AHP approach: a case of Korean home shopping company. Journal of the operational research society 2008;59:322-330.
[4] Sun YH, Ma J, Fan ZP, Wang JA. Group Decision Support Approach to Evaluate Experts for R&D Project Selection. IEEE Transactions on engineering management 2008;55:158-170.
[5] Linton JD, Walsh ST, Morabito J. Analysis, Ranking and Selection of R&D Projects in a Portfolio. R&D Management 2002;32:139-148.
[6] Eilat H, Golany B, Shtub A. Constructing and evaluating balanced portfolios of R&D projects with interactions: A DEA based methodology. European Journal of Operational Research 2006;172:1018–1039.
[7] Anderson P, Peterson NCA. Procedure for Ranking Efficient Units in Data Envelopment Analysis. Management Sciences 1993;39:1261–1264.

An Exploration Study of Data-mining Driven PLM Implementation Approach

Chia-Ping Huang[1]

Department of Information Management, Hsing-Wu College, Taiwan

Abstract. This paper presents an exploration study of data-mining driven Product Life-cycle Management (PLM) implementation approach within a Taiwan steel making company. The research aims to validate a theoretical framework to structure published knowledge on how to effectively implement PLM system in a strategic way, and also on how project management approach and data-mining methods can be harmonized within the business decision making formulation process. Working with the case company, the author used a novel manufacturing strategy formulation framework incorporating both the context and content approaches. The primary research results reveal that the issues and factors could be systematically considered as a strategic guide for successful PLM implementation. This is because data-mining driven PLM implementation project using traceable reasoning mechanism to guide the management in formulating its PLM implementation action plan. Furthermore, the knowledge and understanding already captured in existing management practice is thus made available to industries in a coherent way, so that they can evaluate and improve their business performances. If you have any questions regarding the paper, please contact the author by email: 096013@*mail.hwc.edu.tw*

Keywords. PLM、Data-mining、Project Management

1 Introduction

There is a large body of existing theories and practices for enterprise information systems implementation. They are seldom succeeding in practice because management cannot easily make use of the published work to address their business problems. On the one hand, the volume of literature suggests the practitioners to consider a large number of issues that may or may not be relevant in their context. On the other hand, published implementation process and tools prescribe the steps but does not provide context appropriate suggestions. The wealth of the research is being locked up by the difficulty of its application.

In particular, there is currently very little attention paid to the link between manufacturing strategy and related enterprise information system analysis, design and implementation. This paper presents a data-mining driven PLM implementation approach in the context of existing theories and practices within

[1] Corresponding Author Email : 096013@mail.hwc.edu.tw

national defense industry. This work focuses on the situation when a steel making business unit sets up and operates a PLM system. The main objective is to empirically validate a theoretical framework making strategically project management and key performance indicators (KPI) studies usable to enterprise information system deployment aspect. A practical strategic analysis tool is then presented to help the management to formulate strategic actions that are aligned with the KPI and corporate strategy, based on the insights gained from the PLM implementation process.

The two main research objectives are:

- to structure existing enterprise information system strategic issues in a form usable to managers, in the context of PLM project management , and
- to establish a data-mining performance measurement oriented deployment approach that the management can use to set up PLM improvement and operations actions that are aligned with the business objectives.

2 Strategic PLM implementation

Most of the business treats enterprise information system implementation as an I.T. project or business-wide program [1]. They use different implementation methods (step by step, big-bang, and roll-out) to fulfill business objects [CERP]. The further suggestion implies that detailed system analysis, planning process, and effective project control are critical success factors to implementation enterprise system [8,13].

Based on these existing findings, the authors try to use a novel strategic analysis framework to developed manufacturing enterprise strategic issues which are usable and traceable as guidelines for management to execute enterprise information system implementation work.

The purpose of the authors' strategy deployment framework is to try to provide a focused, simple, but powerful analytical mechanism to enable a logical process to identify: what are the key tasks, who are the process owners and which operating functions are responsible. The analytical steps in the framework include the recognition/statement of the business objectives and the critical success factors (CSFs). From these, the relevant detailed policy, competitiveness, resources, operation, tasks and performance measures are identified as a set of strategic guidelines. The improvement tasks and performance measures plan to deliver the stated objectives are then derived.

The analytical process is supported by an integrated database that would present the relevant issues to the management team to select or use as reference benchmark. The database is built up from an extensive research into manufacturing strategy literature.

In brief, these 7 viewpoints and 14 worksheets help management receive a set of strategic actions. Then, these issues could effectively merge into the accelerated SAP (ASAP) model which consists of project preparation, business blueprint, realization, final preparation, life and support or concisely throw the light of project management life cycle which includes initiation, planning, execution, monitoring and closure stages.

These strategic issues may clarify the information system implementation dilemma. The author would like to emphasis the relationship between enterprise-wide information system implementation and strategy deployment. In fact, the number of issues and the possible relationships illustrate the difficulties in the making decisions within company-wide project management process and information system implementation.

3 Existing theories and models differences

This section explains the differences among existing theories and models which mainly involve manufacturing system strategy, performance measurement, project management and data-ming and may prove to be appropriate in particular situations.

Firstly, strategic enterprise information deployment requires a strategic combination of the different viewpoints to achieve integrated operations throughout the manufacturing enterprise. These different viewpoints need to be structured into an integrated strategy formulation process. The success of the integration is in the alignment of the individual project's performance with the company's business objectives. This is still a challenge for industrial practitioners and academics.

In fact, it is still difficult to represent the strategic decision in both cognitive and practicable terms for industries to follow regarding PLM implementation. This is because manufacturing enterprise strategies are exercises in the unknown and further complicated by the presence of a large number of other driving forces. In this paper, the authors regard manufacturing strategy studies as a deductive and top-down approach, or a prescriptive know-how process without the reasoning, or a detailed content and context at a strategic level. From the literatures, many researchers have attempted to outline the specific pattern, model, actions, and performance measures that comprise the manufacturing strategy and performance measurement [12,7,2]. This school of work proposes manufacturing configurations that has been very useful to our understanding of manufacturing strategy and operations. However, these configurations no longer provide the right way to view a modern enterprise. For example, traditional tools may provide warnings to management about serving I.T. policy, but they do not necessarily communicate the reasons for the problems that are useful to the high level management. Recent research into manufacturing systems integration has identified the need for effective deployment of business objectives through the organisation and the subsequent measurement of performance in critical areas.

The correlations of the company functional performance measures to the business strategic plans and critical success factors are also important. This is the means to ensure that functional responses are aligned with business objectives. As a result of E-business, current theories and practices such as increasing use of outsourcing, lean manufacturing principles, relaxing centralised control and more functional autonomy all have a role in integrating the overall manufacturing enterprise.

Usually, a performance measurement model is goal-directed and uses inductive inference to align measures to strategic goals.

Performance is not an easy subject. There is no universal agree model for industry to follow. For instance, SMART – Strategic Measurement Analysis and Reporting Technique [3]; Performance Measurement Questionnaire [4]; Cambridge Performance Measurement Design Process [5]; Balanced Scorecards [9], these models should be pointed out that other performance measurement systems have also been suggested, and these approaches have addressed many of the shortcomings of past performance measurement systems.

There are two key reasons to select in performance measurement systems:

1. Strategic Alignment: performance measurement programs can be used clearly to communicate and reinforce messages about what is important in the organisation, and to engage people in aligning to objectives.

2. Strategic Learning: performance measurement programs can promote learning as managers sift through performance data and discover pockets of excellence or weakness where more information about what works and what doesn't can be found and reapplied. Learning also occurs as performance data is analysed to discover trends and relationships between measures.

Project management life cycle which includes initiation, planning, execution, monitoring and closure stages. The relationships between 5 stages life cycle which includes initiation, planning, execution, monitoring and closure, and the 9 knowledge fields such as integration and scope, etc. and the designated actions such as scope of work (SOW) and a total of 44 actions are assigned in different stages which shown in table 1. [5]

Table 1. The 5 stages and 9 knowledge field relations table. I:initiation, P:planning, E:execution, M:monitoring and C:closure

	I	P	E	M	C	Actions
Integration	2	1	1	2	1	7
Scope		3		2		5
Time		5		1		6
Cost		2		1		3
Quality		1	1	1		3
HR		1	2	1		4
Communication		1	1	2		4
Risk		5		1		6
Procurement		2	2	1	1	6
Total	2	21	7	12	2	44

A brief comparison of the manufacturing strategy, performance measures project management and data-mining theory is presented in Table 2.[15,16,17]

Table 2. The comparison of manufacturing strategy, project management and performance measures theory.

	MFG strategy	Performance measures	Project Management	Data-mining
Management role	Strategic objectives	strategic alignment	Business objective	KPI
Implementation reasoning	Deductive	Inductive	Step by step	Star/ Snowflake /Constellation
Content logic	Causes and effects	Goal directed	Goal directed	Rational
Context comprehensive	Function based	Model based	Knowledge based	Function based
Process adaptability	Know-why	Know-how	Know-how	Know-how

4 Research Methods

The proposed framework used a spiral approach to validate and refine the content and context of itself that structures existing research results into an applicable format [5]. The research method is illustrated in figure 1.

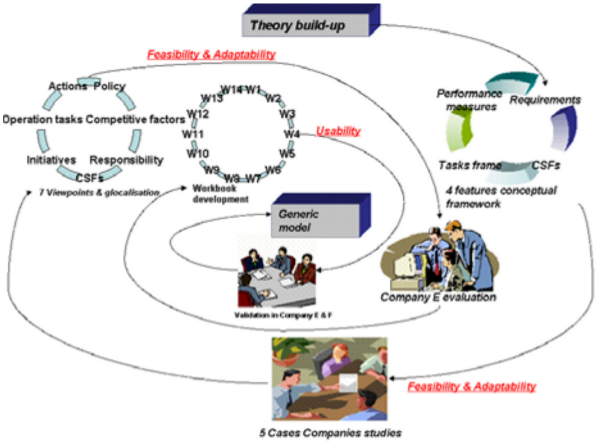


Figure 1. The research method of strategic PDM implementation.

The framework has structured manufacturing strategy deployment methodology which includes the requirements identified and develops the detail framework and its application process. The detail framework uses seven viewpoints to structure knowledge gathered from over 150 publications. The resultant framework consists of a set of relationship tables that correlate 165 foci and 620 key issues. The large volume of information captured within the framework has been formulated into a structured and efficient an embryonic workbook for the company to follow.

Based on precious literature reviews, the author uses data-mining and Online Analytical Process (OLAP) approach transferring 165 foci and 620 key issues into a fact table and 4 dimension tables. The rational connections are illustrated as follows.

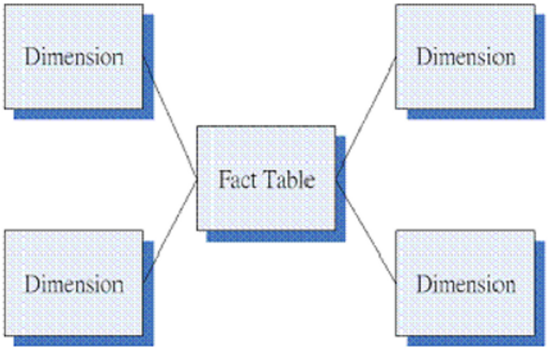


Figure 2. Star schema for fact table and dimention tables.

Based on the proposed approach, the field studies focus is determined and defined as PLM implementation relevant issues. A semi-structure management interviewing is adopted to gather and collect data. Afterwards, the data and information have been analyzed through the proposed framework which produced a set of strategic worksheet which guides management to execute PLM project effectively. The effects of this data-mining driven PLM implementation are summarized in a qualitative way in the following section.

5 The framework reasoning structure and primary findings

This section explains the framework scope, reasoning structure and the primary finding. A substantial body of work has been distilled into seven issues tables and nine relationship tables and the operation is illustrated as follows.

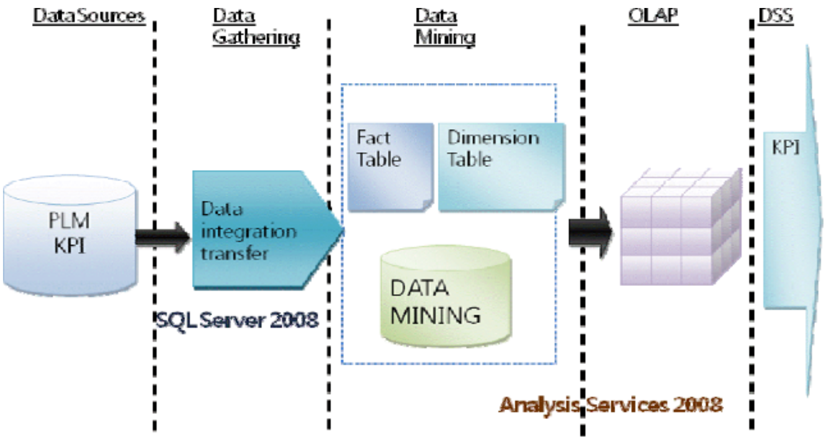


Figure 3. Data mining and olap operations.

It has been structured for the users to select relevant issues and add their expertise as well. An example of policy viewpoints is illustrated in table 3.

Table 3. An example of policy viewpoints

Policy Area	Focus	Key issues
Capacity	Capacity planning	Demand location
		Plant space
		Equipment
		Labour
	Expansion method	New facility
		Sub-contracting agreement
		Outsourcing
		Additional shift
	Variation satisfaction	Cyclical stock level satisfaction
		Increment costs
		Long term contract
		Reduce inventory level
Vertical	Supplier chain ownship	Supplier network
Integration		Customer network
		Distributor relationships
	Expansion	Acquirement
		Make vs buy
		Joint venture

The proposed approach aims to guide management for strategy formulation. It takes the management through an analysis of its business strategy and systematically translates that into a set of improvement actions and associated performance measures. The case company can use the knowledge in these databases to support the strategic PLM implementation process. At each stage of the process, the management is presented with a range of options from the databases. These options help the management to rationalise their choices. The options could be selected if appropriate. For choice not in the option list, the framework has an open structure to capture the decisions of the management team as additional options.

Table 4 illustrated the main activities within the proposed framework in total of 20 working items. Table 5 illustrates an example of actions selection from case company which is derived from company high level management's opinion.

Table 4. The main activities within the proposed framework.

	I	II	III	IV	Activity
Policy	1	1			2
Competitive factors	1			1	2
CSFs	1	2			3
Responsibility	1	2			3
Initiatives	1	2			3
Operations tasks	1	1	1		3
Actions	1		1	2	4
Total	7	8	2	3	20

Table 5. An example of actions selection from case study

Actions	Gaps	Policy areas	Competitiveness						
			Q	DL	DP	DF	VF	C	I
Decision layers	Resist changes	V.L.		X	X			X	X
B.P. Transform	Resist changes	Org					X		
		H.R.					X		
Upgrade existing facility	H.Q. control	Production control		X				X	
		Vendor relation	X					X	
Improve communications	H.Q. control	Organisation						X	
Office automation	Cost control	Facility		X				X	
Inert-functional work team	H.Q. agree	Human resources	X					X	
Worker training	Time limits	Human resources	X					X	
Financial control	Cost limits	Financial	X					X	

These primary strategic actions become the main input source for the further project development and management. The detail imput and output of the analysis process are described in figure 4.

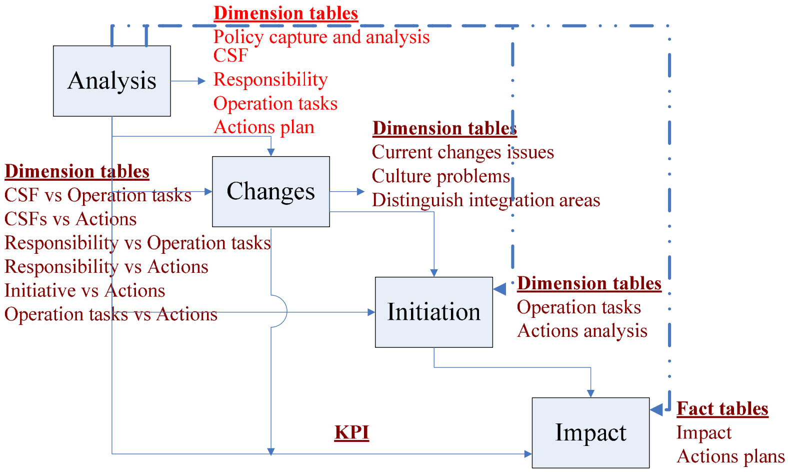


Figure 4. The detail input and output of the analysis process

After these strategic analysis and actions capture, these input issues and considers help management get clear vision and execute project management effectively. Since the project management development stages, initiation and planning needs clarifying the project nature and scope from customers or high level management, the research finds these strategic analysis and action plans are usable and maintainable. The mainly findings are indicating as follows.

- Policy: Capacity, facility, vertical integration, product scope, organization structure, and HR allocation are main considerations for management trade-off.
- Competitive: Cost, delivery lead-time, reliability, and design flexibility are main issues for this PLM system.
- CSFs: Agility, integration, and adaptability are main CSFs for this project.
- Responsibility: The MIS planning, and R&D take project responsibility.
- Initiatives: Introduce step improvement is the guideline for project management.
- Operations: R&D process redesign and functional activities integration are necessary within project execution and monitoring stages.
- Actions: The R&D process redesign and its impacts are main working items.

6 Conclusions

Strategy may be interpreted in different ways. In this paper, strategy is considered to have the purpose of directing operations to deliver business results. It is the statement of direction after a process of establishing an awareness of the business situation and constructing a purposeful response. The primary results reveal that an effective strategy should be feasible to realise and has ways to measure success. Furthermore, the related results show that the issues and factors could be systematically considered as a strategic guide for successful PLM implementation. This is because data-mining driven PLM implementation project using traceable reasoning mechanism to guide the management in formulating its implementation action plan which fact tables could be composed of several dimension tables.

As more steel making companies have seen competition and customer's demands increase in Taiwan, the case company has realised the importance of strategic driven I.T. system implementation in attaining sustainable competitive advantage. Therefore, to remain competitive they must select the right approach to attach themselves to prominent leading edges. From the viewpoint of subject boundaries, data-mining driven PLM implementation fits between and across strategic I.T. system implementation and management. The contributions to knowledge of this work are as follows:

- a novel content and context framework to relate issues and factors for strategic PLM deployment
- a data mining oriented PLM deployment framework that captures the implementation issues of change, culture and integration
- a OLAP process that uses the traceable reasoning of the framework and a set of fact, dimension tables to guide a steel making company in its formulation of strategic action plan
- an embryonic PLM deployment workbook prototype that was tested in a Taiwan steel making company.

7 References

[1] Al-Mashari, M.., Al-Mudimigh, A. and Zairi, M., "Enterprise resource planning: A taxonomy of critical factors", European Journal of Operational Research, vol. 146, pp. 352-364, 2003.

[2] Buffa, E. S., "Meeting the Competitive Challenge: Manufacturing Strategies", Dow Jones-Irwin, Homewood, 1984.

[3] Cross K. F. and Lynch R. L., "The SMART way to define and sustain success", National Productivity Review, 9(1), 23-33, 1988-1989.

[4] Dixon J., Nanni A. J., Vollmann T. E., The new performance challenge: Measuring operations for world class competition. Dow Jones-Irwin, Homewood, 1990.

[5] Fan, I-S, Huang, C., "Aligning Local Office Management Plan to Global Corporate Strategy", 7th International Manufacturing Symposium in Cambridge, 2002.

[6] Ferdows, K., Making the Most of Foreign Factories. Harvard Business Review, Mar-April. 1997.

[7] Hayes, Robert H., and Wheelwright, Steven C., Restoring Our Competitive Edge: Competing Through Manufacturing, John Wiley, New York, 1984.

[8] Holland, C. and Light, B., "A critical success factors model for ERP implementation", IEEE Software, May/June, pp. 30-36, 1999.

[9] Kaplan R. S. and Norton D. P., The balanced scorecard: Measures that drive performance, Harvard Business Review, Jan-Feb., 1992 and 1996

[10] Krasner, H., "Ensuring e-business sussess by learning from ERP failures", IT Pro, March/April, pp. 37-41, 2000.

[11] Project Management Institute(PMI), “A Guide to the Project Management Body of Knowledge”, (PMBOK Guide), Pennsylvania: PMI Inc, 2000.
[12] Skinner, W., Manufacturing- Missing link in corporate strategy, Harvard Business Review, 47, May-June, 136-145, 1978.
[13] 蔡文賢, 簡世文, 許秉瑜, 范毅文, 鄭明松, 呂俊德, “台灣地區企業導入 ERP 系統的管理議題之實證調查”, 中華企業資源學會 2004 年 ERP 學術與實務研討會, pp. 379-389, 2004.
[14] 施清壬，以線上決策分析程序(OLAP)模式探討製造業 ERP 的資料倉儲應用，碩士論文，長庚大學，2004。
[15] 董和昇，運用 OLAP 技術建構商業智慧，商業現代化，49 期，48-53，2001。
[16] 蔡淑如，以 OPAL 模式建構製造業之採購決策支援系統，碩士論文，立德管理學院，2006。
[17] 戴建耘、袁熒助、蔡志宏，OLAP 線上分析處理在資料倉儲的應用-以技職課程資源網站爲例(上)，生活科技教育，35 期，32-38 頁，2002。

Exploring the Links between Competitive Advantage and Enterprise Resource Planning (ERP) Upgrade Decision: A Case Study Approach

Celeste See-Pui Ng [a] and Pei-Chann Chang[1]

[a]Department of Information Management, Yuan Ze University

Abstract. The complexities in making an ERP upgrade decision have been widely cited in the trade press. To the client-organizations, the issues of maintenance and upgrade are continuously existing and requiring extensive attentions and efforts as far as continuous business improvement and benefit-realization are concerned. However, there is very few studies dedicated to provide better understanding of how a firm considers an ERP upgrade decision. Competitive advantage has often been cited and associated with upgrade decision. To what degree and how an ERP upgrade is perceived as providing competitive advantage to a firm? In order to gain in-depth understanding of whether competitive advantage affects upgrade decision and how, an evaluative case study is conducted in a firm that is making an ERP upgrade decision. We find that out of the four characteristics of competitive advantage proposed in the resource view theory, value is perceived as the foremost important, and other characteristics as unimportant. With this, we suggest that the type of competitive advantage perceived and expected from an ERP upgrade is non-sustainable (in the long-run) cost leadership-oriented advantage. However, this finding is likely constrained by how the ERP is used and the scope of its implementation by this case.

Keywords. enterprise resource planning, resource-based theory, information systems resources, IT resources, software upgrade, upgrade decision, competitive advantage

1 Introduction

Enterprise resource planning (ERP) system has gradually become a must-have information system for most large and medium size organizations in or der to more efficienly manage, organize, operate and control basic business functional areas. It is also a neccesity to compete and collaborate with employees, business partners and all players along the supply chains.

The AMR market research [15] estimates an 11% CAGR for the ERP market through 2011. In light of the steady market maturity and economy slowdown, ERP revenues are no longer mainly driven by the software sales or license but ERP related services such as consulting, training, maintenance and upgrade. For example, ERP maintenance and upgrade represented a large percentage – 70% of

[1] Corresponding Author Email : iepchang@saturn.yzu.edu.tw.

the total ERP revenues in 2003 [9]; and SAP, with about 38,000 customers in 2006, expects more than two-thirds of the customer base will be on Enterprise SOA by 2010 (excluding SAP BusinessOne users) [36].

ERP maintenance and upgrade is an important issue to the client-organizations and requires extensive attentions and efforts as far as continuous business improvement and benefit-realization are concerned. While are quite a few drivers for upgrade such as vendor support termination, platform and technology improvement, a large portion of them are related to voluntary business improvements aiming to enhance ones competitive power [40]. Competitive advantage has often been citied and associated with upgrade decision and driver [10, 12]. However, how competitive advantage is linked to or associated with ERP upgrade decision remains unknown. To date we observe very few researches focusing on how a firm considers an ERP upgrade decision. In order to explore this area of research, we adopted the resource-based view (RBV) theory. Wade et al. [42] argue that RBV provides “a cogent framework to evaluate the strategic value of information systems resources” (p. 109).

‘An upgrade decision’ is defined as a decision made which results in the installed old ERP version being replaced by a newer and superior version either for the same or different vendor’s product. We assumes that upgrading an existing ERP system to a more advanced and superior version will generate a set of new resources or capabilities and when assembled, organized deployed and use in firm-specific organizational and business processes will create (competitive) advantages to the ERP client-organization. This assumption is supported in the writings by other researchers [6, 34]. Our research question is: to what degree and how an ERP upgrade is perceived as providing competitive advantage to a firm? In order to gain in-depth understanding of whether competitive advantage affects upgrade decision and how, an evaluative case study is conducted in a firm that is making an ERP upgrade decision.

The paper begins with a discussion of ERP upgrade in section 2. Section 3 provides the theoretical background for the study. This is followed by a discussion of the research method and results in section 4. Section 5 gives the implications of this study and future studies.

2 ERP upgrade project

ERP upgrades are still perceived as expensive, risky and multiyear projects by many CIOs [37]. According to the same study, although large SAP customers acknowledge that upgrade is inevitable and recognize that the technology may bring substantial competitive advantage to them, they are daunting to be the first. As a result, client-organizations postpone ERP upgrades [26]; and tend to gather a lot more information on the values, costs, and skills needed to take advantage of the new software [37]. This situation is also the same as to the Oracle's large customers [36].

On the other hand, forgoing ERP upgrade may result in paying exorbitant maintenance fees after the official support period [26]. New ERP releases by and large have a number of applications enhancements, bug fixes, statutory updates, platform changes and technical enhancements [39, 41]. (Technical enhancements include system and data security, processing speed, system performance, compatibility with other software and hardware, and leveraging new database and operating system features which are mostly needed to enable installation of new modules [41].) In other words, giving up upgrade means client-organizations have to face different type of risks such as system development costs, obsolete technology, recruiting related experts for handling the system, system maintainability and system connectivity with other business partners, new software and hardware. In overall, upgrade decision has to be justified by its value, usefulness, and contribution to the business of the client-organizations [41].

In general, the advantages bring about by ERP upgrade motivations (distilled from the existing literature) can be classified into basically four main categories as shown in Table 1. All together, these advantages can facilitate four main competitive strategies, i.e. cost leadership, growth, innovation, and alliance [28] as illustrated in Table 1.

Table 1. Major advantages of ERP upgrade

Advantage of ERP upgrade	**Reference**	**Strategy supported**
Keeping the system operational at certain level of performance, cost-effectively maintainable, avoiding the system from vendor support termination. Driven by: compliance to the vendor's standard code or keeping the system up-to-date, expired maintenance support, installation of new technical features, synchronizing existing systems, third-party application/ hardware no longer supporting existing version, removal of high maintenance costs of previous enhancements or workarounds, retirement of legacy system and bolt-ons, and expensive maintenance costs.	[19], [30]	-
Realizing business benefits from the system. Driven by: TCO savings, increased efficiencies, obtaining competitive advantages, operational and strategic business benefits of new functionality	[16], [30], [6], [1]	Cost leadership, innovation
Adapting the system to comply with new business requirements, business process, and government's regulations. Driven by: Compliance to new business environment, and stayed competitive.	[40], [41]	Growth
Meet business partners' needs. Driven by: Pressure from the value chain.	[16], [6]	Growth, alliance

3 Theoretical Background

According to O'Brien and Marakas [28], competitive advantage is defined as "developing products, services, processes or capabilities that give a company a superior business position relative to its competitors and other competitive forces" (pg. 591). ERP system is an IS resource. IS resources can be broadly divided into IS assets (technology-based) and IS capabilities (system-based) [42]. While IS assets are easy to copy and most fragile source of sustainable competitive advantage [22], competitive advantage bring about by a firm's superior deployment of IS capabilities is not [11]. ERP system upgrade will generate an additional set of new resources or capabilities and when assembled, organized deployed and use in firm-specific organizational and business processes will create (competitive) advantages to the ERP client-organization [6, 34].

Distilling from existing literature, Wade and Hulland [42] classify IS resources into eight major resources: external relationship management, market responsiveness, IS-business partnerships, IS planning and change management, IS infrastructure, IS technical skills, IS development, and cost effective IS operations. The resources or capabilities that an ERP system or a new ERP version can provide are: IS infrastructure, cost-effective IS operations, market responsiveness and external relationship management [27]. According to Wade et al. [42], each of these four IS resources (IS infrastructure, cost effective in IS operations, external relationship management, market responsiveness) has the four resource-based view (RBV) attributes, i.e. value, rarity, inimitability, and non-substitutability.

a. Resource-Based View

Resource-based view is a theory explains how firm competes based on firm's resources [31]. According to Wernerfelt [43] and Barney [3], firm resources can be important factors of sustainable competitive advantage and superior firm performance. Resources are defined to be "all assets, capabilities, organization processes, firm attributes, information, knowledge, etc. controlled by a firm that enable the firm to conceive of and implement strategies that improve its efficiency and effectiveness" ([3], pg. 10). Firm can use these resources for creating, producing, and/or offering its products (goods or services) to a market [38].

However, these resources themselves have to possess special characteristics that make firm competitively advantage by having the resources. These characteristics are valuable, rare, inimitable and non-substitutable [3]. The definitions for these characteristics distilled from major literature are summarized in Table 2. With these resource characteristics, firm is capable of creating and sustaining competitive advantage that affords the accrual of superior long-term performance [3, 31, 42].

Table 2. Definitions of RBV characteristics

RBV characteristic	Definition
Value	A resource or capability enables a firm to implement strategies that improve efficiency and effectiveness [3]. A resource or capability that has little value has a limited possibility of contributing a sustained competitive advantage on the possessing firm [42].
Rarity	A resource or capability is scarce and not simultaneously available to a large number of firms [2]. A resource or capability that is not rare is not likely to create a strategic benefit.
Inimitability	A resource or capability is not quickly duplicable. Three factors

	contributing to inimitability, i.e. *unique firm history*, *causal ambiguity*, and *social complexity* [3, 23]. A resource or capability is related to *unique firm history* when it can only be developed over long periods of time [25]. *Causal ambiguity* happens when the link between a resource or capability and the particular competitive advantage is unclear or poorly understood. As a consequence, it is extremely hard for other firm to duplicate the resource or capability [35]. This can include firm's culture [4] and/or tacit attributes [35]. On the other hand, *social complexity* describes the multifarious relationships within the firm and between the firm and key stakeholders [14]. This can include a firm's organization's culture [4], reputation [18], and trustworthiness [5].
Non-substitutability	A resource or capability is both rare and inimitable [7].

Alternatively, a firm can also uniquely assemble and deploy the resources to create unique organizational capabilities, e.g. better business efficiency and effectiveness, that once embedded in unique organizational or business process generate firm's competitive advantage [24]. This is crucial if these resources by themselves are not rare, imitable and non-substitutable. However, via firm-specific assembly and deployment of the resources and their utilization in specific firm or business processes, they will eventually create organizational capabilities and/or organization and business processes that are valuable, rare, inimitable and non-substitutable.

An upgraded ERP system can provide a better IT backbone or IS infrastructure for integrating internal cross-functional area business processes, which in turn allows cost effective IS or business operations. Together with other bolt-ons such as supply chain management (SCM), customer relationship management (CRM) and data warehousing, the upgraded system permits a firm to manage external relationship with suppliers and customers, and ensure market responsiveness in a changed or changing environment. Also, by incorporating idiosyncrasies business processes, standard code modifications (to meet specific system requirements), unique organization culture, and customized way of doing business with suppliers and customers will make the new upgrade version inimitable by other competitors. This is because each firm's suppliers and customers are unique in some ways, and the type of information needed for market responsiveness is not necessarily the same. Moreover, both of these ERP capabilities will evolve uniquely for each firm [42].

4 Research methodology

We adopt the case study approach as the topic in ERP upgrade is still understudied and there is no existing literature examining the RBV theory in the context of ERP upgrade. This approach will allow us to collect natural-setting and in-depth data. The scholar in (Eisenhardt, 1989) also suggested that the result of a case study is highly empirically valid and testable as the data are all evidence supported.

A series of interviews with six different interviewees is conducted and is targeted at the upper level management who directly involves in the ERP upgrade decision. In each interview, notes are taken during and immediately after the interviews. The interview transcripts are then sent back to the interviewees for confirmations and verifications to avoid misunderstanding and misinterpretations. Each interviewee is asked the same interview questions. Multiple interviews are meant to serve as multi-source of evidence for the interview findings. Prior to the actual interview, the definitions for each RBV attribute as given in Table 2 and Table 3 are provided to the interviewees for preview.

Table 3. RBV attribute's measurement

Attribute	**Measures**	**Examples**	**Source**
Value	Reduce cost	Save cost by reducing the workforce Reduce cost in information distribution Reduce cost marketing Reduce R&D cost	[8], [21], [29]
	Increase revenue	Increase return on financial assets Better organizational goals achievement Increase market share Increase annual sales revenue	[21], [32]
Rarity	Become first movers (become innovative)	Being the first to provide new products/services to customers Being the first to provide better products/services to customers Being the first to provide easier customer access to information	[8], and [21], [29]
	Able to differentiate from other competitors	Change the traditional way of conducting business Allow customer to customize products/services	
Inimitability	Establish linkages with business partners (causal ambiguity)	Keep close contact with business partners Help establish useful linkages with other organizations	[8], [21], [29]
		Enable easier access to information by business partners Increase efficiency in supply chain	
	Enhance business reputation with partners (social complexity)	Enhance brand distinguish ability Enhance the credibility and prestige of the organization Enable faster retrieval delivery of information or reports by business partners	
Non-substitutability	Has uniqueness	The ways we use the data from the system are different from others. Have a lot of idiosyncrasies business processes, and modifications done.	[42]

a. The case – Apacer company background

Apacer was founded in April 1997. The capital is over ten hundred million US dollars with approximately 500 staff members. The business volume is NTD120 millions in year 2003 and reached NTD140 million in 2004. The head-quarter is situated at the Nankang Software Park, Taipei. The firm currently has offices in USA, Netherlands, France, Tokyo, Middle East, India, Sydney, Hong Kong, Thailand, Singapore and Malaysia. Apacer is a manufacturing company that develops dynamic RAM (DRAM). Apacer offers various types of sale services to its clients based on the size of the order and the size of the client's company.

Apacer started to use JDE OneWorld 7.332 ERP system in 2001. The ERP has been used mainly in financial and stock controlling. However, many problems began to appear after the fifth year and the company experienced a hard time when their financial returns did not overcome their expenses, especially in the stock controlling for over NTD30 million. At the same time, the firm's market position did not expand. The upper management perceived that the existing ERP system was not performing well enough and did not provide specific and accurate reports during the crisis. This was when the ERP upgrade decision was made and business process reengineering was required. The firm decided in year 2006 to upgrade their ERP system to JDE EnterpriseOne 8.11 (JDE is now bought by Oracle). The upgrade was meant to mostly focus on bug fixing, financial reports, re-engineering processes, cost controls, and integrations. During our data collection, the firm was making their upgrade decision.

b. Data analysis

The job title of the six interviewees involved in making the ERP upgrade decision are given in Table 4. Their ages range from 35 to 48 years old; and they have between 1.5 to 10 years of working experience at the Apacer company. The ERP modules used in the firm are (in precedence order): stock management, ordering, customer services, production line scheduling, purchasing, quality management, invoicing, product delivery, product design and development, and human resource management.

The interview survey with six top management shows that (among the four attributes of value, rarity, inimitable and non-substitutable), *value* is the most important RBV attribute, which all of them unanimously agree with affecting their firm's ERP upgrade decision. The IT manager says that the firm is most concerned about its overall revenue income; this is the key reason for upgrading the system. On top of this, the finance manager states that "*Value* attribute aims to achieve the same goals that are set by our firm, i.e. constantly reducing costs and making a higher profit margin." According to the production and material control manager, "Getting a valuable resource means a better performance firm which everyone strives to become." Table 4 shows the viewpoints of the six interviews on the reason why *value* is the most important factor driving their ERP upgrade decision.

Table 4. The importance of the RBV value-attribute

Question	
Among the <u>four</u> attributes of **value**, **rare**, **inimitable** and **non-substitutable**, which are the most important factor(s) in evaluating your company's ERP upgrade decision? Why?	
IT manager	Value Financial income is the utmost important element and consideration for our firm. The capabilities of the information system to optimize the operation processes, obtain real-time information, and provide better integration to all of our branches are the valuable elements that may affect ERP upgrade. Most importantly, if it allows us to gain profit.
Financial manager	Value ERP system enhances the accounting process such as reducing the period of credit accounts. Value is created through reducing time and costs in our business transactions and business processes. It makes the completion of financial report easier and faster. If our company needs to make decision on upgrading ERP, value attributes will definitely be the main driving force.
Manufacture manager	Value ERP helps to transfer manufacturing information effectively so it minimizes the mistakes and delays that could be caused by operators and machineries. It calculates costs accurately and allows transparency in WIP (work in process). All these improvements I mentioned are ways to reduce costs.
Production and Material control manager	Value It makes producing the management reports and WIP control easier. ERP creates value by providing better business processes which is the most important contribution.
Director	Value The main reason ERP upgrade perceived as valuable is that it can provide a better decision making process by controlling the stock flow and updating reports regularly; and can facilitate the firm to earn profit and enhance our business processes.
Sales manager	Value A valuable resource, for example the ERP upgrade improves internal efficiency, the rate of getting sales contract, lowering stock taking, financial accuracy and information transparency. These are the reasons for ERP upgrade, which is to obtain valuable results such as lowering costs in the firm's business process.

The top management does not perceive *rarity* attribute as important as *value* in affecting the upgrade decision. The reason we find out from the interviews is as follows. "Simply talking about the ERP system, the rarity of the system is low and we are not so sure about competitor's system ability. … We do not put much effort on creating rarity but to increase its correctness and functionalities," says the IT manager. Both director and the finance manager add "…it does not really matter if ERP is rare or not. It is much more important to reduce the costs and increase profit margin by eliminating mistakes and delays." The manufacturing manager further comments, "What we concern with is ways to produce goods with the lowest costs. Therefore, attribute *rare* becomes less important than maximizing value." Sales manager adds, "A rare system may cause more efforts in employee training and system maintenance. Thus, it may waste more resources." The production and material control manager emphasizes, "All we need is a system that can record reports and provide necessary tools. It doesn't really matter if it is rare or not."

In general, the attribute *inimitability* is also perceived as not as important as attribute *value* in affecting the upgrade decision, according to manufacture manager, production and material control manager, and the director. In the interviews, we find out from the IT manager that "The main concern in upgrading the system for our firm is to provide better stock management, and to provide the required and necessary management reports. Although some customizations are done to our ERP system due to our company's unique business culture, which may create the inimitability of the system for other competitors, as a whole, it is not as important as *value* attribute." Sales manager states, "In implementing or upgrading our ERP system, we wish it to be more general and has less customizations. Sometime, we even adapt our operation to match the system's working pattern, for the system to perform better. If our system is built in a way that is difficult to imitate, we may also face more problems on future upgrades. Therefore, to reduce these problems, we try to maintain a more general system that is easy to upgrade, to reduce maintenance costs."

As for the attribute *non-substitutable*, according to the finance manager, "*Non-substitutable* is important but not as vital as the *value*. We can calculate precisely the advantages that *value* creates in reality, but *non-substitutable* attribute is difficult to measure." The manufacturing manager comments, "I can not say that specialized functions are not important in our operating process but they are less important than making money. In the context we work in, ERP functions are difficult to be non-substitutable; it's easily replaceable by other software." "There are some parts of ERP that can not be substituted, for example, some components of WIP. However, comparing to *value*, it still has stronger impacts," explains the production and material control manager. Sales manager mentions, "The system itself needs not to be *non-substitutable*. From the cost point of view, an easily substitutable system is easier to lower the maintenance costs." The director adds that *non-substitutable* is not an important factor affecting their ERP upgrade decision.

5 Conclusion

A few market researches claim that advanced technologies (such as ERP system) allow firm to gain competitive business advantage by preparing the firm for future challenges [13], and for some, competitive advantage is one of the top three business drivers for companies adopting ERP [6]. In light of the IS/IT capabilities provided by an ERP system, in this study, we attempt to explore how competitive advantage is related to ERP upgrade decision. In this effort, we have adopted the resource-based view to measure the competitive advantage of an ERP system.

Based on the data analysis of this case organization, we observed that cost reduction and revenue increment are perceived as the fundamental, foremost attractive driver in ERP upgrade decision. The top management considers the *financial* performance of the firm and *measurable* benefits the most. For example, provide better business processes and data integration; shorten time in producing financial / accounting reporting; provide high quality information; minimize errors and delays; and is able to produce better reporting and ease of process control. The role of *value* in competitive business environment is recognized by the case firm. Thus, in order to promote future upgrades, the vendor may consider incorporating and emphasizing more on the *value* delivered to clients in their future upgrade versions.

On the contrast, this firm does not perceive that completely *rare*, *inimitable*, and *non-substitutable* as important as *value* in their ERP upgrade decision-making. The firm does not see that an upgraded ERP system can become a resource or capability which is completely *rare*, *inimitable*, and *non-substitutable*. The top management thinks that the rarity, inimitability, and non-substitutability are hard to achieve in an ERP system. It is obvious that the firm is expecting to realize the operational benefits from the ERP investment. Operational benefits are also recognized as wave one and wave two of IS/IT innovation [33]. From the Michael Portal's classical model of competition perspective, the type of competitive advantage perceived and expected from an ERP upgrade in this case is non-sustainable (in the long-run) cost leadership-oriented advantage.

We believe most firms can possibly easily obtain operational benefits from an ERP system. However, this is likely to provide short-term competitive advantage as observed in this case study. In order to strike for more long-term competitive advantage, this case is likely to be constrained by how the ERP is used and the scope of its ERP implementation (c.f. [17]), and the alignment between the ERP implementation objectives and firm's business objectives (c.f. [20]). While operational benefits are internally focused, ERP client-organizations need to take advantage of the next waves of ERP innovations by using the ERP resources in synergizing with other resources to enhance or create new product and services, enhance executive decision-making and/or reach the customers and suppliers. This can result in differentiation and innovation in their products and services, business growth and establishment of good alliance with customers and suppliers.

Underutilizing the ERP system will undoubtedly result in limited business values. Firm can consider improving the rarity of their ERP system (wherever necessary) by creating differentiation in using the ERP system, not only they can do the daily work accurately, but to enhance the system so that it out-performs other competitors such as extracting business intelligence from the data provided by the ERP system to better manage existing business process and relationship with supplier and/or customer. Spending more effort to explore and use the state-of-the-art ERP system, together with existing knowledge, relationship with the stakeholder and infrastructure resources that a firm has, to perform daily business operation in a way which has not been thought before. As suggested in [17], firms' ability to integrate and combine existing resources in order to build ERP capabilities can strengthen or weaken their business process outcomes.

This study, like all empirical studies, has its limitation. We are using a single case study method to explore a previously unknown and inaccessible data on ERP upgrade decision. Different firm under different firm background, culture and organization size can have different outcomes. Therefore, more firms should be involved in order to strengthen the research outcome or to do case comparison.

We are using vertical cutting data set in which we only have a single point of research, however, ERP upgrade involves in different stages (for example: prior the upgrade, during the upgrade and after the upgrade). For future study, researches on the topic can be conducted in a time line and with followed up studies. The time series of data of the phenomenon can be collected to provide more analysis of how the initial competitive advantages provided by ERP upgrade evolve over time. In addition, it is recommended that the future research can involve other related theories to understand more key factors influencing ERP upgrade decision. One of the possibilities is the institutional theory as upgrade decision may also be affected by senior management supports, user involvement, and organizational fit of the ERP system (c.f. [20]). More empirical studies are warranted to identify what existing theories can better explain and provide more sufficient understanding in ERP upgrade decision.

6 References

[1] Allesch A. SAP NetWeaver Fundamentals, in *Managing SAP Projects 2004*. 2004; Wellesley Information Services: Las Vegas.

[2] Amit R, Schoemaker PJH. Strategic Assets and Organizational Rent. Strategic Management Journal 1993; 14: 33-46.

[3] Barney J. Firm Resources and Sustained Competitive Advantage. Journal of Management 1991; 17: 99-120.

[4] Barney JB. Organizational Culture: Can It Be a Source of Sustained Competitive Advantage. Academy of Management Review 1986; 11: 656-665.

[5] Barney JB. Competitive Advantage from Organizational Analysis. Texas A&M University, 1994.

[6] Beaumont J. Upgrading Your ERP System: Cognizant's Top 10 List. Cognizant Technology Solutions, 2004.

[7] Black JA, Boal KB. Strategic Resources: Traits, Configurations and Paths to Sustainable Competitive Advantage. Strategic Management Journal 1994; 15: 131-148.

[8] Bloch M, Pigneur Y, Segev A. On the Road of Electronic Commerce - A Business Value Framework, Gaining Competitive Advantage and Some Research Issues. University of California, 1996.

[9] Clouther S. Industrial ERP Market Turns Positive after Years of Declining Revenues. ARC Advisory Group, 2004. Available at: <http://www.arcweb.com/Newsmag/ent/erp081204.asp>. Accessed on: September 27 2004.

[10] Davenport TH. A "Bifocal" Approach to Enterprise Solutions. Accenture Institute for Strategic Change, 2002.

[11] Day G. The Capabilities of Market-Driven Organizations. Journal of Marketing 1994; 58: 37-52.

[12] Dunn A. Competitive Edge, in *SAP INFO*. 2003; SAP AG. 74-75.

[13] Franke J. MySAP ERP Technical Upgrade ROI: Look to the Future. SearchSap, 2007. Available at: Accessed on: 29 September 2007.

[14] Hambrick D. Top Management Teams: Key to Strategic Success. California Management Review 1987; 30: 88-108.

[15] Jacobson S, Shepherd J, D'Aquila M, Carter K. The ERP Market Report, 2006-2011. AMR-R-20495, AMR Research, Inc., 2007.

[16] Jahnke A. Are ERP Upgrades Worth the Cost and Trouble? CIO.com, 2002. Available at: <http://comment.cio.com/soundoff/061302.html>. Accessed on: March 11 2004.

[17] Karimi J, Somers TM, Bhattacherjee A. The Role of Information Systems Resources in ERP Capability Building and Business Outcomes. Journal of Management Information Systems 2007; 24: 221-260.

[18] Klein B, Crawford RG, Alchian AA. Vertical Integration, Appropriate Rents, and the Competitive Contracting Process. Journal of Law & Economics 1978; 21.

[19] Kremers M, van Dissel H. ERP System Migrations. Communications of the ACM 2000; 43: 53-56.

[20] Law CCH, Ngai EWT. ERP System Adoption: An Exploratory Study of the Organizational Factors and Impacts of ERP Success. Information & Management 2007; 44: 418-432.

[21] Lederer AL, Mirchandani DA, Sims K. The Link Between Information Strategy and Electronic Commerce. Journal of Organizational Computing and Electronic Commerce 1997; 7: 17-34.

[22] Leonard-Barton D. Core Capabilities and Core Rigidities: A Paradox in Managing New Product Development. Strategic Management Journal 1992; 13: 111-129.

[23] Mahoney JT, Pandian JR. The Resource-Based View within the Conversation of Strategic Management. Strategic Management Journal 1992; 13: 363-380.

[24] Makadok R. Toward a Synthesis of the Resource-Based and Dynamic-Capability Views of Rent Creation. Strategic Management Journal 2001; 22: 387-401.

[25] Mata FJ, Fuerst WL, Barney JB. Information Technology and Sustained Competitive Advantage: A Resource-Based Analysis. MIS Quarterly 1995; 19: 487-505.

[26] McMahon S. Beating the Clock on ERP Upgrades. Jupitermedia Corporation, 2004. Available at: <http://itmanagement.earthweb.com/erp/print.php/3364581>. Accessed on: December 3 2004.

[27] Ng CSP. A Resource-Based Perspective on Enterprise Resource Planning (ERP) Capabilities and Upgrade Decision. In: Pacific Asia Conference on Information Systems (PACIS): Kuala Lumpur, Malaysia. CD-ROM.

[28] O'Brien JA, Marakas GM. Management Information Systems. Boston: McGraw-Hill, 2008; 628.

[29] Orli RJ, C. TJ. If It's Worth More Than It Costs, Buy It! Journal of Information Systems Management 1987; 4: 85-89.

[30] Ostrom T. Advice We Wish We'd Had Before We Upgraded: Time- and Money-Saving Lessons from Recent mySAP ERP and R/3 Enterprise Upgrades, in *Managing SAP Projects 2004*. 2004; Wellesley Information Services: Las Vegas.

[31] Peteraf MA. The Cornerstones of Competitive Advantage: A Resource-based View. Strategic Management Journal 1993; 14: 179-191.

[32] Premkumer G, King W. An Empirical Assessment of Information Systems Planning and the Role of Information Systems in Organizations. Journal of Management Information Systems 1992; 9: 99.

[33] Primozic K, Primozic E, Leben J. Strategic Choices: Supremacy, Survival, or Sayonara. New York: McGraw-Hill, 1991.

[34] Ragowsky A, Gefen D. What Makes the Competitive Contribution of ERP Strategic. The DATA BASE for Advances in Information Systems 2008; 39: 33-49.

[35] Reed R, DeFillippi R, J. Causal Ambiguity, Barriers to Imitation, and Sustainable Competitive Advantage. Academy of Management Review 1990; 15: 88-102.

[36] Richardson B. SAP: A Tale of Two SOAs. AMR Research, 2006. Available at: <http://www.amrresearch.com/>. Accessed on: December 11 2006.
[37] Richardson B. Dissecting SAP's Three Roads to Growth. AMR Research, 2007. Available at: <http://www.amrresearch.com/>. Accessed on: January 29 2007.
[38] Sanchez R, Heene A, Thomas H. Introduction: Towards the Theory and Practice of Competence-Based Competition. Oxford: Pergamon Press, 1996.
[39] Swanton B. Oracle Upgrade Projects: An Ongoing Quest for Functionality. AMR Research, 2004. Available at: <http://www.amrresearch.com/>. Accessed on: December 2 2004.
[40] Swanton B, Samaraweera D, Klein E. Minimizing ERP Upgrade Costs Requires Synchronizing with Business Improvement Projects. AMR Research, 2004.
[41] Thompson O. The Old ERP Dilemma - Should We Install the New Release? TechnologyEvaluation.Com, 2002. Available at: <http://www.technologyevaluation.com/Research/ResearchHighlights/Erp/2002/06/>. Accessed on: June 15 2002.
[42] Wade M, Hulland J. Review: The Resource-Based View and Information Systems Research: Review, Extension, and Suggestions for Future Research. MIS Quarterly 2004; 28: 107-142.
[43] Wernerfelt B. A Resource-based View of the Firm. Strategic Management Journal 1984; 5: 171-180.

Development of a Cost Estimating Framework for Nanotechnology Based Products

Yuchun Xu[a,1], Rajkumar Roy[a], Gianluca Cassaro[b] and Jeremy Ramsden[c]

[a] Decision Engineering Centre, School of Applied Science, Cranfiel University, Cranfield, Bedford, MK43 0AL, UK.
[b] Department of Mechanical Technology, Production and Management Engineering, University of the study of Palermo, Viale delle Scienze, Building 8, Palermo (PA), 90128, Italy.
[c] Microsystems and Nanotechnology Centre, School of Applied Science, Cranfiel University, Cranfield, Bedford, MK43 0AL, UK.

Abstract: Nanotechnology materials and products are more and more used in different industry sectors. Due to the novelty and complexity, cost of nanotechnology materials and products are not well understood, which in many cases implies risk in business. This paper develops a framework to estimate the cost of nanotechnology materials and products, it aims to help manufactures to better understand the cost of nanotechnology materials or products at design stage. Expert interviews were conducted to validate the cost estimation logic. It is concluded that the approach and framework developed is able to estmate the cost of nanotechnology based products reasonably but reliable industry standard data will be needed to achieve that.

Keywords. Nanotechnology, Cost Estimation

1 Introduction

Nanotechnology is defined as both a technology for fabricating ultrasmall materials and devices, and a concept in which everything in the world is considered from the viewpoint of atomic or molecular building blocks [1]. It encompasses the production and application of physical, chemical, and biological systems at scales ranging from individual atoms or molecules to submicro dimensions, as well as the integration of the resulting nanostructures into larger systems [2]. More and more nanotechnology materials and products are used in a wide range of applications. For example in aerospace industry, nanotechnology can be used for producing different materials, such as super light airframe materials, sensor material, multi-functional materials. Nanotechnology based MEMS can also be used in different systems, e.g. sensor systems. Figure 1 shows an application of MEMS in an optical displacement sensor as an accelerometer [3]; and figure 2 shows potential applications of nanotechnology for aircraft structural health monitoring [4].

Nanotechnology based MEMS devices are also widely used in magnetic storage systems, where they are being developed for super-compact and ultrahigh recording – density pursued storage system [2].

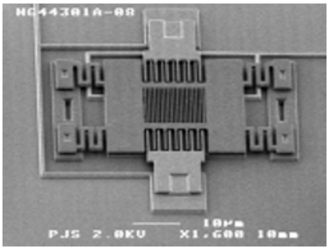


Figure 1. MEMS accelerometer

Figure 2. Potential application of nanotechnology for aircraft structural health

Although nanotechnology materials and products can provide advanced properties and performance, the cost of nanotechnology materials and products is another key issue which needs to be better understood, especially at design stage where "Go or No-Go" decision will be made normally [5]. However cost estimation for nanotechnology mateirals and products is mostly not even carried out at all, or not well carried out.

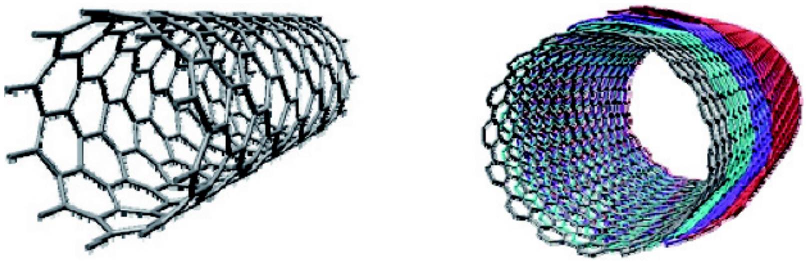
Cost Estimation is well conducted for large scale products in various sectors, especially aerospace and defense industry. A number of approaches are developed and applied for cost analysis, and there are also some commercial software packages used for cost estimation, e.g. SEER®-H and SEER®-DFM from Galorath, and TruePlanning® Tools from PRICE® Systems, and Vanguard Studio® from Vanguard Software®. Among researches, some research were conducted for costing new technology intensive products, e.g. Roy et al developed an approach for costing new technology intensive product in automotive industry [6]. Research have been seen on MEMS costing, e.g. Lawes uses bottom-up approach to estimate the manufacturing cost of a high aspect ratio micro structures [7]; and a commercial MEMS cost model has been developed by IC Knowledge to estimate the MEMS cost based on its categorisation, materials, and fabricating process [8].

However there are no approaches which can be easily adapted for estimating the cost of nanotechnology materials and products. This paper selects Carbon Nanotubes as a case study, to develop an approach and framework to estimate its cost. The approach developed can be benchmarked to estimate the cost of other nanotechnology based products.

2 Background

2.1 Carbon Nanotubes

Carbon nanotubes are nanomaterials or to be more precise nanofibres characterized by a nanostructure that can have a length-to-diameter ratio greater than 10^4. Carbon nanotubes are classified in two categories: multi walled nanotubes (MWNTs) and single walled nanotubes (SWNTs). MWNTs were discovered in 1991 by Iijima using transmission electron microscope (TEM) to examine carbon samples found needle-like tubes. The observation found several concentric tubes of carbon nested inside each other like the Russian dolls [9], as shown in figure 3 (b). SWNTs were discovered independently in 1993 by Iijima and Ichihashi at NEC [10] and Bethune and co-workers at IBM [11] and consisted in a graphene sheet wrapped into a seamless cylinder, as shown in figure 3 (a).



(a) (b)

Figure 3. Single Wall Nanotubes and Multi Wall Nanotubes

An important feature of carbon nanotubes is the chiral vector C_k, as shown in figure 4.

$$C_k = na_1 + ma_2 \tag{1}$$

This vector describes the disposition of the hexagonal rings of the carbon nanotube or in other words describes how the graphene sheet is wrapped. Depending on the value of the chiral vector there are three different morphologies of carbon nanotube:

- If $n = m$, then the carbon nanotube is called armchair.
- If $m = 0$, then the carbon nanotube is called zig-zag.
- Otherwise the carbon nanotube is called chiral.

Figure 4 shows the chiral vector and three morphologies of carbon nanotubes described above.

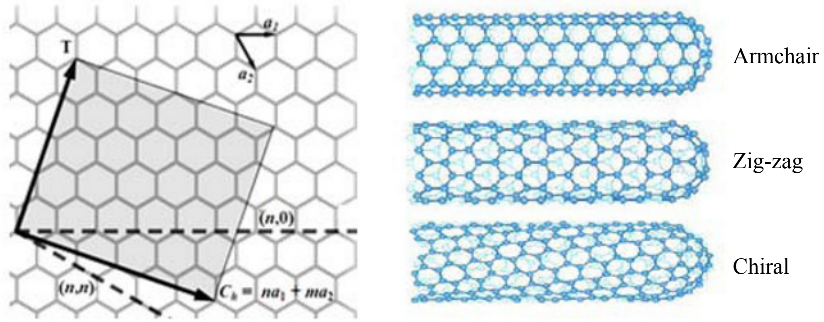


Figure 4. Chiral vector and morphologies of carbon nanotubes

2.2 Production of Carbon Nanotubes

There are four steps to produce carbon nanotubes: acquisition of the equipment and raw materials, synthesisation, purification and quality control. This process can be expressed in figure 5, where the input and output in each step are expressed clearly.

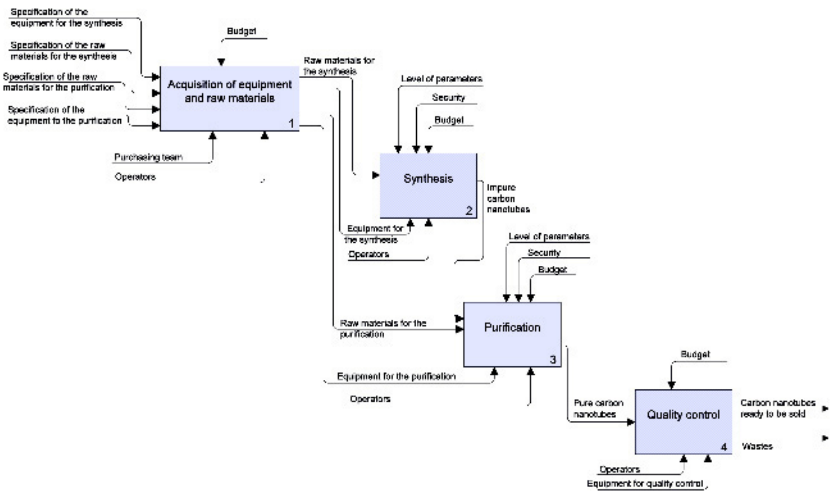


Figure 5. Manufacturing process of Carbon Nanotubes

Sythesisation

There are three different ways to synthesise carbon nanotubes:

- Arc discharge: This method creates nanotubes through arc vaporization of high purity graphite electrodes placed end to end, separated by approximately 1mm, in an enclosure that is usually filled with inert gas (helium, argon) at low pressure (between 50 and 700 mbar) [12]. With this method it is possible to produce both SWNTs and MWNTs.

- Laser ablation: A high-power pulsed laser vaporizes pure graphite targets inside a furnace at 1200°C, in an air atmosphere. This technique can make MWNTs if pure graphite targets are used while SWNTs can be obtained if metallic particles are added as catalysts to the graphite targets. The main problems with this technique are reproducibility and the maximum output of carbon nanotubes achievable that makes this method economically ineffective [13].
- Chemical vapour deposition: This process involves thermal decomposition of hydrocarbons (e.g. methane, benzene, acetylene, naphthalene, ethylene, etc) over metal catalysts (e.g. Co, Ni, Fe, Pt, Pd) that are deposited at predefined locations on a substrate. Of the various means for nanotube synthesis, CVD shows the most promise for industrial scale deposition in terms of its price/unit ratio [14].

Purification

There are a number of purification methods used in industry. A list of the purification methods actually used is as below:

- Oxidation
- Acid treatment
- Annealing
- Ultrasonication
- Magnetic purification
- Micro filtration
- Cutting

Two main purification methods in the list are considered in this paper: the oxidation and the acid treatment. The main disadvantages of oxidation are that not only the impurities are oxidised, but also the SWNTs. Luckily the damage to SWNTs is less than the damage to the impurities. The acid treatment will remove the metal catalyst, i.e. first of all, the surface of the metal must be exposed by oxidation or sonication. The metal catalyst is then exposed to acid and solvated. When using a treatment in HNO_3, the acid only has an effect on the metal catalyst.

Quality Control

Two main approaches for quality control are cconsidered in this paper: the Raman spectroscopy and the Thermogravimetric analysis. Raman spectroscopy is a technique used to study vibrational, rotational, and other low-frequency modes in a system. By Raman spectroscopy different carbon nanotubes can be analysed.

Thermogravimetric Analysis (TGA) is a type of test that is performed on samples to determine changes in weight in relation to change in temperature. In the particular case of carbon nanotubes, the weight change in an air atmosphere is typically a superposition of the weight loss due to oxidation of carbon into gaseous carbon dioxide and the weight gain due to oxidation of residual metal catalyst into solid oxides [15].

3 Cost Estimation Framework Development

The cost estimation framework is established based on the design specification, manufacturing process, and the relationship between them. Cost estimation is carried out in below steps:

- Identify the cost drivers that occur in the manufacturing steps and create a Cost Breakdown Structure;
- Establish the cost estimation for each element of Cost Breakdown Structure;
- Identify the parameters that define the nanotubes material specifications at design stage.
- Identify rules and establish the links between specification and manufacturing process.

Through this apporach, the cost of carbon nanotubes material can be estimated based on design specification data.

3.1 Cost Breakdown Structure

As Carbon Nanotubes is a relatively new nanomaterial, there is no available cost data can be used as a reference. This paper uses bottom-up approach to estimate the cost of Carbon Nanotubes, i.e. the cost of carbon nanotubess is dependant on the process information, e.g. how much raw materials are needed for producing a certain amount of nanotubes, and how long the fabrication time is etc. A hierarchical Cost Breakdown Structure (CBS) for Carbon Nanotubes are established as shown in figure 6.

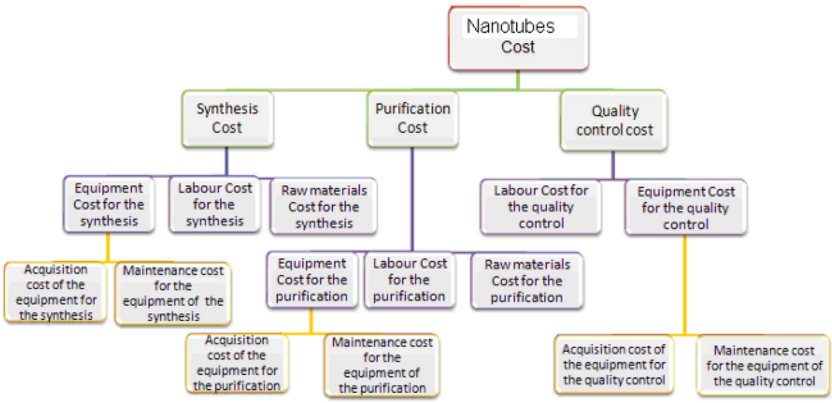


Figure 6. Cost Breakdown Structure for Carbon Nanotubes

3.2 Detailed cost estimation approach

Direct cost elements within CBS can be mainly classified into below categories and their cost estimation are explained as below:

- Raw material cost: the amount of raw materials needed for fabricating a unit of nanotubes multiplies the unit raw material price. The material amount and its unit prices need to be provided by or captured from industry experts and literature.
- Labour cost: committed labour hours multiply the labour rate. Labour hours for a process need to be based on industrial practice provided by or captured from industry experts.
- Machine cost: capital cost of machine tool, and likely life duration of that machine and the proportion of maintenance cost will be used to calculate the amortised machine cost for a unit of Carbon Nanotubes. Again, those data need to be provided by or captured from industry experts.

3.3 Specification

At design stage, some parameters are used to specify specific nanotubes, mainly they are manufacturing process related parameters, and those parameters are identified as the major contributors to the cost of nanotubes materials:

- Volume production: expressed in gram(s).
- Type of carbon nanotubes: either SWNTs or MWNTs.
- Synthesis method: Arc Discharge, CVD, or Laser Ablation.
- Purity of carbon nanotubes: the percentage of carbon nanotubes in the material.
- Quality control method: the Raman spectroscopy and the Thermogravimetric analysis.

3.4 Links between specification and manufacturing data

It's possible to recognise and represent the rules for choosing manufacturing process data based on specification parameters. The framework follows "if, then, else" logic to decide some manufacturing data. For example for arc discharge process, depending on carbon nanotube type and purity level, different amount of raw materias and constituents are needed, the selection of raw material set in the framework will follow below logic as expressed in Microsoft Excel®:

```
IF(AND(Purity<=Level_of_purity,Type_Of_Carbon_Nanotube="SWNT"),
Mat_Set1,IF(AND(Purity<=Level_of_purity,Type_Of_Carbon_Nanotube="MWN
T"),Mat_Set2,IF(AND(Purity>Level_of_purity,Type_Of_Carbon_Nanotube="SW
NT"),Mat_Set3,IF(AND(Purity>Level_of_purity,Type_Of_Carbon_Nanotube="M
WNT"), Mat_Set4)))).
```

The developed framework analyses cost through twelve possible paths. The steps to establish the logic of cost estimation based on specification is as shown in figure 7. As quality control process doesn't change in this case study, it is not taken into account for rule establishment but used for cost estimation.

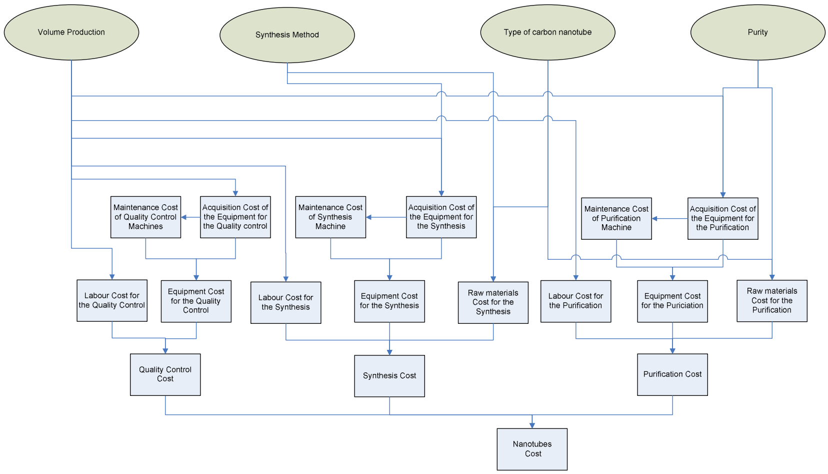


Figure 7. Logic of carbon nanotubes cost estimation based on specification

4 Validation of cost estimation logic

The cost estimation framework developed is preliminarily validated through expert interviews. Two experts working on nanotechnology and two working on cost engineering were invited to assess the cost estmation framework, the approach developed for cost estimation is thought to be a logic way to estimate the cost of carbon nanotubes based on its specification.

Due to non-availability of industrial standard data with the study duration, the framework could not be tested on a particular company, but the validation assesses the logic and the factors that contribute to cost of nanotechnology based products.

5 Discussion and Conclusion

This paper developed a framework for nanotechnology materials and products costing, based on the case study of carbon nanotubes material cost estimation. Limited possible manufacturing paths were considered in the study for demonstrating the approach. In next step, more case studies should be carried out to make the developed approach more generic for nanotechnology product costing.

The cost estimation logic and framework were preliminarily validated though expert interviews. However reliable industry standard data are needed to estimate the cost using this framework. Involvement of industry partners in this process is necessary.

It is concluded that the approach proposed and developed in this paper represents a reasonable logic for cost estimation for nanotechnology based products at design stage. The approach can be benchmarked for cost estimation for other nanotechnology based products.

6 References

[1] J. Ramsden, What is nanotechnology, Nanotechnology perceptions, 2005(1), pp3-17

[2] Bhushan, b. (2003), Springer handbook of nanotechnology

[3] Photo courtesy of Sandia National Laboratories, SUMMiT^{TM} Technologies, www.mems.sandia.gov

[4] Clemens Bockenheimer, Airbus Needs from Nanotechnology for Structural Health Monitoring (SHM), presentation on NANOTECH 2006, Boston, 7th-11th May 2006

[5] R. Roy, Cost Engineering, Why, What and How? Decision Engineering Report Series, Cranfield University, July 2003

[6] R. Roy et al. Estimating the cost of a new technology intensive automotive product: a case study approach, International Journal of Production Economics 97 (2005): 210-226

[7] R. A. Lawes, Manufacturing costs for microsystems/MEMS using high aspect ratio microfabrication techniques, Microsystems Technology (2007) 13: 85–95

[8] http://www.icknowledge.com/our_products/MEMSmodel.html (last accessed on 8th April 2009)

[9] S. Iijima, Helical microtubules of graphitic carbon, Nature 354 (1991), 56-58

[10] S. Iijima and T. Ichihashi, Nature 363 (1993), 603.

[11] D.S. Bethune, C.H. Kiang, M.S. de Vries, G. Gorman, R. Savoy, J. Vazquez and R. Beyers, Nature 363 (1993), 605

[12] V. Hotea, 2007. Present and future with carbon nanotube. Available: http://imtuoradea.ro/auo.fmte/files-2007/TCM_files/Vasile_Hotea_1.pdf (last accessed on 22nd December 2008)

[13] M. Da Silva, Synthesis of carbon nanotubes. Available: http://cobweb.ecn.purdue.edu/~mdasilva/CVDnanotubes.shtml (last accessed on 5th January 2009)

[14] M.J.M. Daenen, R. de Fouw, B. Hamers, P.G.A. Janssen, K. Schouteden, M.A.J. Veld, 2003. Wondrous world of carbon nanotubes. Available: http://students.chem.tue.nl/ifp03/default.htm (Accessed 10th January 2009)

[15] S. Freiman, S. Hooker, K. Migler, S. Arepalli, 2008. Issues in single walled nanotubes. Available: http://www.nist.gov/public_affairs/practiceguides/NIST%20SP960-19.pdf (Accessed 13th March 2009)

An Analogy Based Estimation Framework for Design Rework Efforts

Panumas Arundacahawat, Rajkumar Roy and Ahmed Al-Ashaab

Decision Engineering Centre, Cranfield University, UK

Abstract. This paper aims to reveal the development of design rework effort estimation by an analogy based framework. Design rework is focused in the repetion effort due to the interactions among subsystems or components within a considered system. Factors influence on design reworks are synthesised to be drivers impact on probility of occurence and effort required to resolve design rework issues both of which are verifiied by three industrial case studies. An Analogy based estimation approach is applied to develop a design rework effort estimation framework. The framework is practiced on the water pump development project from an engine manufacturer. The ultimate goal of this development is to develop a framework to estimate for extra time and effort required in the product development projects in the early stage of product development planning.

Keywords. Concurrent engineering, design rework, effort

1 Introduction

Interaction among design activities is a major influence on estimation effort and time required in the product design and development projects. If design activies have strong relationships, any changes from upstream or feedback from downstream could spend a lot of effort and time to resolve. Moreover, the interactions are limited to the estimation by traditional techniques such as CPM and PERT [1]. In product design and development, there two types of planning considered in product design and development [2]. The first type is the vertical planning, which is the planning to achieve product development by coordination among functions, eg. design function, prototyping, testing, manufacturing process design. The second type is the horizontal planning, which the interactions among subsystems or components are considered of which concerened in this research. In general, design reworks are part of iteration in every product development project, which are considered as negative iteration and stochastic in nature[3]. In this paper design rework is defined as unnecessary repetition of design effort due to influences from other design tasks [4]. Therefore, the interactions among subsystems or components are major concerned. However, there are other factors influenced on design rework. A product design and development project is considered as a network of activities. Activities are arranged in sequence. The activity considers in the early order call “upstream”, while the later activity is called “downstream”. The upstream and downstream design tasks are applied in “vertical”, while upstream and downstream subsystems or components are applied for “horizontal” directions as defined in Clark and Fujimoto [2]. Design reworks are classified to be rework due to dependency and interdependency design tasks [5]; and literatures could be divided to be a framework for two and multi design tasks [4]. An overlapping is common practice to reduce lead time in concurrent product development, which is the early incomplete exchanged information between upstream and downstream activities. Design reworks could be severed, if the optimisation of early exchanging information and consideration of dependency or interdependency among tasks are not met [5]. Krishnan et al. [6] proposed that early exchanged information can cause unnecessary iteration due to dependency of preliminary information between upstream and downstream tasks [6]. Loch and Terwiesch [7] viewed downstream rework occurrences as overlapping tasks makes downstream design task to rely on preliminary information. The strength of dependency between upstream and downstream design tasks and overlapping period are used to estimate design rework time. There are other works contributed to design rework for dependency design tasks in vertical direction as reviewed in [4]. For interdependency design activities, work from Smith and Eppinger [8] is a key literature contribute on design rework for interaction among subsystems, however, further literatures are view in [4]. There are two major gaps found in literatures. Probability of rework is either obtained from experts or from assuming. There are not clearly stated what kinds of factors are influenced on probability or effort required. The conclusion of factors impacts on design rework are shown in section two, where drivers on probability of occurrence and effort required are developed. The design structure matrix (DSM) is used to decomposition of products and evaluated the impacted on design reworks, then the analytical hierarchy process (AHP) is applied to evaluation the strength of drivers on probability of design rework occurrence and effort required as well as estimation revealed in section three. Finally the conclusions and future are discussed in section four.

2 Design Rework Effort Estimation Framework

Factors impacted on design reworks are extracted from literatures presented in column 1 of table 1. The factors are classified into probability of occurrence and effort required as shown in last two column of table 1. The drivers will be used as cost estimation relationship for effort estimation purpose [9]. The first column is factors concluded from literatures [4], while the second column is the summary of current practices from three case studies. The current practices are captured from 18 hours of four experts by semi structure interview. The industrial case studies are from an automotive manufacturer, an automotive seat supplier and an engine manufacturer. Upstream changes, faults found by downstream and amount of overlapping are not applicable in this paper, because they are common practices in vertical planning direction on which is not the focus in this paper. Supplier expertise and Constraint to deliver the project on time with acceptable cost are captured in company C, because there is strong evidence from interviews and company's internal document stated that supplier expertise has strong relationship with prototype failures in the validation phase. The drivers in table 1 are validated with experts in an engine manufacturer. Furthermore, all drivers are used to evaluate their strength on design rework as present in section 2.1 and 2.2.

Table 1. Synthesis of drivers impact on design reworks

Factors from literatures	Drivers captured	Company			Probability	Effort required
		A	B	C		
Project complexity (Integration Issues)	Exploitation of CAE software to resolve the critical issues	✓		✓	✓	
	Lessons learnt from previous project	✓	✓	✓	✓	
Dependency	Interaction among subsystems/components of critical issues	✓	✓	✓		✓
Upstream changes	n/a					
Faults found by downstream	n/a					
Amount of overlapping tasks	n/a					
Pre-communication	Coordination across the development from the early stage of development	✓		✓	✓	
	Clear specification			✓	✓	
Crashing	Constraint to deliver the project on time with acceptable cost			✓	✓	
	Supplier expertise			✓	✓	

2.1 Critical design dependency issues

Dependency factor is synthesised to be interaction among subsystems or components of critical issues. This factor needs to be looked at the decomposition of a system in detail. Therefore design structure matrix (DSM) is applied for a water pump development system of an engine development project from company C. A DSM displays the relationships among components of a system in a compact, format [10]. A DSM is a square matrix, which row and column labels are identical as shown in figure 1. The water pump system is decomposed into 13 components and the relationships are signified. The components are sequenced based on knowledge of engineers within the company C. An off-diagonal cross presents the dependency of one component on another. Reading across a row reveals dependency of a component in row against each component in column. The crosses below diagonal are signified the dependent of the component located in each column (upstream components), while the crosses above the diagonal signified the feedback from the component located in each column (downstream component). For example, to select a proper needle bearing needs to know the size and load carried by helical gear, while the possibility of the shaft dimension needs to "guess" before hand. When designing the shaft in the later step and the shaft dimension is not complied with the first guess. Feedback to negotiate with selection needle bearing is unavoidable. From interview, the critical issue is the interaction of torsional activities among ball bearing impeller and helical gear.

		1	2	3	4	5	6	7	8	9	10	12	13	13
Helical gear	1		X			X	X	X						
Impeller	2	X		X			X	X		X				
Pump Body	3					X		X	X	X		X	X	X
Ball Bearing	4	X	X	X			X			X	X			
Needle Bearing	5	X					X							
Shaft	6	X	X		X	X		X	X		X			
Water seal	7	X	X		X		X							
Oil Seal	8			X			X							
Outer clip	9													
Inner clip	10					X		X						
Gasket	12				X								X	X
Back plate	13				X							X		
Bolt	13				X							X	X	

Figure 1. Capturing relationship among components with DSM

The torsional activities create the axial forces from an impeller and a helical gear transferred to ball bearing through a shaft. The impeller transferred axial force due to load while it drives water, and the rotational force is partially converted to axial force due to the helical shape of the gear. Finally all these axial forces interactions make ball bearing fail many time in the validation test. There were 68 failures out of 180 pumps supplied in the validation test, while 23 failures were from ball bearing failed. Therefore, the ball bearing selection is strongly depended on a helical gear and an impeller, while dependency on water pump body and feedback from shaft, inner and outer clips are neglected. The failure due to the interactions has highest probability of occurrence compared with the other causes. Furthermore, each failure requires simulation by CAE software to find out the problem, redesign and sign off with CAE software, which is considered as high amount of rework efforts. The interactions among impeller, helical gear and ball bearing are accounted as drivers impacted on design rework efforts required. The drivers are structured and scored to evaluate the strength, and the details are shown in the next section.

2.2 Analytical hierarchy process (AHP) of the design rework drivers

The analogy based estimation method is selected due to the advantage of manipulating with multi factors [11]. The drivers which have impact on design rework have been identified through a case study. Moreover, the drivers influenced on the probability of design rework occurrence are structured by the analytical hierarchy process (AHP) which is one technique in analogy based estimation, figure 2. The hierarchy is structured followed by the principle from Saaty [12].

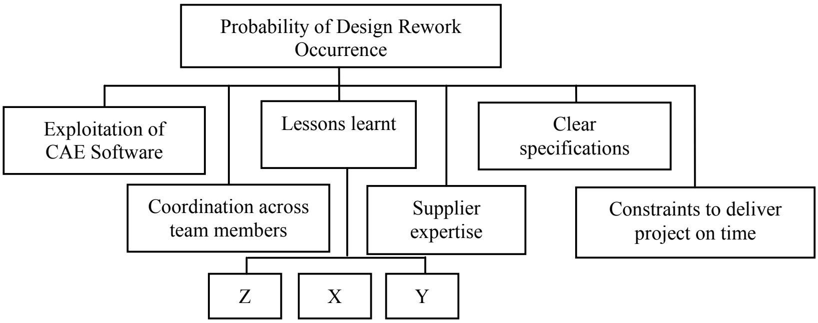


Figure 2. Drivers influenced on probability of design rework occurrence hierarchy

From figure 2, the first level is defined in the AHP technique as the goal needs to be evaluated, while the second level is structuring factors impact on the first level. There is the third level, where projects are compared under each factor; in this case, Z is new project, while X and Y are previous projects which have similarity

		1	2	3	4	5	6	Driver weighted averages
Exploitation of CAE Software	1	1	6	1/8	1/9	1	1/4	0.52
Coordination across team members	2	1/6	1	1/7	1/4	1	1/2	0.31
Lessons learnt	3	8	7	1	2	4	3	2.20
Supplier expertise	4	9	1/4	1/2	1	4	5	1.77
Clear specification	5	1	1	1/4	1/4	1	1	0.64
Constraints to deliver project on time	6	4	2	1/3	/1/5	1	1	0.56

Figure 3. Pair wise comparison for driver impacts on probability of design rework occurrence

Drivers influence on design rework efforts required are structure in the similar principle. Pair wise comparison is conducted to evaluate the strength for each driver by scoring in pair as shown in figure 3. All of the scores are from an expert in company C. Each driver needs to compare across the row against each factor. Score 1 means equally important, while 9 is highest important. The fraction means the driver in row has less important compared with the driver in the considered column. For example, exploitation of CAE software to resolve design issues has six times higher impact than coordination across team members, while it has eight times less impact than lessons learnt from previous projects. The comparison is conducted for the upper diagonal part only as illustrated in figure 4, while the lower diagonal scores are reciprocal. The pair wise comparison matrixes need to check for inconsistency [12]. If the consistency ratio (CR) for each matrix is higher than 0.1, it mean the scoring is inconsistence. Therefore, a re-scoring or other consistency improvement techniques need to be applied [12]. The driver weighted averages are followed from [13], where the score in column are sum up and the numbers in the matrix are divided by their respective column, then the total number in row are calculated to calculate the driver weighted averages in the last column; furthermore, the last column is a 6×1 matrix, figure 3. For the scoring among projects under each driver, the weighted averages are shown in figure 4. With similar principles the drivers influenced on design rework efforts required are perform and helical gear has higher impacts on design rework effort required compared with impeller.

	Exploitation of CAE software	Coordination across team members	Lesson s learnt	Supplier expertise	Clear Specification	Constraints to deliver project on time
Z	1.70	1.81	1.79	1.44	1.71	0.86
X	0.29	0.25	0.17	0.17	0.43	0.86
Y	1.00	0.95	1.04	1.39	0.86	1.71

Figure 4. The project weighted matrix under consideration of each driver

3 Perform Design Rework Estimation

To perform estimation the total weighted average matrix needs to be calculated by equation 1. The total weighted average matrix for probability of occurrence and efforts required of design rework are in figure 5.

$$[\textit{Total weighted}] = [\textit{Project weighted}] \times [\textit{Driver weighted average}] \qquad (1)$$

Projects	Total weighted averages	
	Probability of occurrence	Efforts required
Z	8.37	0.09
X	2.90	0.78
Y	6.89	0.13

Figure 5. Total weighted average matrixes

To estimate design rework probability of occurrence is used equation 2; however, the w_{es}/w_i needs to be replace w_i/w_{es} in equation 2 for estimating design efforts required. Equation 2 is adopted from Bashir and Thomson [1].

$$V_{es} = \frac{1}{n_r} \sum_{i=1}^{n} \frac{w_i}{w_{es}} \times V_i \qquad (2)$$

where V_{es} is the variable being estimated; V_i is the referenced variable i; n_r number of reference project; w_i weighted average for variable i; and w_{es} is the weighted average for variable being estimated.

For preliminary framework validation, the mean magnitude of relative error (*MMRE*) is applied. This evaluation is taken from software effort estimation [1]. The framework is considered to be accurate, if its *MMRE* is equal to 0.25 or less. In this case, project X and Y are use to calculate *MMRE*, because there are real values from company's record. The *MMRE* is shown in equation 4.

$$MMRE = \frac{1}{n} \sum_{j=1}^{n} \left| \left(\hat{V}_j - V_j \right) / V_j \right| \qquad (4)$$

where $\hat{V}_j$ is the predicted variable of design rework occurrence for project j; V_j is the actual probability of project j; n number of project.

The *MMRE* for the framework to estimate design rework probability of occurrence and efforts required are summarise in estimation results in table 2.

Table 2. *MMRE* for estimation frameworks and estimation results

	Estimated Project	Ref. Projects	w_{es}/w_i / w_i/w_{es}	Estimated Values ($\hat{V}_j$)	Actual Values (V_j)	$\frac{(\hat{V}_j - V_j)}{V_j}$	*MMRE*
Probability	X	Y	2.38	0.178	0.138	0.40	0.34
	Y	X	0.42	0.054	0.075	-0.28	
	Z	X	0.35	0.53	-		
		Y	0.83				
Effort Required	X	Y	0.17	1372.48	1610	-0.21	0.24
	Y	X	6.06	265.70	210	0.27	
	Z	X	0.71	168.97	-		
		Y	0.13				

4 Conclusion and future work

The factors driven design rework in this paper are divided to be drivers for probability of occurrence and effort required. All drivers for probability of design rework occurrence are inversely proportional to probability; on the other hand, drivers impact design rework efforts required are directly proportional. All drivers are reflected each company's context. It could be implied from the driver weighted average of which drivers has high priority. If there are new projects to start with, each of them could be compared previous projects by structuring in the third level as shown in figure 2. So, companies can evaluate the readiness of new projects against each driver and planning to mitigate probability of design rework occurrence. In similar way of drivers impacted on design rework efforts required, it allows designer to evaluate the strength of subsystems or component which have high impact and lower the impact. The limitation for this framework is that it requires at least one historical project to perform comparison. Furthermore, it bases on the judgments of experts in scoring; so it has high risk, if list of drivers are long. Drivers captured are depended on company's context, which could be more than this exercise. The framework for design rework efforts required is valid with the criteria $MMRE \leq 0.25$, while the *MMRE* for probability of design rework occurrence is exceeded ($MMRE = 0.34$). However, the scoring in this practice is based on one expert, so the multi-experts scoring is required to resolve this discrepancy.

The future works are defined in two directions. The developing a user friendly pair wise comparison system is required to mitigate the discrepancy of scoring. Implementing the framework with more product design projects is in order to expand the list of drivers impacted on design rework and strengthen the validity of the estimation.

5 References

[1] Bashir H A, Thomson V, An analogy-based model for estimating design effort. J Design Studies 2001; 22:157-167

[2] Clark K B, Fujimoto T, Lead Time in Automobile Product Development Explaining the Japanese Advantage. J Engineering and Technology Management 1989; 6: 25–60.

[3] Ballard G, Positive VS Negative Iteration in Design. In Lean Construction 8th, Brighton, UK, 2000

[4] Arundachawat P, Design Rework Prediction in Concurrent Design Environment: Current Trends and Future Research Directions. CIRP design 2009

[5] Yassine A A, Chelst K R, Falkenburg D R, A Decision Analytic Framework for Evaluating Concurrent Engineering. IEEE Transactions on Engineering Management 1999; 46; 2:144-157

[6] Krishnan V, Eppinger S D, Whitney D E, A Model-Based Framework to Overlap Product Development Activities. J Management Science 1997; 43; 4: 437-451

[7] Loch C H and Terwiesch C, Communication and Uncertainty in Concurrent Engineering, J Management Science 1998; 44; 8: 1032-1048

[8] Smith R P, Eppinger S D, Identifying Controlling Features of Engineering Design Iteration. J Management Science, 1997; 43; 3: 276-293

[9] Roy R., Kelvesjo S.; Forsberg S.; Rush C. Quantitative and qualitative cost estimating for engineering design J. Engineering Design 2001, 12, 2, 147–162

[10] Browning T R, Applying the Design Structure Matrix to system Decomposition and Integration Problems: A Review and New Directions. IEEE Trans on Engineering Management, 2001; 48; 3: 292-306

[11] R. Roy, Cost Engineering, Why, What and How? Decision Engineering Report Series, Cranfield University, July 2003
[12] Saaty T L. Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process: 6, RWS Publications, Pittsburgh, PA. 1994; 1-93
[13] Render B., Stair, R. M. Quantitative Analysis for Management, 7th ed., A Pearson Custom Publication, 2000; 519-536

Design for Sustainability

Greening Economy as a Key Solution to the Economic Crisis

Peter Yang[a,1] and Injazz Chen[b]

[a] Department of Modern Languages and Literatures, Case Western Reserve University
[b] Department of Operations Management, Cleveland State University

Abstract: Although the UN called for a "Global Green New Deal" as a response to the current financial and economic crisis and the "American Recovery and Reinvestment Act of 2009" (ARRA) includes "green" stimulus incentives, there is unfortunately far from a consensus on whether such green incentives will serve as a solution to the current financial and economic crisis. To answer this fundamental question, we will analyze and identify one of the more deeply rooted structural problems of the U.S. economy, the wasteful depletion of finite fossil fuels, and its relationship with three of the major problems facing the U.S. economy – global financial crisis, skyrocketing energy prices, and climate change. We argue that the triple predicament of the existing carbon-based U.S. economy can be ultimately cured by the development and implementation of a comprehensive visionary "green" economic strategy to accelerate the transition to a greener economy.

Key words: carbon-based economy, CO_2 emissions, financial crisis, green economy, renewable energy sources, energy efficiency, climate change

1 Introduction

The UN explicitly called for a "Global Green New Deal" as a response to the current financial and economic crisis and a genuine opportunity for all mankind to achieve "real and long term growth, genuine prosperity and job creation" [1]. The ARRA also sets out the investment of fifteen billion dollars a year to develop technologies like wind power, solar power, advanced biofuels, clean coal, and more fuel-efficient cars and trucks [2]. However, in the midst of the U.S. and global economic downturn and oil price plunge, stimulus packages earmarked for massive bailouts may seem more appealing than those intended for a greener economy. Distrusts about the "Green New Deal" include first of all if the green transition is really a key to solving the economic crisis [3]. In addition, there is disbelief that the U.S. economy really lags in terms of international competitiveness in such a green transition. Clearly, doubts on these crucial issues

[1] Corresponding Author Email : pjy52@hotmail.com

prevent the employment of these green stimulus incentives as a catalyst for the green transformation and, therefore, deserve a thorough investigation.

This paper attempts to answer these questions. We will first investigate the nature of the current economic crisis. We argue that the economic crisis is, just like skyrocketing energy price and climate change are, a result of the existing carbon-based economy. We claim that accelerating the transition to a greener economy provides a key to ultimately solving the more fundamental problems of this seemingly mere mortgage-crunch related crisis. We argue that the economy's energy intensity, car efficiency standard, and installations of renewable energy sources (RES) can be used as its green economic indicators.

The remainder of this paper is organized as follows. Section 2 investigates the fundamental structural problems of the U.S. economy and analyzes the relationship of three major problems facing the U.S. economy – global financial crisis, skyrocketing energy prices, and climate change. Section 3 explores the U.S. economy's international competitiveness in the green transition and examines the sectoral losers and winners during the economic downturn and the legislative and regulatory factors that might have impacted this competitive outcome. This section is followed by conclusions that the accelerated green transition can be a key answer to the triple predicament of the existing carbon-based U.S. economy.

2 Fundamental Structural Problems of the U.S. Economy

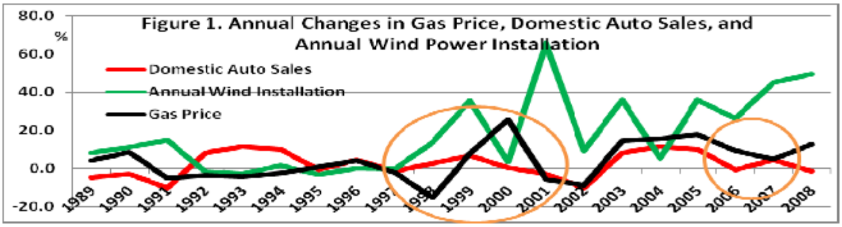
We propose that the fundamental structural problems (weak manufacturing in international competition) of the existing U.S. economy are closely related to its wasteful and inefficient carbon-based energy structure—fossil fuels such oil, gas, and coal—and its manufacturing of fuel-inefficient products that rely heavily on fossil fuels. To test this hypothesis, we examined several related key indicators, including energy intensity of selected economies, world GDP, world energy consumption, world oil price, and U.S. delinquency rate.

2.1 Examinations and Results

First, we compared the U.S. energy intensity with that of some other economies. Our findings reveal that, although the U.S. energy efficiency has somewhat improved over the last quarter century, it remained one of the most wasteful and inefficient major economy [4]. This was true in terms of both rising energy consumption per capita and energy consumption per dollar of GDP. As a drastic contrast, the German economy both improved energy efficiency and reduced energy consumption per capita between 1990 and 2005.

Second, we reviewed the statistical data of the growth in world GDP, world oil consumption, the increase in world oil price, and U.S. delinquency rate between 1981 and 2008 from international and U.S. organizations [4, 5, 6, 7]. Results reveal a high correlation between the growth in world GDP and the change in world oil consumption (0.62) on the one hand and a high negative correlation between changes in world oil consumption and U.S. delinquency rate (-0.57) on the other during this period. These correlations indicate a clear oil dependency of the existing world economy and the consumers' ability to consume oil and to make mortgage payments. Also, the oil dependent world economy was often affected by the "overreaction" of the supersensitive world oil price movement and surges in oil price were often followed with delay by increases in U.S. delinquency rate.

Third, we examined the relationship of the two industrial sectors, the automobile sector and the wind energy sector, with the rising gas price between 1989 and 2008. We use statistical data on annual changes in domestic auto sales, annual wind power installation and gas price. While the annual wind power installation in general responded excitedly to the rising gas price, the sales of domestic vehicles often reacted depressively to it. While these reactions were also affected by other factors such as government incentives, the "invisible hand" of market forces cannot be ignored [Figure 1]. With the help of this "invisible hand," the U.S. wind power capacity soared since 2006. In 2008, the U.S. even overtook Germany as the world's number one wind power generator [8].



2.2 Discussion

Over the last 50 years, the world economic growth caused accelerated environmental decline. Although the global GDP more than doubled from 1981 to 2005, "60 percent of the world's ecosystems were degraded or used in an unsustainable manner" [1]. As the world's dominating economic power and most wasteful and inefficient carbon-dependent economy, the U.S. economy left far more environmental and climate footprint than other developed and developing economies on the planet. It bears a major responsibility for these environmental problems and needs to look for and work on their solutions.

The wasteful and inefficient carbon energy structure of the U.S. economy had a "hard budgetary constraint," which kept the U.S. manufacturing, especially the energy sector and oil-dependent automotive sector, largely uncompetitive on international markets and dependent on foreign oil and products. The wasteful depletion of the exhaustible carbon resources has serious cost effects on the existing U.S. economy. As fossil fuels are hitting their finite limits, any expansion of the U.S. and global economy and the wasteful and inefficient use of these resources inevitably translate into their accelerated depletion and higher prices.

In addition, the enormous amounts of pollutants and CO_2 discharged from the wasteful production and consumption of the fossil fuels also add enormous economic strains to the ailing U.S. economy. Since the pollution adversely impacts our living space and environment (e.g. air, land, and water) and our health, the environmental pollution and its abatement have been and will remain extremely costly in the U.S. Moreover, the unprecedented massive impacts on global climate caused by CO_2 emissions and the perceived enormous costs associated with their reduction have become a new fear factor for the already troubled economy.

Certainly the current mortgage crunch was primarily triggered by the subprime lending because of "shoddy" subprime lending practices. However, we should not ignore a deeper reason for this financial crisis. As the expansion of the world GDP and the related increase in energy demand caused recent rising oil price, this skyrocketing oil prices puts, in turn, significant strains on the U.S. manufacturing in terms of rising production and transportation costs [6], the related job markets (rising unemployment rate), and financial markets (rising interest rate). Therefore, the hard budgetary constraint of the carbon-based U.S. economic model ultimately caused the mortgage crunch in the U.S. [9]. This relationship is also confirmed by the correlation of the rising oil price with the delinquency rate, which demonstrates the significant lagged relationship of the skyrocketing oil price with mortgage markets; the rising delinquency rate followed the rising oil price with delays.

To be sure, the recent drastic drop in oil prices is only a temporary phenomenon as a result of drastically reduced energy demand in response to the contracted economy. Once the existing carbon-based economy is back to "business as usual," the hard budgetary constraint will again haunt this economy. If we mistake the temporary drop in oil prices as the turning point of long-term rising oil demand and oil prices, we will miss the opportunity of switching over to a greener economy.

Therefore, although the current global financial crisis was primarily triggered by the U.S. mortgage crunch, it was ultimately the hard budgetary constraint of the U.S. economy's prolonged reliance on the wasteful and inefficient carbon-based energy structure and oil-dependent manufacturing that caused the threefold dilemma of the U.S. economy – soaring energy prices, pollution and climate change, and the global financial mess. As Pavan Sukdhev, head of global markets at Deutsche Bank, pointed out, the carbon-based "economic models of the 20th century are now hitting the limits of what is possible" [1].

3 The U.S. Economy's International Green Competitiveness

An economy's international competitiveness in the green transition is directly related to the understanding of the importance of green strategies and their execution at both the corporate and the government levels. Empirical studies demonstrate that proactive corporate green strategies generally have a positive impact on firms' competitive performance [10, 11].

However, environmental economics has long pointed out the market failure related to environmental and climate protection as a result of high future commercial rates and associated high initial costs for R&D, and marketing of green energy sources and green products, which prevent firms from being voluntarily interested in reducing environment-unfriendly activities [12, 13, 14, 15].

This is especially true for innovations related to processes, which have clearly a less positive competitive impact if exclusively driven by legislations than if driven by market incentives because of economic and technical risks of environmental innovations (e.g., high investment costs and uncertain profitability) [10].

Therefore, a key to resolving green competitive problems lies in providing, in conjunction with environment-conscious standards and regulations, necessary market incentives to green businesses. Such market incentives can help reduce the economic and technical risks of innovations and make the green economy, including clean, renewable energy and energy efficient products such as cars and trucks appealing and affordable for consumers and profitable for developers and producers. The combination of green regulations and standards and necessary market incentives can motivate firms to develop corporate social responsibility and innovate in terms of designing and developing products that both meet the needs of consumers and have minimum external costs to environment and climate [16, 17]. It can also help them develop strategic competitiveness at low cost [10, 18, 19].

Certainly, when oil prices continued to skyrocket, U.S. companies started speeding up investment in renewable energy sources (RES) and this allowed the U.S. to become the world's biggest wind power generator. The contrasting results of the skyrocketing oil prices in terms of plummeting car sales and record high wind energy installation in the U.S. can serve as a wake-up call for decision makers while they choose a priority for stimulus packages. The trouble of the "Big 3" U.S. automakers–GM, Ford, and Chrysler–should not merely be seen as of temporary, external or financial nature. Their deeper ailment in terms of the lack of more competitive, more fuel efficient products must be diagnosed and treated. The triumphant surge of the wind power installation points out the U.S. economy's future direction, and it shows where government stimulus packages should invest.

However, it was also the powerful "invisible hand" that soon deepened financial and economic crises, which in turn caused oil price to plunge and diminished green market opportunities [20]. The mere reliance on market forces faces a dilemma; if we exclude the government role in transition to a greener economy, the market failure for RES due to plummeting oil prices will mislead us to water down or even postpone the green transformation until the next round of energy price hikes force us to take going green seriously again. Now that the "invisible hand" failed to sustain the rapid growth of RES, it is more important for the government incentives to kick in to sustain or even accelerate the RES growth.

The underlying problem of the carbon-based U.S. economy is that it does not have a consistent, coordinated long-term strategic plan and an accompanying legal, regulatory, and incentive-based support system to facilitate the necessary changes for a green economy despite the worsening environmental and climate conditions and the weakening U.S. competitive position in global markets.

Like some other governments, the U.S. Government often sent wrong market signals to businesses by subsidizing without a clear strategic goal in mind. On the one hand, it provides huge subsidies to fossil fuels, the carbon-based automobile manufacturing, and agriculture. On the other, it has done little compared with some European economies to foster an economy based on a green energy structure. Federal energy related funds between 1950 and 2003 biased disproportionally (72 percent) to fossil fuels (approx. 47 percent for oil, 13 percent for natural gas, and 12 percent for coal), while Federal incentives for RES such as wind, solar photovoltaic (PV), hydro, and geothermal powers accounted for only 17 percent. The subsidies for fossil fuels between 2007 and 2012 will total US$32 billion.

In addition, the U.S. failed to set high fuel efficiency standards for its products. The current average U.S. car fuel efficiency standard (25 mpg), for instance, is not only far below those in developed economies such as Europe and Japan (40 mpg), but even below that in a developing country such as China (36 mpg). Even the new fuel efficiency standard for vehicle models in 2020 (35 mpg) will still be below China's current standard, not to mention China's higher fuel efficiency standard (43 mpg) for car models in 2009 and Europe's and Japan's even higher fuel efficiency standards (48.9 mpg) for car models in 2012 and 2013 respectively.

This kind of wrong market signals, along with other structural problems such as perceived low quality, unreliability, and high maintenance cost of the products, misled U.S. manufacturers of gas guzzling products. U.S. automakers, for instance, continued to complacently output unpopular products such as wasteful and inefficient automobiles that few consumers are willing to purchase, especially when they faced the hard budgetary constraint of skyrocketing oil prices, related high interest rate, and tightened financial resources.

As a result, they lacked the motivation to whet competitive edge on international markets, compared with other developed countries such as Germany and Japan, and the necessary appeal to the U.S. consumers, especially in the wake of the skyrocketing oil price and the global financial crisis. This, in turn, helped lead to a huge and growing trade deficit with other trading nations and tremendous losses on the U.S. markets for years. No wonder that the "Big 3" were among the first giant companies that were hit hard by the skyrocketing oil prices and tightened financial resources, and begged the government for massive survival aid.

Because of the lack of environment-friendly regulations and market incentives, many U.S. manufacturers failed to react to the public concerns about global warming related damages by "greening up," i.e. transition to a renewable energy structure and upgrade to more energy efficient products. This can be seen as one of the main reasons for the chronicle disease of the world's number one carbon-based economy, which, in turn, has a serious impact on the global economy.

Greening up is an environmentally and economically responsible decision. It is also an economically meaningful upgrade and structural change required by the socioeconomic development for developed economies compared with developing economies in line with various economic theories on comparative labor-capital cost structure and cost base, product cycle, and comparative advantage. Switching to RES and more energy efficient products and services can attract investments into new technologies and green products and services, revive manufacturing, create energy efficient products and markets, as well as new jobs [21, 22].

If the Big 3 had been focusing on making fuel-efficient, compact and reasonably priced vehicles for the past 15 years, they would not be losing business to Honda and Nissan among others, which had a much better competitive edge because they focused on the green economics of their products. Because the Japanese and German competitors have many renewable energy technologies and fuel efficient products or "gas sippers," they did very well at soaring gas prices, were less hit by the current economic crisis, and will recover much faster after the crisis. Without strict energy efficiency standards and related market incentives, the corporate America seemed to be less interested in greening up its products, and more interested in lobbying the Congress for retaining the low fuel efficiency standards or greening "out" by outsourcing its outdated ineffective and wasteful manufacturing to developing countries, especially China, and renewable energy production to climate-friendly countries such as Germany. For instance, 97% of the equipment for making solar panels produced by Applied Materials in Santa Clara, California, "goes to foreign manufacturers, who then sell panels in the U.S." [22].

Since the U.S. Government did not play a pivotal role as the German Government did in providing a long-term, consistent green economic policies in conjunction of well designed market incentives for a strategic green transformation, the corporate America lacked motivation and ability to green up, which in turn caused U.S. manufacturers to further lose their international competitive edge to more ambitious green competitors such as German and Japanese manufacturers and thus expand the massive U.S. trade deficits.

The U.S. is now struggling to bail out the ailing automakers by providing them with hefty rescue packages using tax payers' money. However, any rescue package without addressing the fundamental problem of weak international competitiveness cannot ultimately resolve the deeply rooted structural problems of the U.S. economy in a long run. This indicates that the U.S. Government needs a visionary, transformative economic strategy to reenergize its economy, motivate businesses to invest in R&D of renewable and clean energy and energy efficient production processes and products, such as cars and buildings. If the U.S. Government fails to do so, the U.S. economy runs the risk of sidestepping the opportunity of strengthening its manufacturing by increasing its current output share of only 20 percent in GDP compared with its major share of 76 percent for the service sector.

4 Conclusions

This paper linked the current financial crisis in the U.S., skyrocketing energy prices, and climate change with the constraint and the structural weakness of the carbon-based economy, along with the insufficient market incentives to facilitate a strategic transition to a greener economy. Through this linkage, we demonstrate that the current U.S. and global financial and economic crisis has its structural problems more deeply rooted in its carbon dependency and its wasteful depletion of the finite carbon resources. We also analyze the U.S.'s competitive issues in terms of a green transition from the carbon based economy to the green economy, by comparing the U.S. energy development policy and energy efficiency standards with those of other countries. Based on such a premise and analysis, the green transition can be considered the only way to save the planet from catastrophic climate change and the ultimate solution to the constraint and the structural weakness of the carbon-based economy in the long run.

5 References

[1] UNITED NATION ENVIRONMENT PROGRAMME. Global green new deal - environmentally-focused investment historic opportunity for 21st century prosperity and job generation. London/Nairobi, 22 October 2008.

[2] AMERICAN RECOVERY AND REINVESTMENT ACT OF 2009. Pub.L.111-5, H.R. 1, S. 1.
[3] NICKLAUS, D. Economists say stimulus won't work, St. Louis Post-Dispatch, 29 January 2009.
[4] ENERGY INFORMATION ADMINISTRATION. International Energy Annual 2006, 31 December 2007.
[5] INTERNATIONAL MONETARY FUND. World Economic Outlook Database, October 2007.
[6] WORLD BANK. GEP2009 - Global Economic Prospects 2009: Commodities at the Crossroads, Washington, DC, 2008.
[7] MORTGAGE BANKERS ASSOCIATION. MBA National Delinquency Survey, 2008.
[8] AMERICAN WIND ENERGY ASSOCIATION. Wind energy grows by record 8,300 MW in 2008. Smart policies, stimulus bill needed to maintain momentum in 2009, 9 January 2009.
[9] CARR, F.M.; BEESE, J.A. "How raising oil prices caused home foreclosures from 2005-2008," Proceedings of the Allied Academies 15, no. 2(2008): 70-73.
[10] TRIEBSWETTER, U.; WACKERBAUER, J. Integrated environmental product innovation in the region of Munich and its Impacts on company competitiveness. *Journal of Cleaner Production* 16 (2008): 1484-1493.
[11] ARAGÓN-CORREA, J. A.; HURTADO-TORRES, N.; SHARMA, S.; GARCÍA-MORALES, J. Environmental strategy and performance in small firms: A resource-based perspective, Journal of Environmental Management 86, no. 1(2008): 88-103.
[12] REINHARDT, F. Market failure and the environmental policies of firms: economic rationales for 'beyond compliance' behavior. Journal of Industrial Ecology 3, no. 1 (1999): 9-21.
[13] KING, A.; LENOX, M. Does it really pay to be green? Journal of Industrial Ecology 5(2002): 105-117.
[14] WAGNER, M. Does it pay to be eco-efficient in the European energy supply industry? Zeitschrift für Energiewirtschaft 27, no. 4(2003): 309-318.
[15] WAGNER, M.; SCHALTEGGER, S.; W. WEHRMEYER, S. The relationship between environmental and economic performance of firms: what does the theory propose and what does the empirical evidence tell us? Greener Management International 34(2002): 95-108.
[16] PORTER, M.E.; VAN DE LINDE, C. Green and competitive, Harvard Business Review (September–October 1995): 120–134.
[17] WALKER, H.; DI SISTO, L.; MCBAIN, D. Drivers and barriers to environmental supply chain management practices: lessons from the public and private sectors, Journal of Purchasing and Supply Management 14(2008): 69–85.
[18] SHARMA, S.; ARAGON-CORREA, A. A Corporate Environmental Strategy and Competitive Advantage (Cheltenham: Edward Elgar Publishing, 2005).
[19] ORSATO, R. Competitive environmental strategies: when does it pay to be green? California Management Review 48, no. 2(2006): 127–143.
[20] SCOTT, M.; TSCHANG, C.-C. Clouds over the Solar Power Industry. *Business Week Online* (December 18, 2008): 12.
[21] UNITED NATION ENVIRONMENT PROGRAMME. Green jobs: Towards decent work in a sustainable, low-carbon world, UNEP/ILO/IOE/ITUC, London/Nairobi, September 2008b.
[22] CAREY, J. A hazy forecast for green jobs: Can Obama's energy plan turbocharge America's eco-friendly businesses? Business Week (8 December 2008): 30-32.

A Study on Evaluation of Environmental Effectiveness of Manufacturing Processes

Nozomu Mishima[1], Shinsuke Kondoh, Keijiro Masui, Masayoshi Yasuoka, Yuji Hotta and Koji Watari

Advanced Manufacturing Research Institute, AIST, Japan

Abstract: To achieve sustainable society, many efforts in reducing environmental impact of manufacturing have been tried. However, since the quality is the most important issue, it is necessary to take manufacturing quality into account in evaluating environmental effectiveness of a manufacturing process. For the purpose, we have proposed a new index to evaluate products and manufacturing processes named "total performance indicator (TPI)." TPI shows a balance of manufacturing quality versus environmental impact and cost. The bottleneck process in enhancing product quality can be clarified, by calculating TPI of each process. This paper analyses an actual product and allocates quality characteristics to functional requirements of the product. Then, it quantifies the contribution of each process in creating the product value. A process which doesn't contribute much in creating value and generates considerable environmental impact and cost, should be improved. It is shown that a designer can evaluate and redesign manufacturing processes, based on the result of this approach.

Keywords: Environmental effectiveness, manufacturing process, manufacturing quality

1 Backgrounds

To achieve sustainable society, it is important to reduce environmental impact of manufacturing processes. For manufacturing engineers, enhancing manufacturing quality has long been the most significant goal. Therefore, it is necessary to consider manufacturing quality. Sometimes there are trade-offs between quality and environmental impact of manufacturing processes. A slight increase of the environmental impact of a manufacturing process may reduce total lifecycle environmental impact. That is why we should consider products and manufacturing processes to fabricate the products simultaneously. There have been many eco-indexes such as "eco-efficiency [1]." However, those indexes cannot evaluate each component of the product, or each segment process of the total manufacturing process, since the existing indexes often evaluate the product or process as an inseparatable unit. This strategy is powerful in order to label "green"

[1] Corresponding Author Email : n-mishima@aist.go.jp

products/processes. However, they usually do not mention about individual components or segment processes. Thus, it is difficult to know design improvement strategies. We propose a new environmental effectiveness index named "total performance indicator [2] (TPI)." TPI can be a powerful tool in determining design strategies for "green products," because it tries to consider a product value as a weighed sum of many functional requirements. Each functional requirement is allocated to components. Through this procedure, it is possible to know which components contribute in creating product value greatly, and which components are generating large costs and impacts. It means that the analysis can calculate "efficiency" of each component in creating product requirements. . In this paper, we apply TPI to manufacturing processes. By calculating TPI of each segment process, bottleneck processes in enhancing quality can be clarified. This paper takes ceramic diesel particulate filter (DPF) as an example. The approach quantifies the contribution of each segment process in creating the product value. A segment process which doesn't contribute much to create value and generates large environmental impact should be improved. It is expected that a designer can determine which processes are really environmentally effective.

2 An Index for Environmental Effectiveness

Usually manufacturing processes are combinations of many segment processes, such as material processing, forging, finish machining, etc. In addition, there are many ways to combine processes and boundary conditions. Therefore, it is very important to evaluate which manufacturing process is really environmental conscious comparing to alternative options. In present research, we propose an index to evaluate real environmental effectiveness of products, by considering product's utility value, cost and environmental impact. Efficiency index is defined by (1) and is named total performance indicator (TPI). Since in existing evaluation indexes the "value" is usually a fixed value, it cannot consider change of the value throughout the product lifecycle. The proposed index is the simplest combination of environmental and economical aspects. In our proposal, because the utility value of the product can be expressed by integration of occasional values throughout the lifecycle, it can simulate value decrease due to obsolescence and physical factor.

$$TPI = \frac{\sum_{i=1}^{n} UV_i}{\sqrt{LCC}\sqrt{LCE}} \qquad (1)$$

TPI: Total performance indicator, *UV*: Utility value of the product
LCC: Lifecycle cost, *LCE*: Lifecycle environmental impact

3 Extension to Manufacturing Process Evaluation

It is important to take quality into account to evaluate whether the manufacturing process is really environmentally effective. Since the purpose of manufacturing is achieving a certain product quality, low quality process will not be used. That is the reason why design engineers and manufacturers have concentrated in improving the quality and reducing cost of manufacturing. They might not accept an indicator that does not evaluate cost and functionality, properly. The idea of TPI of manufacturing process is based on product TPI. To evaluate the performances of manufacturing processes, the same idea can be applied. We define the total performance of the manufacturing process by (2). The equation expresses the balance of the product value created by the process, versus the cost and environmental impact necessary to fabricate a product. But in the manufacturing stage, it is usually difficult to know the lifecycle of the product such as what is the obsolescence rate, etc. Thus, in order to simplify the expression, we propose to replace “utility value” by the occasional “value” of the product. The simple “value” can be measured by the market price, when the product is commercially available. Then, (3) shows the simplified TPI of each segment process.

The numerator “Vi” of the equation may vary due to process quality. For example, a manufacturing process with higher profile accuracy may have a higher value than a similar manufacturing process with lower quality. Manufacturing quality also has a strong relationship between cost and environmental impact of the process. For example, it is known that cost and environmental impact may vary due to the cutting conditions [4] and usually they are larger when the manufacturing quality is higher. In addition, for these reasons, in evaluating manufacturing processes, it is necessary to consider value of the segment process versus cost and environmental impact concurrently. We can quantify how the target manufacturing process is environmentally effective, by calculating (3).

$$TPI_{processg} = \frac{UV}{\sum_{i=1}^{i=n} \sqrt{MCE_i \cdot MCC_i}} \tag{2}$$

$TPI_{process}$: Total performance indicator of the manufacturing technology
$MCCi$: Cost of ith segment process, n: Number of processes
$MCEi$: Environmental impact of ith segment process

$$TPI_{segment} = \frac{V_i}{\sqrt{MCE_i \cdot MCC_i}} \tag{3}$$

$TPI_{segment}$: Total performance indicator of the segment process
Vi: Value of the product added by ith segment process

4 Example of Manufacturing Process Evaluation

To show the actual procedure of process analysis, a ceramic diesel particulate filter (DPF) which is shown in Figure 1, was chosen as the example. Ceramic DPF is used frequently because of its high thermal endurance and high specific strength.

Roughly speaking, the function of a DPF can be separated into 5 functional requirements (FR) shown in Table 1. Then, the 5 FRs can be related to 11 quality characteristics also shown in Table 1. The price of the filter unit is assumed to be 20 kJPY. Applying QFD [5], [6], it is possible to clarify the importance of each FR of a DPF. Table 1 shows how each FR is allocated to the 11 quality characteristics. By considering the importance of each FR, it is possible to determine the value. The analysis suggests that each FR occupies 1/5 of the total value of the DPF. Then, the value of the quality characteristics can be calculated by distributing the value of FR accordingly to the importance of each characteristic using (4).

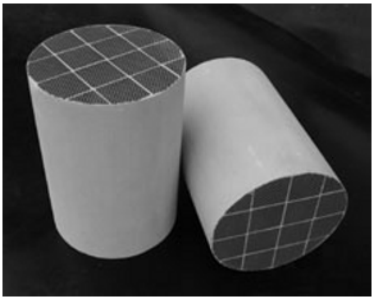


Figure 1. Diesel Particulate Filters for Automobiles

$$QV_k = \sum_{i=1}^{5} V_i \cdot (w_{i,k} / T_i) \tag{4}$$

QV_k: Value of kth quality characteristic, V_i: Value of ith FR
$w_{i,k}$: sum of importance of kth quality characteristic on ith FR
T_i: sum of importance of all the related quality characteristics

Table 1. Relationship between FRs and quality characteristics

		Functional requirements					
		PM capturing capability	Fuel loss due to pressure loss	Fuel loss due to reproduction	Life	Reliability (crack-free)	Value (kJPY)
Value of functional requirement (kJPY)		4	4	4	4	4	-
Quality characteristics of DPF	Thermal conductivity					9	1
	Coefficient of thermal expansion					9	1
	Thermal endurance		3	3		9	2.8
	Pore rate	9	9				3.5
	Specific heat capacity			9			3
	Uniformity of pore distribution	3	3				1.2
	Average pore diameter	3					0.4

	Mechanical strength				3		1
	Profile accuracy (length)	9				3	1.4
	Profile accuracy (section)	9				3	1.4
	Uniformity of the material				9	3	3.3
	Sum of the importance	33	15	12	21	36	

The second step of the analysis is to determine the contribution of each segment process to the value. By identifying the relationship between segment processes of the total process and the quality characteristics, it is possible to calculate the value of the segment processes. We dismantled the total manufacturing process into 6 processes. The value of each segment process is expressed by (5). Table 2 shows the results of the calculation of the segment process value.

$$PV_j = \sum_{k=1}^{11} QV_k \cdot (x_{j,k} / S_k) \tag{5}$$

PV_j: Value of jth segment process, QV_k: Value of kth quality characteristic
$x_{j,k}$: importance of kth quality characteristic have on jth segment process
S_k: sum of importance of all the related segment processes

Table 2. Relation between quality characteristics and manufacturing processes

			Segment process						
			Material mixture	Ball milling	Injection moulding	Binder removal	Sintering	Bonding of units	Total
Quality characteristics of DPF	Thermal conductivity	1	9	3					12
	Coefficient of thermal expansion	1	9	3					12
	Thermal endurance	2.8		9	3	1	1		14
	Pore rate	3.5	9	3	1	1	1		15
	Specific heat capacity	3	9						9
	Uniformity of pore distribution	1.2	1	3	3	3	1		11
	Average pore diameter	0.4	9	3					12
	Mechanical strength	1		3	3	3	3	1	13
	Profile accuracy (length)	1.4	9	3				1	13
	Profile accuracy (section)	1.4	9	3				1	13
	Uniformity of the material	3.3		3		3	1		7
	Value of the process (kJPY)		9.0	5.7	1.4	2.3	1.3	0.3	20
	Yield rate of process		0.99	0.6	0.8	0.95	0.95	0.95	-
	Real value of process (kJPY)		8.9	3.4	1.1	2.2	1.2	0.3	17.1

In an actual manufacturing process, an output of a certain process is usually an input of the next process. These intermediate relations often do not affect the quality of the final product, but they do affect each other. For example, ball-milled slurry often has high viscosity and causes relatively large shrinkage during "sintering." Although the viscosity of the slurry does not affect the final product, it strongly affects the quality of "sintering." To express the interaction, "yield rate" was introduced. In the example process, "ball milling" has relatively low yield rate. This is because there are some uncertainties and some of the intermediate products of "ball milling" do not satisfy the requirements. The low yield rate is reflected in the table as the "real value" of the segment process. The yield rates were roughly estimated based on the knowledge of manufacturing engineers in this field.

As the next step, it is necessary to quantify cost and environmental impact of each process. Not only the process itself but also facility is an important option. The purpose of the evaluation is to determine what product should be made by which facility through which manufacturing process, concurrently. Therefore, machine cost, material cost, energy cost and labour cost should be considered by the term "cost." And environmental impact of machine, manufacturing processes, material of the products and energy should be all totalled. Machine cost of the processes can be calculated by allocating initial cost of the machine to the corresponding process time in which the machine is under operation, by assuming the length of the life of the manufacturing facilities. Table 3 shows the total cost and environmental impact of the processes to fabricate one DPF unit.

Table 3. Total cost and environmental impact of segment process

	Segment process					
	Material mixture	Ball milling	Injection moulding	Binder removal	Sintering	Bonding of units
Environmental impact of process (kg-CO2)	5.0	1.0	1.4	9	10	0.1
Cost of the process (kJPY)	5.0	3.0	1.2	5.0	1.5	2

5 Redesign of Manufacturing Process

Using value, cost and environmental impact shown in the table in the former section, a TPI graph can be drawn. Figure 2 is the TPI graph of the manufacturing process. The solid line indicates the unadjusted value. The dotted line shows the adjusted value, when yield rate is considered. The inclination of a segment line shows the TPI of the corresponding segment process. The inclination of a virtual line connecting the starting-point and the end-point indicates the TPI of the total process. Compared to the TPI of the total process, segment processes "binder removal" and "sintering" have a lower TPI, and the other processes have a relatively higher TPI. This is because "binder removal" and "sintering" require temperature rise using a furnace, which consumes a large amount of energy.

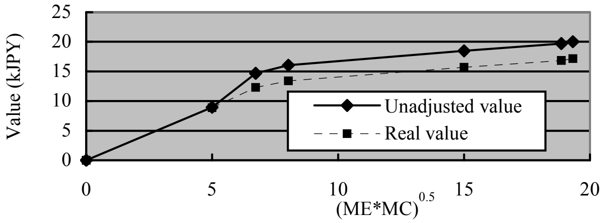


Figure 2. TPI of the manufacturing process of a DPF

To improve the TPI of the total process, we should focus on a segment process having a shallow inclination. Although all the segment processes are targets to consider process improvement options, there are some limitations in the actual manufacturing process. In this example, materials to be mixed are strictly determined in order to ensure overall performance of the filter. Secondly, "injection moulding" is not very efficient according to Figure 3. But, currently there is no candidate for an alternative process. Thirdly, "bonding of unit" should be also removed, because the value created by the process is very small. However, since the cost and the environmental impact of this process are very small, big effect of improvement cannot be expected. Because of the reasons, "ball milling," "binder removal" and "sintering" are identified as the improvement targets.

Improvement of the DPF manufacturing process is an ongoing research topic. Some methods for enhancing the performance or reducing the process time have already been studied. A new technique [7] to replace organic binder by inorganic binder which is environmentally effective is an alternative technique for "binder removal." A technique called "wet jet milling [8]" was also implemented. Raw ceramic body using jet-milled slurry that had low viscosity and low re-flocculation properties, showed small shrinkage during sintering. Because of this, the yield rate was greatly improved. Table 4 shows the value, yield rate, cost and environmental impact of the new process. Improved processes contributed in reducing the cost and environmental impact, and enhancing the value. Figure 3 is the TPI graph of the improved process. The graph tells us that the TPI of the total process was greatly improved and the new process is more environmentally effective.

Table 4. Value, cost, environmental impact of the new process

	Segment process						
	Material mixture	Wet jet milling	Injection moulding	Improved binder	Sintering	Bonding of units	Total
Value of the process (kJPY)	9.0	5.7	1.4	2.3	1.3	0.3	20
Yield rate of the process	0.99	0.95	0.8	0.95	0.95	0.95	-
Real value of the process (kJPY)	8.9	5.4	1.2	2.2	1.2	0.3	19.2
Environmental impact (kg-CO2/unit)	5	1.3	1.3	6.7	9.6	0.1	
Cost of process (kJPY)	5	2.6	1.1	0.5	1.3	1.6	

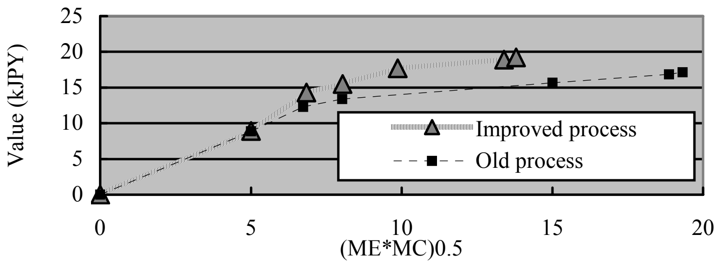


Figure 3. TPI of the improved manufacturing process of DPF

6 Conclusions

In this paper, a new method was proposed to evaluate and redesign manufacturing processes. As a result of applying the method to the manufacturing process of a ceramic diesel particulate filter, it was suggested that the environmental effectiveness could be improved by replacing certain processes by more efficient processes. An analysis of the actual process explained “wet jet milling” and “improved binder removal” were effective in reducing the cost and environmental impact. These improvements also contributed to enhance the manufacturing quality by improving the yield rate which strongly affects the value of the total manufacturing process. Precisely speaking, “wet jet milling” improved the yield rate greatly, and replacement of organic binder to inorganic binder was effective in reducing the environmental impact and cost. Through this approach, it is concluded that the proposed design approach is helpful in evaluating and redesigning environmentally conscious high quality manufacturing processes.

As future work, it is necessary to consider how quantification of value enhancement is possible when the “yield rate” is same and the quality of the final product is improved. In addition, a totally new process improvement should be analysed by this approach and put into practice in order to prove the suggestion is useful in determining new process improvement options.

7 References

[1] Tahara.K, Yamaguchi K., Sagisaka M. and Inaba A. CO2 Efficiency: A New Evaluation Method of Companies Environmental Performances, Proc. of the 6th International Conference on EcoBalance, Tsukuba, Japan, Oct. 2004, pp.269-270.

[2] Kondoh S., Masui K., Hattori M., Mishima N. and Matsumoto M. Total Performance Analysis of Product Life Cycle Considering the Deterioration and Obsolescence of Product Value, Proc. of CARE INNOVATION 2006, Vienna, Austria, Nov. 2006.

[3] Ernzer M., Lindahl M., Masui K. and Sakao T. An International Study on Utilizing of Design Environmental Methods (DfE) -A pre-study-, Proc. of the Third International Symposium on Environmentally Conscious Design and Inverse Manufacturing, IEEE Computer Society, 2003, pp.124-131.

[4] Narita H., Kawamura H., Norihisa T., Chen L., Fujimoto H., and Hasebe T. Development of Evaluation System of Environmental Burden for Machining Operation (2nd report), Proc. of JSPE 2004 Autumn Conference, Sept. 2004, pp.111-112. (In Japanese)
[5] Akao K. Quality Function Deployment, Productivity Process, Cambridge, M.A., 1990.
[6] Kondoh S., Umeda Y., and Togawa H. Development of redesign method of production system based on QFD, Journal of Advanced Mechanical Design, Systems, and Manufacturing, Vol. 1, No. 1 (2007), pp.181-192.
[7] Sato K., Hotta Y., Nagaoka T., Watari K., Asai M. and Kawasaki S. Mutual Linkage of Particles in a Ceramic Green Body through Photoreactive Organic Binders, Journal of the Ceramic Society of Japan, 113 [10] (2005), pp.687-691.
[8] Omura N., Hotta Y., Sato K., Kinemuchi Y., Kume S. and Watari K., Fabrication of Stable Al2O3 Slurries and Dense Green Bodies Using Wet Jet Milling, Journal of Am. Ceramic Society, 89 [9] (2006), pp.2738-2743.

Understanding the Waste Net: A Method for Waste Elimination Prioritization in Product Development

Marcus V. P. Pessôa [a,1], Warren Seering [b], Eric Rebentisch [c] and Christoph Bauch [d]

[a] Departamento de Controle do Espaço Aéreo – DECEA, Brazil.
[b] Division of Engineering Systems – MIT, EUA.
[c] Lean Advancement Initiative – LAI/MIT, EUA.
[d] Porsche Consulting, Germany.

Abstract. This paper describes a method for prioritization of waste reduction efforts during a Product Development System (PDS) improvement project. Waste refers to activities that absorb resources and increase cost without adding value. Although waste identification and elimination along the value chain has proven useful to PDS improvement, previous work focuses on the waste identification, rather than on its prioritization. This work aims to fill this gap by proposing a "waste coupling" based method to define this sequencing. The method uses a waste set composed of 10 waste types specified into 28 subtypes. The approach is illustrated by an example of possible causal dependencies among subtypes in which both direct and indirect (reinforcement) waste relations are included. This paper's main contributions are: (1) defining a PD specific waste set covering all the PDS elements; (2) describing a procedure to calculate the impact of coupling between wastes through indirect relations; and (3) presenting a heuristic to prioritize from highly coupled waste subsets.

Keywords. waste reduction, product development system, product development process, product development process improvement, lean product development.

1 Introduction

The Product Development System (PDS) is an organizational system that manages both the product portfolio and each individual product development [1]. The PDS is on the interface between the enterprise and the market; it has the duty of identifying (or even anticipating) the market's needs, and proposing how to fulfill these needs [2]. A high performance PDS, therefore, is capable of consistently articulating market opportunities that match the enterprise's competencies [1], and executing the Product Development Process (PDP), guaranteeing that progress is made and value is added by creating useful and timely results [3].

[1] Corresponding Author : Process Engineer, CISCEA, Rio de Janeiro, RJ, Brazil; Tel: +55 (21) 2123 6305; Fax: +55 (21) 2123 6305; Email: mvppessoa@gmail.com; http://www.decea.gov.br/

The PDP itself is a creative, innovative, interdisciplinary, dynamic, highly coupled, massively parallel, iterative, communication based, uncertain, and risky process of intensive planning and activity [4]. Consequently, a wide spectrum of variables can affect its success. The list of product development (PD) project success factors, though, is far from universal, being indeed contingent upon the specific type of project [5].

On "lean terms", low performance is the consequence of waste. In the Toyota Production System (TPS), "waste refers to all elements of production that only increase cost without adding value" [6, p. 54] or "any human activity that absorbs resources but creates no value" [7, p.16]. In fact, the TPS is "a method to thoroughly eliminate waste and enhance productivity" [6, p. 54].

Previous work on PDS improvement through waste elimination [8, 9, 10] focuses on identifying waste throughout the value chain, rather than prioritizing the waste elimination. Attempts at ordering [10, 11] fall short by: (1) not considering waste that occurs in all the PDP elements, notably the external environment; (2) considering crude first-order effects that may fail to address deeper causes of waste in a complex socio-technical system (i.e. PICK charts [12]). This work aims to fill this gap by proposing a method to define a sequencing plan based on the waste coupling. The method uses a waste set composed of 10 waste types specified into 28 subtypes. A cause-consequence map between the subtypes determines which to eliminate first.

The first part of the paper presents the waste set used in this work. In sequence, the proposed method is described through an example. Next, the implications of this work are discussed. Finally, some conclusions and suggestions to future work are presented.

2 Waste in product development

Despite the fact that the waste typology was formulated for manufacturing and in the context of the shop floor, it can be generalized to other domains such as product development, order taking, and the office [7, p. 355; 13, p.28]. In fact, there are several adaptations from the original 7 wastes, most of them varying on the descriptions and including additional wastes to the set. Instead of being a deviation from the original thinking, these changes much more reflect the different expected use of the proposed sets.

The authors decided to use a waste set that is a merging of those mentioned previously [6, 7, 10, 11, 13, 14, 15]. The choice and organization of the waste drivers do not greatly differ from what is presented in the literature. Whenever possible, the original waste nomenclature was maintained, in order to avoid misinterpretations and misunderstandings. The most relevant contribution on the set is the inclusion of "Happenings", as a waste type rooted in the external environment (as identified by [16]). Each of the 10 waste types has subtypes (Figure 1) that better define their scope (see Pessôa [17] for a detailed description of all the presented waste types and subtypes), and can be summarized as:

1. **Overproduction** means producing process outputs at a higher rate or earlier than the next process can use; its subtypes are "unnecessary processes (1.1)"

and “unsynchronized processes (1.2)”. The former includes any forced process output that is not necessary (different from defective) and has to be sorted from useful work. The latter means that the delivered process outputs will not be promptly used because of lack of capacity or because other elements needed in order to proceed are unavailable.

2. **Waiting** is the part of processing time when the creation of value remains static, hence the value stream is considered as ‘non-flowing’ due to the lack of necessary inputs, resources or controls. Waiting can be scheduled during planning when buffers are added or tasks are sequenced instead of done in parallel due to the lack of resources. More visible waiting, though, is the unscheduled waiting, when the execution differs from the planned flow. Waiting subtypes are “scheduled wait (2.1)”, and “unscheduled wait (2.2)”.

3. **Transportation** happens whenever information or materials change ownership or have to overcome structural barriers. Transportation also occurs when information has to be “loaded and unloaded in a person” due to knowledge barriers (need to learn) or to continuity barriers (interruptions, multitasking, etc.). Transportation subtypes are “due to change of ownership (3.1)”, “due to structural barriers (3.2)”, “due to knowledge barriers (3.3)”, and “due to work continuity barriers (3.4)”.

4. **Overprocessing** includes completing unnecessary work during a process. Overprocessing can be divided into “over engineering (4.1)” (beyond what the specifications require), “data conversion (4.2)” (converting data between information systems or between people), and the “re-invention (4.3)” of anything that could be readily reused or adapted.

5. **Inventory** includes raw, in-process or finished build up of information, knowledge, or material such as prototypes that is not being used. Inventories can be found in the company as equipments or data storage, between processes as work in process, or inside the deliverables as excessive information, components, or design options. Inventory subtypes are “in process inventory (5.1)”, “in product inventory (5.2)”, and “in company inventory(5.3)”.

6. **Motion** is any unnecessary movement of people or activity during non-transformation task execution in a process. Motion can be typified as unnecessary human motion due to “bad information systems (6.1)”, “remote locations (6.2)”, and not optimized use of equipments, tools and techniques by not understanding them (due to “complex equipment, tools and techniques (6.3)”).

7. **Defects** consider the creation of defective outputs as a result of the development process. Defects are perceived as “deficient physical deliverables (7.1)”, “deficient information (7.2)”, or information that becomes obsolete while in process (“obsolete deliverables (7.3)”).

8. **Correcting** is the redoing or scrapping, due to feedback. Correcting subtypes are “repairing/reworking (8.1)”, “scrapping (8.2)”, and “inspecting (8.3)” to find problems.

9. **Wishful thinking** means making decisions without the needed inputs (data) or operating according to incorrect controls. Subtypes of wishful thinking include decisions made using “information wrongly perceived to be complete (9.1),” decisions biased by “bounded rationality (9.2)”, or the execution of “poor tests and verifications (9.3)” that do not guarantee the value delivery.

10. **Happenings** include all reactions to unexpected happenings in the environment. Happenings result from failing to forecast the changes in the market and in the business, or from changes of the internal environment (structure, rules, etc.). Its subtypes are “bad forecasting (10.1)”, and “enterprise happenings (10.2)”.

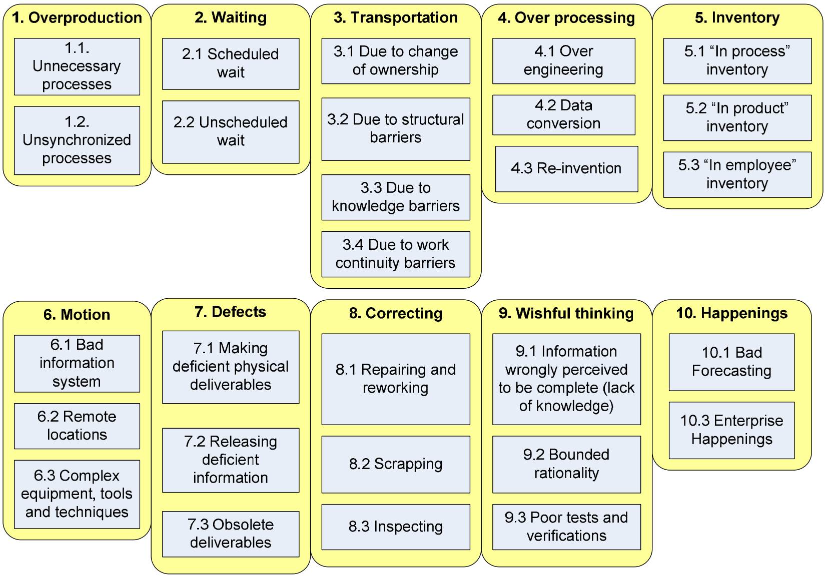


Figure 1. Our waste types and subtypes.

3 The waste prioritization process

Wastes themselves affect each other, creating an intricate net. Understanding this net requires knowing how each of the wastes is influenced by (passive role) or influences (active role) the others in the set. The proposed method models this net for prioritizing the waste reduction, and is composed of three steps:

(1) mapping the direct causes;

(2) calculating the impact of coupling; and

(3) sorting according to the defined priorities.

The authors emphasize that the proposed method is not a VSMA technique, since it does not deal with process mapping and waste plotting. The method complements the VSMA by helping to prioritize which wastes to attack first, in order to have a more effective improvement. It goes beyond the commonly used PICK charts [12], which is four categories 2x2 matrix (Possible, Implement, Challenge and Kill). Although similar to the research presented by Bauch [11], this work does a crisp true/false direct mapping to calculate the indirect reinforcement, instead of grading the relationship through a multi-leveled criteria (high influence, medium influence, etc.). Crisp values create less ambiguity while filling the matrix [10]. Note that the coupled active and passive sums (Table 1) will provide the necessary mechanism to discern the impact on/from the waste subtypes.

Proposing a final and universal mapping among the waste subtypes is beyond the scope of this paper, Table 1 shows an example of a possible mapping, which will be used while explaining the method's steps and discussing the results. Consequently, this paper presents a generic process that can be adopted for PDS improvement, where individuals or organizations should do this mapping based on local conditions and constraints.

Table 1. Example of wastes interrelationships

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	Direct Passive Sum	Total Passive Sum
1.1	1																								1	1			1	3	12
1.2	2	1			1	1	1	1	1	1			1			1	1	1				1	1	1	1	1			1	17	63
2.1	3																								1	1				2	11
2.2	4		1																											1	18
3.1	5																												1	1	1
3.2	6														1														1	2	3
3.3	7																												1	1	1
3.4	8																												1	1	1
4.1	9															1									1	1				3	13
4.2	10																												1	1	1
4.3	11															1	1								1	1				4	16
5.1	12	1	1																											2	22
5.3	13																								1	1				2	11
5.4	14																												1	1	1
6.1	15																												1	1	1
6.2	16														1														1	2	3
6.3	17																												1	1	1
7.1	18											1				1		1							1	1	1			6	27
7.2	19									1	1	1				1	1	1							1	1	1			9	36
7.3	20			1	1																									2	5
8.1	21																		1	1	1									3	20
8.2	22	1																	1	1	1									4	24
8.3	23	1				1	1	1	1	1	1								1	1	1	1					1			12	50
9.1	24	1	1													1	1	1		1							1	1		8	47
9.2	25																								1					1	9
9.3	26													1				1		1	1				1	1				6	29
10.1	27																													0	0
10.2	28																													0	0
Direct Active Sum		5	3	1	2	2	2	2	2	3	2	2	1	1	2	6	4	5	3	5	4	2	1	1	10	9	4	1	11		
Total Active Sum		21	16	5	9	6	6	6	6	12	8	10	4	5	8	32	24	30	7	23	12	6	4	4	41	31	23	11	41		

3.1 Mapping the direct causes

This step captures and quantifies the direct causal relation among the waste subtypes themselves. A particular subtype was considered a direct cause to another when the occurrence of one (cause) triggers the other (consequence).

As shown in Table 1, the authors determined themselves the dependencies between waste subtypes. They mapped the causal relations into a square (28x28) design structure matrix (DSM). For a formal introduction to DSM refer to Steward [18] and Yassine [19]. The identically labeled rows and columns of this matrix name the 28 waste subtypes, each cell of the DSM received the values "0" (empty) meaning "not directly related" or "1", meaning "directly related" respectively.

Hence, “unscheduled wait” is a consequence of “pushing unsynchronized work”, but is not directly related to the several wastes that are themselves causes of “pushing unsynchronized work”.

The degree of how a particular waste might influence and might be influenced by the others is captured by the direct active and direct passive sums, respectively. “Pushing not useful work”, for instance, is influenced by three wastes (direct passive sum) and influences five other wastes (direct active sum). Direct active sums are shown at the bottom of the table, and direct passive sums are shown on the right. The total active and passive sums are explained in the next section.

3.2 Calculating the impact of coupling

Unfortunately, the direct sums create an unreal picture, because two wastes with the same direct passive sum can not necessarily be considered equally easy to eliminate, nor can two wastes with the same direct active sum be considered equally important to be attacked. With the example, both “unscheduled wait” and “due to change of ownership” have the same direct passive sum (equal to 1) but the cause (“pushing unsynchronized work”) to the former has itself much more causes than the cause (“enterprise happenings”) to the latter, making “unscheduled wait” much more difficult to eliminate. Therefore, the indirect relations should be taken into account.

To address this limitation, the method calculates the “total sums” which take into account the coupling. The total passive sum for each waste equals the number of wastes that directly cause the waste in question plus 0.7 times the sum of the numbers of wastes that cause the direct causes of the waste in question. For example, for the case of “unscheduled wait,” there is just one direct cause, “pushing unsynchronous work.” But “pushing unsynchronous work” has 17 direct causes. So the total passive sum for “unscheduled wait” will be 1 + (0.7 * 17) or 12.9. “Unscheduled wait” causes “pushing unsynchronized work” and “obsolete deliverables.” So the total active sum for “unscheduled wait” is 2 + [0.7 * (3 + 4)] or 6.9.

The factor 0.7 is arbitrary, but allows for accounting of second order effects of related wastes in the complex interdependent PDS (interdependency impacts beyond second order are not assessed in this approach.) The factor 0.7 was chosen by the authors after comparing the results using candidate factors from 0.00 to 1.00 (0.01 steps). With this factor, all active reinforcement is taken into account, and most of the passive reinforcement is considered (leaving only the influence on a couple of subtypes that swap positions with factors close to 1.00).

3.3 Sorting to define the priorities

A standard algorithm [19] was employed to lower triangularize the DSM as a first step to order the causal relations. After the operation, the first two rows held focal wastes. That is, the wastes in the first two rows were not caused by any other wastes, and are identified as being in Band #1. It is proposed here that initial efforts be made to eliminate the wastes in band one. If these can be eliminated, then the next six wastes in the DSM (Band 2) become focal wastes. The wastes clustered at the bottom of the DSM are tightly coupled and consequently can not be lower triangularized.

To achieve the final ordering, the wastes within a band or inside a cluster have been prioritized according to two heuristics. First they are prioritized by ascending total passive sum so that those which are caused by the fewest others are addressed first. These wastes are least interdependent with other wastes and therefore are assumed to be eliminated with lesser difficulty. When adjacent sets of wastes have equal total passive sums, these wastes are then sorted by descending total active sum because the elimination of the more active ones will provide higher return by reducing the influence on several others. Table 2 presents the final waste reduction sequence after partitioning the DSM and applying the two criteria.

Table 2. Final waste sequence and the impact after eliminating Bands #1, #2, and #3.

		Final sequence			Band#1 elimination				Band#2 elimination				Band#3 elimination			
#		Waste Subtype	AS	PS	New Waste Sequence	AS	PS	Red. Passivity	New Waste Sequence	AS	PS	Red. Passivity	New Waste Sequence	AS	PS	Red. Passivity
1	Band #1	**10.2**	32.0	0	**10.2**	0.0	0		**10.2**	0	0		**10.2**	0.0	0	
2	Band #1	**10.1**	8.0	0	**10.1**	0.0	0		**10.1**	0	0		**10.1**	0.0	0	
3	Band #2	**6.1**	24.2	1	**6.1**	24.2	0	100%	**6.1**	0	0		**6.1**	0	0	
4	Band #2	**6.3**	22.5	1	**6.3**	22.5	0	100%	**6.3**	0	0		**6.3**	0.0	0	
5	Band #2	**4.2**	6.2	1	**4.2**	6.2	0	100%	**4.2**	0	0		**4.2**	0.0	0	
6	Band #2	**5.4**	6.2	1	**5.4**	6.2	0	100%	**5.4**	0	0		**5.4**	0.0	0	
7	Band #2	**3.1**	4.8	1	**3.1**	4.8	0	100%	**3.1**	0	0		**3.1**	0.0	0	
8	Band #2	**3.3**	4.8	1	**3.3**	4.8	0	100%	**3.3**	0	0		**3.3**	0.0	0	
9	Band #2	**3.4**	4.8	1	**3.4**	4.8	0	100%	**3.4**	0	0		**3.4**	0.0	0	
10	Band #3	**6.2**	18.0	2.7	**6.2**	18.0	1	63%	**6.2**	18.0	0	100%	**6.2**	0.0	0	
11	Band #3	**3.2**	4.8	2.7	**3.2**	4.8	1	63%	**3.2**	4.8	0	100%	**3.2**	0.0	0	
12	Cluster	**7.3**	9.6	4.1	**7.3**	9.6	4.1	0%	**7.3**	9.6	4.1	0%	**9.2**	24.4	3.8	42%
13	Cluster	**9.2**	24.4	6.6	**9.2**	24.4	5.9	11%	**9.2**	24.4	4.5	32%	**7.3**	9.6	4.1	0%
14	Cluster	**2.1**	3.8	8.3	**1.1**	16.2	7.6	18%	**1.1**	16.2	6.2	25%	**1.1**	16.2	5.5	41%
15	Cluster	**5.3**	3.8	8.3	**2.1**	3.8	7.6	8%	**4.1**	9.3	6.2	38%	**4.1**	9.3	5.5	34%
16	Cluster	**1.1**	16.2	9.3	**5.3**	3.8	7.6	8%	**2.1**	3.8	6.2	33%	**4.3**	7.6	5.5	56%
17	Cluster	**4.1**	9.3	10	**4.1**	9.3	8.6	14%	**5.3**	3.8	6.2	25%	**2.1**	3.8	5.5	34%
18	Cluster	**4.3**	7.6	12.4	**4.3**	7.6	10.3	17%	**4.3**	7.6	7.2	42%	**5.3**	3.8	5.5	45%
19	Cluster	**2.2**	6.9	12.9	**2.2**	6.9	12.2	5%	**2.2**	6.9	8.7	33%	**2.2**	6.9	7.3	43%
20	Cluster	**8.1**	4.8	14.9	**5.1**	3.1	14.6	9%	**5.1**	3.1	11.1	26%	**5.1**	3.1	9.7	39%
21	Cluster	**5.1**	3.1	16	**8.1**	4.8	14.9	0%	**8.1**	4.8	11.4	29%	**8.1**	4.8	10.7	28%
22	Cluster	**8.2**	3.1	18	**8.2**	3.1	17.3	4%	**7.1**	5.8	13.8	33%	**7.1**	5.8	12.4	31%
23	Cluster	**7.1**	5.8	20.7	**7.1**	5.8	18.6	10%	**8.2**	3.1	13.8	23%	**8.2**	3.1	13.1	37%
24	Cluster	**9.3**	17.3	22.1	**9.3**	17.3	20.7	6%	**9.3**	17.3	16.2	27%	**7.2**	17.6	14.8	47%
25	Cluster	**7.2**	17.6	27.9	**7.2**	17.6	24.4	13%	**7.2**	17.6	17.2	38%	**9.3**	17.3	14.8	33%
26	Cluster	**9.1**	31.7	35.3	**9.1**	31.7	30.8	13%	**9.1**	31.7	21.8	38%	**9.1**	31.7	18.7	47%
27	Cluster	**8.3**	3.1	38.6	**8.3**	3.1	34.4	11%	**8.3**	3.1	24.8	36%	**8.3**	3.1	23.1	40%
28	Cluster	**1.2**	12.1	49.2	**1.2**	12.1	41.9	15%	**1.2**	12.1	30.6	38%	**1.2**	12.1	27.2	45%

4 Discussion

The objective of this section is to make a step-by-step analysis of the results, by simulating the waste elimination, analyzing the changes on the waste net, and verifying the soundness of the process. The considered steps were:

(1) Elimination of the waste subtypes on the first band;

(2) Elimination of the waste subtypes on the second band;

(3) Elimination of the waste subtypes on the third band; and analysis of the strategies to deal with the cluster.

4.1 First band elimination

The first band included the waste subtypes that have no causes within the waste set, and represent the environment's influence on the system. Their existence and relevance document the difficulty of having a waste-free PDS. Market and business risks are difficult to predict and are the causes of "bad forecasting". It is unlikely that the first band wastes will be completely eliminated, though it should be possible to reduce them. After having put forth a good effort to eliminate these wastes, it is reasonable to proceed to the wastes in the next band.

For the sake of the example, the elimination of Band #1 is assumed, causing most of the remaining subtypes to perceive some reduction on their passivity (100% on Band #2) (Table 2). The priority of some wastes also changed due to the new total passive sum.

4.2 Second band elimination

Continuing the improvement, the elimination of Band #2 promotes changes not only to the total passive sum, but also to the total active sum of some of the remaining waste subtypes. Similarly to the previous step, the subtypes on Band #3 had 100% of reduction on their passivity (see Table 2).

4.3 Third band elimination and dealing with the cluster

Since Band #3 was the last band before the cluster, its elimination did not reduce any of the remaining total passive sums to "0" (see Table 2). A cluster characteristic is that its elements are coupled, thus assumptions have to be made about which subtype to consider first. Two approaches can be pursued: first working on the more active wastes and expecting a higher impact to result from the improvement process, or first working on the less passive wastes and expecting a more stable improvement (less risk on recurring symptoms). Note on Table 2 that the waste subtype with higher total active sum ("information wrongly perceived to be complete") also have an expressive total passive sum, showing that it is influenced by several others. For these cases, we have chosen to employ the second approach, according to the previously presented heuristics.

5 Conclusion

The method presented in this work provides a systematic approach to prioritizing waste reduction during the improvement of a product development system (PDS). This paper has three contributions. First, it defines a PD-specific waste set covering all the PDS elements. Second, it describes a procedure to calculate the coupling between wastes through indirect relations. Finally, it presents heuristics to prioritize from highly coupled waste subsets.

The proposed method can be summarized in three steps: 1) map the direct causal relation between wastes, 2) calculate the impact of coupling, and 3) sort according to the defined priorities. This method is particularly applicable to improvement projects after the current state of the value stream has been drawn and the waste occurrences identified, and when the improvement team is defining the strategy to move to a future state.

By studying an example of a possible mapping between the waste subtypes, useful insight was gained about the implications of direct and indirect relations to the definition of a PDS improvement strategy. However, these findings cannot be generalized before completing a comprehensive survey on the subtypes' relationship.

The type of analysis illustrated here may outline the study of other issues related to the coupling of product development waste subtypes and its root causes on particular business environments and development projects. A challenge for future research is to extend this method to explore the consequences of acting on these specific causes and the expected implications on the waste net.

This study uses an example with connections defined by the authors' discretion. Another future challenge is to capture an empirically grounded linking from lean thinking experts and from practitioners, and change agents, in order to both validate a general model, as well as assess its usability in PDS change interventions.

6 References

[1] Cheng, L. C., 2003, "QFD em Desenvolvimento de Produto: Características Metodológicas e um Guia para Intervenção," Revista Produção Online, Florianópolis, v. 3, n. 2.

[2] Rozenfeld, H. et al, 2006, "Gestão de Desenvolvimento de Produtos," São Paulo, Editora Saraiva.

[3] Murman, E. et al., 2002, "Lean Enterprise Value: Insights," New York, Polgrave, MIT's LAI.

[4] Negele, H. et al., 1999, "Modelling of Integrated Product Development Processes," In: Annual Symposium of Incose, 9th., Proceedings…[S.l.]: [s.n.].

[5] Dvir, D., Lipovetsky, S., and Shenhar, A. J., 2003, "What is really important for project success?," International Journal of Management and decision Making, v.4, n. 4, p. 382–404.

[6] Ohno, T., 1998, "Toyota production system," New York: Productivity Press.

[7] Womack, J. P., and Jones, D. T., 2003, "Lean Thinking," New York, Free Press.

[8] Millard, R. L., 2001, "Value Stream Analysis and Mapping for Product Development," Thesis (S.M.) - Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA.

[9] Mcmanus, H., 2004, "Product Development Value Stream Mapping (PDVSM) Manual," Beta draft, Cambridge: MIT Lean Aerospace Initiative.

[10] Kato, J., 2005, "Development of a Process for Continuous Creation of Lean Value in Product Development Organizations," Thesis (Master), Massachusetts Institute of Technology, Cambridge.

[11] Bauch, C., 2004, “Lean Product Development: Making waste transparent,” Thesis (Diploma) – Massachusetts Institute of Technology, Cambridge.
[12] Michael. L. G., 2006, “Lean Six Sigma for Services,” New York, McGrawHil.
[13] Liker, J. K., 2004, “The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer,” New York, McGraw-Hill.
[14] Morgan, J. M., and Liker J. K., 2006, “The Toyota product development system,” New York, Productivity Press.
[15] Ward, A., 2007, “Lean product and process development,” Cambridge, The Lean Enterprise Institute.
[16] Gershenfeld, J. and Rebentisch, E., 2004, “The Impact of Instability on Complex Social and Technical Systems,” Proceedings from MIT Engineering Systems Division External Symposium
[17] Pessôa, M. V. P., 2008, “Weaving the waste net: a model to the product development system low performance drivers and its causes,” Lean Aerospace Initiative Report WP08-01, MIT, Cambridge, MA.
[18] Steward, D.V., 1981, “The Design Structure System: A Method for Managing the Design of Complex Systems,” IEEE Transactions on Engineering Management, Vol. 28, pp. 71-74.
[19] Yassine, A., 2004, “An Introduction to Modeling and Analyzing Complex Product Development Processes Using the Design Structure Matrix (DSM) Method,” Quaderni di Management, n.9, (Italian Management Review), www.quaderni-di-management.it .

The Green Product Eco-design Approach and System Complying with Energy Using Products (EuP) Directive

Amy J.C. Trappey[a, b, 1], Meng-Yu Chen[b], David W. Hsiao[b] and Gilbert Y.P. Lin[b]

[a]Department of Industrial Engineering and Management, National Taipei University of Technology, Taiwan
[b]Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan

Abstract. Owing to the enforcement and implementation of green environmental directives, e.g., the Restriction of Hazardous Substances (RoHS), Waste Electronic and Electrical Equipment (WEEE), and Energy using Products (EuP) in European Union, the "green" concept is fast becoming the global market trend and the base of industrial development. Consequently, green product design capabilities become the key competence for enterprises to stay globally competitive. Due to the shortened life span of consumer products, information-based tools or methods are critical to support enterprise R&D in the stages of ecologically sound product design. In this paper, an integrated green product design methodology and system are developed to support environmental conscious, energy-using products development. The proposed methodology adopts LCA, QFDE, theory of TRIZ and the projections of green conceptual design improvement based on ANN approach. LCA and QFDE are developed for assessing and comparing the environmental impacts of (human) activities for the production and provision of goods and services. TRIZ is used to support R&D to create innovative product design ideas effectively and efficiently during the concept design stage. An ANN model is developed for the R&D decision support of environmental conscious product design. Finally, this research will present a green conceptual design case study applying the proposed eco-design methodologies and system.

Keywords. Eco-design Requirements for Energy Using Products (EuP), Eco-design, Life Cycle Assessment (LCA), Quality Function Deployment for Environment (QFDE), Theory of Inventive Problem Solving (TRIZ), Back Propagation Network (BPN).

1 Introduction

Nowadays, enterprises face great challenges to maintain its competitiveness in the global marketplace. The green product design capabilities become very critical for companies to stay competitive globally. In the EU green directives, RoHS and

[1] Corresponding Author : Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei (10608), Taiwan, E-mail: trappey@ntut.edu.tw. Tel: +886-2-2771-2171 Ext. 4541; Fax: +886-2-2776-3996

WEEE are specific regulations developed for specific purposes of environmental concerns. The EuP directive is developed for the requirement of eco-design, which means EuP aims at integrating environmental consciousness into product design, development, production and distribution. Because of the strict environmental regulations and directives and the short lifespan of consumer products, enterprises need IT-based tools or methods to effectively and efficiently support R&D during the stages of product conceptual designs. In this research, an integrated green product design methodology is proposed to support environmental consciousness of energy-using product design. This research focuses on the following key tasks. Firstly, we will construct the green product/part information management platform to enhance the functions of data collection and management in a PLM system. Meantime, a kernel decision support system for green innovative eco-design is built by integrating product design processes and incorporating Life Cycle Assessment (LCA), Quality Function Deployment for Environment (QFDE), Theory of Inventive Problem Solving (TRIZ) and artificial neural network (ANN) approaches. Based on the proposed methods and models, an information system is constructed. To verify the feasibility of the proposed methodology and information system, a case study of the electric adaptor eco-design is discussed in the paper. Finally, the outcomes of this research are summarized in conclusion.

2 Literature Review

When talking about EuP directive, we should consider the requirement of eco-design which integrates environmental aspects into product design and development. In this research, we will integrate LCA, benchmarking, checklist and QFDE methods suggested by ISO14062 [5], as well as TRIZ and ANN methods, to develop a unique eco-design methodology.

EuP is a framework directive, which specifies not only energy consumption but also environmental performance extending from Integrated Product Policy (IPP) [10]. The EuP framework directive helps design of energy using products complying with eco-design requirements. Furthermore, EuP framework directive is the latest announced directive of EU targeting at electronic products, e.g., personal computer and liquid crystal display (lot 3), consumer electronics (lot 5) and power supplies (lot 7). EuP causes great impacts on the countries and regions with huge electrical and electronic equipment industries, e.g., Taiwan, China, Korea and Japan [3].

LCA is the investigation and valuation of the environmental impacts of a given product or service caused or necessitated by its existence. Rebitzer et al. [14] introduced the LCA framework which outlines the procedures to define and model a product’s life cycle. They also provide an overview of available methods and tools for tabulating and compiling associated emissions and resource consumption data in a life cycle inventory (LCI). Pennington et al. [11] highlighted the key features, summarized available approaches, and outlined the key challenges of assessing the aforementioned inventory data in terms of contributions to the environmental impacts. This process is named as life cycle impact assessment (LCIA). LCA consisting of LCI and LCIA is defined in ISO14040 series [6][7].

QFD used to analyze functions required for a product or the product structure to realize these functions is a design tool applicable to the early stage of product developments [9]. Zhang et al. [17] developed a method called green quality function deployment II (GQFD-II), which integrates LCA into QFD. As a result, when evaluating environmental aspects of a product using GQFD-II, qualitative LCA must be performed. According to previous research about QFD, it is concluded that the method combining QFD and LCA is named QFDE.

The TRIZ methodology constructed in the former Soviet Union by Altshuller [1] [15] is a well-structured inventive problem-solving process. The application of TRIZ thinking tools is used in the industries successfully and replaces the unsystematic trial-and-error method in the search for solutions [8]. For most people, the most commonly applied tool is the contradiction matrix, which is composed of 39 engineering parameters and 40 inventive principles. However, some other TRIZ tools (e.g., Su-Field analysis with 76 standard solutions) are considered more useful than the contradiction matrix with 40 inventive principles for achieving breakthrough innovation [4].

3 System Framework and Green Product Eco-design Approach

In this section, the methodology is described in four sub-sections to clarify the proposed idea. Moreover, the functional modules of the design decision support system platform are also depicted.

3.1 Functional modules of the platform

In this section, the modules of proposed system platform are defined. As shown in Figure 1, the system framework is divided into three key modules, including the green information management module, the eco-design approach module and the design assessment and decision support module. The major concept of the proposed platform is to integrate environmental aspects into product design and development and conduct the design and evaluation with efficiency and efficacy.

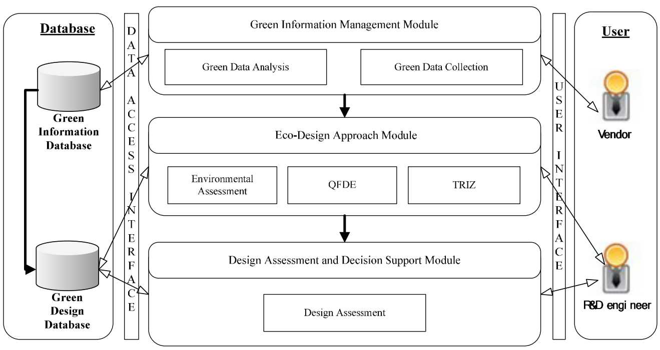


Figure 1. The framework of the eco-design platform with three key modules

The modules in the information platform are discussed in detail as follows. Green information management module is built to gather green information for the use of part and component in our product and also collect green information of product during its life cycle. Eco-design approach module consists of green data analysis, QFDE method and the TRIZ method to construct the proposed environmental design methodology. The design assessment and design decision support module is used for measuring product environmental efficiency and to improve design based on ANN. In the following sections, we will discuss all methods and modules incorporated and integrated in this research, which are the bases of constructing the environmental design information platform.

3.2 Green Information Management Module

In this section, we discuss the process of this module based on LCA method used for measuring and comparing the environmental impacts of human activities for the provision of goods and services. There are three steps in the process of this module. Firstly, vendors fill green information of part/component by accessing the web-based interface of proposed system. This step is named LCI defined in ISO14040 series for environmental data collection using the green data collection function of this module. Meantime, R&D staffs also use the same interface to fill green information of product. Finally, green data analysis function is used for calculating environmental impact and the process named as LCIA in ISO14040 series. After completing the green information collection process of this module, we can get the environmental performance information of current product design and use the result to design new and improved product(s). In the following sections, we will discuss how to use the analysis result in eco-design.

3.3 Eco-design Approach Module

In this section, an eco-design approach is developed to help R&D staffs design products complying with environmental regulations. Park et al. [12] proposed an integrated environmental assessment method composed of LCA, benchmarking, and checklist is proposed to search the environmental weak points of current product design. This step is called the bottom-up environmental assessment method. A product life cycle is divided into five stages, i.e., raw material acquisition, manufacturing, distribution, use and disposal. During the key life cycle stage (i.e., the manufacturing stage), the checklist approach is used to find weak points of present products. In other stages, we use environmental benchmarking method. In Step 1, the parameters of benchmarking and the items of a checklist are defined using inter-enterprise brainstorming based on the product environmental report. Thus, there are different benchmarking parameters and checklist items in specific industry. In Step 2, we modify the method proposed by Park et al. [12] and use the advised eco-design strategies in Annex 1 of EuP directive [3]. The eco-design strategies are set as the voice of environment (VOE), and QFD technique is used to find the relation between eco-design strategies and environmental weak points from present product and calculates the weights on VOE. The voice of customer (VOC) and weights are analyzed based on the market survey.

Sakao [16] proposed a QFD-centred design methodology for environmentally conscious product design. In this research, we use QFDE to find the design improving factors in Steps 3. After design improving factors are found in the process, we start to discuss how to use design improving factors to develop the conceptual product design. According to Ahn [2], contradiction matrix and innovative principles are used to develop the conceptual design. And Chang [4] indicates that Su-Field method is useful to analyze and solve product design problems. In this research, we integrate the two main methods of TRIZ theory to construct the process of innovate creation of product design. Step 4 analyzes design improving factors for realizing the relations between engineering parameters and parts specifications and transfers the real design problem to the generic problem. TRIZ theory such as contradiction matrix and Su-Field are used to solve the problem. When solving the technical contradiction problem, we choose contradiction matrix with 40 inventive principles to solve the problem. Nevertheless, Su-Field analysis method is suitable for modeling the problem and the 76 standard solutions can be applied to solve the problems. In the proposed system, the two problem solving methods are constructed for design improvement. Finally, the results of finding design improving factors and innovative creation design are transformed into the real solution for new product design concepts. Section 3.4 describes the decision support module for selecting the best eco-design.

3.4 Design Assessment and Decision Support Module

This section explains the intelligent design decision support using BPN approach, an ANN method. The BPN algorithm is a multi-layered decision inference network [13]. The algorithm is also called a supervised learning method that can be used to solve non-linear problems. As described in the previous section, the green information of previous products is collected and analyzed by green information management module and the result of analysis is stored in the database for design assessment and decision support module. In this research, we use the high level properties of electronic products, e.g., the size of a printed circuit board (PCB), the amount of materials, the number of capacitors, the number of inductors and the number of chips and so on, as input nodes. Their impact value is considered as the output node to construct the BPN decision support module. Afterward, we use the defined attribute values as input to estimate the result of new product design concepts created by the eco-design approach module. In this paper, we will use the methodology to help R&D staff to decide which design has much better environmental performance in product concept design stage. In following section, we will demonstrate a case study of adaptor using the proposed methodology for an eco-product concept design.

4 Case Example and Experiment

According to EuP directive, an AC (or called switching) adaptor is usually a must-have component of major electronic products. However, the deficient design of an AC adaptor, not complying with EuP guidelines, may potentially cause great impacts on the countries and regions with electrical and electronic equipment industries. Thus, this research uses adaptor as the case study to verify and demonstrate the proposed eco-design approach and the information platform.

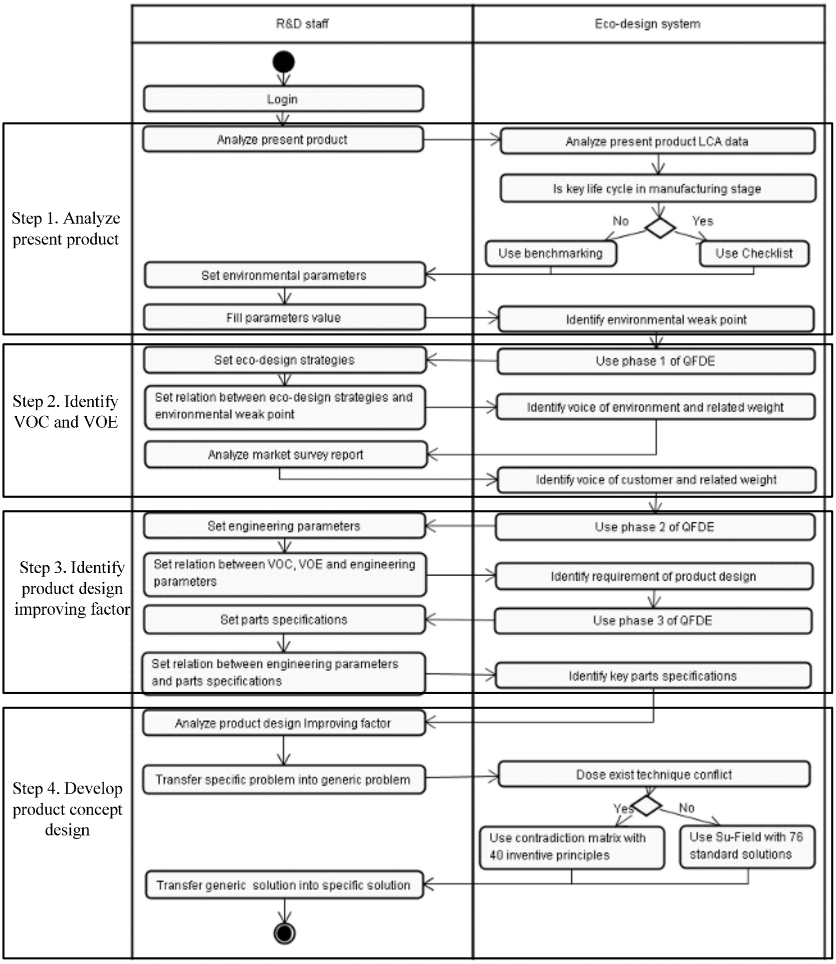


Figure 2. UML activity diagram shows application processes of the proposed eco-design approach

During the product deign processes, R&D staffs can get assistances from the proposed information system embedded with proposed green design methodologies in four steps, i.e., analyze present product, identify VOC and VOE, identify critical design improving factors, and develop product design concept. Firstly, the green information data is collected by the green information management module.

Afterward, the product design concept is generated by the eco-design approach module (from Step 1 to Step 4). Finally, the design assessment and decision support module help R&D staff select design concepts, which meet the actual market requirements. The detailed UML activity diagram used to model the process of the proposed eco-design approach is shown in Figure 2.

4.1 Analyze present product

After analyzing the LCA results of the current (as-is) AC adaptor design from the green information management module (depicted in Step 1 of Figure 2), the environmental impacts of the raw material acquisition stage is 52.3% of total environmental impact for the AC adaptor. Furthermore, we also found that the weight value of the raw material acquisition stage is the highest, since the use stage is 43.6%, the manufacturing stage is 2.1%, the disposal stage is 1.8%, and the distribution stage is 0.2%. Thus, the raw material acquisition stage is considered to be the critical life cycle stage of the AC adaptor. Afterward, benchmarking was conducted based on six parameters to analyze the differences between the AC adaptor and the competitor's products. In Table 1, we identify the environmental weak points and their corresponding weights for the phase 1 deployment of QFDE for the environmental benchmarking.

Table 1. Identify environmental weak points

Product / BP	Our AC adapter	Competitor 1	Competitor 2	Improvement target (%)	Relative weight
Number of parts			V	24	24/122=0.19
Number of kinds of material		v	V	14	14/122=0.11
Weight of a product			V	13	13/122=0.11
Weight of packing material			V	12	12/122=0.1
Weight/length of cable			V	51	51/122=0.42
Surface area of PCB		v		8	8/122=0.07
	0	2	5	122	1
	Total number of mark			Total	

4.2 Identify VOC and VOE

In this section, the voice of customer and the voice of environment are identified (in Step 2 of Figure 2). To better understand the requirements from the target customers, market research is conducted and found that "less noise" and "reduce electromagnetic wave" are two major concerns from the target customers. Consequently, these two parameters are considered the major voices of customer. Moreover, the weights are also determined based on VOC (Table 2).

Table 2. Identify voice of customer (VOC)

Voice of customer (VOC)	Corresponging weight
Less noise	0.5
Reduce electromagnetic wave	0.5

After identifying the environmental weak points, we combine bottom-up and top-down eco-design method using QFDE shown in Table 3. The phase one (I) deployment of QFDE is used to calculate the relations between environmental weak points and eco-design strategies, which are derived from eco-design parameters in Annex 1 of EuP directive. As shown in Table 3, the voices of environment and the corresponding weights are identified for the phase two (II) deployment of QFDE.

Table 3. Identify the voice of environment (VOE)

QFDE (I)	weighting	Less resource consumption	Less energy consumption	Less weight and volume	Less waste and emission	Easy for reuse and recycling	Easy to process and assemble	Easy to disassemble	Safe emission
Number of parts	0.19	9	5	5	9	9	9	9	3
Number of kinds of material	0.11	3	3	3	3	9	3	3	3
Weight of a product	0.11	9	3	9	3	3	3	3	3
Weight of packing material	0.1	9	3	9	9	3	3	3	3
Weight/length of cable	0.42	9	9	9	9	9	3	3	3
Surface area of PCB	0.07	9	3	9	9	9	3	3	3
Row Score (RS)		8.34	6.32	7.58	7.68	4.8	4.14	4.14	3
Relative weight (Ratio of RS/ Total RS)		0.17	0.11	0.16	0.16	0.16	0.09	0.09	0.06

4.3 Identify product design improving factors

In this section, we discuss how to identify product improving factors from the present product (in Step 3 of Figure 2). Firstly, engineering parameters are identified to ensure the requirements of environment and customer are met. Afterward, parts specifications are identified by product analysis to detail the improving factors. Among the identified parameters in the previous section, eight parameters belong to VOE fields and two parameters belong to VOC fields. In Table 4, we calculate the relations between the VOEs and the VOCs with engineering parameters based on the phase two (II) deployment of QFDE. The process identifies the importance of engineering parameters in the case. Finally,

Table 5 shows the key parts/components reflecting the 10 design improving factors for new eco-design in the phase three (III) deployment of QFDE.

Table 4. Phase two (II) of QFDE

QFDE (II)	weighting	Energy consumption	Weight	Volume	Vibration	Electromagnetic wave
Less resource	0.17*0.8=0.14	3	9	9	3	3
Less energy consumption	0.11*0.8=0.08	9	3	3	3	5
Less weight and volume	0.16*0.8=0.13	3	9	9	5	5
Less waste and emission	0.16*0.8=0.13	3	5	5	5	5
Easy for reuse and recycling	0.16*0.8=0.13	3	5	5	5	5
Easy to process and assemble	0.09*0.8=0.07	3	5	5	5	5
Easy to disassemble	0.09*0.8=0.07	3	5	5	5	5
Safe emission	0.06*0.8=0.05	3	3	3	9	5
Less noise	0.5*0.2=0.10	3	3	3	9	3
Reduce electromagnetic wave	0.5*0.2=0.10	3	3	3	3	9
Row Score (RS)		3.48	5.42	5.42	4.96	4.92
Relative weight (Ratio of RS/ Total RS)		0.14	0.23	0.23	0.2	0.2

Table 5. Phase three (III) of QFDE

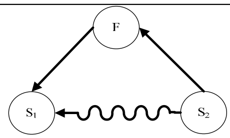
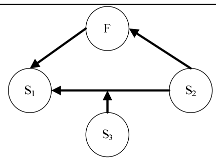
QFDE (III)	weighting	Capacitor	Diode	IC	**Transformer**	Mechanical parts	PCB	Resister
Energy consumption	0.14	9	9	9	**9**	3	3	9
Weight	0.23	3	3	3	**9**	5	3	3
Volume	0.23	3	3	3	**9**	9	9	3
Vibration	0.20	3	3	3	**9**	5	3	3
Electromagnetic wave	0.20	3	3	3	**9**	9	3	3
Row Score (RS)		3.84	3.84	3.84	**9**	6.44	4.38	3.84
Relative weight (Ratio of RS/ Total RS)		0.11	0.11	0.11	**0.26**	0.18	0.12	0.11

4.4 Develop product concept design

In Step 4 of Figure 2**,** after analyzing the as-is AC adaptor in the phase III of QFDE (in Table 5), the transformer, with the highest row score, is the most critical component for an adaptor re-design. By improving the transformer design, we can both meet the requirements of customer specifications and increase the adapter's environmental performance. In this case study, it is desired to eliminate the vibration of a transformer for noise reduction. However, the direction of design improvement is still unknown in this stage. Thus, Su-Field analysis of TRIZ approach is applied to help us identify the direction of design improvement.

In our case study, the problem can be disclosed via the construction of the initial Su-Field model (ISM). In the ISM, S_1 is a core made of stainless steel, S_2 is the coil of the transformer and F is the electric and magnetic fields. Owing to the change of electric and magnetic field according to frequency alternation, the "hum" sound caused by the change of electric and magnetic field of the transformer core is produced. In the ISM, the harmful action is caused by a field, which is generated by S_2 to influence S_1. We analyze the problem using Su-Field analysis with 76 standard solutions. Afterward, the standard solution 1.2.3 is proposed to solve the problem. An element S_3 is introduced to absorb the harmful effects from the 76 standard solutions. Afterward, we use the proposed solution to construct the desired Su-Field mode (DSM) shown in Table 6. According to the solution from Su-Field analysis, we use S_3 varnish to absorb the harmful action caused by electric and magnetic fields.

Table 6. The Su-Field Analysis of the adapter case study

	ISM (Initial Su-Field Model)	DSM (Desired Su-Field Mode)
Su-Field model		
Content	The harmful action is caused by a field.	The harmful action is caused by a field. Introduce an element S_3 to absorb the harmful effects.

After using Su-Field method of TRIZ, multiple design concepts are developed to create and improve new eco-designs. In the final phase of our proposed methodology, a design decision support module based on the back-propagation neural network (BPN) method is used to evaluate the environmental performance and help choose the final eco-design direction.

5 Conclusion and Future Work

In this research, we developed a methodology and system for eco-design including the green information management module (green data collection and analysis), eco-design approach module (integrating multiple design methods), and design assessment and decision support module (evaluating the eco-design concepts). The proposed methodology embedded in the proposed information platform can be summarized in six steps. Green information collection, including product, part, component and materials from vendors, is analyzed using LCA to find the key life cycle stages of product for eco-design improvement. After analyzing the key life cycle stages of product, benchmarking and checklist is used to find the environmental weakness points from the present product. If the key life cycle stage is the "manufacturing" stage, the checklist method is adopted. Contrarily, if the key life cycle stage is identified as any other stage, the benchmarking method is used to find the environmental weakness points. Moreover, a systematic eco-design method is developed to combine the top-down approach from the analysis result of key life cycle stage and the bottom-up approach from the eco-design strategies to integrate environmental consciousness into the design process. Meanwhile, QFDE is a powerful method to identify the priority of design improving factors for environmental conscious design. Furthermore, TRIZ is an innovative problem solving approach deployed for creating eco-design improving concept. Finally, an intelligent decision support system based on BPN is built to support R&D staff to evaluate the performance of design improving concept for the final design improvement.

The major contributions of this research consist of integrating commonly applied design approaches to a total solution/methodology and constructing an information platform based on the proposed methodology. With the case study demonstrated, it is found that the methodology is well applicable for green product design. Therefore, while facing the impact of EU and other trading regions environmental directives, this research provides prime product providers, particularly in electronic industry, a very useful IT solution for designing environmental conscious products effectively.

6 References

[1] Altshuller, G., 1996, And Suddenly the Inventor Appeared - TRIZ, the Theory of Inventive Problem Solving, Massachusetts: Technical Innovation Center.

[2] Ahn, Y.J., and Lee, K.-W., 2006, "Application of Axiomatic Design and TRIZ in Ecodesign," The TRIZ Journal.

[3] CEC, 2005, Directive 2005/32/EC of the European Parliament and of the Council of 6 July 2005 establishing a framework for the setting of Eco-design requirements for Energy-using Products and amending Council Directive 92/42/EEC and Directive 96/57/EC and 2000/55/EC of the European Parliament and of the Council, Official Journal of the European Union, European.

[4] Chang, H.-T., 2005, "The Study of Integrating Su-Field Analysis Modeling with Eco-Innovative Concept for Product Design," Environmentally Conscious Design and

Inverse Manufacturing, 2005. Eco Design 2005. Fourth International Symposium on, pp. 663-670.
[5] ISO, 2002, ISO 14062 Environmental management – Integrating environmental aspects into product design and development, Technical Report, Genvea.
[6] ISO, 2006, ISO 14040 Environmental management - life cycle assessment - principles and framework, Technical Report, Geneva.
[7] ISO, 2006, ISO 14044 Environmental management - life cycle assessment - requirements and guidelines, Technical Report, Geneva.
[8] Lin, C.-S., and Su, C.-T., 2007, "An Innovative Way to Create New Service Applying the Methodology," Journal of the Chinese Institute of Industrial Engineers, Vol. 24, No. 2, pp. 142-152.
[9] Masui, K., Sakao, T., and Inaba, A., 2001, "Quality function deployment for environment: QFDE (1st report)-a methodology in early stage of DfE," Environmentally Conscious Design and Inverse Manufacturing, 2001. Proceedings EcoDesign 2001: Second International Symposium on, pp.852-857.
[10] Nissen, N.F., Stobbe, L., Schischke, K., Muller, J, and Reichl, H., 2007, "European Environmental Legislation - Insights into the EuP Process," Electronic Manufacturing Technology Symposium, 2007, IEMT '07, 32nd IEEE/CPMT International, pp.325-331.
[11] Pennington, D.W., Potting, J., Finnveden, G., Lindeijer, E., Jolliet, O., Rydberg, T., and Rebitzer, G., 2004, "Life cycle assessment Part 2: Current impact assessment practice," Environment International, Vol. 30, No. 5, pp. 721-739.
[12] Park, P.-J., Lee, K.-M., and Wimmer, W., 2006, "Development of an Environmental Assessment Method for Consumer Electronics by combining Top-down and Bottom-up Approaches," The International Journal of Life Cycle Assessment, Vol. 11 , No. 4, pp. 254-264.
[13] Rumelhart, D., Hinton, G., and Williams, R., 1986, Learning Internal Representations by Error Propagation, Parallel Distributed Processing, Vol. 1. Cambridge, MA: MIT Press.
[14] Rebitzer, G., Ekvall, T., Frischknecht, R., Hunkeler, D., Norris, G., Rydberg, T., Schmidt, W.-P., Suh, S., Weidema, B.P., and Pennington, D.W., 2004, "Life cycle assessment: Part 1: Framework, goal and scope definition, inventory analysis, and applications," Environment international: a journal of environmental science, risk and health, Vol. 30, No. 5, pp. 701-720.
[15] Savransky, S.D., 2000, Engineering of Creativity-Introduction to TRIZ Methodology of Inventive Problem Solving, New York: CRC Press.
[16] Sakao, T., 2007, "A QFD-centred design methodology for environmentally conscious product design," International journal of production research*,* Vol. 45, No. 18, pp. 4143-4162.
[17] Zhang, Y., Wang, H.P., and Zhang, C., 1998, "Product concept evaluation using GQFD-II and AHP," International Journal of Environmentally Conscious Design & Manufacturing, Vol. 7 No. 3, pp. 1-15.

Developing an ISO 14048-Based EuP Integrated Service Platform for Evaluating Environment Impacts and Supporting Eco-Design in Taiwan

Tzu-An Chiang[a], Hsing Hsu[b], Ping-Yu Chang[c,1] , Hung-chia Wei[d]

[a] Department of Commerce Automation & Management, National Pingtung Institute of Commerce – Pingtung, Taiwan.
[b] Taiwan Electrical and Electronic Manufacturers Association – Taipei, Taiwan.
[c] Department of Industrial Engineering and Management, Mingchi University of Technology – Taipei county, Taiwan.
[d] Avectec.com, Inc – Hsinchu, Taiwan.

Abstract. Due to Environmental pollution and natural resource exhaustion, environmental protection has become the critical issue. Hence, the European Commission published the Energy Using Products (EuP) Directive, which creates an architecture for product requirements of the eco-friendly design on all environmental aspects from cradle to grave. In order to help electronic companies execute the life cycle assessment (LCA) and apply the environmental product declarations, this research develops an ISO 14048-based EuP integrated service platform (EuPISP). The EuPISP applies the web service technology and the eco-spold XML data exchange standard to share the LCI (life cycle inventory) data between the members of the EuPISP. For ensuring the consistency of the EPDs, the EuPISP provides the electronic companies with the product category rules that explain which the information must be reported in the EPDs. In addition, the EuPISP invites the related domain experts to review and verify the correctness of LCI data, and provides the electronic map to let the electronic companies quickly search the LCI data of the specific common parts or components. Through the employing the EuPISP, the cost and time of LCA for electronic products can be diminished significantly

Keywords. Life cycle assessment, Energy-using product (EuP), life cycle inventory

[1] Corresponding Author Email: pchang@mail.mcut.edu.tw

1 Introduction

Over the past decades, owing to the global warming and ecological destruction, the issues of environmental damages are paid more attentions. Hence, there is a significant increasing awareness of the environment protection. In 2003, European Union (EU) enacted the Restriction of Hazardous Substance (RoHS) and Waste Electrical and Electronic Equipment (WEEE) directives. Consequently, firms are requested to design and produce green products. However, the environmental protection consciousness should be throughout the whole product lifecycle. In order to achieve the aim of sustainable development, the Energy-using Products (EuP) directive became law in 2007 to force companies to re-evaluate the environment performance of their products, and to reduce environmental burdens. The aim of the EuP directive is to introduce eco-design concepts for improving the environmental performance of energy-using products. As a result, the EuP directive is expected to be a key driver in reducing the environmental burdens. Accordingly to the results of the life cycle inventory (LCI) analysis, the companies execute the life cycle assessment (LCA) for understanding the environmental impacts of products. Finally, the eco-profile report can be completed. However, the LCI analysis and the LCA are very complex and time-consuming works. In order to assist electronic companies in efficiently executing the above-mentioned tasks, this research develops ISO 14048-based EuP integrated service platform (EuPISP). This platform provides the service of the product-category rules (PCR) and environmental product declarations (EPDs). With regard to the LCI, this platform provides the review and verification of the LCI data. In addition, the members of this service platform can present the requests to obtain LCI data of the standard parts or components by applying web service and the eco-spold XML data exchange standard so as to establish the LCI of their products. Then these companies can execute the LCA and then complete the eco-design investigation.

The remainder of this paper describes as follows: The section 2 reviews the related literature and section 3 introduces the ISO-14048 standard. The section 4 explains the current circumstances of implementing the eco-design requirements for the EuP in the Taiwan electronics industry. In the section 5 the architecture of the EuPISP is proposed. Finally, the section 6 gives the conclusions.

2 Literature review

In this decade, environmental protection has been an important issue in many areas. For an enterprise to contribute in protecting environment, three strategies, pollution prevention, cleaner production, and product stewardship should be taken in an enterprise’s long-term planning. However, production stewardship not only focus on contamination during process but also integrating pollution prevention into manufacturing process and product development. Hence, Design for Environment (DfE) is introduced into product and process design. DfE considers repair, remanufacturing, and reuse within manufacturing process so that the impact of a product can be evaluated in the designing phase.

To ensure the effectiveness of implementing DfE, LCA is usually adapted to evaluate elements (i.e. raw materials, energy consuming, and pollution) involved in manufacturing, using, and disposing. However, the rapid changes in the consuming market shorten product life cycle which increases the numbers of discarded product. Such abandoned electronics, containing toxic substances, goes to landfill and contaminate the environment [7]. Hence, governments are forced to enact related laws to control the numbers of disposed product.

EU enacted the laws of WEEE and RoHS in February, 2003. Furthermore, the RoHS directive it self is deceptively simple-sounding: electrical and electronic products sold in EU must be substantially free of six chemical substances. The WEEE, on the other hand, imposes the responsibility for the disposal of waste electrical and electronic equipment on the manufacturers of such equipment. The WEEE and RoHS directives, which go into effect in 2006, both require that manufacturers track and report the material composition of their products [9]. However, WEEE and RoHS only concentrate on discarded product and hazardous substances and didn't consider the effect of energy consuming for using the product. Therefore, EU proposed the directive of Eco-Design Requirements for EuP in 2004 to limit the impact of energy consumed in using electronic product on the environment.

2.1 Energy-using Products (EuP)

EuP directive requires manufacturers to limit the consumed energy to produce, transport, sell, use and dispose in every product. The manufacturers are required to calculate energy in every stage so that the impact on the environment can be examined closely. According to EU, the specification of eco-design should be defined for each specific product and parts. The eco-design of EuP is the primary principle of Integrated Product Policy (IPP) which considers the impact of product on environment in different stage of product life cycle. Therefore, firms in every supply chain echelon have to face the impact of information sharing, technical specification, managing supply chain, eco-design, and cost. Hence, firms should introduce strategies such as developing eco-design procedure, enhance green supply chain collaboration and green innovation, and construct life cycle management (LCM) principle to strengthen their competitiveness under EuP regulations.

2.2 Green supply chain

The importance of environmental protection has drawn lots attentions in recent years. Firms and countries devote their funds and time in research to develop green productivity, eco-design product, or prevent hazardous substances. Also, logistics and supply chain manager have to balance efforts to reduce costs and innovate while maintaining good environmental (ecological) performance [6]. Therefore, green supply chain management (GSCM) has emerged as an approach to balance these competitive requirements [5]. However, most of these research studies tend to focus on single aspects such as green purchasing, internal environmental operations management or green logistics in stead of empirical, theoretical grounded approach [3].

Therefore, to explore environmental issues in GSCM, a variety of related practices such as green supply, environmental purchasing, product stewardship, life cycle assessment, and reverse logistics has been investigated [12]. Zhu and Cote [10] conducted an empirical study in analyzing the Guitang Group and its evolving integrated green supply chain management model. They discovered that if the supply chain integration begins to form webs instead of following the path of a single product, and if the ecological dimensions can be extended over product life cycle, environmental protection could be more fulfilled. Zhu et al. [11] explored the GSCM pressures/motivators, initiatives, and performance of China automotive supply chain using an empirical analysis. Their results show that the Chinese automobile supply chain enterprises have experienced high and increasing regulatory and market pressures from GSCM practice adoption such as oil pollution, energy consuming, investment recovery, eco-design, and internal environmental management.

Kainuma and Tawara [4] extended the range of supply to include re-use and recycling throughout the life cycle of products and services. They proposed the multiple attribute utility theory method for assessing and evaluating a supply chain. Ferretti et al. [2] evaluated economic and environmental effect of aluminium industry using mathematical models. Their models determine the supply aluminium mix, capable of balancing the economic benefits as well as environmental requirements. Sheu [8] presented a multi-objective optimization programming approach to address the issue of nuclear power generation and the corresponding induced waste reverse logistics. The results indicated that using the proposed approach, the induced environmental impact including the corresponding costs and risks can be improved up to a certain level.

3 14048 description

Since LCA is usually adapted to evaluate elements involved in manufacturing, using, and disposing, it is necessary to discuss international technical specification ISO/TS 14048 environment management and data documentation format (ISO, 2001). The main purpose of 14048 is to standardize LCA data format so that a specified form can be used to report LCA data within industries, institutes, consultancy agencies. Charlson and Tivander [1] also specified some other purposes of ISO/TS 14048, that is, (1) can be used as a template for standard compliant and compatible designs; (2) can support transparent reporting, interpretation and review of data collection, data calculation, data quality and data reporting, as well as facilitates data exchange; (3) is developed for LCA, but can also be used to support environmental data management. Furthermore, ISO/TS

14048 describes the data documentation format for LCI (life cycle inventory) as described in ISO 14041 and enables fully transparent, and hence fully reviewable and verifiable data documentation for LCI, individual plants, and production lines. Due to the wide international acceptance of the semi structured data language XML (Extensible Markup Language), this language has been chosen for the implementation of a data exchange format, and due to the wide use of relational databases, the language of relational database models has been used for the database format specification – ISO/TS 14048 [1].

To demonstrate a data format of EuP based on the specification of ISO/TS 14048, an example is generated. Table 1 shows the process and process descriptions of this specific example. In Table 1, process information is categorized into nine fields. The first five fields include basic information of process, that is, name, class, quantitative reference, technical scope, and aggregation type. The sixth field contains various information of process technology, that is, description of technology, function and picture of technology, product flow, input and output source, input and output destination, and mathematical model of process. The last 3 fields include time span of technology, geography of sites using this technology, and required data to evaluate this technology.

Table 1. Data format for process

<table>
<tr><td rowspan="16">1.1. Process Description</td><td colspan="2">1.1.1. Name</td></tr>
<tr><td>1.1.2. Class</td><td>1.1.2.1. Reference to nomenclature</td></tr>
<tr><td rowspan="4">1.1.3. Quantitative reference</td><td>1.1.3.1. Type</td></tr>
<tr><td>1.1.3.2 Name</td></tr>
<tr><td>1.1.3.3 Unit</td></tr>
<tr><td>1.1.3.4 Amount</td></tr>
<tr><td colspan="2">1.1.4. Technical scope</td></tr>
<tr><td colspan="2">1.1.5. Aggregation type</td></tr>
<tr><td rowspan="6">1.1.6. Technology</td><td>1.1.6.1. Short technology descriptor</td></tr>
<tr><td>1.1.6.2. Technical content and functionality</td></tr>
<tr><td>1.1.6.3. Technology picture</td></tr>
<tr><td>1.1.6.4. Process contents</td></tr>
<tr><td>1.1.6.5. Operating conditions</td></tr>
<tr><td>1.1.6.6. Mathematical model</td></tr>
<tr><td colspan="2">1.1.7. Valid time span</td></tr>
<tr><td colspan="2">1.1.8. Valid geography</td></tr>
<tr><td></td><td colspan="2">1.1.9. Data acquisition</td></tr>
</table>

In Table 1, basic information such as manufacturing process, process description, technology, mathematics formula, time span, and geography are presented. However, to realize the impact of these manufacturing process on environment, a more specific information should be provided. Therefore, Table 2 describes the specific information of technology and process. In Table 2, detailed descriptions of input and output are specified within fourteen fields. The first nine fields contain basic information of input and output while the following five fields include name, property, amount, mathematical relation, and documentation of input and output.

Table 2. Inputs and outputs

Identification number	
Direction	
Group	
Receiving environment	
Receiving environment specification	
Environment condition	
Graphical location	
Related external system	Origin or destination
	Transport type
	Information reference
Internal location	
Name	Name text
	Reference to nomemclature
	1.2.10.3. Specification of name
Property	
Amount	
Mathematical relation	
Documentation	

Tables 1 and 2 describe the process and the calculation of the impact on the environment. However, these presented formula should be verified before implementation. Therefore, Table 3 shows the procedures of model validation and verification. In Table 3, seven fields are specified to demonstrate model and model validation. The first two fields show the application and information source of this specific model. The third field demonstrates model principles which include data selection principles, adaptation principles, and model constants. The fourth field shows model choices which contain criteria of product flows, allocation, and process explanation. The fifth field is data quantity statement while the sixth field states model validation method.

Table 3. Model validation and verification

<table>
<tr><td colspan="3">Intended application</td></tr>
<tr><td colspan="3">Information sources</td></tr>
<tr><td rowspan="3">Modelling principles</td><td colspan="2">Data selection principle</td></tr>
<tr><td colspan="2">Adaptation principles</td></tr>
<tr><td colspan="2">Modelling constants</td></tr>
<tr><td rowspan="7">Modelling choices</td><td colspan="2">Criteria for excluding elementary flows</td></tr>
<tr><td colspan="2">Criteria for excluding intermedia product flows</td></tr>
<tr><td colspan="2">Criteria for externalising</td></tr>
<tr><td rowspan="2">Allocation performed</td><td>Allocated co-products</td></tr>
<tr><td>Allocation explanation</td></tr>
<tr><td rowspan="2">Process expansion</td><td>Process included in expansion</td></tr>
<tr><td>Process expansion explanation</td></tr>
<tr><td colspan="3">Data quality statement</td></tr>
<tr><td rowspan="4">Validation</td><td colspan="2">Method</td></tr>
<tr><td colspan="2">Procedure</td></tr>
<tr><td colspan="2">Result</td></tr>
<tr><td colspan="2">Validator</td></tr>
<tr><td colspan="3">Other information</td></tr>
</table>

4 The current circumstances of implementing the eco-design requirements for the EuP in the Taiwan Electronics Industry

In order to reduce resource consumption and pollutants, the European Union (EU) publishes the EuP directive to clearly display the concern regarding the environmental impact of a product across the entire product life cycle, including development, production, usage, and disposal. New environmental regulations have increased the pressure on companies, especially the small- and medium-sized enterprises (SMEs) in Taiwan. Figure 1 shows the operational process of the as-is model for complying with the eco-design requirements for EuPs. Firstly, the SMEs need to develop the product-category rules (PCRs). According to the PCRs, SMEs execute the life cycle inventory (LCI) and the life cycle assessment (LCA). The result of the LCA needs to be verified by the third party. Finally, the eco-profile report is constructed. However, the electronics industry lacks common PCRs in the current operational situation (the as-is model). In addition, the LCI and the LCA are complex and time-consuming works for SMEs. Therefore, if electronics companies can share the LCI and the LCA data for common parts, the companies can significantly reduce the cost of constructing eco-profile reports and enhance the competitiveness of the Taiwan's electronics industry in the market. Moreover, the result of a LCA needs to be verified its accuracy by an independent third party for selling their electronic products on the EU market. Finally, the companies use the LCA data to construct the Eco-profiles of the products for EPDs.

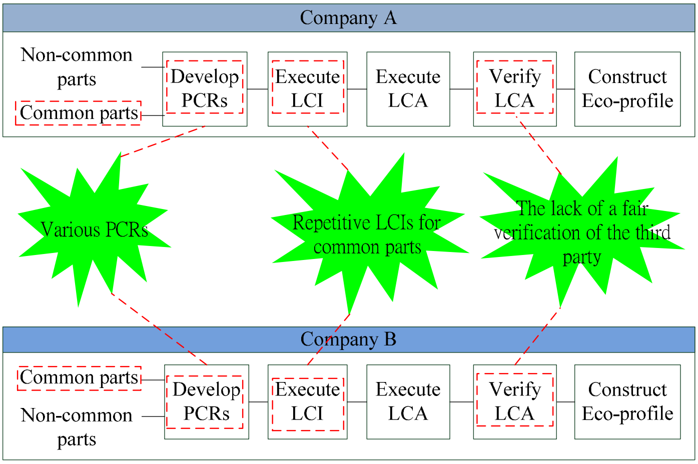


Figure 1. The as-is model for achieving the eco-design requirements for EuPs

5 The framework and the operational model of the ISO 14048-based EuPISP

Facing above-mentioned problems and challenges, this research develops the system framework of the ISO 14048-based EuPISP as shown in Figure 2. The EuPISP platform includes eight parts, which are (a) PCR/EPD services, (b) LCA calculation, (c) LCI information exchange center, (d) LCI verification and review, (e) LCI questionnaire forms, (f) industrial electronic map, (g) application for LCI consultation, and (h) authority management.

The following explains each part of the EuPISP.

(a) PCR/EPD services: In order to assist the companies to establish the objectivity and consistency of the product environmental declaration. The EuPISP provides an uniformity in the definition of product information, the requirements and rules of the LCA calculation, the scenarios of product use and disposal, as well as the parameters to be declared in the EPDs. In addition, the EuPISP applies the PCRs and LCA data to construct the EPDs for the EuPISP members.

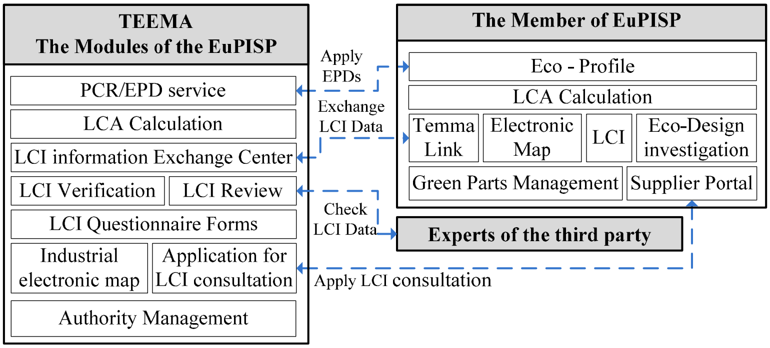


Figure 2. The system framework of the EuPISP

(b) LCA calculation: Figure 3 shows the procedure of LCA calculation. Firstly, product lifecycle management (PLM) system provides the bill of material (BOM) for the product to the EuPISP. The EuPISP can use the part-level information to understand how to constitute the products, and then search the LCI data for common parts. With regard to non-common parts, the EuPISP can utilize the industrial electronic map to gain the LCI data of elementary flow for raw materials. The suppliers are responsible for the LCI of the parts. Then the EuPISP can calculate the product-level LCI information. Through applying the above-mentioned information and the analytical model of environmental impact assessment, the EuPISP can produce the environmental impact report of a specific product.

(c) LCI information exchange center: The center of the information exchange complies with the ISO 14048 standard to express LCI data. In order to exchange LCI information between the LCI information exchange center and the member of the EuPISP, the center adopts the Eco-spold formatted XML as the information exchange standard, and employs the web service technology. Hence, the members of the EuPISP can easily obtain the LCI data for the common parts.

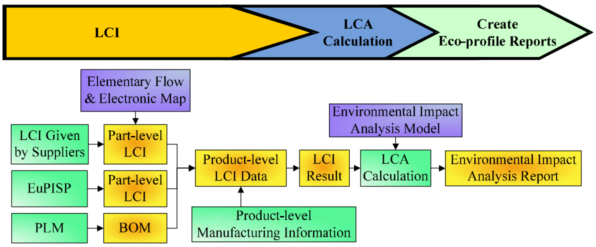


Figure 3. The procedure of LCA calculation

(d) LCI verification and review. To ensure the objectivity and correctness of the LCI information, the EuPISP requires the domain experts to review and verify the LCI data. Figure 4 displays the procedure of the LCI verification and review. Firstly, the company needs to finish the LCI questionnaire table and explains how to execute LCI data model. The domain experts review and verify the correctness of the LCI data. If the LCI data is correct, the EuPISP utilizes the eco-spold XML data exchange format to transfer the LCI data to the LCI database of the EuPISP. The members of the EuPISP can use the LCI data to create an eco-profile file. The above-information can share other members of the EuPISP through LCI information exchange center.

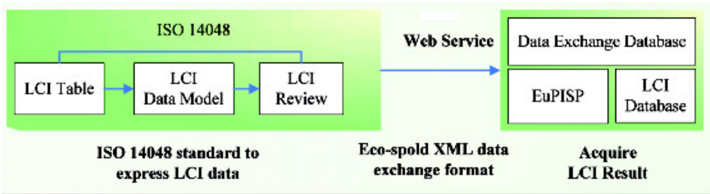


Figure 4. The procedure of the LCI verification and review.

(e) LCI questionnaire forms. In order to ensure the consistency of LCI data format, the EuPISP provides the LCI questionnaire forms to the members of the EuPISP.

(f) Industrial electronic map. The industrial electronic map employs the tree structure to establish the classification standard of electronic parts. By using the interface of the industrial electronic map, the members of the EuPISP can efficiently search the needed LCI data.

(g) Application for LCI consultation: Because the system integrators do not have the LCI data of the parts, this function of the EuPISP can assist these companies to collect the needed LCI information of the parts and to evaluate the LCA. In addition, the LCI method complying with the ISO 14048 format is very complex and time-consuming work. Hence, each member of this platform can use the application for LCI consultation to reduce the cost and time of the LCI data collection.

(h) Authority management. The EuPISP has three authorities, i.e., system administrators, LCI reviewers, members of the EuPISP. Each role is assigned different authorities so as to decide which functions of the EuPISP can be used. The following explains each authority's role of the EuPISP.

- System administrators. The role has the top authority to execute any function of the EuPISP. In addition, the system administrators can decide which functions can be given for each role.
- LCI reviewers. The EuPISP can provide the LCI reviewers with the friendly interface to review and verify the correctness of the LCI data.
- The members of the EuPISP. The members of the EuPISP (i.e., the suppliers and the system integrators) can upload the LCI data to the

EuPISP platform, then the experts check the correctness of the LCI data. In addition, the members of the EuPISP can exchange LCI data and apply the function of LCI consultation to assist LCA work for acquiring the EPDs.

6 Conclusion

EU enacted the laws of WEEE and RoHS in 2003 and proposed the directive of Eco-Design Requirements for EuP in 2004 to limit the impact of energy consumed in using electronic product on the environment. The EuP directive is expected to be a key driver in reducing the environmental burdens in which LCI, LCA, and eco-profile report should be developed for product design. To assist electronic companies in efficiently executing the above-mentioned tasks, this research develops ISO 14048-based EuP integrated service platform (EuPISP) to provide the service of review and verification of the LCI data, PCR, and EPD so that LCA and eco-design investigation can be finalized. Furthermore, this research develops the system framework of the ISO 14048-based EuPISP with eight parts, which are (a) PCR/EPD services, (b) LCA calculation, (c) LCI information exchange center, (d) LCI verification and review, (e) LCI questionnaire forms, (f) industrial electronic map, (g) application for LCI consultation, and (h) authority management. It is our belief that the demonstrated system will assist industries to conduct their eco-profile of product for the requirements of EuP directive.

7 References

[1] Carlson R, Tivander J, Data definition and file syntax for ISO/TS 14048 data exchange: with data storage format based on ISO/TS 14048. (A report by Center for environmental assessment of Product and material Systems, Chalmers University of Technology, Goteborg, Sweden), 2001.

[2] Ferretti I, Zanoni S, Zavanella L, Diana A, Greening the aluminium supply chain, International Journal of Production Economics, 2007: 108(1-2): 236-245.

[3] Gavagjan K, Klein RC, Olson JP, Pritchett T, The Greening of the Supply Chain, Supply chain management Review, 1998: 76-86.

[4] Kainuma Y, Tawara N, A multiple attribute utility theory approach to lean and green supply chain management, International Journal of Production Economics, 2006: 101(1): 99-108.

[5] Narasimhan R, Carter JR, Environmental supply chain management. The Center for Advanced Purchasing Studies, Arizona State University, Tempe, AZ, USA, 1998.

[6] Pagell M, Yang CL, Krumwiede DW, Sheu C, Does the Competitive Environment Influence the Efficacy of Investments in Environmental Management, Journal of Supply Chain Management, 2004: 40(3): 30 – 39.

[7] Sarkis J, Manufacturing strategy and environmental consciousness, Technovation, 1995: 15(2): 79-97.
[8] Sheu JB, Green supply chain management, reverse logistics and nuclear power generation, Transportation Research Part E: Logistics and Transportation Review, 2008: 44(1): 19-46.
[9] Van Hoek R, The contribution of performance measurement to the expansion of third party logistics alliances in the supply chain, International Journal of Operations & Production Management, 2001: 21(1): 15.
[10] Zhu Q, Cote RP, Integrating green supply chain management into an embryonic eco-industrial development: a case study of the Guitang Group, Journal of Cleaner Production, 2004: 12(8-10): 1025-1035.
[11] Zhu Q, Sarkis J, Lai KH, Green supply chain management: pressures, practices and performance within the Chinese automobile industry, Journal of Cleaner Production, 2007: 15(11-12): 1041-1052.
[12] Zsidisin GA, Siferd SP, Environmental purchasing: A framework for theory development, Europe Journal of Purchase & Supply Management, 2001: 7(1): 61–73.

Systematic Lean Techniques for Improving Honeycomb Bonding Process

Chiun-Ming Liu[a,1,] and Min-Shu Chiang[b]

[a]Professor, Department of Industrial Engineering and Systems Management, Feng Chia University, Taichung, Taiwan.
Email: cmliu@fcu.edu.tw
[b]I.E. Engineer, Aerospace Industrial Development Corp.

Abstract. Honeycomb sandwich panel, which is mostly made of composite materials, is the major component constituting airplane body structure. The composite materials can greatly reduce production and transportation costs and so are widely used in the aerospace industry. Traditional honeycomb sandwich panel companies are facing competitive challenges such as shorter delivery time, high quality, and quick response to customer demand. Lean technology has been effectively applied to various industries for improving production efficiency. In this study, systematic lean techniques are proposed and implemented in an aerospace company for improving the honeycomb bonding process. A series of lean production methods are organized and presented. Implementation results suggest the overall productivity of the honeycomb bonding process increase due to elimination of bottlenecks, reduction of cycle time, and decrease of WIP inventory. The proposed approach can be expected to apply to other manufacturing processes for improving their productivity.

Keywords. Lean technique, honeycomb bonding process, productive capability

1 Introduction

Composite materials, which are the major component for producing honeycomb sandwich panel, are made of two more single materials through different synthesis and fabrication methods. The advantages of composite materials lie in the capability to hold the merits of original materials while preventing from their weakness. Compared with traditional metal materials, composite materials possess many special features, such as lighter mass density, stronger structure, better anti-corrosiveness, better anti-fatigue, and the like. So, the utilizations of composite materials are widely seen in many industries, for instance, aerospace industry, vehicle industry, construction industry, etc. In particular, composite materials are gradually replacing metal materials for manufacturing in aerospace industry.

[1] Corresponding Author Email: minshuchiang@ms.aidc.com.tw

For many domestic honeycomb sandwich panel manufacturers, their production facility and manufacturing process have been designed for small-to-medium production scale with push production systems. Due to market competition, the customers always demand shorter delivery, cost effectiveness and high quality. Therefore, those domestic manufacturers must somehow improve their manufacturing techniques and efficiency in order to strength their competitive.

Lean techniques are originated and developed from Japanese Toyota Production Systems. The original idea of lean techniques is to eliminate waste or non-value-added activities in production process and reduce production cost. The core methods are to apply value stream mapping to analyze the enterprise value chain, to use takt time to accommodate customer demand, to implement line balancing and heijunka scheduling for production planning, and to apply pull-system to replace push-system.

In the early 1980s and 1990s, several scholars dedicated to the lean production research and obtained significant results. Sugimori et al. [8], Monden [6], and Pegels [7] studied multi-function work design, just-in-time production process, small batch size, pull-type kanban systems, and quick changeover production line. Wantuck [10] dedicated to continuous improving, just-in-time production process, small batch size, pull-type kanban systems, and quality control management. Jones et al. [3] stressed on the importance and benefit of using value stream mapping for analyzing production process. Lee and Ebrahimpour [4] applied the value stream mapping tool to find roots of production problems and provide necessary information for improvement. Mabry [5] studied methods of applying small batch size, eliminating non-value-added activities, relocating facility layout and material flow, and designing efficient work study for improving auto parts suppliers efficiency.

Recently, Sullivan and Aken [9] summarized 7 tools from value stream mapping and types of major wastes. Womack and Jones [11] pointed out that by using lean thinking way the company can benefit gaining more outputs with less inputs, transforming waste into valuable activities, quick response to feedback, and making people feel satisfactory. Beachum [2] pointed out that under an ideal lean environment, the company can shorter lead time, reduce work-in-process inventory, improve product quality, and increase the on-time delivery rate if one piece flow with pull systems is implemented. Abdulmaleka and Rajgopalb [1] succeeded in reducing work-in-process inventory and production lead-time by applying value stream mapping to analyze non-value-added activities and visual management tools for improvement.

Lean technology has been effectively applied to several industries such as vehicle, electronic and shoes industries. For the aerospace industry, due to its special demand for manufacturing process, change of engineering, order quantity, and cost control, lean technology can be expected to gain much more benefits if we properly design and implement lean techniques and methods.

In this study, systematic lean techniques are proposed to improve the manufacturing process for the composite materials. The proposed approach systematically develops a series of procedures and methods from lean techniques. The proposed techniques are implemented in a domestic aerospace company for improving the honeycomb bonding process. Results suggest the overall productivity of the honeycomb bonding process significantly increase by eliminating bottlenecks, reducing cycle time, and decreasing inventory level. The developed approach can be expected to apply to other various manufacturing processes for improving productivity.

2 Development of Lean Systematic Techniques

The methodology of lean techniques consists of a series of continuous improvement procedures and methods. In this study, these series of lean methods are reorganized in a systematic way that can be effectively applied to improve manufacturing process. The proposed systematic lean techniques may include five phases: 1) value stream mapping, 2) formula developing and computing, 3) 5S, kanban, and cellular-layout planning, 4) line balancing, and 5) heijunka scheduling. The detailed procedures are described as follows.

Phase 1. Value Stream Mapping

The aim of value stream mapping is to analyze the value chain within the enterprise. The value chain includes material flow and information flow. The analysis of value chain provides details of current operational state within the enterprise and opportunities for improvement. The entire procedures for value stream mapping are given as follows.

Step 1. Analyze suppliers, input, processes, output, and customers within the value chain.

Step 2. Select product family. The selection of major product family can be done using product flow analysis to group products that share common flow patterns into families and value streams.

Step 3. Collect process data. The process data include manual cycle time, down time, flow time, changeover time, work-in-process inventory, and yield.

Step 4. Perform Gemba. Gemba is go see where the operations are and quantify value-added, non-value-added, and necessary non-value-added components of each particular activity.

Step 5. Draw the current-state value stream map. The drawing of current state value stream map can be done using some articulate tools, icons, and techniques.

Value stream mapping can help an enterprise focus on managing the value chain for all products and services from suppliers, input, process, output, and customers. Through the analysis of enterprise value stream, a company can grasps the customer demand and provides value-added activities for meeting customer demand.

Phase 2. Formula Developing and Computing

In order to identify and quantify the bottleneck workstations, several formulas are developed and used to provide necessary information. The development of equations is given as follows.

Step 1. Define variables. The variables defined in this study are: 1) operation time (OT): OT is the sum of primal time (PT) and additional time (AT); 2) standard time (ST): ST is an operation time (OT) with some additional allowance (r); 3) standard capacity (SC): SC is the number from ST divided by 3600 seconds; 4) needed persons (NP): NP is the number from dividing the sum of ST and order quantity (D) by total working hours (TH); 5) workload (WL): WL is the number from NP divided by setting persons (SP); and 6) takt time (TT): TT is the number from total working hours divided by order quantity (D).

Step 2. Formulating equations. The equations for computing the variables are: 1) $OT = PT + AT$; 2) $ST = OT \times (1 + r)$; 3) $SC = 3600 / ST$; 4) $NP = (ST \times D) / TH$; 5) $WL = NP / SP$; and 6) $TT = TH / D$.

Step 3. Obtaining the number for workload (WL), and takt time (TT).

Phase 3. 5S, Kanban, and Cellular-Layout Planning

5S is a set of basic concepts that help enterprises ensure a clean and organized workplace. Those five basic concepts are: sorting, sweeping, standardizing, simplifying, and self-discipline. Kanban describes a variety of pull mechanisms that draw work through an entire value stream according to demand and downstream capacity. Kanban is a useful approach to simplify and manage the flow of production and control inventory. A kanban system can be seen as a loop of information and materials circulating between two stores of inventory. There are three distinct types of kanban systems: product-specific kanban, generic kanban, and hybrid kanban. A scenario with multiple stores of inventory at various stages of production should be carefully planned. The cellular-layout is a product-oriented layout that places various machines in the exact sequence required to process a family of parts. The cellular-layout planning can be done using the following procedure.

Step 1. Applying U-shape cellular layout principle.

Step 2. Designing material flow.

Step 3. Allocating facility and hardware.

Phase 4. Line Balancing

Line balancing is to balance the assignment of operations to workstations so that idle time and the number of people working on the line are minimized. The procedures for line balancing are:

Step 1. Compete work measurement and record all data.

Step 2. Obtain takt time and compare with cycle time. Takt time is calculated by dividing total effective working time for the period by total demand.

Step 3. Depict a balance analysis sheet.

Step 4. Analyze production line using takt time and cycle time to identify the location of bottleneck stations.

Step 5. Improve the bottleneck stations by relocating or redesigning workstations. A balanced cell is preferred, where the cycle time of all operations are within 30% of one another.

Phase 5. Heijunka Scheduling.

The heijunka method is also referred to as rate-based, level, or campaign scheduling. Heijunka scheduling works as a shock absorber, buffering variations in supply and demand, providing the shop floor and suppliers with a stable short-term production plan. A heijunka schedule requires a forecast of quantity and mix for each value stream. Then some batching is required. The batch size can be determined by calculating a changeover interval, which is the period of time required to produce one full cycle of a product family. The interval can be calculated with a method called the every-part-every interval (EPE). EPE interval = (effective working time in period – run time × quantity for period) ÷ (number of products in mix × setup time). The heijunka scheduling can be summarized as follows.

Step 1. Break down the total volume of orders for a given planning period into scheduling intervals.

Step 2. Define a repetitive production sequence for that scheduling interval by a heijunka calculation.

Step 3. Dictate the model mix scheduled on a given line.

Step 4. Use kanban cards or signals for the mix of products to put that schedule into operation.

3 Implementation of Systematic Lean Techniques

3.1 Analysis for Honeycomb Sandwich Panel Fabrication Process

Honeycomb sandwich panel is one of the most important airplane body structure components. The configuration of honeycomb sandwich panel is type of sandwich-type structure and fabricated by using metal panel or carbon fiber pre-fitting cloth to serve as outer covering, filling with honeycomb structuring material in-between, and gluing those components together by adhesives. The fabrication process for honeycomb sandwich panel consists of material cutting, lay-up, autoclave curing, remolding, bench work, pre-fitting, sanding and filling, painting, assembly, marking, and stocking. Figure 1 displays the sequence of fabrication process for honeycomb sandwich panel. The detailed fabrication process is described as follows.

1. Material Cutting: Before releasing of materials, task of materials and mold preparation is needed. The preparation of materials includes thawing procedure and cutting procedure.
2. Lay-Up: Lay-up operation includes glue injection, laying, and bagging for honeycomb apertures.

3. Autoclave Curing: After bagging, insert molds into the heat furnace for autoclave curing operation. Autoclave curing operates under the holding condition of high temperature and pressure.
4. Remolding: when autoclave curing operation is finished, remove molds from the heat furnace, tear down bagging material, and take parts out of mold.
5. Component Mechanical Processing: Fix parts at the winding stand and polish broken filaments using tools.
6. Prepare: Fix parts at the drilling stand and drill bores needed.
7. Wet Grinding: Remove impurities and pollutants on surface by using abrasive paper.
8. Painting: Before painting, cover up the unpainted surface area by screen and then remove it after painting.
9. Assembly: Apply glue on the connecting surface between parts, hinges, and binder bolt, and then rivet with a rivet.
10. Marking: According to the indication by a blueprint, mark the part number using printing ink and then pack for stocking.

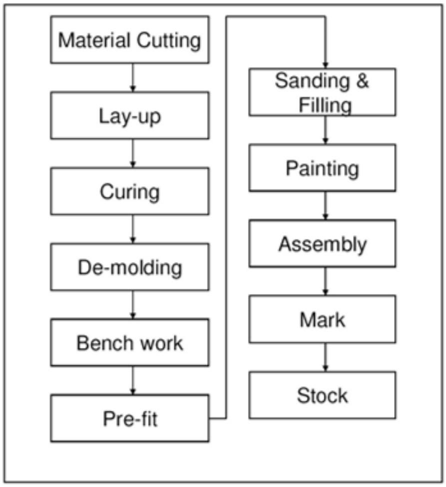


Figure 1. Fabrication process for composite materials

After completing the analysis of honeycomb manufacturing process, several production problems are diagnosed. First, the plant space is getting crowded due to an increasing customer order and inappropriate plant layout and material flow design. Secondly, lots of WIP inventory can be found in-between each workstation due to bottleneck stations. Thirdly, manufacturing quality is getting worse due to pollutants and waste activities scattered within production lines. Finally, production capability is going down due to unbalance of production lines. Traditional improvement methods have been utilized to somehow tackle the production problems, but no significant improvement is obtained.

3.2 Value Stream Mapping

After analyzing suppliers, input, processes, output, and customers within the value chain, selecting product family by using product flow analysis, collecting process data of cycle time, down time, flow time, changeover time, and work-in-process inventory, going to observe where the operations are and quantify value-added, non-value-added, and necessary non-value-added components of each particular step, a current-state value stream map is done using some articulate tools, icons, and techniques. From Figure 2 showing the obtained current-state value stream map one can see the whole honeycomb bonding process, grasp the connection between material flow and information flow, and identify the sources of waste and bottleneck stations.

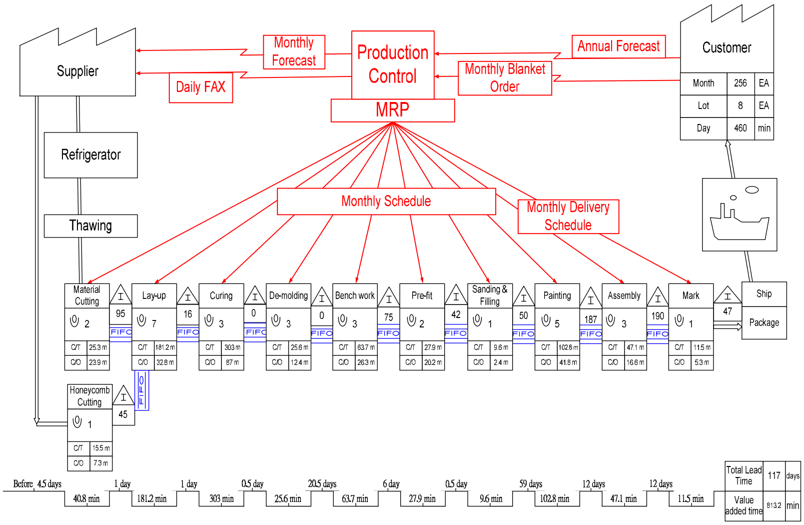


Figure 2. Value stream map before implementation

The takt time, workforce, and process cycle efficiency are calculated as follows. This information can be used to determine a suitable batch size and balance a production line.

1. Daily Demand = Monthly Demand ÷ Monthly Work Days

   = 256 ÷ 20 = 12.8 pieces

2. Total Daily Available Time = 460 minutes
3. Takt Time = Total Daily Available Time ÷ Daily Demand

   = 460 ÷ 12.8 = 35.9 minutes

4. Material Cutting Workforce = 49.2 ÷ 35.9 = 1.4 persons
5. Honeycomb Cutting Workforce = 22.8 ÷ 35.9 = 0.6 persons
6. Lay-Up Workforce = 214 ÷ 35.9 = 5.9 persons
7. Autoclave Curing Workforce = 39 ÷ 35.9 = 1.1 persons
8. Remolding Workforce = 38 ÷ 35.9 = 1.1 persons

9. Component Mechanical Processing Workforce = 90 ÷ 35.9 = 2.5 persons
10. Prepare Workforce = 48 ÷ 35.9 = 1.3 persons
11. Wet Grinding Workforce = 12 ÷ 35.9 = 0.3 person
12. Painting Workforce = 144.6 ÷ 35.9 = 4.0 persons
13. Assembly Workforce = 63.7 ÷ 35.9 = 1.8 persons
14. Marking Workforce = 15.6 ÷ 35.9 = 0.4 person
15. Process Cycle Efficiency = 813.2 minutes ÷ 117 days = 1.99%.

3.3 5S, Kanban, and Cellular-Layout

When implementing systematic lean techniques, 5S is the fundamental discipline for employees and should be enforced into the routine operations in workplace. The requirements for 5S are:

1. Unnecessary tools, equipment, and procedures are removed from the workplace.
2. Unsafe conditions or damaged equipment are dealt with early in the process.
3. Define how a task should be done and let everyone involved know the best way to perform tasks.
4. Put everything in its place and organize material according to how frequently it is used.
5. Ensure that all housekeeping policies are adhered to.

The location of kanban is selected mainly at the connection between suppliers and production line, and between the customer and final assembly workstations in order to facilitate the pull system and visual workplace. The kanban container is used in-between each workstation to control the work-in-process inventory.

The cellular layout principle is used to improve the material flow in shop floor. Figure 3 displays the facility layout and material flow before the implementation of the cellular layout principle. Under such a layout environment, too many waste spaces cause congestion for people, materials, and equipment. The movement of people and materials is unsmooth and hindered. Figure 4 is the proposed cellular layout for shop floor. The proposed cellular layout principle reduces the space in use and provides more free space. The travel distance of people and materials is significantly reduced.

3.4 Line Balancing and Heijunka Scheduling

Figure 5 shows the production schedules and the associated cycle time in minute for each operation before implementation. The times for value-added and non-value-added activities in each workstation are identified. Also, the calculated takt time and observed cycle time from each workstation are used to identify the location of bottleneck stations. The difference between the largest cycle time and the smallest cycle time is significant. This big gap results in serious bottleneck stations and large number of work-in-process inventory.

In order to balance the production line, the batch size is reduced to 12.8 pieces from the current batch size 256 pieces by estimating every-part-every interval (EPE). Then the production line is reallocated by combination and simplification of workstations. The obtained balanced production schedule is depicted in Figure 6. The number of workstations is reduced to 6 from 11. The difference between the largest cycle time and the smallest cycle time is negligible.

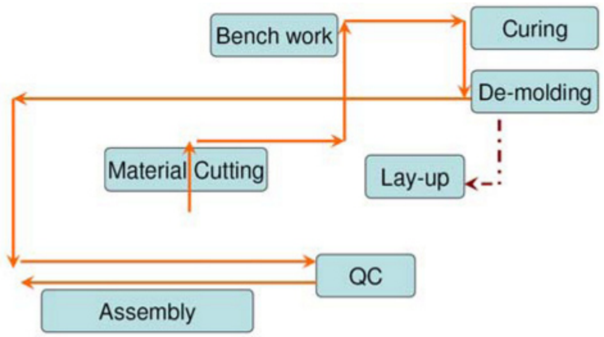


Figure 3. Material flow layout before implementation

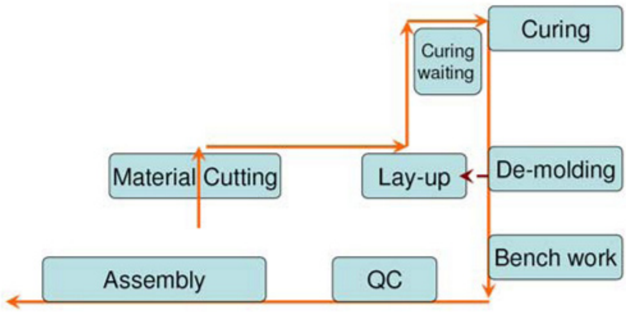


Figure 4. Material flow cell-layout after implementation

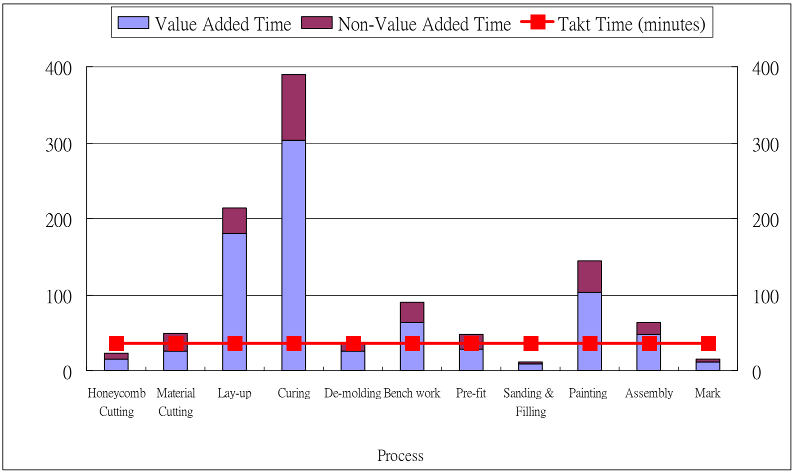


Figure 5. Production schedule before implementation

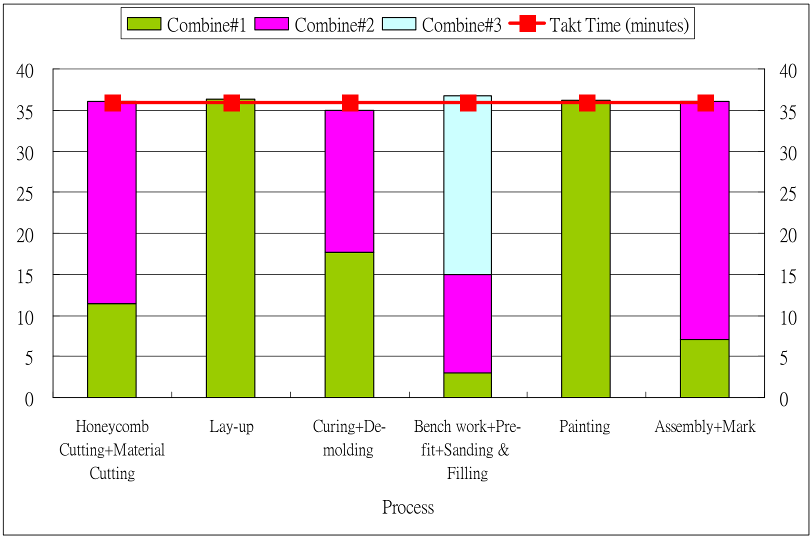


Figure 6. Production schedule after implementation

The value stream map after implementing the proposed systematic lean techniques is shown in Figure 7. Major improvements are: reduction of production cycle time, adoption of pull-type kanban system, reduction of workforce and WIP inventory, and decrease of fabrication time.

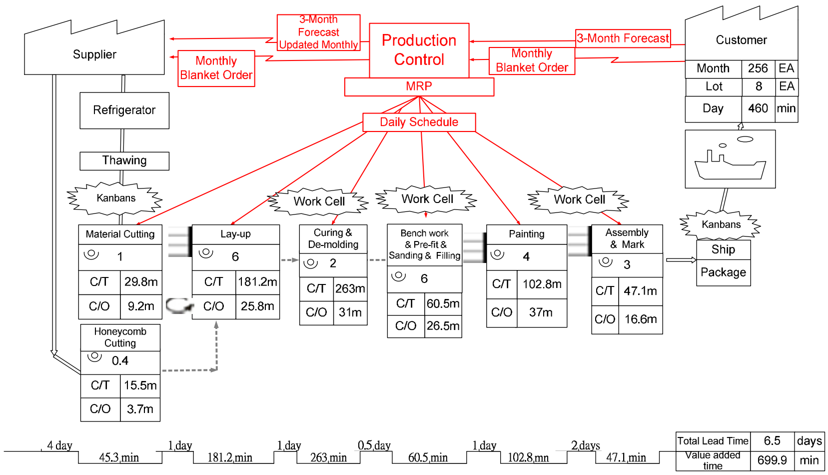


Figure 7. Value stream map after implementation

3.5 Benefit Analysis

The benefits obtained can be summarized as follows.

1. The batch size is reduced from 640 pieces to 16 pieces according to the takt time computation and the heijunka schedule.
2. The shape of material flow is changed to U-shape flow resulting from the cellular-layout design.
3. The travel distance decreases to 258 meters from 504 meters.
4. The work-in-process inventory is reduced to 424 pieces from 802 pieces.
5. The production cycle time decreases to 43 days from 117 days.
6. The fabrication time per piece is reduced to 15.31 hours from 18.13 hours.
7. The workforce decreases to 23 persons from 30 persons.
8. The process cycle efficiency increases to 4.97% from 1.99%.
9. Production cost is reduced to 214,000 NTD from 225,000 NTD.
10. The quality failure cost is reduced to 0.099% from 0.399%.
11. The throughput per run increased to 16 pieces from 8 pieces.

The major improvements after implementation of the proposed systematic lean techniques are shown in Table 1.

Table 1. Comparison of performances before and after implementation.

Indicator	Before	After
Batch Size	640 pieces	16 pieces
Material Flow	Back and forth flow	U-shape flow
Travel Distance	504 meters	258 meters
Work-in-Process Inventory	802 pieces	424 pieces
Production Cycle Time	117 days	43 days
Fabrication Time	18.13 hours per piece	15.31 hours per piece
Workforce	30 persons	23 persons
Process Cycle Efficiency	1.99 %	4.97 %
Production Cost	225, 000 NTD	214, 000 NTD
Quality Failure Cost	0.399%	0.099%
Throughput	8 pieces per run	16 pieces per run

4 Conclusions

In this study, systematic lean techniques is developed and implemented for improving composite materials manufacturing processes. The developed lean techniques encompass six phases and in each phase detailed steps are designed and provided for displaying how lean techniques are applied. The proposed approach is implemented in a composite materials plant owned by a local aerospace company. Indicators from production lines are collected and used to compare the performance before and after implementation. Significant improvements in indicators result in streamline of material flow, reduction of waste and cost, increase of productive capability, and flexibility of order fulfillment. Results from this study indicate that the proposed systematic lean techniques can be implemented in the composite materials manufacturing process and also in different manufacturing processes.

5 References

[1] Abdulmaleka FA, Rajgopalb J. Analyzing the benefits of lean manufacturing and value stream mapping via simulation: a process sector case study. International Journal of Production Economics 2007; 107, 223–236.
[2] Beachum D. Lean manufacturing beefs up margins pull systems, takt time, and one-piece flow benefit the operation of a powder coating system. Organic Finishing Group, Walgren Co. 2005.
[3] Jones DT, Hines P, Rich N. Lean logistics. International Journal of Physical Distribution Logistics management 1997; 27(3/4): 53-73.
[4] Lee SM, Ebrahimpour M. Just-in-time production system: some requirements for implementation. International Journal of Operations and Production Management 1984; 4(4), 3–15.
[5] Mabry BG. Transformation to lean manufacturing by an automotive supplier. Computers and Industrial Engineering 1996; 31(112), 95-98.

[6] Monden Y. What makes the Toyota production system really tick. Industrial Engineering 1981; 13(1), 13–16.
[7] Pegels CC. The Toyota production system: lessons for American management. International Journal of Operations and Production Management 1984; 4(1), 3–11.
[8] Sugimori Y, Kusunoki F, Cho F, Uchikawa S. Toyota production system and kanban system: materialization of just-in-time and respect for human systems. International Journal of Production Research 1977; 15(6), 553–564.
[9] Sullivan WG, van Aken EM. Equipment replacement decisions and lean manufacturing. Robotics and Computer Integrated Manufacturing 2002; 18, 255-265.
[10] Wantuck KA. The Japanese approach to productivity. In: Chase R, Aquilano NJ. (Eds.). Production and Operations Management: A Life Cycle Approach, 6th ed. Irwin, Homewood, IL. 1983.
[11] Womack JP, Jones DT. Lean thinking: banish waste and create wealth in your corporation, 2nd Ed. 2003.

Expanding Environmental Information Management: Meeting Future Requirements in the Electronics Industry

Eric Simmon[a,1] , John Messina[b]

[a]Electrical Engineer, National Institute of Standards and Technology (NIST), USA.
[b]Computer Scientist, NIST.

Abstract. The environmental impact of product manufacturing, use, and disposal has become a worldwide concern. Laws and regulations designed to protect human health and the environment are being established throughout the global community. In addition to regulatory needs, manufacturing networks have begun to respond to market driven eco-efficiency and sustainability requirements. To optimize manufacturing systems for these requirements, information must flow freely upstream and downstream during a product's lifecycle. This information includes regulatory compliance information, material content, energy use, and test data among others types of information. To modify systems to effectively manage these data requires an understanding of what information is required. This paper discusses both regulatory and voluntary information needs with a focus on electronics industry efforts for environmental information management and exchange. The information needs are broken down into regulatory information (compliance information, material content, etc.) and voluntary information (lifecycle assessment, cradle to cradle, etc).

Keywords. Data Management, Environment, Sustainability, Electronics, Material Composition, Life Cycle Assessment, Manufacturing

1 Intersection of Concurrent Engineering and Environmental Concerns

Governments and industries have begun to embrace the idea of reducing the impact our manufacturing has on the global environment. Taken as a whole, the global environment can be considered a single system containing smaller internal systems including those focused product manufacturing. From a system perspective, determining total environmental impact of a product requires looking at the impact from every stage of its lifecycle including manufacturing, use, and reclamation. This "cradle-to-cradle" approach, where all outputs from the product lifecycle are

[1] Corresponding Author : Electrical Engineer, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD, 20899, US; Tel: 301-975-3956; Fax: 301-975-6021; Email: eric.simmon@nist.gov; http://www.nist.gov

inputs into other manufacturing processes, is a change from a traditional cradle-to-grave approach where material is extracted from the earth, products are manufactured and used, and then the resulting waste is returned into the environment. Reducing the environmental impact for a given product will require optimizing this extremely complex system (see figure 1). Any such effort will require detailed information about resources, environmental data, design choices, etc. to flow through the different segments of a product's lifecycle.

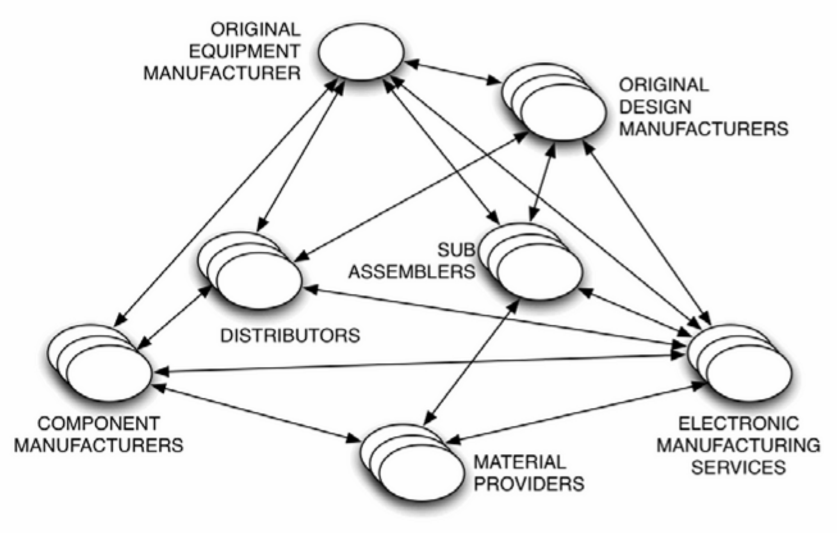


Figure 1: Modern Manufacturing Network

One key type of information that will need to be tracked by manufacturers is a compendium of the resources used during a product's lifecycle. Resources, in this case, can be usual broken down into material resources and energy resources. Material and energy are input into the product's manufacturing, during product usage, and during reclamation at its end of life. Ideally, these resources can be reclaimed at the end of life and could be used as input to other systems (but often ends up as waste).

The electronics industry has been a leader in the change towards sustainable design and manufacturing. One reason for this is that electronic products have a large impact on the environment due to their high volume production and quick product obsolescence. This has created the drive for the electronics industry to move to a true lifecycle-based approach. This paper will discuss some of the specific information that should flow through the manufacturing network to create an environmentally positive product lifecycle from the perspective of the electronics industry.

2 Information Management

To achieve the goal of cradle-to-cradle sustainability, the different stages of a product's lifecycle will need to adapt and the entire system optimized. Information must flow throughout the entire system so that each piece can be optimized relative to the whole system. Some examples include material content information going from the raw material producers to the product designers, compliance information going from the suppliers to the environmental specialists, and disassembling instructions going from the producers to the recyclers. Due to the complexity of modern electronics manufacturing networks, where one company rarely controls every aspect of production, ensuring this flow of information between the many participants is a difficult task.

Over time the manufacturing network has evolved into a complex heterogeneous system where each company has its own internal data systems and approach to doing business. At the same time the suppliers must interact with the other stakeholders exchanging goods, services, and information. An example of this is the methods companies use to respond to the environmental regulation of hazardous substances. Based on differing risk assessments some companies chose to track only yes/no compliance information, while others are tracking basic substance information and still others track the complete material composition of the products. These different management philosophies and the different data systems used to implement them can create a barrier to information flowing through the supply network. Regardless of philosophy, the first step of the information flow is to identify what information needs to be exchanged. Once there is an understanding of the necessary information, suitable information management systems and networks can be established. Most of the required environmentally related information can be divided into two categories, regulatory and voluntary. Regulatory covers the direct and indirect information needs imposed by governmental regulation. Voluntary information covers additional information that can be used for business goals such as design for sustainability and lifecycle assessment.

3 Environmental Regulations

Environmental regulations established some of the first requirements for the exchange of environmental data within the electronics supply network. Originally, most electronics manufacturing companies did not store or track material information as it was considered not functionally relevant. When these regulations go into effect, new data management systems need to be developed and existing systems modified to handle this new class of data.

These new environmental regulations cover a wide range of issues going from the specific (such as banning hazardous substances) to the general (such as promoting energy conservation). To truly understand the overarching informational requirements imposed by these new regulations requires taking a closer look at both key current regulations and potential future regulations.

3.1 Current Regulations

The focus of many recent environmental regulations has been on protecting both human health and the health of the environment. This has largely taken two forms: reducing risks associated with key substances and working to reduce waste from entering landfills (such as through recycling). To understand the basic data needed to support these two activities, this paper will look at two different European Union laws: The Restriction of Hazardous Substances (RoHS) Directive and the Regulation, Evaluation, Authorization and Restriction of Chemical Substances (REACH) Regulation.

EU RoHS

EU RoHS Directive (2002/95/EC) [1] sets restrictions on the use of certain hazardous substances in electrical and electronic equipment thereby banning non-compliant products. Determining compliance is a combination of calculating the amount of the restriction substances in the product comparing the levels to a predetermined maximum concentration value (based on the Directive) and identifying whether the product or substance has an exemption from the Directive. There are currently no direct declaration informational requirements; however it is the simple act of bringing a product into the EU that acts as a “self declaration” of compliance by the final producer/importer. For a final producer to know a product is in compliance generally requires tracking material content information through the supply chain. At a minimum this is done through the use of a compliance statement whereby the supplier simply states the product is RoHS compliant. The compliance statement places the burden on supplier, and the customer must trust that the supplier has provided accurate information. By gathering a summary of information on the material composition, companies have better idea of compliance which can protect against legal action due to accidentally importing non-compliant products (the process of performing checks to ensure the data are accurate is known as performing due diligence). Potential types of information to track to verify compliance status include full material composition declarations, material testing results, and auditing procedures. Still, there is no right answer to how much information is needed, and each company must determine its own needs based on its risk assessment. Only then can a company properly invest in compliance and information systems.

EU REACH

REACH (EC NO 1907/2006) [2] focuses on the chemical industry and chemical safety and was created to improve Europe’s existing chemical regulatory system. The previous system suffered from being a patchwork system created over the course of the EU’s history. One of the major issues the previous system had was that it codified differences in safety requirements for chemical substances produced before and after 1981. Requirements differed for chemicals based on the date they were being used in products rather than the chemicals’ physical properties. The new system was designed to address those concerns by creating a registration system that is applied equally to all substances new and old. While REACH was not specifically targeted at the electronics industry some of its provisions do impact the industry.

The registration process requires that the manufacturers and importers of substances provide the information needed to assess the risks associated with each substance and how those risks are managed. Information needs to flow both ways in the supply chain and to a new central European Chemicals Agency (ECHA) [3]. For the electronics industry, one of the key pieces of information that will need to be reported is again related to material content. Substances on the REACH Substances of Very High Concern (SVHC) list contained within a manufactured product must be reported. Specifically, manufacturers or importers of articles (defined as products whose physical properties are more important to its function than any other chemical properties) must notify their customers of the presence of any SVHC substance in excess of 0.1% weight by product.

3.2 Future Regulations

Evaluating information needs for future environmental regulation is naturally not an exact science. However, by looking at pending legislation and customer concerns it is possible to identify two areas that are likely to be codified into future regulations. Two such areas are energy and water usage.

Energy Usage

Most countries are limited in their ability to increase energy supplies thus leaving energy conservation as one of the few viable approaches to meeting their energy requirements. One area in which countries are looking to promote energy conservation is by regulating energy using products such as electronics. The idea being that if electronic products are designed to be more energy efficient it would automatically lead to reduced energy demand. To encourage companies to be more focused on producing products that conserve energy, some countries are now exploring legislation that will require companies to report the energy used by a product during it's lifecycle and to reduce the energy used during the product lifecycle.

An example of this type of legislation is the European Union's Energy Using Products (EuP) Directive [4] that establishes a framework by which eco-design requirements for energy-using products can be established. Information gathered in various studies will then be used to create future Implementing Measures (IM) that will encourage industry to move to more eco-friendly product designs. While the ongoing studies are looking at energy consumed during product usage, the EuP covers a product's entire lifecycle leaving open the possibility of future IMs covering manufacturing processes.

Using the EuP as a guide, it is easy to predict some of the future informational needs concerning energy. Information will need to be collected by industry to calculate energy consumption, both in product manufacturing and product usage. Without that information it will be impossible for companies to make informed decisions to design energy efficient products. To achieve this reduced energy consumption, usage information must flow from each step in the product manufacturing and use to the product designers. This information could include how much energy is used during each manufacturing step and what type of energy is used.

Water Usage

Water scarcity (when water demand outpaces supply) often caused by natural (droughts) and artificial (civil projects, consumer use, industry use) is a growing worldwide concern [5]. Even countries typically considered water rich such as the USA [6] and the European Union are facing these issues. It is estimated that 10% of the European Union's population and 20% of its territory are affected [7]. Concerns that climate change will significantly both increase droughts and reduce water supplies have countries looking for ways to recover and stretch existing water supplies.

Without specific regulations to look at, it is hard to predict the informational needs on an issue that is only now being studied in depth. Still there have been several first steps in field of water conservation (reports, commissions, directives) that have attempted to frame some guiding principles that will be need to address the problem. For an industry looking to identify its future information needs concerning water usage, a good starting point for companies would be to start tracking water usage data, such as total water needed for a given manufacturing process and water-borne waste.

3.3 Voluntary Information - Sustainable Lifecycle Design and Manufacturing

Moving beyond the need to comply with environmental regulations, many organizations have begun to see sustainable manufacturing as a competitive advantage [8]. While there is no uniformly accepted definition of sustainability in terms of manufacturing, the concept is that companies will be able to manufacturer products that do not deplete resources and therefore do little damage the environment. This would require a new paradigm in how products are designed and manufactured where inputs to the manufacturing process will come from the outputs of other systems. Moving toward this goal will require a substantial amount of information to be exchanged beyond that which is currently required by law. The bulk of this information will focus on quantifying the matter and energy inputs and outputs of a particular system in order to understand the impact of the product lifecycle "system" on the global ecosystem.

Life Cycle Assessment

One technique being used to make sustainability decisions is Lifecycle Assessment (LCA) [9], a method for compiling and assessing the inputs and outputs of a manufacturing system to determine the potential environmental impact of a product's lifecycle (i.e., cradle-to-grave). LCA consists of three distinct phases: scope determination, cataloguing the system's inputs and outputs, and impact assessment. Ideally, once the LCA data for each piece of a system (a system in its own right) has been compiled, manufacturers would have the information to make informed decisions on how to integrate the product lifecycle into the bigger global system that is our environment. Potential benefits of using LCA include; reducing the total environmental impact for a given product or process, improving resource utilization, and identifying product inefficiencies. Still, it is important to realize that performing a LCA does not automatically produce a better product or process; rather it is up to the producers to use the information wisely.

The biggest problem associated with LCA is it is an information-intensive process, and gathering the required data (with enough accuracy) can be very difficult and time consuming. As manufacturing environmental information management systems are developed and improved, the information needed to perform a LCA during final production will be captured and made available. This will allow the goal and scope information from stage one of LCA and the Life Cycle Inventory (LCI) information from stage two of LCA to be transferred up the supply chain. The final producer can then perform a lifecycle impact assessment for stage three of the LCA.

The goal and scope information covers the reasons for the study, including the intended audience (in this case it should be the supply chain stakeholders), what system is being studied, the functional unit being studied, and the boundaries placed around this unit. The flexibility companies have in defining the goals and scope is another major problem with LCA. This is a problem that is difficult to solve, but as industry gains more experience with LCA consensus should be achieved and a unified approach to defining stage one will be developed. The core of LCA may be the LCI. This stage is where the matter and energy inputs and outputs are quantified. The LCI catalogues all of the system's inputs (raw materials, ancillary inputs including and process chemicals and energy, other inputs (such as transportation energy), etc.) and all of the system's outputs (products, co-products, solid waste, emissions, etc.).

Cradle-To-Cradle (C2C)

The next step for true sustainability is to move beyond the cradle-to-grave philosophy to the cradle-to-cradle approach. To achieve this, LCA information must flow not only within a manufacturing network, but between manufacturing networks and between different industries. This will require collaboration between industry organizations to develop standardized LCA methodologies and data exchange standards.

4 Conclusion

Environmental concerns are some of the major driving forces that are effecting the electronics manufacturing sector. Whether due to regulation or market forces, companies are changing the way they do business to lessen the impact on the environment. To minimize environmental impact and create sustainable manufacturing systems, information must flow through the entire manufacturing, repair, and recycling network. One way to organize this information is to recognize that all this environmental information can be placed into two groups: matter and energy. This grouping includes materials used for manufacturing and maintenance and energy used during manufacturing, product usage, and recycling.

Once these basic building blocks are understood, the bigger problem of modernizing the information systems infrastructure within the supply chain can be dealt with. A perfect example of this situation is illustrated by the electronics industry's response to the EU RoHS Directive. In addition to identifying ways to modify manufacturing processes in order to comply with RoHS, a whole new information infrastructure was needed in order to make sure this information could easily flow through the supply chain. Over a two-year period the electronics industry raced to have a supporting data exchange standard in place before RoHS went into effect. In the end several organizations created solutions for RoHS (such as IPC 1752 [10]). All this effort was expended in response to just one new environmental regulation. This approach of creating a new data exchange standard for each new piece of regulation is not feasible in a global environment where every country is creating its own environmental regulations.

As companies move beyond just trying to comply with regulations and work towards a cradle-to-cradle approach, the information management system will again have to adjust. To integrate these new goals into the manufacturing network, data from every part of the lifecycle will need to be communicated throughout the manufacturing information network and even across traditional industry boundaries.

A feasible method moving forward is for the electronics industry to take a more concurrent approach by developing information management solutions that are capable of dealing with multiple pieces of regulations at once. This will require shifting to a concurrent engineering approach, treating the environment as a single system and connecting the stages of manufacturing not only within a specific industry but all industries. The environmental information management systems built to achieve this can categorize the environmental information into the two areas mentioned earlier: energy and materials. Focusing on these two sets of information would cover most information needs required by both current and expected future environmental regulations. By using good systems design and modeling practices, it should be possible to create system models that will help industry understand and design the information management systems and ease the development of new data exchange standards [11].

While these changes may appear daunting, the observations in this paper should help facilitate the move toward a comprehensive environmental information management infrastructure. Once that infrastructure is in place, it will assist the move towards greater sustainability in manufacturing.

5 Citations and References

[1] RoHS Directive website available at: <http://www.rohs.eu/english/index.html>. Accessed on: February 11, 2009.

[2] REACH Home Page Available at: <http://ec.europa.eu/environment/chemicals/reach/reach_intro.htm>. Accessed on: February 11, 2009.
[3] ECHA Home Page. Available at: < http://echa.europa.eu/home_en.asp >. Accessed on: February 11, 2009.
[4] Eco-Design of Energy Using Products. Home Page. Available at: <http://ec.europa.eu/enterprise/eco_design/index_en.htm>. Accessed on: February 11, 2009.
[5] L. Keating, H. Passell, "Addressing Our Global Water Future," Center for Strategic and International Studies and Sandia National Laboratories, September 30, 2005.
[6] Calif. Drought Prompts State of Emergency. (2009, February 28). *The Washington Post.* p. A-14.
[7] Water Information System For Europe (WISE) Note 5. Available at: <http://ec.europa.eu/environment/water/water-framework/pdf/water_note5_economics.pdf >. Accessed on: February 11, 2009.
[8] C. Laszlo, "Sustainability for Competitive Advantage" (2008, May 14). *The San Francisco Chronicle*. p. B-9.
[9] Environmental Protection Agency's Life-cycle Assessment Research Page. Available at: < http://www.epa.gov/nrmrl/lcaccess/>. Accessed on: February. 11, 2009.
[10] IPC 1752 Home Page Available at: <http://members.ipc.org/committee/drafts/2-18_d_MaterialsDeclarationRequest.asp>. Accessed on: February 11, 2009.
[11] J. V. Messina, E. D. Simmon, "Managing Material Information in the Supply Chain," FAB Engineering and Operations, No. 5, pp. 70-74 (01-Nov-2008).

Rule-Based Recursive Selective Disassembly Sequence Planning for Green Design

Shana Smith[1] and Wei-Hsiang Chen
Department of Mechanical Engineering, National Taiwan University

Abstract. Disassembly sequence planning not only reduces product lifecycle cost, but also greatly influences environmental impact Many prior green design research studies have focused on complete disassembly of an end-of-life product to recover valuable components. However, complete disassembly is often not practical or cost effective if only a few components will be recovered and recycled from a given product. Selective disassembly sequence planning focuses on disassembling only one or more selected components from a product for reuse, recycling, maintenance, and remanufacturing. This paper presents a rule-based recursive method for finding an optimal heuristic selective disassembly sequence for green design. Most prior methods either enumerate all solutions or use a stochastic method to generate random solutions. Enumerative or stochastic methods often require tremendous computational resources while, at the same time, they often fail to find realistic or optimal solutions. On the contrary, the proposed method establishes certain heuristic disassembly rules to eliminate uncommon or unrealistic solutions. In addition, rather than considering geometric constraints for each pair of components, the developed method only considers geometric relationships between a part and its neighboring parts. As a result, the developed method can effectively find an optimal heuristic solution while greatly reducing computational time and space.

Keywords. Selective disassembly sequence planning, rule-based, recursive, green design.

1 Introduction

Selective disassembly sequence planning is a powerful and an efficient tool for solving de-manufacturing (DM) problems. DM involves separating certain components and materials from a product for reuse, recycling, replacement, and maintenance to reduce product lifecycle cost [6]. However, finding an optimal selective disassembly sequence is a very difficult and complex problem when multiple factors are involved, e.g., disassembly time, cost, reorientations, tools, and environmental regulations.

Some prior studies have utilized advanced searching algorithms to find optimal selective disassembly sequences. For example, Srinivasan et al. applied a wave propagation method to solve selective disassembly problems [19, 20, 21]. Their method includes disassembling one component or more components, as well

[1] Corresponding Author E-mail: ssmith@ntu.edu.tw

as total selective disassembly [19]. However, they only evaluate each disassembly sequence by the total number of removed components. They consider the optimal selective disassembly sequence to be the sequence with the minimum number of removals [19, 20, 21]. Although their evaluation function is simple, it might not satisfy the demands of many realistic product design problems. Therefore, Chung and Peng added more evaluation criteria, e.g., disassembly time, tool changes, and cost, to selective disassembly sequence planning using the wave propagation method [7].

Ant colony optimization (ACO) algorithms and genetic algorithms (GA) have also been used for optimal selective disassembly sequence planning [7, 23, 24, 25, 27]. Prior methods often consider all geometric constraints of the assembly and evaluate each selective disassembly sequence, with respect to number of reorientations and number of removed components, to find the optimal solutions.

Kara et al. reversed and modified assembly sequences and used a liaison diagram to evaluate geometric connections to find selective disassembly sequences [14]. Aguinaga et al. used a rapid-growing random tree method to find optimal sequences. However, their method generates many paths, and, thus, it takes a significant amount of time to find optimal sequences. In addition, their results might not be consistent [2]. Shyamsundar and Gadh developed a recursive method that considers both separation direction and disassembly direction to disassemble a target component [18]. Since their method allows a component to be disassembled from any direction, the input information required regarding separation and disassembly directions is relatively complicated. Their method also requires input information concerning sequential disassembly and additional disassembly constraints, which makes their method difficult to use for general products. Srinivasan and Gadh also developed a global selective disassembly method, which includes spatial constraints and user-defined constraints [22].

Most searching methods use specific information in their searching processes: geometric constraints [2, 11, 15, 22, 23, 26], topological positions [7, 15, 22, 26], liaison relationships [14], AND/OR graphs [2, 11], precedence graphs [11], fastener accessibility [5], and component accessibility [22]. Criteria used to evaluate disassembly sequences include the number of removed components [1, 7, 11, 15, 22, 23, 24, 26], disassembly time [2, 6, 13, 14, 25], reorientations [23, 25, 27], and tool changes [7, 25, 27]. To reduce disassembly cost, most selective disassembly sequence planning methods focus on minimizing the number of removed components, disassembly time, and reorientation time.

Selective disassembly sequence planning research aims to find optimal solutions to the selective disassembly planning problem. However, finding an optimal solution is a difficult problem. Most prior methods either enumerate all solutions or use stochastic methods to generate random solutions. Methods which enumerate all solutions can find optimal solutions. However, they might require a tremendous amount of computational resources. Therefore, they are generally not practical for solving realistic product design problems. As a result, most recent methods aim to find near-optimal or heuristic solutions. Stochastic random methods, such as ACO and GA, might generate solutions which meet geometric and topological constraints. However, the given solutions might not be practical for use in reality.

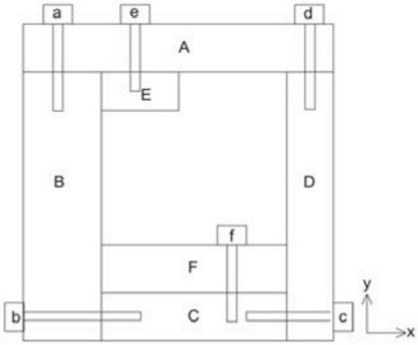


Figure 1. Sample example

This paper presents a rule-based recursive method for obtaining optimal heuristic selective disassembly sequences. The method uses certain disassembly rules to eliminate uncommon or unrealistic solutions. In addition, rather than consider geometric constraints for each pair of components, the developed method only considers the geometric relationship between a part and its neighboring parts. The topological information and fastener accessibility of a product is examined from inner to outer components. As a result, the developed method can effectively find an optimal heuristic solution while greatly reducing computational time and space. Our method can handle both single-component and multiple-component disassembly problems. The evaluation function includes disassembly time, reorientations, and number of removed fasteners and components.

2 Definitions

2.1 Disassembly Parameters for Fasteners:

We define a disassembly parameter matrix for fasteners, *DF*, which records the disassembly directions of each fastener. For each fastener, the tuples of valid disassembly directions in *DF* are set to 0, while other tuples are set to *N*. For example, in Figure 1, the only valid disassembly direction for fastener *a* is in the +y direction. Thus, DF_a $(+x : –x : +y : -y) = (N : N : 0 : N)$. Similarly, DF_b $(+x : –x : +y : -y) = (N : 0 : N : N)$, DF_c $(+x : –x : +y : -y) = (0 : N : N : N)$, DF_d $(+x : –x : +y : -y) = (N : N : 0 : N)$, DF_e $(+x : –x : +y : -y) = (N : N : 0 : N)$, and $DF_f = (+x : –x : +y : -y) = (N : N : 0 : N)$. Finally, $DF = [\mathrm{D}F_a\ DF_b\ DF_c\ DF_d\ DF_e\ DF_f]^{\mathrm{T}}$:

$$DF = \begin{bmatrix} DF_a \\ DF_b \\ DF_c \\ DF_d \\ DF_e \\ DF_f \end{bmatrix} = \begin{bmatrix} N & N & 0 & N \\ N & 0 & N & N \\ 0 & N & N & N \\ N & N & 0 & N \\ N & N & 0 & N \\ N & N & 0 & N \end{bmatrix}$$

In this study, we assume all fasteners can only be disassembled in one direction. If a fastener is welded to another component, it becomes an integral part of that component.

2.2 Disassembly Parameters for Components:

We define a disassembly parameter matrix for components, *DC*, which records the immediately touched components and fasteners which constrain the motion of a target component in only one direction of a principal axis. Parts constrain the motion of a target component in both directions of a principal axis will be considered in the next session .We use lower case letters to represent fasteners and upper case letters to represent components.

For example, in Figure 1, $DC_A\,(+x : –x : +y : -y)$ =(0 : 0 : *ade* : *BD*). Since fasteners *a*, *d*, and *e* constrain component *A* in both the +*x* and –*x* directions, we do not include them in the component disassembly parameter matrix *DC*. Similarly, $DC_B\,(+x : –x : +y : -y) = (CEF : b : Aa : 0)$, $DC_C\,(+x : –x : +y : -y) = (Dc : Bb : Ff : 0)$, $DC_D\,(+x : –x : +y : -y) = (c : CF : Ad : 0)$, $DC_E\,(+x : –x : +y : -y) = (0 : B : Ae : 0)$, $DC_F\,(+x : –x : +y : -y) = (D : B : f : C)$. Finally, $DC = [DC_A\ DC_B\ DC_C\ DC_D\ DC_E\ DC_F]^T$:

$$DC = \begin{bmatrix} DC_A \\ DC_B \\ DC_C \\ DC_D \\ DC_E \\ DC_F \end{bmatrix} = \begin{bmatrix} 0 & 0 & ade & BD \\ CEF & b & Aa & 0 \\ Dc & Bb & Ff & 0 \\ c & CF & Ad & 0 \\ 0 & B & Ae & 0 \\ D & B & f & C \end{bmatrix}$$

2.3 Motion Constraint Parameters:

We define two motion constraint parameter matrices, one records motion constraints for fasteners (*MF*), and the other records motion constraints for components (*MC*). Before *MF* and *MC* can be defined, “first-level components” need to be defined. First-level components are components which do not immediately touch the target components or fasteners but which are the first components beyond the immediately touching components which would block movement of fasteners or target components in given moving directions.

The *MF* matrix records both the first-level components of each fastener and any immediately touching components of the fastener in a given disassembly direction. For example, in Figure 1, fastener *f* can only be disassembled alone the +*y* direction. However, since component *A* is the first component which fastener *f* would collide with, in the given disassembly direction, component *A* is a first-level component of fastener *f*. Thus, $MF_f(+x : –x : +y : -y) = (0 : 0 : A : 0)$.

The disassembly parameter matrix for components, *DC*, only records immediately touching components and fasteners which constrain motion of a target component in one direction of a principal axis. In contrast, *MC* records only the first-level components of a target component, omits fasteners, and includes immediately touching components and fasteners which constrain motion of the target component in both directions of a principal axis. For example, in Figure 1, the first-level component of component *F* is component *E*, and component *F* is also constrained by fastener *f* in both directions of the x-axis. Thus, $MC_F (+x : –x : +y : -y) = (f : f : E : 0)$.

Figure 2 shows a second example. The assembly in Figure 2 includes two components, *A* and *B*, and one fastener *a*. For the given assembly, $DC_A (+x : –x : +y : -y) = (0 : 0 : aB : 0)$, $MC_A (+x : –x : +y : -y) = (Ba : Ba : 0 : 0)$, $DC_B (+x : –x : +y : -y) = (0 : 0 : a : A)$, and $MC_B (+x : –x : +y : -y) = (Aa : Aa : 0 : 0)$.

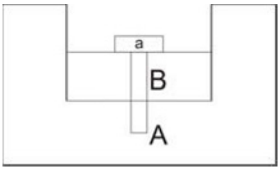


Figure 2. Second example

3 Selective Disassembly Planning Process

We establish six rules which define our recursive selective disassembly planning process. In the rules below, *n* is the target component which has been selected for disassembly, and a parent component is a component which has already been selected for disassembly, for the given incomplete disassembly sequence.

Rule 1: IF (there is any fastener attached to a new target component) THEN (the fastener needs to be disassembled first)

Rule 2: IF (there are corresponding tuples in both DF_j and MF_j which are 0) THEN (disassemble fastener *j* along the direction associated with the tuples)

Rule 3: IF (there are no corresponding tuples in both DF_j and MF_j which are 0) THEN (place the component which is in MF_j but which is not a parent component of *n* in the new target components queue)

Rule 4: IF (there are corresponding tuples in both DC_n and MC_n which are 0) THEN (one disassembly sequence can be found by disassembling component *n* and all of its parent components in reverse order

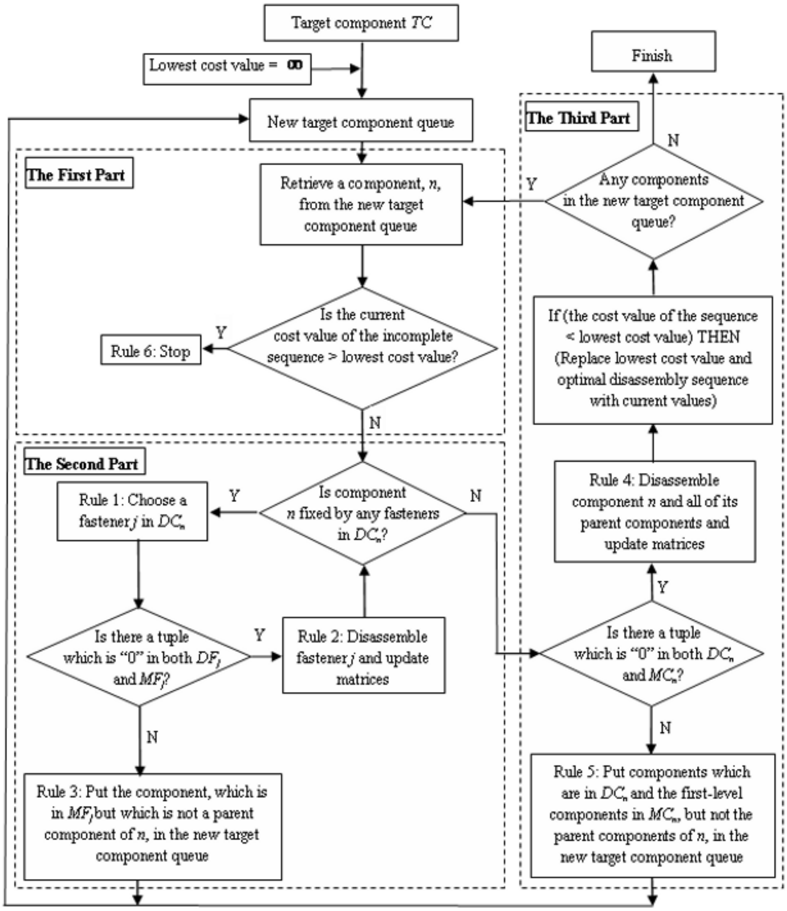


Figure 3. The complete flowchart of the rule-based recursive disassembly

Rule 5: IF (there are no corresponding tuples in both DC_n and MC_n which are 0) THEN (place components which are in DC_n and the first-level components which are in MC_n, but not the parent components of n, in the new target components queue)

Rule 6: IF (the cost value of an incomplete disassembly sequence is greater than an existing one) THEN (stop searching)

As shown in Figure 3, our recursive selective disassembly planning method includes three basic parts. The first part adds one additional component to the current incomplete disassembly sequence and calculates a cost value for the current incomplete sequence. If the cost value is greater than the lowest cost value of a previous searched sequence, the search process stops.

The second part determines if the new target component *n* is fixed by any fasteners and disassembles all the fasteners first. If a fastener is disassembled, it will be removed from the *DC, MC,* and *MF* matrices. If a selected fastener is constrained by any component in *MF*, place the component which is in *MF* but which is not a parent component of *n* in the new target components queue. The first step repeats until all the fasteners attached to the target component are disassembled.

For example, in Figure 2, if component *B* is a target component, since DC_B $(+x : -x : +y : -y) = (CEF : b : Aa : 0)$, fasteners *a* and *b* need to be disassembled first. Since $DF_a\ (+y) = MF_a\ (+y) = 0$, fastener *a* can be disassembled in the +y direction. After fastener *a* is removed, $DC_B\ (+x : -x : +y : -y) = (CEF : b : A : 0)$. Since DF_b (-x) = MF_b (-x) = 0, fastener *b* can be disassembled in the -x direction. After fastener *b* is removed, $DC_B\ (+x : -x : +y : -y) = (CEF : 0 : A : 0)$, and all the fasteners attached to component *B* have been removed.

From Figure 3, the third part of the selective disassembly process generates a disassembly sequence, if *n* is not fixed by any fasteners and *n* can be disassembled in any direction. The disassembly sequence is generated by disassembling component *n* and all of the parent components of *n*, which is accomplished by updating the *DC, MC,* and *MF* matrices. If the cost value of the sequence is lower than the lowest cost value of any previously generated sequence, the sequence is recorded as the most optimal sequence and the cost values is recorded as the lowest cost value. If *n* cannot be disassembled in any direction, the components in DC_n and the first-level components of *n* in MC_n are added to the new target component queue. The process repeats until there are no new target components.

For example, in Figure 2, after fasteners *a* and *b* are removed, DC_B (-x) = MC_B (-x) = 0, Thus, component *B* can be disassembled in the -x direction. Component *B* is then deleted from *DC, MC,* and *MF.*

The process is rule-based and recursive. Thus, all possible solutions are not generated and checked. However, the method generates reasonable and near-optimal heuristic solutions both efficiently and effectively. The given rules reduce searching time by eliminating unrealistic and uncommon solutions.

4 Cost Function

Our cost function for evaluating disassembly sequences includes disassembly time, reorientations, and number of components and fasteners removed.

$$Cost\ value = w_1 \times time + w_2 \times reorientations + w_3 \times parts \qquad \text{Eq. (1)}$$

In Equation 1, we can choose weight values *w1*, *w2*, and *w3* to establish the weighted importance of each of the cost parameters in determining the outcome of the search process. The final best or optimal selective disassembly sequence which is found has the lowest cost value.

4.1 Time

Some prior studies use time as an evaluation parameter. However, the time values cannot be easily verified. Boothroyd et al. [3] reported experimental times, which we use for comparison purposes. Boothroyd et al.'s reported time values consider the effects of part symmetry, grasping or manipulating with hands or with the aid of grasping tools, and parts which are either inserted without being immediately secured, as well as parts which are immediately secured by screws fastened with power tools [27].

4.2 Reorientations

During disassembly, if the number of disassembly direction reorientations is reduced, disassembly time is also reduced [23, 24, 25, 27]. Since we only consider principal disassembly directions, each reorientation requires either a 90-degree or a 180-degree direction change. For example, if the disassembly direction changes from $+x$ to $+y$, $-y$, $+z$, or $-z$, the reorientation requires a 90-degree direction change, for which we increase the reorientations cost function parameter by 1. However, if the disassembly direction changes from $+x$ to $-x$, the reorientation requires a 180-degree direction change, for which we increase the cost function parameter by 2. When no reorientations are needed, we set the reorientations cost function parameter to 0.

4.3 Parts

Many research studies consider the problem of reducing the number of components which need to be removed. Most prior methods include number of components in their cost functions for evaluating the quality of a disassembly sequence [2, 11, 15, 21, 22, 23, 24, 26]. With our cost function, if fewer parts need to be removed to disassemble a target component, the cost of the disassembly sequence is lower and less time is also required to complete the disassembly process.

5 Case Study

5.1 Example 1

We used two examples to test our rule-based recursive selective disassembly method. Figure 4 shows the first example. Figure 5 shows the corresponding *DC* and *MC* matrices. For target component *C*, there are no fasteners. Since $DC_C = (0 : 0 : D : B)$ and $MC_C = (A : A : 0 : 0)$, there is no tuple which is 0 in both DC_C and MC_C, Therefore, Rule 5 is executed. Since component *A* is not a first-level component of *C*, only components *D* and *B* are put into the new target component queue.

If component *D* is retrieved next, there is no tuple which is 0 in both DC_D and MC_D. Therefore, Rule 5 is executed. Since components *C* is a parent component of component *D*, only component *E* is put into the new target component queue. If component *E* is retrieved next, $DC_E = (0 : 0 : 0 : D)$ and $MC_E =$

(A : A : 0 : 0). Therefore, component E can be removed in the +y direction. According to Rule 4, one disassembly sequence can be generated by removing component E, and all of its parent components, in reverse order. After component E is removed, DC_D is updated to (0 : 0 : 0 : C), and MC_D is updated to (A : A : 0 : 0). Thus, component D can be removed in the +y direction. Finally, one disassembly sequence, *E-D-C*, can be generated. Similarly, a second disassembly sequence, *A-B-C*, can be found. The two disassembly sequences are shown in Figure 6.

The assembly in Figure 4 has five components and, therefore, there are 5! = 120 possible disassembly sequence possibilities. However, our rule-based recursive selective disassembly planning can eliminate considering many unrealistic or uncommon solutions. The developed method can find optimal heuristic selective disassembly sequences quickly and effectively.

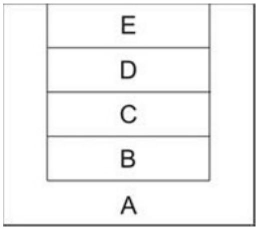


Figure 4. Example assembly

$$DC = \begin{bmatrix} DC_A \\ DC_B \\ DC_C \\ DC_D \\ DC_E \end{bmatrix} = \begin{bmatrix} 0 & 0 & B & 0 \\ 0 & 0 & C & A \\ 0 & 0 & D & B \\ 0 & 0 & E & C \\ 0 & 0 & 0 & D \end{bmatrix} \quad MC = \begin{bmatrix} MC_A \\ MC_B \\ MC_C \\ MC_D \\ MC_E \end{bmatrix} = \begin{bmatrix} BCDE & BCDE & 0 & 0 \\ A & A & 0 & 0 \\ A & A & 0 & 0 \\ A & A & 0 & 0 \\ A & A & 0 & 0 \end{bmatrix}$$

Figure 5. The *DC and MC* matrices for the example assembly in Figure 4

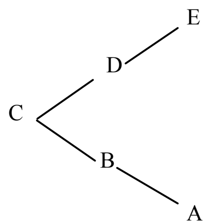


Figure 6. Disassembly sequences

5.2 Example 2

Figure 7 shows a gear reducer assembly from Scrinivasan and Gadh (2000). To simplify the problem, we only considered disassembly in the *x* direction.

5.2.1 Single Target Component

Following the process shown in Figure 3 to disassemble target component C_5, after fasteners f_1, f_2, f_3 are removed, DC_2 is be updated to ($C_1 : C_{3,8}$:0 : *0* : 0: 0 : 0 : 0), and MC_2 is updated to (0 : 0 : $C_{1,3,7}$: $C_{1,3,7}$: $C_{1,3,7}$: $C_{1,3,7}$).

By continuing to follow the process diagram, one disassembly sequence $f_1, f_2, f_3, C_1, C_2, C_3, C_4, C_5$, can be found. The disassembly sequence includes 3 fasteners and 5 components. Since Scrinivasan and Gadh (2000) did not consider time and reorientations in their study, to compare our results with their results, we set our cost function weights for time and reorientations to 0. Thus, our cost for the sequence = 0 + 0 + 8 = 8.

If, in our process, component C_6 is chosen first instead of C_4, the disassembly sequence found is: $f_6, f_7, f_8, f_9, C_{23}, C_{22}, C_{21}, C_{20}, C_{19}, C_6, C_5$, and the cost value = 0 + 0 + 11 = 11.

Comparing the two disassembly sequences, the first sequence is better than the second because the first only removes 8 elements, while the second sequence removes 11 elements. For single-component disassembly, in their example, Scrinivasan and Gadh (2000) chose component C_3 as the target component, and they did not consider disassembly of fasteners, in which case, C_1, C_2, C_3 is the obvious best disassembly sequence solution.

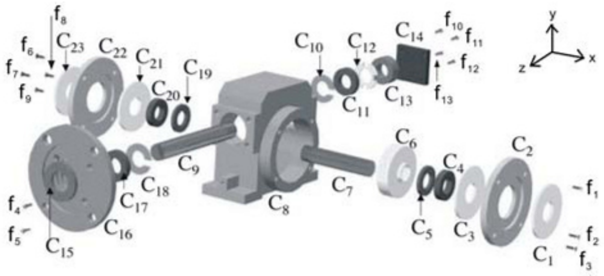


Figure 7. Gear reducer assembly.

6 Conclusions

In this paper, we present a rule-based recursive selective disassembly sequence planning method. The method can be used to solve selective disassembly sequence planning problems with only simple geometric and topological information concerning designs supplied by a user. The method is based upon six disassembly rules which are used to eliminate unrealistic and uncommon disassembly sequences. With the given search rules, the searching process can effectively and efficiently find reasonable and near-optimal selective disassembly sequence solutions to complex selective disassembly sequence planning problems. In addition, with the proposed rule-based recursive approach, users only need to supply information concerning the geometric constraints of a component with respect to its adjacent components and fasteners. Compared to methods which consider geometric constraints between each pair of components, the developed method greatly reduces required information storage space and searching complexity. The method can solve both single-target component and multiple-target component selective disassembly sequence planning problems. In the future, method for defining and disassembling subassemblies needs to be investigated. Compared to most existing methods, our method is much easier to implement for general products.

Acknowledgement

Support from the National Science Council, Taiwan, (NSC 97-2221-E-002-158-MY2) is gratefully acknowledged.

7 References

[1] Aguinaga, I., Borro, D., and Matey, L., Automatic Selective Disassembly and Path Planning for the Simulation of Maintenance Operations. IEEE Virtual Reality Conference 2007 March 10-14; 283-284.

[2] Aguinaga, I., Borro, D., and Matey, L., Parallel RRT-based path planning for selective disassembly planning. Int J Adv Manuf Technol 2008; 36: 1221-1233.

[3] Boothroyd, G., Dewhurst, P., and Knight, W., Product Design for Manufacture and Assembly Second Edition Revised and Expanded. Published by CRC Press Taylor & Francis Group. Boca RatonLondon New York, 2002; 96-103.

[4] Chen, S-F., Oliver, J. H., Chuo, S-Y., and Chen, L-L., Parallel disassembly by onion peeling. Transactions of The ASME, Journal of Mechanical Design 1997; 119: 267-274.

[5] Chung, C., and Peng, Q., An integrated approach to selective-disassembly sequence planning. Robotics and Computer-Integrated Manufacturing 2005; 21: 475-485.

[6] Chung, C., and Peng, Q., A hybrid approach to selective-disassembly sequence planning for de-manufacturing and its implementation on the Internet. Int J Adv Manuf Technol 2006; 30: 521-529.

[7] Chung, C., and Peng, Q., Sequence planning for selective disassembly in de-manufacturing. International Journal of Computer Integrated Manufacturing 2006; 19: 3: 278-286.

[8] Directive 2002/95/EC of the European Parliament and of the Council of 27 January 2003 on the restriction of the use of certain hazardous substances in electrical and electronic equipment. Official Journal of the European Union, 2003.

[9] Directive 2002/96/EC of the Europoean Parliament and of the Council of 27 January 2003 on waste electrical and electronic equipment. Official Journal of the European Union, 2003.

[10] Directive 2005/32/EC of the European Parliament and of the Council of 6 July 2005. Official Journal of the European Union, 2005.

[11] García, M. A., Larré, A., López, B., and Oller, A., Reducing the Complexity of Geometric Selective Disassembly. Proceedings of the 2000 IEEE/RSJ International

Conference on Intelligent Robots and Systems. 1474-1479.

[12] Giudice, F., and Fargione, G., Disassembly planning of mechanical systems for service and recovery: a genetic algorithms based approach. J Intell Manuf 2007; 18:313–329.

[13] Kara, S., Pornprasitpol, P., and Kaebernick, H., A selective disassembly methodology for end-of-life products. Assembly Automation 2005; 25: 2: 124–134.

[14] Kara, S., Pornprasitpol, P., and Kaebernick, H., Selective Disassembly Sequencing: A Methodology for the Disassembly of End-of-Life Products. Annals of the CIRP 2006; 55: 1: 37-40.

[15] Mascle, C., and Balasoiu, B. -A., Algorithmic selection of a disassembly sequence of a component by a wave propagation method. Robotics and Computer Integrated Manufacturing 2003; 19: 439-448.

[16] Mascle, C., and Hong, P -Z., Integrating environmental consciousness in product/process development based on life-cycle thinking. Int. J. Production Economics 2008; 112: 5-17.

[17] Pitipong, V., and Surendra M., G., Optimal analysis of lot-size balancing for multiproducts selective disassembly. International Journal of Flexible Automation and Integrated Manufacturing 1998; 6:3: 245-269.

[18] Shyamsundar, N., and Gadh, R., Selective Disassembly of Virtual Prototypes. Proceedings of the IEEE International Conference on Systems, Man and Cybernetics 1996; 4: 3159-3164.

[19] Srinivasan, H., and Gadh, R., Selective disassembly: Representation and Comparative Analysis of Wave Propagation Abstractions in Sequence Planning. Proceedings of the 1999 IEEE International Symposium on Assembly and Task Planning Porto, Portugal, 1999; 129-134.

[20] Srinivasan, H., Figueroa, R., and Gadh, R., Selective disassembly for virtual prototyping as applied to de-manufacturing. Robotics and Computer Integrated Manufacturing 1999; 15: 231-245.

[21] Srinivasan, H., and Gadh, R., Efficient Geometric Disassembly of Multiple Components from an Assembly Using Wave Propagation. Jouranl of Mechanical Design 2000; 122: 179-184.

[22] Srinivasan, H., and Gadh, R., A non-interfering selective disassembly sequence for components with geometric constraints. IIE Transactions 2002; 34: 349-361.

[23] Wang, J, -F., Liu, J. -H., Li S. -Q., and Zhong, Y. -F., Intelligent selective disassembly using the ant colony algorithm. Artificial Intelligence for Engineering Design, Analysis and Manufacturing 2003; 17: 325–333.

[24] Wang, J. –F., Li, S. –Q., and Liu, J. –H., Selective disassembly planning for product green manufacturing. Computer Integrated Manufacturing Systems June 2007; 13: 6: 1097-1102.

[25] Xue, J. –F., Qiu, C. –H., and Xiang, D., Planning of Selective Disassembly Sequence Based on Ant Colony Optimization Algorithm. Journal of Computer Aided Design & Computer Graphics 2007; l119: 16: 742-747.

[26] Yi, J., Yu, B., Du, L., Li, C., amd Hu, D., Research on the selectable disassembly strategy of mechanical parts based on the generalized CAD model. Int J Adv Manuf Technol 2008; 37: 599-604.

[27] Zhang, X. -H., Li, S. -Q., Wang, J. -F., and Li, X., Single object selective disassembly sequence planning based on ant colony algorithm. Computer Integrated Manufacturing Systems 2007; 13: 6: 1110-1114.

Design Knowledge Utilization

Investigation on Evaluation of Design Decision for Door-Shaped Structure by Using Systematic Knowledge Analysis

Zone-Ching Lin[a,1], Chen-Hsing Cheng[b]

[a] Department of Mechanical Engineering, National Taiwan University of Science and Technology – Taiwan, R.O.C.
[b] Graduate Institute of automation and Control, National Taiwan University of Science and Technology –Taiwan, R.O.C.

Abstract. The paper employs the concept of relational knowledgebase and revised analytical hierarchy process (AHP) to carry out systematic analysis on evaluation of design decision for door-shaped structure. The knowledge audit proposed by the study includes structure kind, structure subsystem, technique function, boundary condition, limitation condition and applied engineering principle method. By means of the relativity of classified hierarchical knowledge, the study establishes relational engineering knowledgebase. The paper proposes calculating the ratios of occurrence numbers of related keywords from the relative patent and engineering knowledge documents on each hierarchy so as to determine the weighting value. Furthermore, combining with analytical hierarchy process, and taking the 3 techniques like local reinforcement and so on, as well as the 3 functions like light-weighting and so on for example, the paper finds the ratios of occurrence numbers of these technique functions and eigenvector, then carries out consistence inspection and importance evaluation. It effectively provides a priority order of design decision evaluation.

Keywords. door-shaped structure, knowledge audit, relational knowledgebase, systematic analysis, analytical hierarchy process.

1 Introduction

By using traditional trial-and-error calculation mode, it is difficult to acquire effective decisions quickly, thus increasing the time spent on design. With the evolution of computer in processing speed, both calculation and presentation abilities have been prosperously developed. In addition, the advancement of network technology and knowledge engineering practice makes it more practical and workable in the integration of structural design and professional knowledge. The combination of knowledge framework with systematic analysis established by this paper has the advantage that it can provide designer with flexible needs and

[1] Corresponding Author E-mail: zclin@mail.ntust.edu.tw

services, achieving the effects of knowledge integration and reuse. It will effectively offer the analysis and application of knowledge.

Currently, the application of knowledge-based engineering gradually becomes a prevailing practice. For example, Chapman [1] discusses about computer-aided design and the application of knowledge engineering. By means of rules, he establishes an environment for framework classification and modeling. Peter [2] adopts knowledge engineering method, to construct an organizational model, working model, agent model and knowledge model. He establishes alarm surveillance and provides suggested information to reduce the operation load. Miyamoto et al. [3] develops K-KARDS system and establishes knowledgebase of rules to deal with the structural design problem of vessel in order to shorten the working time. There are also some other applications matching with CAE and assisting in design [4-6]. Analytic Hierarchy Process (AHP) was developed by Thomas L. Saaty in 1971 [7]. It is mainly applied to uncertain situations and decision problems with multiple evaluation standards. Other methods, such as genetic algorithm [8], decision making[9,10], and so on are also commonly seen in the literature of optimization and decision making.

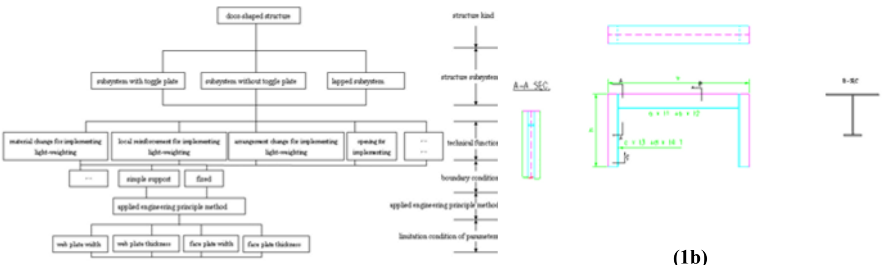
Currently, the studies relating to knowledge management [11-12] emphasize more on management classification, creation and knowledge sharing. The study proposes the way of systematic engineering knowledge audit and establishes relational engineering knowledgebase. The paper also proposes the calculation of the ratios of occurrence numbers of related keywords on different hierarchies, and then uses revised AHP to find the ratios of occurrence numbers of technique and function terms, thus providing a priority order for design decision evaluation.

2 Door-shaped structure knowledge audit

Knowledge audit is a planned design and review of procedures. Focusing on the internal professional domain and external customer needs of a corporate, hierarchical division method is employed to carry out systematic investigation and analysis, and look for the gap of knowledge. In order to offer convenience for the analysis of parameters as well as the establishment and application of relational knowledgebase, the study firstly undergoes classification audit of the related knowledge of door-shaped structure, and uses hierarchical framework classification to carry out systematic accumulation and analysis. Through the audit results, the study establishes references step by step for the application of knowledge framework. The door-shaped structure knowledge audit proposed by the study includes structure kind, structure subsystem, technique function, boundary condition, limitation condition and applied engineering principle method. By using the relativity of classified hierarchical knowledge, relational engineering knowledgebase is established.

Composed of transverse beam and side frame, door-shaped structure is the foundation of a multi-layered structure. The geometric shape of door-shaped structure includes width, height, opening and thickness. Having collected related referential data and design cases, we can divide the subsystem of general door-shaped structure components into subsystem with toggle plate, subsystem without toggle plate and lapped subsystem according to the difference in function and design pattern. Subsystem with toggle plate is always most commonly considered in the improvement of stress concentration. When there is a request of local reinforcement, strengthened plate is also additionally adopted. The bracket plate of subsystem can be designed with different patterns according to different functional requirements, construction difficulties and costs.

The functions of door-shaped vessel structure include support structure, loading, convenience, light-weighting, avoidance of resonance, and compliance with strength requirements. In order to achieve the related functions, door-shaped structure has to consider the techniques of local reinforcement, material change, arrangement change and hole opening. The classified and audited engineering knowledge and relativity of door-shaped structure are shown clearly in Figure 1. According to the classification and audit results, it is matched with hierarchical knowledge. Based on the knowledge framework of Figure 1, the establishment of relational engineering knowledgebase and interface program has been successfully completed. The matrix of related technique function of door-shaped structure developed by the paper is considered for further analysis.



(1a) **(1b)**

Figure 1. (1a) Door-shaped structure hierarchical engineering knowledge framework (1b) Schematic diagram of door-shaped structure

When designing a door-shaped structure, we have to consider its appearance and design pattern according to the actual space need and loading situation in order to meet the requirements. Speaking of the engineering principle method of door-shaped structure, when designing and analyzing the transverse beam vibration of door-shaped structure, the main consideration factors include mass, stiffness, displacement and exciting force. The government equation is:

$$[M]\{\ddot{D}\}+c\{\dot{D}\}+[K]\{D\}=F \tag{1}$$

When damping effect and exciting force are not considered, the above equation is simplified as:

$$[M]\{\ddot{D}\}+[K]\{D\}=0 \tag{2}$$

where [M], [K] and {D} are mass, stiffness and displacement matrix respectively. Equation (2) is a eigenvalue problem. Its eigenvalue is just the natural frequency of structure.

Suppose that the component is a steel-made simple support beam, then the natural frequency fn can be calculated by the following equation:

$$fn = 512\ c\ \sqrt{\frac{I}{A}} \cdot \frac{1}{l^2} \quad (hz) \tag{3}$$

where c denotes the boundary condition parameter, I denotes the moment of inertia (m^4), A denotes the section area (m^2), and l denotes the length of beam (m). Besides, to a simple support beam with uniform load, its bending moment distribution is:

$$M = -\frac{wlx}{2}(1 - \frac{x}{l}) \tag{4}$$

As to the simple support beam with central-concentrated load(W), when $0 \le x \le \frac{l}{2}$, the bending moment M can be expressed as follows:

$$M = -Wx/2 \tag{5}$$

By using bending moment M and section modulus Z, the corresponding stress σ can be acquired, that is:

$$\sigma = M / Z \tag{6}$$

From the abovementioned knowledge framework, it can be found that knowledge framework is adopted to solve the vibration and stress problems systematically and rapidly. The framework of the knowledge application system established by the study according to this idea covers the procedures of relational knowledgebase, decision evaluation, and relevant system interface.

3 Establishment and application of relational knowledgebase

Relational database saves data by two-dimensional matrix, forming relativity among the data saved in columns and rows. Among the files of database, there exists certain specific relationship. When user revises one of the files, other related file data will also be revised accordingly. By using the key columns of different relational files, mutual relational module is established to link all the related data.

The paper combines door-shaped structure components with the related physical properties. Through classification of engineering knowledge frameworks, mutual subordination relationship can be found. The classified techniques and functions as well as the related data of their corresponding cases are collated as the related knowledgebase of technique function with the meaning of knowledge analysis. Basically, the framework of relational knowledgebase established by the paper applies the concept of relational database, and then combines the door-shaped structure knowledge mentioned below, makes further revision, and constructs door-shaped structure relational knowledgebase. The system framework of the paper adopts JAVA language and combines with MySQL software to undergo the construction and application of related interface. The first hierarchy is formulated with structure subsystem and technique function hierarchy. Structure subsystem covers the subsystem with toggle plate, subsystem without toggle plate and lapped subsystem. Technique function engineering knowledge includes the material changes, local reinforcement, arrangement change, and hole opening for implementing light-weighting. The second hierarchy is boundary condition hierarchy, including the entries of selective items of boundary conditions. The third hierarchy is knowledge hierarchy of applied principles, including the source of various applied knowledge, their mutual relativity and advanced description. The fourth hierarchy is limitation condition of applied principles. The limitation condition hierarchy covers the restraint conditions of selective items, including the conditions like upper and lower limits of dimensions, and permissible stress. The first interface of door-shaped structure relational engineering knowledgebase program established by the paper is shown in Figure 2. Through the conversational programming interface indicated in Figure 2, relational engineering knowledgebase is established sequentially.

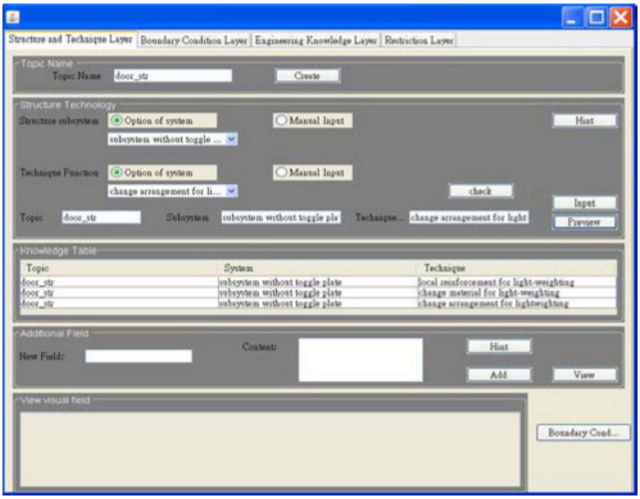


Figure 2. Interface of door-shaped structure relational engineering knowledgebase

Table 1 Probability statistics and micro-adjustment value of related techniques of support structure

Function description	Technique keywords	the ratios of occurrence numbers	adjustment of relationship between occurrence number and weighting value
support structure	local reinforcement	0.333	2
	arrangement change	0.245	1
	material change	0.422	3

4 Using revised AHP to establish ratios of occurrence numbers of techniques and functions, and make decision evaluation

In general, we should know the whole important weighting values in front of making evaluation and analysis. In order to evaluate and improve the door-shaped structure design, a revised AHP method is proposed in this paper. AHP is mainly applied to uncertain situations, and has a decision problem that it contains multiple evaluation standards. Therefore, AHP [7] is developed with the purpose to systematize complicated problems, disintegrate hierarchies from different aspects, and make subjective judgment become quantitative. In the past, weighting value was the selective value from 1 to 9 according to its importance [7]. But weighting value is always determined by experienced people, or by the survey method of question and answer. Just as what the paper has mentioned, this is a difficult method to a user with less experience in design. Hence, the revised AHP method used in this article suggests considering the ratios of occurrence numbers of the related technical terms corresponding to different functions according to some technical documents and patent literature. It takes the 3 techniques such as local reinforcement, arrangement change and material change, as well as the 3 functions such as support structure, light-weighting and strength requirement. It then establishes pair-wise comparison matrix, and calculates the maximum eigenvalue. Finally, referring to the ratios of occurrence numbers of these technical and functional terms, according to the evaluation criterion of adjustment and revision of relationship between occurrence number and weighting value, it calculates the hierarchical importance values of essential factors on different hierarchies. The main procedures of the revised AHP proposed by this paper are generally summarized as follows:

(1) Establish pairwise comparison matrix

First of all, programs are used to search different patents and technical documents. Find the occurrence numbers of different technical terms under a certain function, and calculate its ratio by using equation (7). The paper supposes that the support function is (k=1), light-weighting function is (k=2), strength requirement function is (k=3), local reinforcement is (i=1), arrangement change is (i=2), and material change is (i=3), then:

$$e_{ik} = f_{ik} / \sum_{i=1}^{n} f_{ik} \qquad (7)$$

Here, f_{ik} denotes the occurrence number of technical term, and n denotes the number of techniques. For example, in Table 1, under the function of support structure, the ratios e_{ik} for the occurrence numbers of the corresponding techniques of local reinforcement, arrangement change and material change are 0.333,0.245 and 0.422 respectively.

When a certain essential factor is the evaluation standard, adjustment and revision are made, just like the relationship among d_{11}, d_{21},...d_{n1}.

Use equation (8) to calculate the corresponding hierarchical weighting values s_{11}, $s_{21}, \ldots s_{n1}$, i.e. $s_{ik} = d_{ik} / \sum_{i=1}^{n} d_{ik}$ (8)

where s_{ik} denotes the hierarchical weighting value, n denotes the number of techniques, and the definitions of i and k are mentioned above.

Establish a pairwise comparison matrix, $A=[a_{ij}]|_k$ according to equation (9):

$$[A]_{nxn}|_k = [a_{ij}]|_k = \begin{bmatrix} 1 & s_1/s_2 & \ldots. & s_1/s_n \\ s_2/s_1 & 1 & \ldots. & s_2/s_n \\ \ldots & \ldots & \ldots & \ldots \\ s_n/s_1 & s_n/s_2 & \ldots. & 1 \end{bmatrix}_k \quad (9)$$

Each row of the matrix is formed by the proportion of the weighting value of a single essential factor to the weighting value of other essential factors. And $a_{ij}=s_i/s_j$, $a_{ji}=1/a_{ij}$.

(2) Calculation of maximum eigenvalue and eigenvector to perform consistence inspection.

The purpose of calculating the maximum eigenvalue and eigenvector is to inspect whether the pairwise comparison matrix [A] meets the consistence requirement.

The equation of eigenvector $[W]_{nx1}|_k$ for each function is as follows:

$$[W]_{nx1}|_k = \begin{bmatrix} w_{1k} \\ w_{2k} \\ \ldots. \\ w_{nk} \end{bmatrix} \text{, where } w_{ik} = \left[\prod_{j=1}^{n} a_{ij}|_k\right]^{1/n} \Big/ \sum_{i=1}^{n}\left[\prod_{j=1}^{n} a_{ij}|_k\right]^{1/n} \quad (10)$$

where n denotes the number of technical factors, and i and k are mentioned above.

First of all, the pairwise comparison matrix [A] is multiplied by the acquired eigenvector $[W]_{nx1}|_k$ to achieve a new vector $[W']_{nx1}|_k$. After that, the maximum eigenvalue $\lambda_{\max}|_k$ can be found by equation (12).

$$[W']_{nx1}|_k = [A]_{nxn}|_k [W]_{nx1}|_k \quad (11)$$

$$\lambda_{\max}|_k = (1/n)*(w_{1k}'/w_{1k} + w_{2k}'/w_{2k} + \ldots + w_{nk}'/w_{nk})|_k \quad (12)$$

The paper proposes converting the ratios of occurrence numbers of technical and functional terms to be suitable numerical value of proportional criterion, which is then substituted in AHP evaluation matrix to perform weighting analysis. In the evaluation process, decision-makers may have inconsistent judgments for the importance of different decision-making factors. Thus, to solve the inconsistent judgments, consistence inspection has to be made for pairwise comparison matrix before confirming whether the judgment result is workable. When consistence ratio exceeds the allowed value, i.e. $C.R. \geq 0.1$, it implies that the consistence degree of the matrix has exceeded the permissible deviation range. The decision-maker has to re-consider the important relationship among different decisions. During this time, it is required to re-make micro-adjustment of the evaluation criterion value of AHP evaluation matrix, and re-inspect the eigenvalue and consistence.

The definition of consistence ratio C.R. is as follows [7]:

C.R.=C.I./R.I. (13)

where C.I. refers to consistence index.

$$C.I. = (\lambda_{\max}|_k - n)/(n-1) \tag{14}$$

(3) Importance evaluation

As mentioned above, focusing on the eigenvalues w_{11},w_{21},w_{31} of the function of support structure load towards the 3 techniques, the eigenvalues w_{12},w_{22},w_{32} of the function of light-weighting towards the 3 techniques, and the eigenvalues w_{13}, w_{23}, w_{33} of the function of strength requirement towards the 3 techniques, the mixed eigenvalue matrix $[W_t]_{nxn}$ is as follows:

$$[W_t]_{nxn} = [w_{ij}] = \begin{bmatrix} w_{11} & w_{12} \ldots. w_{1n} \\ w_{21} & w_{22} \ldots. w_{2n} \\ \ldots\ldots\ldots\ldots & \\ w_{n1} & w_{n2} \ldots. w_{nn} \end{bmatrix}$$

Then the importance of the design technique of door-shaped structure can be expressed by $[X]_{nx1}$ matrix of equation (15).

$$[X]_{nx1} = [W_t]_{nxn} \ [W_f]_{nx1} \tag{15}$$

The paper takes the occurrence numbers of the 3 functions of support structure load, light-weighting and strength for example, and makes importance evaluation of workable techniques, with the calculated results presented in Table 2 and Figure 3. When inspecting the AHP evaluation results indicated in Table 2 and Figure 3, the eigenvalue has fully presented the importance order of the corresponding technical terms of the various functions. Figure 3 shows that from the searching patents and technical documents, the priority of most frequently considered functions is support structure, light-weighting and strengthening requirement. For example, among the corresponding technical terms of light-weighting function, the priority is material change (0.625), local reinforcement (0.25) and arrangement change (0.125). From here, it implies that among the analyzed documents, the most frequently used technical method for the function of light-weighting is material change; and the next one is local reinforcement. And the most frequently used technical method for strength requirement is local reinforcement; and the next one is material change. According to the various patents and technical documents, we can find the corresponding eigenvalues for the 3 functions of support, light-weighting and strength requirement as well as the 3 techniques of local reinforcement, arrangement change and material change. The importance matrix of equation (15) is used to calculate the relative importance, which are sequentially material change (0.580), local reinforcement (0.368) and arrangement change (0.163). It is obviously seen that the prioritized consideration of design techniques is material change; the next one is local reinforcement; and the last one is arrangement change. Such analytic result is very close to the general consideration of structural design, and is thus verified. Therefore, according to this result, we can think about the innovation, research and development directions of technical functions for exploring towards the aspect of mature development, or developing towards the direction of lower sequence. The acquired result can also provide a reference for the decision evaluation of priority order in times of research and development. In the application of knowledge engineering, it is a pioneering work for the paper to propose revised AHP steps for the hierarchical consideration of this kind of combination with engineering knowledge audit. This proposal not only can provide a decision of priority order for innovative research and development, but also can match with the engineering knowledge to establish design improvement by systematic consideration.

Table 2 Revised AHP evaluation matrix of 3 techniques versus 3 functions of door-shaped structure

Evaluation item	support structure	light-weighting	strength requirement	importance evaluation
	0.614	0.368	0.123	
local reinforcement	0.333	0.250	0.579	0.368
arrangement change	0.167	0.125	0.116	0.163
material change	0.500	0.625	0.348	0.580

4.1 Design improvement cases

For the light-weighting of a door shaped stucture with length 5000 mm, effective breadth 400 mm , effective breadth thickness 12 mm, web width 250 mm, web thickness 12.5 mm, and face plate width 150 mm, face thickness 16 mm. When an uniform load 18.5 T/m^2 applys to this structure, after calculation by finite element method, the maximum stress is 220 N/mm^2 as shown in Figure 4, which is beyond the permisible stress 188 N/mm^2. Based on the evaluation of revised AHP from Table 2, the first priority is material change which has greater weighting value of 0.625 and the second priority is local reinforcement which weighting value is 0.250. There are two changes can be applied to improve the design problem. Priority one is the simplest method by material change from mild steel to high tensile steel of grade HT32. The permible stress of HT32 is 251N/mm^2, which is greater than 220 N/mm^2. The second one is improved by local reinforcement which could be by local thicker web plate (15.5mm) or subsystem by bracket (300mmx15mm) as shown in Figure 5 and Figure 6 respectively. After the finite element method improvement calculation, it is found the maximum stress both are 183 N/mm^2 in Figure 5 and 176 N/mm^2 in Figure 6, they are fulfilled the permisible stress.

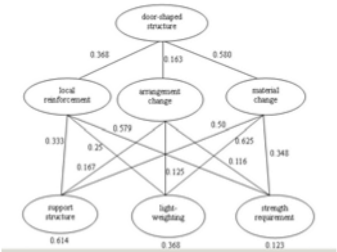


Figure 3. Importance evaluation of door shaped structure AHP matrix

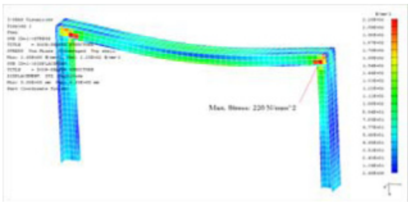


Figure 4. Von Mises stress distribution of original scantling

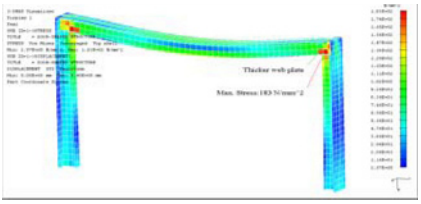


Figure 5. Von Mises stress distribution of local reinforcement by thicker web plate

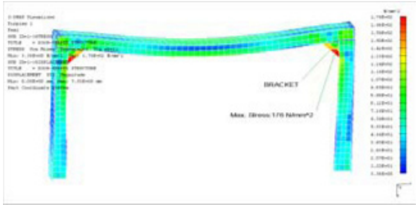


Figure 6. Von Mises stress distribution of local reinforcement by bracket plate

5 Conclusion

The paper uses door-shaped structure knowledge framework to carry out engineering knowledge classification. Employing the hierarchical analysis of systematic techniques proposed by this paper, classified hierarchical engineering knowledge audit are carried out. The hierarchical implementation of knowledge classification audit in this paper includes structure kind, structure subsystem, technique function, boundary condition, limitation condition and applied engineering principle method. By means of the relativity of hierarchical knowledge, the study establishes relational engineering knowledgebase. Comparing with the past audits, the study even combines with the technique functions of engineering knowledge, enabling an effective accumulation and application of related engineering knowledge. The study suggests a revised AHP method which considering the ratios of occurrence numbers of the corresponding related technical terms of different functions mentioned in multiple technical documents and patent literature. By using this revised AHP method, the study takes the 3 techniques like local reinforcement and so on, as well as the 3 functions like light-weighting and so on for example, establishes pairwise comparison matrix, and calculates the maximum eigenvalue. Finally, referring to the ratios of occurrence numbers of these technical and functional terms, according to the evaluation criterion of adjustment and revision of relationship between occurrence number and weighting value, the study calculates the hierarchical importance values of different hierarchical essential factors, effectively providing a reference for making a design decision of priority order for innovative research and development.

According to the engineering knowledge and important parameters, engineering knowledge framework is performed to establish related knowledge base so as to shorten the time for searching data. The systematic design consideration provides a proper direction for improvement. Through the evaluation by revised AHP, suitable structure design can be rapidly refined by finite element method. It can also be applied to engineering analysis, or the evaluation of professional technical subsystem attribute in other domains, to achieve rapid analysis, prediction and application to decision-making.

6 Acknowledgement

The authors thank National Science Council (grant number NSC-96-2221-E-011-106-MY3) supporting this research.

7 References

[1] C.B Chapman, M. Pinfold, "Design engineering-a need to rethink the solution using knowledge based engineering", Knowledge-Based Systems 12, 1999, 257–267.
[2] Peter L. Jones, Alan Harrison," The application of knowledge-based techniques to the monitoring of computers in a large heterogeneous distributed environment", Knowledge-Based Systems 19, 2006, 565–575.
[3] Shuji Miyamoto, Shuji Nonoguchi, Jiro Matsuno and Takuya Matsumura," Application of Knowledge Based Modeling to Detail Structure Design for Shipbuilding", ICCAS, 2002, 717-729.
[4] W.J. Zhang, D. Zhang, "A general Methodology for Developing Intelligent CAE System- a model based reasoning approach", Journal of Materials Processing Technology 61, 1996, 148-153.
[5] Q. Li., W.J. Zhang, "Application of model-based reasoning to the development of intelligent CAE systems", Engineering Applications of Artificial Intelligence 11 ,1998, 327-336.
[6] G. Bravo-Aranda, F. Hernandez-Rodriguez, A. Martin Navarro, "Knowledge-based system development for assisting structural design", Advances in Engineering Software 30, 1999, 763–774.
[7] Thomas L. Saaty, The Analytic Hierarchy Process, New York: McGraw-Hill, 1980.
[8] Tsai S.W., "A Study on Automatic Document Classification By Using Genetic Algorithms," Master Degree Thesis, Department of Information Engineering, Tamkang University, Taipei, 1998.
[9] Thomas L. Satty," Automatic Decision-Making: Neural Firing and Response", Journal of System Science and System Engineering, Volume 13, Number 4, 2004, 385-404.
[10] Christine W Chan, Michael Johnston, "Knowledge modeling for constructing an expert system to support reforestation decisions", Knowledge-Based Systems 9, 1996, 41-59.
[11] Snejina Michailova, Akshey Gupta, "Knowledge Sharing in Consulting Companies: Opportunities and Limitations of Knowledge Codification", Journal of Information & Knowledge Management, Vol. 4, No. 3, 2005, 201–212.
[12] S. Y. Choy, W. B. Lee and C. F. Cheung, " Development of a Knowledge Management Culture Assessment Tool with Applications in Aviation Industry", Journal of Information & Knowledge Management, Vol. 4, No. 3, 2005, 179–189.

Knowledge Extraction System from Reports in Fabrication Workshops

Kazuo Hiekata[a,1], Hiroyuki Yamato[b] and Sho Tsujimoto[b]

[a] Graduate School of Engineering, the University of Tokyo, Japan
[b] Graduate School of Frontier Sciences, the University of Tokyo, Japan

Abstract. In this paper, a method to extract knowledge from a set of surveyor's reports is proposed. Many surveyors from shipping companies are working in shipyards to improve the quality and meet requirements of the ship owner. At present, the reports from surveyors are used only for managing the status of problems of the product. The proposed method here has three steps. First, the description field of the reports which contain information about some trouble is summarized in a pair of component name and trouble name in the failure database; this is made possible thanks to the text-processing technology. Second, synonimous component and trouble names are consolidated into normalized terms using synonym ontology to prepare for statistical analysis. The synonym ontology is developed prior to this procedure. The system evaluates frequent trouble records by classifying the records using component and process ontology. In the last step, the designers extract knowledge from the information generated by the system. The results illustrate common design problems and the method is proven to be useful.

Keywords. Knowledge extraction, failure report, text processing, ontology.

1 Introduction

In shipyards, troubles and failures due to design mistakes are discovered during fabrication in the workshop. When troubles arise, a report about them is written in the natural language and stored in the database as a surveyor's report. Although the report is stored, the same mistakes still occur in fabrication. For the improvement in the efficiency of production, design errors should be collected and reported to the design department to avoid committing the same mistakes. For this purpose, it is important to extract knowledge from the workshop for the needed feedback. The electronic surveyor's report is considered as the most valuable resource. There are numerous existing researches to utilize accumulated text data [1, 2, 3, 4], though it is still difficult to extract knowledge from the natural language text. This paper proposes the method to extract knowledge from surveyor's reports written in natural language text using ontology and text-processing.

---

[1] Corresponding Author E-mail: hiekata@nakl.t.u-tokyo.ac.jp

2 Proposed Method

The overview of the proposed method is shown in Figure 1. The method is composed of four data instances and three steps. The initial input for this method are the surveyor's reports. All the surveyor's reports have many fields to describe problems of the ship under construction. The method focuses on "description" field and also uses "Follow up section" field in the last procedure.

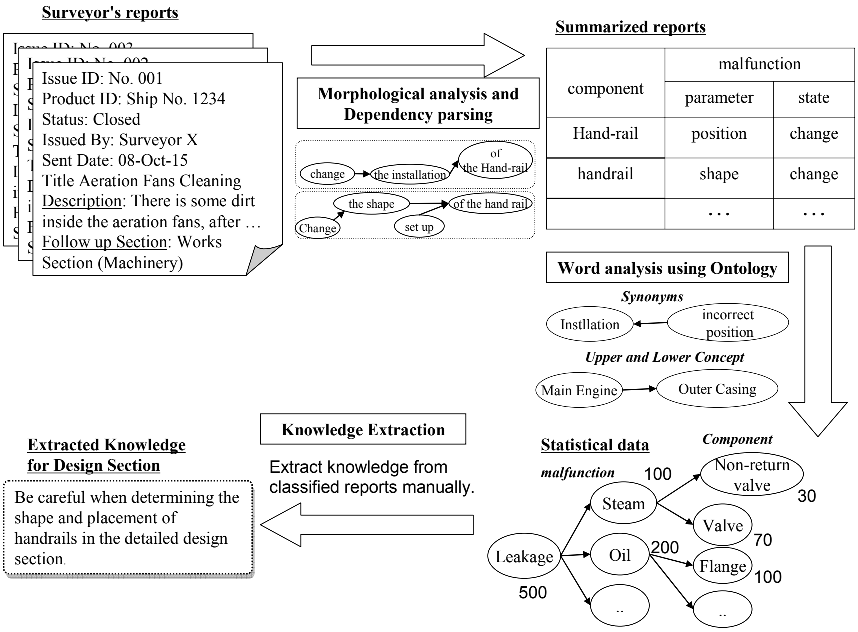


Figure 1. Overview of Proposed Method

In the first step, "Morphological analysis and dependency parsing", the description field of all the reports are extracted into a pair of a component and a malfunction description, and the descriptions of surveyor's reports are summarized into a table with three columns. Then the summarized reports are processed in the next step to get statistical data. Ontology is introduced to improve the accuracy of automatic classification by the system. In using ontology, the system considers synonyms as words having equivalent meaning. The system then takes into account words of upper concept and lower concept when classifying similar reports. The gathered reports are classified in a tree view by components, malfunction and "Follow up section" field. Finally, the design engineers or knowledge engineers evaluate the classified reports with the accompanying statistical data to extract knowledge thereby improve the quality of their design work.

2.1 Morphological Analysis and Dependency Parsing

At first, the word dependency structure is analyzed. The proposed method ignores words not related to the trouble. For example, "Repair welding is required because the position of the step is inappropriate", recognizes that the step was located in the wrong position by using a Japanese text-processing machine, Mecab and CaboCha [5, 6]. Mecab is used as a morphological analyser and CaboCha for dependency analyser. The trouble expression consists of a component name and a trouble name. There is word dependency to the component name word and the trouble name word in Japanese, so this method can generate summarized reports automatically.

By replacing the text-processing machine, this method can be applied to analyze reports written in other languages.

2.2 Word Analysis using Ontology

Reports written in natural language are formatted in a table with several columns. To get more accurate statistical data, ontology is introduced for handling synonyms and inferring upper or lower concept words. Ontology is one of the knowledge modeling methods applied in many fields, such as product engineering [7], modeling company activities [8] and design knowledge [9]. In this research, ontology technology is used as a dictionary for synonyms and concept hierarchy.

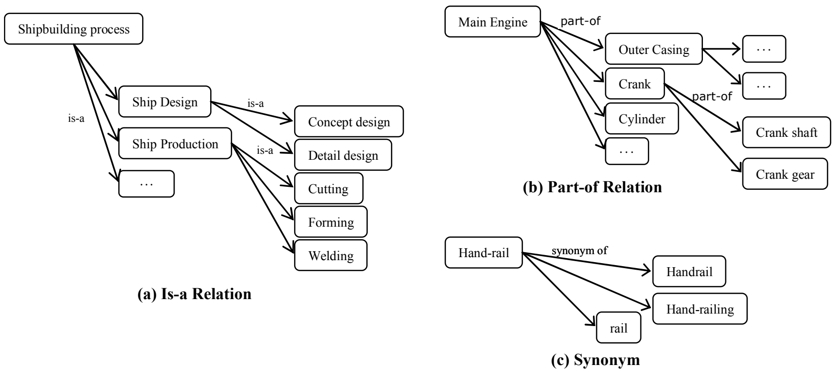


Figure 2. Example of Ontology

For advanced inference, concept hierarchy is utilized to gather similar and related reports. Components of a product and process hierarchy are defined by ontology. An example in using this ontology is as follows. A report may include trouble information such as "Mistake in the welding position" and "Painting mistake." Both are classified in the same "construction" category. "Construction" is an upper concept of "welding" and "painting" so when analyzing "construction" of a ship, reports about "welding" and "painting" should also be considered. As for product components, when retrieving "Main engine" the result should include the trouble information related not only to “Main Engine” but also to “Cylinder,” “Piston” or other sub-components. The relationship of each concept is defined by “part-of” and “is-a” as shown in Figure 2 (a) and (b). Using ontology, the concept can be defined as formal and explicit way [10].

To solve the synonym problem, different notations of the same object/meaning will be replaced with a normalized term defined in the synonyms dictionary based on ontology. The relationship is expressed by “synonym-of” as shown in Figure 2 (c).

2.3 Knowledge Extraction

The proposed method gives attributes such as component name, trouble name and follow-up-section to all the reports. This means that the reports can be classified as a tree similar to FTA using ontology. For example, reports can be classified by the components using the ontology which describes “steam valve is part-of accommodation space” or “Flange is part-of tank.”

In the lower-right portion of Figure 1, the tree view shows that 500 reports have something to do with “Leakage” and 100 reports described “Steam Leakage.” Also, 30 reports of “Steam Leakage” occurred at the “Non-return valve.” This kind of analysis can be performed by using component and process ontology.

The last step of knowledge extraction is manually performed by design engineers or knowledge engineers. In Figure 1, the common knowledge about steam leakage at the non-return valve is expected to be extracted by reading the 30 reports about “Steam Leakage.” The engineer need not read all the reports, but only the classified or related ones. The knowledge is expected to be used in the design section. More essential problems might be revealed during this process, although further analysis depends on the skills of the engineer and is not covered in this paper.

3 Case Study

The proposed method is evaluated by applying to actual 9604 surveyor’s reports in a shipyard. Table 1 shows a sample of a record of actual reports. The proposed method classified the description field of the surveyor’s reports.

3.1 Ontology Development

The Protégé Ontology Editor [11] is adapted to construct the ontology. Protégé is an open source ontology editor. The ontology is constructed based on the ontological normalization theory [12]. Figure 3 (a) shows part of the process hierarchy. “Work related to shipbuilding” is the root of the concept tree and is divided into “work outside a shipyard” and “work inside a shipyard.” Then, “Work inside the shipyard” is subdivided into “the manufacturing site work,” like construction process, and “outside the manufacturing site work” such as design process. There are eight layers here. Figure 3 (b) is an example of the hierarchy of the main engine. There are six layers in this example. These were constructed based on interviews with veteran designers. Figure 3 (c) shows examples of the synonym dictionary. The gray squares and the white squares are defined as normal expressions and synonyms, respectively, so "Main machine" is shown as a normal expression. This ontology must be developed prior to performing the analysis.

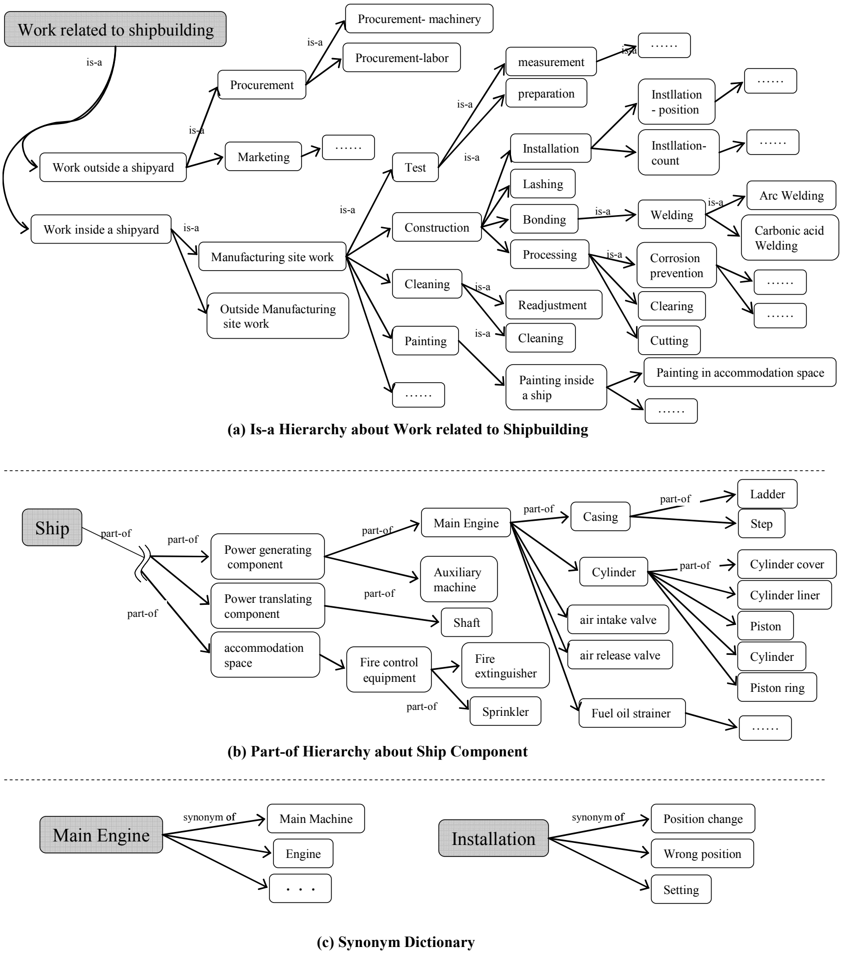


Figure 3. Part of Ontology

3.2 Result

Figure 4 shows the classification result about wrong positioning. 1104 reports about wrong position are classified by the proposed method. The 1104 reports are classified by component name. The proposed method found 112 reports about wrong positions occurring for hand rails. 15 reports of the 112 reports are problems about wrong position at the hand rail in the cargo tank, and the other problems occurred at valves. According to the result, problems about wrong positions often occurred at the hand rails and valves and a lot of them about handrails take place at the cargo tank.

This finding is a type of design knowledge and will contribute to the detailed design section. Improvement in the performance at the design section will prevent the same mistake such as wrong positioning at the handrail or valve.

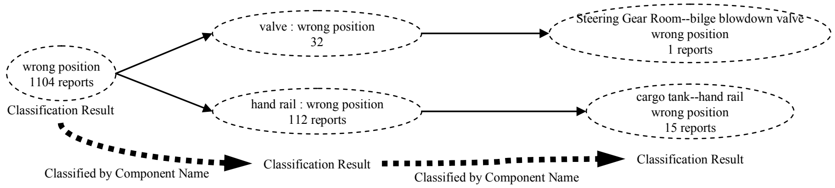


Figure 4. Classification Result about Wrong Position

Table 1. Example of Surveyor's Report

Field Name	Value
Issue ID	No. 004
Product ID	Ship No.1258
Status	Closed
Issued By	Surveyor X
Sent Date	08-Apr-13
Title	Aeration Fans Cleaning
Description	There is some dirt inside the aeration fans. After fitting and producing new air ducts, please clean out.
Follow up Section	Works Section(Machinery)

The reports has "Follow up section" field, so the reports can be classified according to this section. Figure 5 illustrates classification results about abnormal interference showing 103 reports in this example. 18 reports about abnormal interference found by this system were from the Machinery Outfit Designing Section. 3 of the 18 reports described interference at the valve handle. 2 of the 18 reports showed abnormal interference at the hand rail. Finally, the Hull Outfit Designing Section had 19 reports about abnormal interference. Other examples of classified reports are shown in Table 2.

According to the tree shown in Figure 5, Machinery Outfit Designing Section and Hull Outfit Designing Section exhibited many abnormal interference cases, so the design quality in detailed design section can be improved by defining a rule for design to avoid abnormal interference.

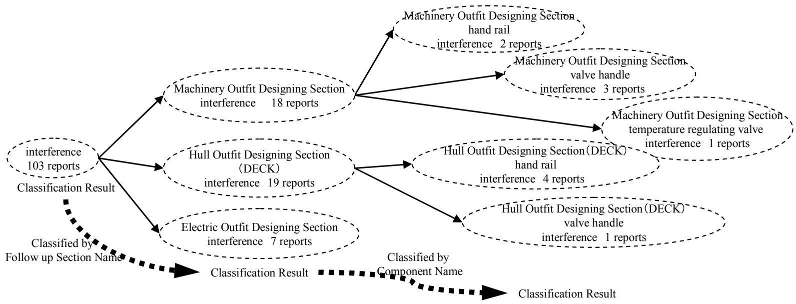


Figure 5. Classification Result about Interference According to Follow up Sections

Table 2. Examples of Classified Reports

Follow Up Section	Trouble Name	Description
Machinery Outfit Designing Section	interference	The hand comes in contact with Hand-rail because the position of the hand rail and the valve steering wheel is adjacent. Please change the position of Hand-rail.
		Temperature gauge TI5055 cannot be detached because of interference with pipes.
	Wrong Position	Handle for the steam valve cannot be attached. Move the handle, stairway or change the layout.
		Please add intermediate valve to steam pipes for the emergency power generator to avoid leakage.

4 Discussion

The extraction system and ontology allowed for evaluation of surveyor's reports making it ready for proper analysis. The extraction system gets a pair of component name and malfunction name from the surveyor's reports, while ontology was used for word analysis.

Figure 6 shows hand rail as a component of cargo tank, and bilge blowdown valve as a component of steering gear room. This means that "the bilge blowdown valve" is parsed as a component of "the steering gear room." Therefore, the failure information about wrong positioning at the bilge blowdown valve is classified as information about the steering gear room as well.

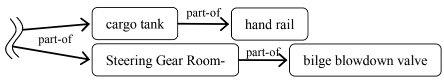


Figure 6. Example of Component Ontology

In Figure 5, the method can classify the surveyor's reports in the order of malfunction name, follow up section and component names. This result was achieved according to the following feature. The surveyor's reports are summarized as a pair of component and malfunction names, and the follow up section as Meta information in addition.

According to the case study, the proposed method can extract the failure and complaint information as a simple word. The information is stored in the database. If appropriate software synchronized with the database is developed, the knowledge of the production process can be extracted.

5 Conclusion and Future Work

The proposed method extracts knowledge from reports to improve the quality of the design. Reports are summarized to pairs of a component and a malfunction, and statistical analysis is applied. Finally, statistical data is extracted as abstract konwledge for the design section to improve the quality of detailed design. An example about the position of handrails is illustrated in this case study.

6 References

[1] TRAPPEY, A., TRAPPEY, C. and C. WU. “A Semantic Based Approach for Automatic Patent Document Summarization.” Proceedings of the 15th ISPE International Conference on Concurrent Engineering, (2008): 485-494.

[2] NASUKAWA, T. and T. NAGANO. “Text analysis and knowledge mining system.” IBM Systems Journal, 40: 4 (2001).

[3] POPESCU, A. and O. ETZIONI. “Extracting Product Features and Opinions from Reviews.” Proceedings of Joint Conference on Human Language Technology (HLT) and Conference on Empirical Methods in Natural Language Proceeding (EMNLP), (2005).

[4] HU, M. and B. LIU. “Mining opinion features in customer reviews.” Proceedings of the 19th National Conference on Artificial Intelligence (AAAI). (2004): 755-760.

[5] KUDO, T and Y. MATSUMOTO. “Japanese Dependency Structure Analysis based on Support Vector Machines.” In Proceedings of the EMNLP/VLC-2000. (2000): 18–25.

[6] KUDO, T. and Y. MATSUMOTO. “Japanese dependency analysis using cascaded chunking.” In Proceedings of the CoNLL-2002. (2002): 63–69.

[7] BRITMAN, K. K. et al. Ontology as a Requirements Engineering Product, In 11th IEEE International Requirements Engineering Conference (RE’03), pp. 309–, Monterey Bay, California, USA, Sep. 2003. Mini-Tutorial.

[8] FOX, M., BARBUCEANU, M. and M. GRUNINGER. “An organisation ontology for enterprise modeling: Preliminary concepts for linking structure and behaviour.” Computers in Industry, 29:1(1996):123-134

[9] KITAMURA, Y., MIZOGUCHI, R. “Ontology-based description of functional design knowledge and its use in a functional way server. “ Expert Systems with Application, 24:2, (2003): 153-166.

[10] GRUBER, T. R. “A Translation Approach to Portable Ontology Specifications.” Knowledge Acquisition, 5:2 (1993):199-220.

[11] NOY, N.F., SINTEK, M., DECKER, S., CRUBEZY, M., FERGERSON, R.W. and M. MUSEN. “Creating Semantic Web contents with Protege-2000. “ IEEE, Intelligent System, 16, No.12, pp.66-71, (2001)

[12] RECTOR, A., DRUMMOND, N., et al. "Modularisation of domain ontologies implemented in description logics and related formalisms including owl. " Knowledge Computer, ACM, pp.121-128, 2003

Knowledge Based Sales Forecasting Model for Non-Linear Trend Products

Kenji Tanaka[a,1], Hideaki Miyata[a] and Shoji Takechi[a]

[a] School of Engineering, The University of Tokyo, Japan.

Abstract. This paper proposes a knowledge based sales forecasting approach for nonlinear trend products. Though forecasting of the future demand is an essential part of business planning and operation, most of major existing forecasting methods are applied to only regular consuming items which show linear sales trend because of seasonal cycles. In this study, using correlation between two sales-date points among similar products group, a new forecasting model is presented. It enables its accuracy advancement by incorporating experts' knowledge as grouping rules. The model is applied to Japanese literature 70 titles to verify its performance. It proved its accuracy by giving 34% error rate which is superior to 273% error rate of the exising exponential smoothing method.

Keywords. knowledge management, forecasting, nonlinear, book retail

1 Introduction

Compare to production efficiency improvement in these decades, the accuracy of demand forecasting remains much room to be designed. Mentzer and Bienstock[1] indicated that forecast of sales influences various decisions at organizational level. Agrawal and Schorling[2] also pointed out that accurate demand forecasting plays a critical role in profitable retail operations. It is, therefore, the often cases that misleading forecast cause amount of business loss, which is not only unsold return goods loss but also chance loss. Especially there are those tendencies among those irregular products industries.

The existing quantitative approaches include heuristic methods such as time series decomposition and exponential smoothing as well as time series regression and autoregressive and integrated moving average (ARIMA) models. Chu and Zhang[3] pointed out that one of the major limitations of the traditional methods is that they are linear methods. Users can not consider the complex relationship in the data into those forecast as far as it is not seasonal one. There are, therefore, researches on nonlinear modeling approach with neural networks. Elkateb et al[4]

---

[1]Corresponding Author : Assistant Professor, Department of Systems Innovation, Graduate School of Engineering, The University of Tokyo, Japan. E-mail: kenji_tanaka@sys.t.u-tokyo.ac.jp, +81-3-5841-6522

reported that neural network model in electric load forecasting showed better performance than ARIMA models. Prybutok et al[5] also reported its superiorness in forecasting ozone concentration. However, Callen et al[6], Kirby et al[7], Darbellay and Slama[8] reported the opposite results. They figured out neural network is not superior to the time series models even if the data are nonlinear. Without certain range self results to tune its forecast, it is difficult to keep their performance. In addition to that point, Faraway and Chatfield[9] reported that neural network model requires experimental efforts and traditional modeling skills to fix each product forecast settings. Those problems mean neural network method is not necessarily suitable forecasting method for those industries whose products have short lifecycle or wide variety. No method has been established as standard forecasting method for those nonlinear products.

One of the typical products of wide variety and shortlifecyle is book retail products. That industry have serious problems of misleading forecast. Book sales is not seasonal nor cyclical, because their customers buy the same title only once. Figure1 shows Japanese book market trend and its book return ratio. In addition to the shrinking market size, there remains high ratio of book return. Approximate 40% of printed books, as a result, have unsold and returned to publishers as surplus-goods, and that causes industry players' sustainability issue. Nagai[10] estimates total amounts of waste books value at $13billion USD in 2003. One of the main reasons for the problem is each player's each misleading forecast. Publishers, wholesale-distributors and shop-chains have each forecast of titles to be printed or to be ordered. According to management members of publisher and wholesale-distributor, existing numerical forecasting method itself is not worth practical using for them. There are too many elements or attributes of one title to be considered. It is necessary to predict not only by numerical forecasting method but also knowledge based approach that is able to deal with each attribute which experts partly know.

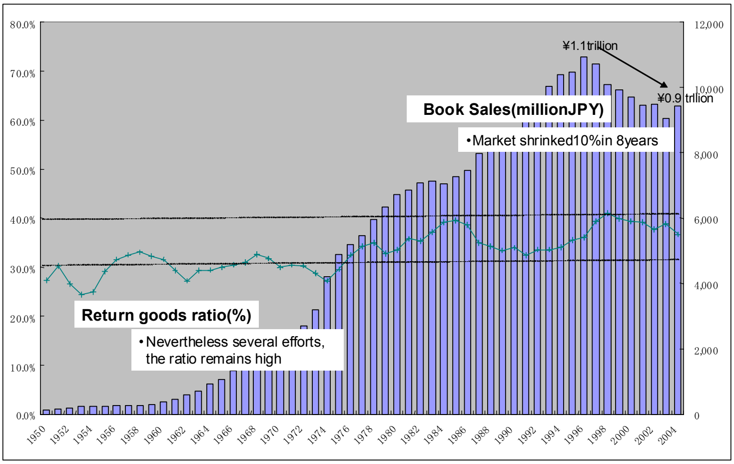


Figure 1. Japanese market trend of book sales and return goods ratio[11]

The main objective of this study is to propose a knowledge based forecasting method for nonlinear items. We verify the performance of our method by applying it to book items forecasting. The evaluation of its accuracy and potential improvement will be stated.

2 Forecasting Model

2.1 Concept

The basic concept of proposed forecasting method is illustrated in Figure2. Using correlation between specific date pair of sales results, the target value of cummulative sales forecast is predicted.

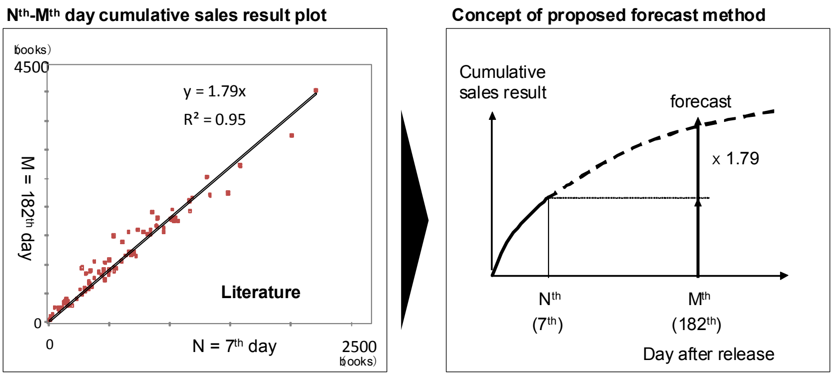


Figure 2. Book sales, 100 novels released in Spt. 2005

We call this forecasting method "NM method". It is the methods which predict M^{th}day accumulated sales forecast from N^{th}day accumulated sales result as follows;

$$X_{(group,N,M)} = k_{(group,M,N)} \cdot R_M \tag{1}$$

$$k_{(group,M,N)} = \frac{n(\sum_i X_{Mi} \cdot X_{Ni}) - (\sum_i X_{Mi}) \cdot (\sum_i X_{Ni})}{n(\sum_i (X_{Mi})^2) - (\sum_i X_{Mi})^2} \tag{2}$$

where, $X_{(group,N,M)}$ is M^{th} day accumulated sales forecast of the group. The symbols of "M"and"N"mean days after item released. The parameter $k_{(group,N^{th},M^{th})}$ is N^{th} day coefficient for calculating M^{th} day forecast of the reference item group. The parameter n means number of items in the group.

There remains high correlation between two date sales results of relevant item group. Some industry experts have the knowledge that they can roughly figure out their forecast with first 3-4days sales result. That knowledge is verified with the analysis of correlation between two date sales result. Figure3 shows convergence transition in multiples of the two date sales result in first 7 days in literature titles. The horizontal axis is for the accumulated sales multiples between two dates of each title, and the vertical for the sum of titles.

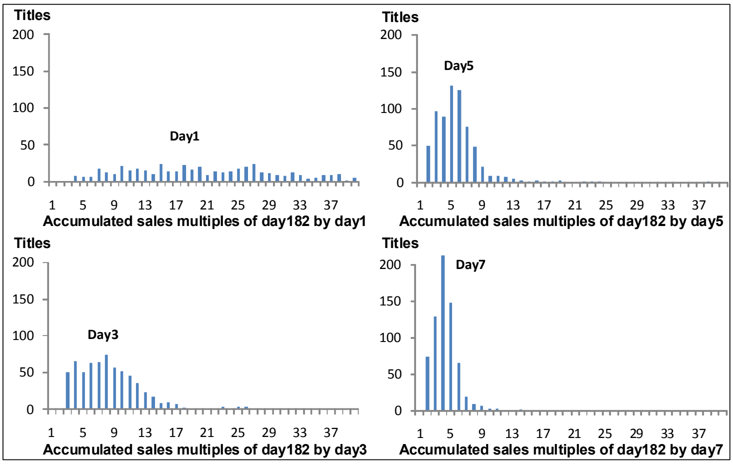


Figure 3. Convergence in multiples of two date sales result

2.2 Forecasting workflow

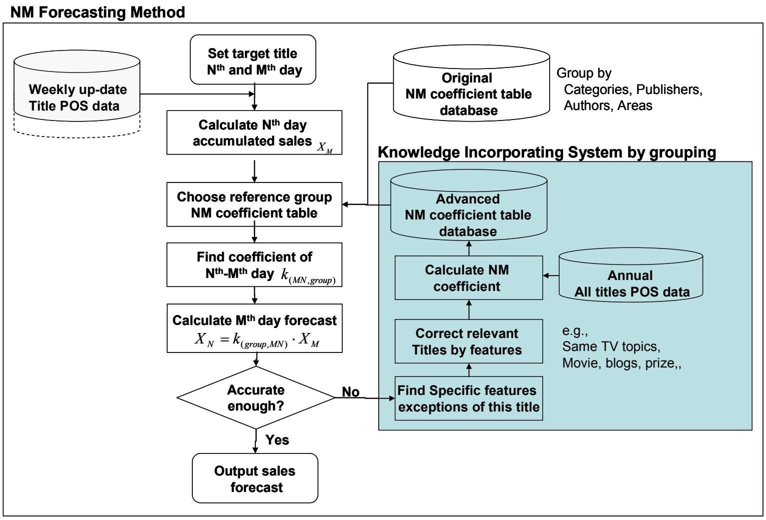


Figure 4. Forecasting workflow with proposed method

Figure4 shows the forecasting workflow with proposed NM methods. Firstly the target title, its N^{th} day and M^{th} day are set. Secondly the accumulated N^{th} day result is calculated with weekly updated sales data. Thirdly the NM coefficient of chosen group is referred. Normally the one of original NM groups which are generated by existing categories, publishers, authors and release date, is used for calculation. If the forecast is not accurate enough, new advanced NM group can be corrected by experts' knowledge (Figure5).

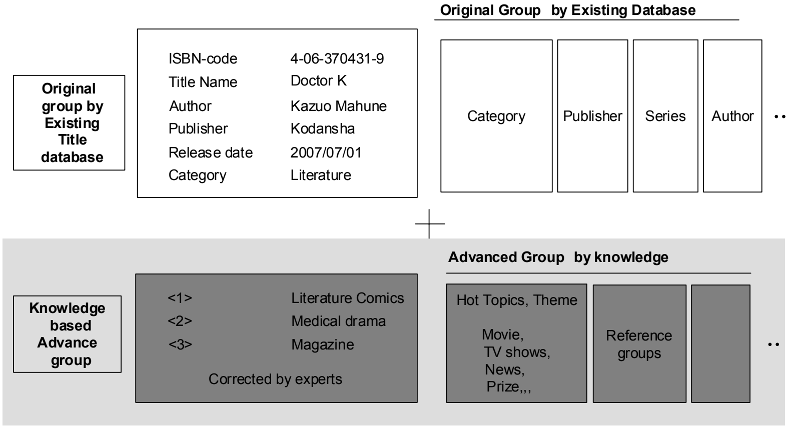


Figure 5. Grouping example by knowledge management

Key words or topics are defined as data elements such hot topics as movie, TV show, news prize and so on. In addition to original grouping, experts are able to take their knowledge into considering in forecast.

3 Result

3.1 Conditions of verification

In this section the performance of the proposed NM forecasting method is verified by applying to literature 70 titles published in Jun 2007 by one publisher, Kodansha. The sales result data of literature 7660 titles published in 2006 is used for preparing NM coefficient tables.

The duration of 3^{rd}-7^{th} day is the first timing for decision of reprinting, because most of titles have their sales peak within 3weeks after release. The publisher has to decide how many books to be printed for not losing its business chance. The number should be decided within title's lifecycle sales result. Usually titles lifecycle end in the 6^{th} month. The forecast of N^{th} (first 14 days) to M^{th} (182^{th} day) is, therefore, used for verification of the performance. The accuracy of NM method and two existing methods, exponential smoothing method(α =0.5), moving average method (3days) are compared.

3.2 Comparison to existing forecast method

Figure 6 shows forecast examples of literature TitleA and TitleB. TitleA is top sales title among the 70 titles. Title B is the 35th sales title (average title).Though the error results of both existing forecasting methods are over twice as much as sales results, the one of the NM methods was within one times. While the sales result of TitleA was 10,999, the moving average method gives more than 80,000 error books on the 7th day. The exponential smoothing method gives more than 60,000 error books. The NM method gives within 1,000 error books.

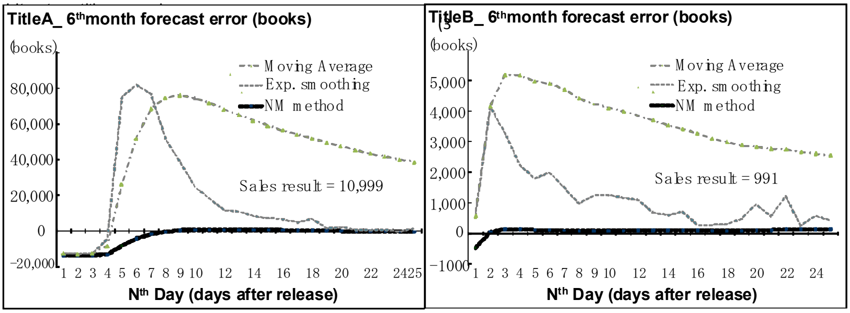


Figure 6. Long term 6th month forecast error result

To clarify the forecasting accuracy quantitatively, the error rate is defined as follows;

$$Error\_rate(\%) = \frac{SalesResult(M^{th}\,day)}{SalesForecast(M^{th}\,day)} - 1 \qquad (3)$$

where, *SalesResult* is accumulated sales result of the title till Mth day after release. *SalesForecast* is accumulated sales forecast of the title till Mth day after release.

Figure7 shows the average error rates of three forecasting methods in the 70 titles. The NM method remains the best performance throughout first 14 days. The average error rates are 509%(moving average), 273%(exponential smoothing method), 56%(NM method). Thus the NM method proved better performance. Over half titles reached within ±50% error rate on the 3rd day(Table1).

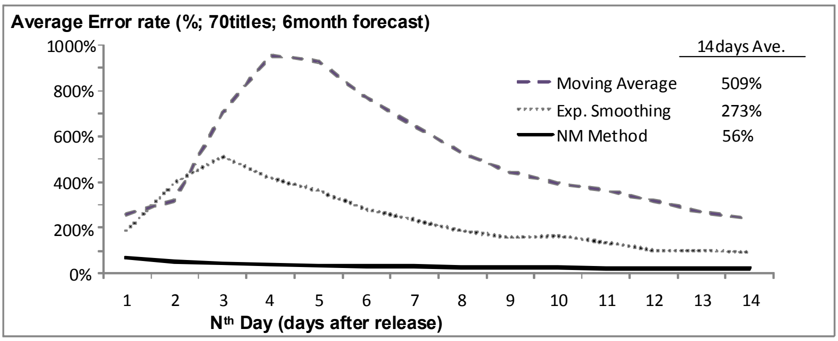


Figure 7. Comparison of average error rate

Table 1. Proportion of titles within ±50% error rate

	day1	day2	day3	day4	day5	day6	day7	day8	day9	day10
NM method	17%	49%	51%	60%	60%	61%	64%	63%	61%	61%
Moving Average	30%	11%	4%	3%	0%	0%	0%	0%	1%	0%
Exponetial	34%	10%	4%	3%	3%	7%	3%	9%	6%	7%

3.3 Grouping improvement

The performance of the NM forecasting method depends on the reference grouping. In this subsection, the existing group and knowledge based group are compared to understand if there is room to improve by grouping. It is possible to consider the necessary knowledge about the complex relationship among the titles.

The knowledge based grouping that is generated by wholesaler experts are incorporated in addition to the existing literature and publisher group. This grouping divides literature category into the detail 5 categories such as literature comics, teenager literature, general literature, academic literature and literature of general arts. It is 4^{th} day that the average error rate of knowledge based grouping reaches within 40%.on the 7^{th} day it reaches within 30%. On the other hand the literature group never reaches within 30%. On the 7^{th} day 90% of titles are within 50% error rate.

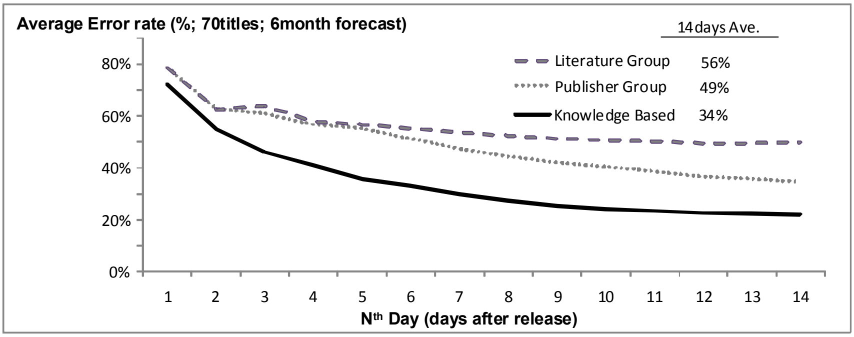


Figure 8. Comparison of average error rate(comparison of group)

Table 2. Proportion of titles within ±50% error rate (comparison of groups)

	day1	day2	day3	day4	day5	day6	day7	day8	day9	day10
Literature Group	17%	49%	51%	60%	60%	61%	64%	63%	61%	61%
Publisher Group	26%	49%	47%	60%	61%	61%	64%	69%	70%	70%
Knowledge Based	41%	56%	66%	69%	77%	83%	89%	90%	90%	90%

4 Discussion

The performance comparison to the existing methods proved the effectiveness of proposed forecasting approach for nonlinear products. While the average error rate of the 70 litareture titles during first 14 days after release of exponential smoothing forecasting method shows 273%, the proposed NM forecasting method performs 34%. It improved forecasting accuarancy by 240%. That means it is possible for publishers to decide the number of reprinting in the first 14 days. Also this forecasting method improved its accuracy by 20% with the knowledge based grouping approach, and that indicated its continuous improvement possibilities.

The verification case studies in this paper, however, are on just the literature 70 titles. It might not be enough to prove the efficiency of every area of this forecasting method. The authors need to work more on further areas on large number of products.

5 Conclusion

The research method presented in this paper provides a useful approach to forecast to irregular and nonlinear items. It showed better performance in literature book sales forecast than the existing linear methods. It also provides practical forecast from very early stage(few days after released) for publishers' reprinting decision. Furthermore by incorporating experts' knowledge as reference products group, this method have the possibility of continuous improvement of its accuracy.

6 References

[1] Mentzer, J.T., Bienstock, C.C., Sales Forecasting, Management. Sage, 1996, Thousand Oaks, CA.

[2] Agrawal, D., Schorling, C., Market share forecasting: An empirical comparison of artificial neural networks and multinomial logit model, 1996, Journal of Retailing 72 (4),383–407.

[3] Ching-Wu Chu, Guoqiang Peter Zhang, A comparative study of linear and nonlinear models for aggregate retail sales forecasting, 2003, Int. J. Production Economics 86, 217–231

[4] Elkateb, M.M., Solaiman, K., Al-Turki, Y., 1998. A comparative study of medium-weather-dependent load forecasting using enhanced artificial/fuzzy neural network and statistical techniques. Neurocomputing 23, 3-13

[5] Prybutok, V.R., Yi, J., Mitchell, D., 2000, Comparison of neural network models with ARIMA and regression models for prediction of Huston's daily maximum ozone concentration, European Journal of Operations Research 122, 31-40

[6] Callen, J.L., Kwan, C.Y., Yip, C.Y., Yuan, Y., 1996. Neural network forecasting of quarteryly accounting earnings. International Journal of Forecasting 12, 255-267

[7] Kirby, H.R., Watson, S.M., Dougherty, M.S., 1997. Should we use neural networks or statistical models for short-term motorway traffic forecasting?, International Journal of Forecasting 12, 43-50

[8] Darbellay, G.A., Slama, M., 2000 Forecasting consumers' expenditure: A comparison between econometric and neural network models. International Journal of Forecasting 12, 255-267

[9] Faraway, J., Chatfield, C., 1998. Time series forecasting with neural networks: A comparative study using the airline data. Applied Statistics 47, 231-250

[10] Shoichi Nagai, Electric Tag Implication for Publishing Industry Logistics, Japan Publisher Infrastructure Centre Survey Report, Japan Publisher Infrastructure Centre, Tokyo, 2004, pp1-22

[11] Publishing Science Research Institute, Annual Report of Publishing Indicator 2007, 2007, Domestic publishing association, Japan

Knowledge Engineering

A Negotiation Strategy of Collaborative Maintenance Chain and Its Multi-Agent System Design and Development

Amy J.C. Trappey[a, b, 1] , Wei-Chun Ni[a] and Chun-Yi Wu[a]

[a] Department of Industrial Engineering and Engineering Management, National Tsing Hwa University, Hsinchu, Taiwan.
[b] Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei, Taiwan

Abstract. The common maintenance methods in the electrical power system relies on the emergent maintenance and periodic maintenance, which is lacking to avoid the equipment shutdown. To enrich the maintenance efficiency and system benefit, this reseach uses the electrical power systems as case example to propose a negotiation model of collaborative maintenance chain. Meanwhile, to flexibly adopt new information technology and new methodologies for expert systems, this research applies multi-agent system (MAS) technology to enable the integration of different systems and databases. Via wrapping the systems and databases into agents with certain degree of authorities, these agents are able to assist humans in information communication, discussion and negotiation, and finally the generation of decisions.

Finally, this research implements a prototype system based on MAS architecture. And the MAS technology with characteristics such as pro-activity, reactivity, social ability, and autonomy into this collaborative maintenance chain is applied to enhance the communication efficiency. Moreover, this research applies game theory to design a negotiation game scenario and provides a platform for asset operation site, system provider and first-tier collaborator to negotiate their maintenance decisions. The objectives of different departments plan their resources allocation efficiently and effectively for the maintenance jobs planned and scheduled to ensure the best utilization. Consequently, this research is a devoted to develop the optimize functions as the basis for further agent-based enterprise resources negotiation.

Keywords. Engineering Asset Management, Maintenance Decision, Intelligent Multi-Agent System, Game Theory, Negotiation Mechanism.

[1] Corresponding Author: Chair Professor and Dean, College of Management, National Taipei University of Technology, Taipei (10608), Taiwan; Tel: +886-2-2771-2171 ext. 4541; Fax: +886-2-2776-3996; Email: trappey@ntut.edu.tw and trappey@ie.nthu.edu.tw.

1 Introduction

Engineering asset management (EAM) is a significant stage of the continuous life cycle, which usually includes stages such as design, construction, use, maintaining, repair, disposal and recycling [1]. Enterprises often focus on the opertating aspect of engineering assets, which may be shutdown due to lack of proper maintenance or other man-made equipment damages that lead to the stop of production/service activities. Therefore, poor EAM may cause great unexpected losses (e.g., the asset value losses and the opportunity losses due to assets dis-functioning). Hence, many academia and practitioners have made efforts in studying asset diagnosis and asset health prognosis based on the asset condition monitoring data, equipment diagnoses and life cycle prediction [2][3][4]. Traditionally, mechanical equipment is maintained only after people are conscious of exceptional events or shutdown. Works can hardly be done intelligently with real-time condition monitoring, prediction, diagnosing and pre-arrangement of maintenance. In view of the EAM challenges, the traditional experience knowledge bases are unable to deal with the complicated prediction issues [4][5]. Therefore, many enterprises cooperate with research institutes, which have professional diagnosis and prediction knowledge, to help enterprises solve complicated engineering asset problems efficiently and effectively. When the demands for high-level maintenance services happen, system providers can participate in maintenance consulting. They will negotiate with service receivers and maintenance chain participants in order to fulfill the maintenance demands on time. Consequently, maintenance scheduling and coordination communication become critical for satisfactory maintenance fulfillment [6]. This research challenges the above-mentioned issues. The problem of how to effectively coordinate the operations among different parties and resources will be the focus of this research. We use the current network system of power transmission and distribution as case study to depict the EAM methodology and IT solutions.

In the case study, this research will develop the preventive maintenance decision support specifically for the large-scale assets of voltage transformers. The multi-agent system technology, which have some characteristics such as autonomous, social, goal-oriented, reactive, rational and self-learning, will be deployed to ensure information exchange, communication and decision supports efficiently [7][8][9][10]. In the core of agent behavioral modeling, this research will design and develop a negotiation mechanism to strengthen the cooperation of the entire collaborative maintenance chain. Moreover, this research will implement an intelligent IT prototype system based on the framework of the collaborative process of maintenance chain and the negotiation mechanisom.

2 Current Practice of Collaborative Maintenance Chain

In this section, we first introduce the case description and, then, analyse the exitsting problems in the current transformer maintenance practice. The proposed to-be model with negotiation game mechanism is depicted.

2.1 Case Description

The electrical power system can be regarded as a huge system network consisting of electrical power generators, voltage transformers and connecting power cables. The critical issue is how to judge and detemine the maintenance times and types of numerous transformers in the system to ensure reliable power supply. Further, how to maintain the profitability of the entire system. The profit of the electric power system is often used to measure the financial status of the utility company. It is also regarded as an indicator of a nation's social welfare and its competitiveness.

However, nowadays, the common maintenance methods in the current electrical power system relies on the facility's staffs to identify the maintenance demands. In most cases, they send mainenance orders to the system providers after the equipment shutdown has occurred. In order to overcome the above problems, this reseach uses the large-scale voltage transformer (Figure 1) as case example to discuss the sudden emergency behaviors and maintenance decision model. We construct a negotiation model of collaborative maintenance chain applying game theory in order to maximizing the entire system benefit.

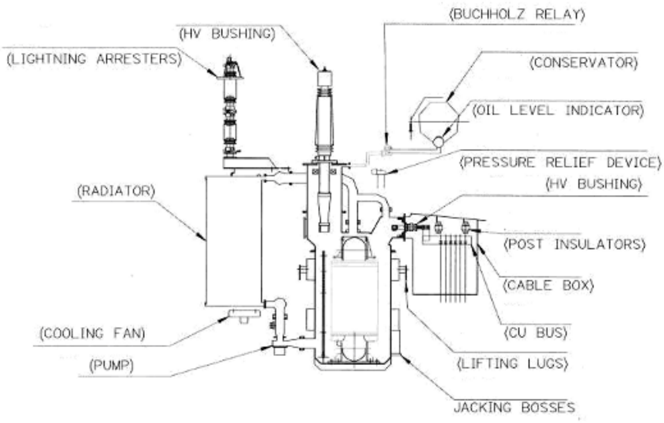


Figure 1. The structure drawing of a large-scale voltage transformer [11]

2.2 Current Maintenance Practice Analysis

The maintenance development course starts from accident repair, going through daily maintenance and then periodic maintenance. However, it is unable to prevent the unexpected shutdown in equipment system. As a results, we will utilize the state-of-the-art condition monitoring and data analysis technique in this research and set up an agent-base coordination predictive and preventive maintenance chain. It includes three main types of participants as shown in Figure 2.

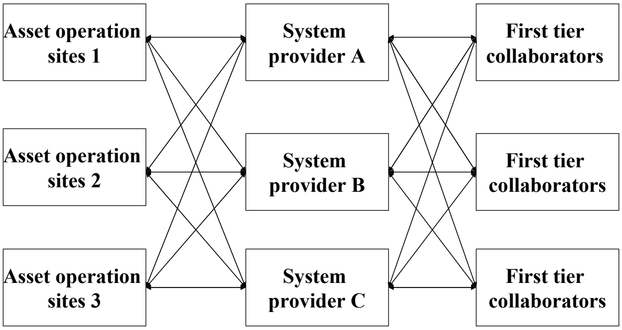


Figure 2. The current maintenance chain structure

Asset operation sites: The stakeholder, user and/or operator of the voltage transformer, who utilizes the engineering asset for production or services, and is the asset maintenance demander (or service receiver).
System provider: The engineering asset producer or maintenance serve provider performing jobs such as routine, emergent and preventive maintenances for asset operation sites.
First-tier collaborators: The first-tier collaborator offers human resources (maintenance crews) and maintenance parts for the jobs.

Therefore, the voltage transformer's current maintenance model is as complicated as shown in Figure 2. In this model, the different asset operation sites and different system providers all consider their own objectives while conducting maintenance tasks, which may not be feasible due to the resource limitations in the entire system. After the current maintenance analysis, we find two key directions to implement this maintenance practice, which includes the routine maintenance and the emergency repair. Figure 3 presents their operational sequence diagram in the current practice.

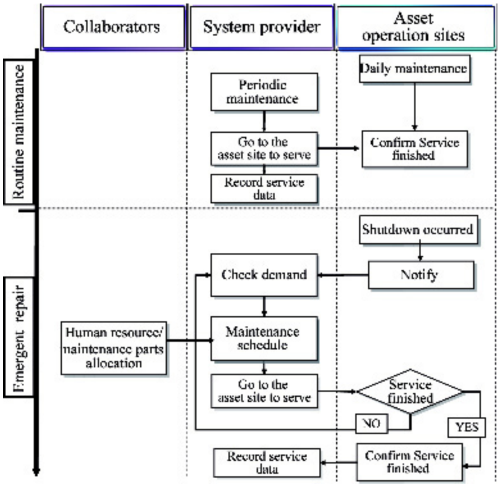


Figure 3. The operational sequence diagram of the current maintenance practice

According to the above analysis, we can generalize several problems in the current maintenance practice.

- The lack of life cycle information of the voltage tranformer. Besides daily maintenance and periodic maintenance, there is no other preventing maintenance method to avoid exceptional causes leading to great damage in the power network system.
- The lack of historical maintenance records and codition monitoring data.
- When asset operation sites notify the system provider of maintenance demand, the coordinating time is long and often risks further damage of the equipment.

2.3 To-be Model of Collaborative Maintenance Chain

In order to solve the above-mentioned problem, this research proposes a service center model. The service center has the dominant expertise of diagnosis and prognosis techniques. It also can monitor engineering assets in the asset operation sites and integrate the supply and demand between asset operation sites and the system providers (Figure 4).

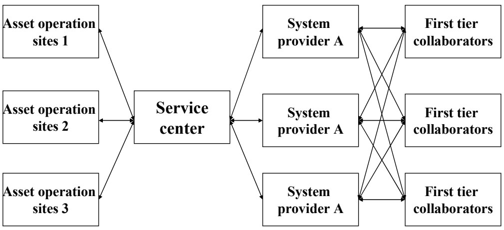


Figure 4. The collaborative (to-be) maintenance chain structure

There are usually three aspects into the maintenance decision-making, i.e., the strategic level, the tactical level and the operational level [12][13].

At the strategic level: This aspect includes the preparation and recovery stages. The decision-marking largely considers the organizational and financial impacts of the decisions. In this research, we first evaluate the maintenance cost according to the reliability-centered asset maintenance method (RCAM) [14].

Total maintenance cost [cost/yr]: $TCPM(S) = CPM_f(S) + CPM_{PM}(S) + CPM_{\text{int}}(S)$

$CPM_f(S)$ The cost of failure (repair costs);

$CPM_{PM}(S)$ The cost of preventive maintenance (planed maintenance or replacement of a transformer in advance of failure);

$CPM_{\text{int}}(S)$ The cost of interruption (power interruption affects the user, and may suffer direct costs and/or be compensated via a penalty payment);

S Preventive maintenance strategy.

After changing into the marginal cost $CPM_{g,t}$ [cost], it will combine with Cournot-Nash equilibrium model (CNE) [15] to form the optimization model considering both equipment reliability and system profit.

$$\textit{Maximize} \sum_t \left(\alpha_t - \frac{1}{2}\beta_t Y_t \right) Y_t - \sum_{g,t} G_{g,t} C_{g,t} - \sigma_{g,t} CPM_{g,t} - \sum_{i,t} \frac{1}{2}\beta_t X_{i,t}^2$$

$$\textit{Subject to} \sum_i X_{i,t} = Y_t \;\; \forall t$$

$$\sum_{g \in \Lambda_{i,g}} G_{g,t} = X_{i,t} \;\; \forall (i,t)$$

$$G_{g,t} \le G_{g,t}^{\max} \times (1 - \sigma_{g,t}) \;\; \forall (g,t)$$

$$\sum_t \sigma_{g,t} = \Omega_g \quad \forall g$$

$$\sum_g \left(G_{g,t}^{\max} \times (1 - \sigma_{g,t}) - G_{g,t} \right) \ge R_t \;\; \forall t$$

$$X_{i,t}, G_{g,t}, Y_t \ge 0$$

$$\sigma_{g,t} = 0 \; or \; 1$$

α_t, β_t Linear demand equation parameters for t;

Y_t Total power supply in t [MW];

$X_{i,t}$ genco I generation [MW];

$C_{g,t}$ marginal cost of transformer g in t [$/MWh];

$CPM_{g,t}$ Total maintenance cost of transformer g in t [cost];

$G_{g,t}$ Transformer g load in t [MW];

$G_{g,t}^{\max}$ Transformer g max load capacity in t [MW];

$\sigma_{g,t}$ Binary maintenance decision variable for g in t;

Ω Number of periods transformer needs to be on maintained;

R_t Reserve requirement in period t.

In this model, we can execute the decision-making of preventive maintenance according to the voltage transformer's loading. In the other words, the model can determine the maintenance time, maintenance period which transformer needs to be serviced (namely $\sigma_{i,t}$) according to whole system losses in that transformer supports coverage (namely $CPM_t^i(S)$). Therefore, this optimization model can reach the best state considering both reliability and system profit (i.e., the strategic views of social welfare and national competitiveness).

At the tactical level: This includes prediction and prevention stages. The decision-marking considers mainly the medium-term problems and associated objectives. In this research, the Java Agent Development Framework (JADE) [16] is deployed to construct multi-agent system for the enabling of collaborative maintenance chain.

At the operational level: This includes the detection and response stages. This stage requires timely decision making. In this research, we provide the maintenance decision support agent (MDSA) and the service system agent (SSA) for constructing the emergency maintenance negotiation mechanism. The steps of the negotiation mechanism are described as follows.

- Step 0: Initialization: When the exceptional situations occurred, MDSA offer the maximal maintenance cost, minimal maintenance cost and compensation to SSA.
- Step 1: Feasibility analysis: Check all restriction, including time restriction, region restriction and technology restriction.
- Step 2: Evaluation of the maintenance cost, time and period:
  - Evaluate dangerous level and recovery time limit.
  - Evaluate maintenance time and cost according to the possible failure cause.
- Step 3: Updates of maintenance decisions: MDSA reacts new decision-making (or accept) according to SSA result, but each new response needs to pay compensation for this action.
- Step 4: Coordination state record: All MDSA have the final result, update SSA and retrun to Step 1.

Finally, the operational sequence diagram of collaborative maintenance chain is shown in Figure 5.

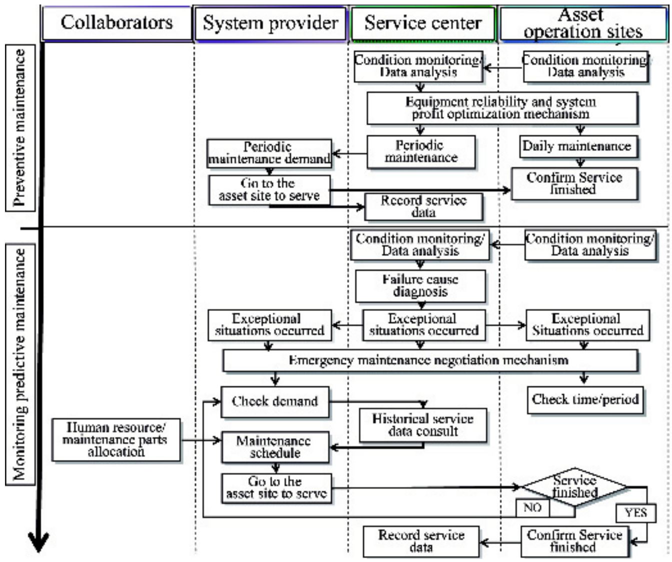


Figure 5. The operational sequence diagram of to-be collaborative maintenance chain

3 Agent-based Collaborative Maintenance Chain Platform

The proposed information platform is supported by software agents to autonomously help communicate and negotiate among maintenance chain participants. In the IT platform design phase, the expected goals, the agent identities and how agents achieve these goals are summarized. Meanwhile, we must identify the agent types required in the system after defining the functions and users of the system. Table 1 briefly describes the agents and the goals in the prototype system.

Table 1. Goals and agents analysis

Goal	Agent type:	How to achieve the goal
Data/signal extraction	Monitoring Agent (MA) Asset Agent (AA)	Condition monitoring system Data/ signal transformation
Accurate diagnosis	Diagnosis Agent (DA)	Collect knowledge from diagnosis experts Integrate expert knowledge into diagnosis knowledge base Provide diagnosis result
Reliable prognosis	Prognosis Agent (PA)	Collect knowledge from prognosis expert Integrate expert knowledge into prognosis knowledge base Provide prognosis result Provide system profit optimization mechanism
Timely and reliable maintenance	Service System Agent (SSA) Maintenance Decision Support Agent (MDSA) System-Provider Maintenance Scheduling Agent (SMSA)	Arrange maintenance provider Provide emergency maintenance negotiation mechanism Provide expected maintenance time Provide expected maintenance start time Have enough maintenance resources
Personalized interface	Human-Resource Agent (HA)	Provide personalized work list
Inventory management	Spare Part Agent (SPA)	Notification of procurement

During the platform development phase, more detailed blueprints of agent conversations are built based on JADE, which follows the Foundation for Intelligent Physical Agents (FIPA) specifications and provides Graphical User Interface (GUI) to enable easy development and debugging of the system. Then, using the integrated agent relationship diagrams, the overall agent-based collaborative maintenance platform is generated as demonstrated in Figure 6.

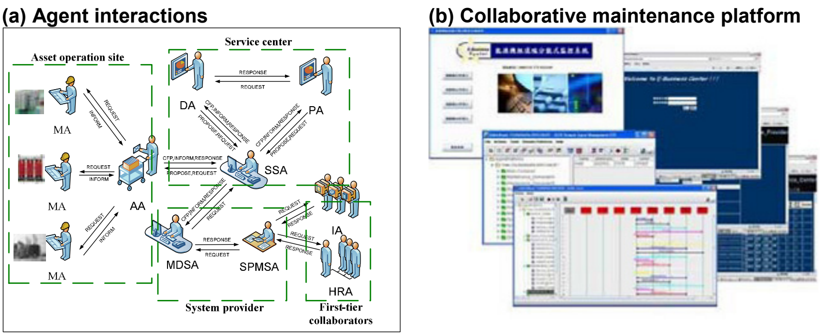


Figure 6. Prototype system implementation: (a) the model of agent interactions and (b) the web-based interfaces

4 Conclusion

The purpose of this research is to provide complete collaborative maintenance chain architecture and implement the framework using multi-agent techniques. A collaborative business model, integrating maintenance chain members with a new intermediary agent (the service center), is proposed. The to-be model depends on accurate diagnosis and reliable prognosis to combine asset operational sites, system providers and first tier collaborators. Meanwhile, the platform also provides the equipment reliability and system profit optimization mechanism and emergency maintenance negotiation mechanism to enable an intelligent and collaborative maintenance practice. This research takes related maintenance resources into consideration so as to reach the optimization of the entire maintenance system benefits.

5 Acknowledgement

This research is partially supported by the research projects funded by the Industrial Technology Research Institute (ITRI) and the National Science Council.

6 References

[1] CIEAM (CRC for Integrated Engineering Asset Management). Available at: <http://www.cieam.com/>, Accessed on: Mar. 15th 2008.
[2] Majidian, A. and Saidi, M.H., 2007, "Comparison of fuzzy logic and neural network in life prediction of boiler tubes," *International Journal of Fatigue*, Vol. 29, pp. 489-498.
[3] Sun, Y., Ma, L., Mathew, J., and Zhang, S., 2006, "An analytical model for interactive failures," *Reliability Engineering & System Safety*, Vol. 91, Issue 5, pp. 495-504.
[4] Yao, Y.H., Lin, G.Y.P., and Trappey, A.J.C., 2005, "Using knowledge-based intelligent reasoning to support dynamic equipment diagnosis and maintenance," *International Journal of Enterprise Information Systems*, Vol. 2, No. 1, pp. 17-29.
[5] Sun, Y., Ma, L., Mathew, J., Wang, Y., and Zhang, S., "Mechanical systems hazard estimation using condition monitoring," Mechanical Systems and Signal Processing, Vol. 20, Issue 5, pp. 1189-1201 (2006)
[6] Huang, C.J., Trappey, A.J.C., and Yao, Y.H., 2006, "Developing an agent-based workflow management system for collaborative product design," *Industrial Management and Data System*, Vol. 106, No. 5, pp. 680-699.
[7] Davidson, E.M., McArthur, S.D.J., McDonald, J.R., Cumming, T., and Watt, I., 2006, "Applying multi-agent system technology in practice: automated management and analysis of SCADA and digital fault recorder data," *IEEE Transactions on Power Systems*, Vol. 21, Issue2, pp. 559 – 567.
[8] Hossack, J.A., Menal, J., McArthur S.D.J., and McDonald, J.R., 2003, "A multiagent architecture for protection engineering diagnostic assistance," *IEEE Transactions on Power Systems*, Vol. 18, Issue 2, pp. 639 – 647.
[9] McArthur, S.D.J., Booth, C.D., McDonald, J.R., and McFadyen, I.T., "An agent-based anomaly detection architecture for condition monitoring," IEEE Transactions on Power Systems, Vol. 20, Issue 4, pp. 1675 – 1682 (2005)

[10] Trappey, A.J.C., Trappey, C.V., and Lin, F.T.L., 2006, "Automated silicon intellectual property trade using mobile agent technology," *Robotics and CIM*, Vol. 22, pp. 189-202.
[11] Fortune, http://www.fortune.com.tw/.
[12] Tien, J. M., 2005, "Viewing urban disruptions from a decision informatics perspective," *J. Syst. Sci. Syst. Eng.*, Vol. 14, No. 3, pp. 257–288.
[13] Hipel, K. W., Jamshidi, M. M., Tien, J. M., and White III, C. C., 2007, "The future of systems, man, and cybernetics: application domains and research methods," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 37, No. 5, pp. 726-743.
[14] Bertling, L., Allan, R., and Eriksson, R., 2005, "A reliability-centered asset maintenance method for assessing the impact of maintenance in power distribution systems," *IEEE Transactions on Power Systems*, Vol. 20, No. 1, pp. 75-82.
[15] Chattopadhyay, D., 2004, "A game theoretic model for strategic maintenance and dispatch decisions," *IEEE Transactions on Power Systems*, Vol. 19, No. 4, pp. 2014-2021.
[16] JADE (Java Agent DEvelopment framework), http://jade.tilab.com, Accessed 28/5/2007.

Develop Non-Exhaustive Overlapping Partitioning Clustering for Patent Analysis Based on the Key Phrases Extracted Using Ontology Schema and Fuzzy Adaptive Resonance Theory

Amy J.C. Trappey,[a, b, 1] Charles V. Trappey,[c] and Chun-Yi Wu[b]

[a]Department of Industrial Engineering and Management, National Taipei University of Technology, Taiwan
[b]Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan
[c]Department of Management Science, National Chiao Tung University, Taiwan

Abstract. Patent documents are an ample source of innovative, technical and commercial knowledge. Thus patent analysis has been considered a useful vehicle for R&D management and knowledge sharing. However, the number of patents in recent years has expanded noticeably. It is a challenge to interpret the knowledge contents and their relationship between patents, because many different types of expression may be presented in patents. Therefore, various clustering methods, e.g., partitioning clustering (K-means, K-medoids) or competitive clustering (Self-Organizing Maps, Adaptive Resonance Theory), have been developed to elucidate the characteristic of expression pattern. Nevertheless, patent documents can be partitioned into a number of clusters, where each cluster represents the documents in a certain area. The traditional clustering algorithms do not allow an object to belong to multiple clusters. In this research, we develop a novel non-exhaustive overlapping partitioning clustering (OPC) algorithm, a type of fuzzy partitioning approach, to the patent documents to overcome the exclusive clustering methods. The proposed algorithm considers the dissimilarity among cluster centers and an object can simultaneously belong to multiple clusters if the distances from this object to the cluster centers are no more than the given threshold value. Finally, the experiment uses radio frequency identification (RFID) technology related patent documents to evaluate the performance of the proposed clustering method, which combine non-exhaustive OPC algorithm with fuzzy adaptive resonance theory (ART) for effective patent clustering.

Keywords. Overlapping partitioning clustering (OPC), Adaptive resonance theory (ART), Patent analysis.

[1] Please send all correspondence to Professor Amy Trappey, Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei (10608), Taiwan, E-mail: trappey@ntut.edu.tw. Tel: +886-2-2771-2171 Ext. 4541; Fax: +886-2-2776-3996

1 Introduction

Facing the challenges from a knowledge-based economy, having a comprehensive understanding of technology development is the basic and necessary requirement to gain competitive edges for enterprises. Patent documents often consist of innovative, technical and commercial knowledge. Through careful analysis, patent documents can show technological details and relations, reveal business trends, inspire novel industrial solutions, and help technology investment strategy [1]. Hence, patent analysis has been considered a useful vehicle for R&D management and knowledge sharing. Patent analysis is to interpret the technical and knowledge interactions between related patents, which are not intutitive due to the huge number of patents and their different ways of expressing and describing the inventions. In order to providing a unified view, clustering with an overview has received considerable attention in recent patent analysis efforts. Clustering is the process of grouping a set of objects into clusters so that objects within a cluster are similar to each other but dissimilar to objects in other clusters [2]. A recent research has classified clustering methods into several categories, i.e., partitioning methods, hierarchical methods, competitive clustering methods, density-based methods, and grid-based methods [3]. In traditional cluster methods, one object can only belong to one cluster and each object is assigned to the nearest cluster. Nonetheless, patent documents can be partitioned into a number of clusters, where each cluster represents the documents with a meaningful focus. Therefore, this research proposed a novel non-exhaustive overlapping partitioning clustering (OPC) algorithm, a type of fuzzy partitioning method, to the patent documents. Thus, each patent now can belong to multiple clusters rather than a single cluster. In our case study, the experiment uses a set of radio frequency identification (RFID) patent documents to evaluate the performance of the proposed algorithm, which combines non-exhaustive OPC with fuzzy adaptive resonance theory (fuzzy-ART) for the most suitable and effective patent clustering.

This paper is organized as follows. First, in this introductory section, the general background of patent analysis and clustering is presented. Next section will provide literature reviews of existing clustering algorithms. Section 3 presents the non-exhaustive overlapping partitioning clustering and fuzzy adaptive resonance theory. In Section 4, the RFID patent documents from USPTO database are searched and analyzed using the proposed clustering algorithm. Comparisons are made between the results obtained from the suggested overlapping partitioning clustering with the competitive cluster methods. Finally, Section 5 presents the conclusion and the contribution of this research.

2 Literature review

The literature review focuses on three key topics. The first topic is the analysis of patent documents, including the importance of patent analysis, the scope of patent strategy and the types of deployment models. The second focus discusses the research in most commonly applied clustering algorithms, such as k-means and k-medoids. The last section discuss the adaptive resonance theory for clustering.

2.1 Patent analysis

Patent analysess have often been employed as economic indicators that measure the linkages between technology development and economic growth [4]. Recently, the strategic importance of patent analysis is highlighted in high-technology sectors as the processes of innovation becomes complex, the cycle time of innovative R&D is shorten, and the market demand becomes volatile. The practical applications of patent analysis are wide among technology-oriented professions, e.g., R&D managers, academicians, patent attorneys, and technology policy makers. They use patent analysis to estimate technological knowledge trends, impacting on productivity and profitability, and comparing innovative performance in the international context [5, 6]. As Granstrand [7] pointed out the types of patent deployment models include specific blocking and inventing around, strategic patent design, blanketing and flooding, fencing, surrounding design, and hybrid strategies. Successful patent analysis not only can save cost of enterprise research and development, but also strengthen their R&D efforts offensively and defensively. Therefore, this research use variant clustering method to rapidly and automatically analyze the features of patents and extract the multi-technological characteristics of patents.

2.2 Clustering algorithm

Clustering is the unsupervised classification of patterns into groups. It is an important data analysis technique, which organizes a collection of patterns into clusters based on their similarity of given attributes. Clustering is useful in several situations including exploratory pattern analysis, grouping, decision making, and machine learning. Thus, clustering is useful in the applications, such as document retrieval, image segmentation, and pattern classification [8]. The analysis clustering is a sequence process, which consists of data collection, measuring the similarity of data, choosing the suitable clustering methods, evaluating the performance of chosen method, and interpreting the results of clustering for meaningful analysis by domain expert as shown in Figure 1. The similarities of data sets are usually represented in terms of the distance in space with the well known methods such as Euclidean distance and Manhattan distance (Equation 1).

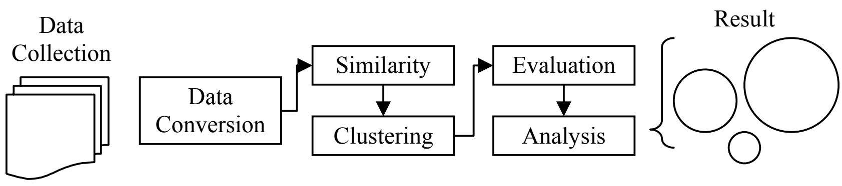


Figure 1. The process of clustering

Euclidean distance $d_E(x_i, x_j) = \left(\sum_{d=1}^{k} \left| x_{id} - x_{jd} \right|^2 \right)^{1/2}$;

Manhattan distance $d_M(x_i, x_j) = \sum_{d=1}^{k} |x_{id} - x_{jd}|$ (1)

The purpose of application usually influences the clustering algorithms to be chosen. The partitioning algorithms are often used to cluster circle based group and the same size of clusters. And hierarchical algorithms or density based algorithms can cluster natural shape group and arbitrarily sizes of clusters. This research focuses on partitioning algorithms. K-means is one of the popular partitioning methods for clustering. The K-means approach iteratively finds the k centroids and assigns the objects to their nearest centroids, where the centroids are the mean of the objects in the given cluster. However, there are four types of problems affiliate with K-means approach [3].

1. The cluster center is not an object of the data set.
2. The result of clustering may not be globally optimal because the clustering results depend on how we initially choose the initial centroids.
3. It does not consider the noise data and outliers.
4. It is not suitable to discover clusters with non-circle shapes.

K-medoids clustering are sometimes applied where representative objects called medoids are considered instead of centroids. Because it is based on the most centrally located object in a cluster, it is less sensitive to outliers in comparison to the K-means clustering. The advantages of this approach over the K-means algorithm contain three folds [9].

1. It is less sensitive to noise data and outliers.
2. It uses a real object to represent the cluster center.
3. It can cluster categorically of non-numerical data.

All the clusters obtained by K-means or K-medoids are exclusive model which allows each object be clustered in only one group entirely. The objective of these algorithms is to minimize the sum of distances between object centers. In this research, we suggest a non-exhaustive clustering method [10] based on overlapping partitioning clustering algorithm. The key purposes considered in this method are to make the number of objects in a cluster as large as possible and ensure the cluster centers as far as possible from each others. Further, the key concept of OPC is that an object can be assigned as a member of multiple clusters.

2.3 Adaptive resonance theory network

Clustering is an important function in data mining. Its typical application includes the analysis of consumer's materials. Adaptive resonance theory (ART) network is very popular for the applications of unsupervised neural network. Type I adaptive resonance theory network (ART-1) deals with the binary numerical data, whereas type II adaptive resonance theory network (ART-2) deals with the general numerical data. Several ART-based information systems also collect the mixed attitudes, including numeric attributes and categorical attributes [11].

Adaptive resonance theory is a neural networks model aiming for real-time prediction, search, learning, and recognition. The topology of the network consists of one layer of input nodes and one layer of output nodes. The input units are connected to all the output nodes and feedback connections are provided from the output nodes to the input nodes.

- b_{ij} = bottom-up weights from input node i to output node j.
- t_{ij} = top-down weights from output node j to input node i.

When the input is presented, the specimen with the highest matching score is selected by competitive learning (Formula 2). If the degree of similarity required for patterns is greater than or equal to vigilance parameter (ρ), the output node J is chosen as shown in Formula 3.

$$y_J = \max_j\{y_j\} = \max_j\{\sum_{i=1}^{n} b_{ij}x_i\} \text{ , } x_i = (x_1, x_2, \sim, x_n) \text{ is the input vector.} \quad (2)$$

$$\frac{\sum_{i=1}^{n} t_{Ji}x_i}{\sum_{i=1}^{n} x_i} \geq \rho \quad (3)$$

The weights of b_{ij} and t_{ij} for the winning cluster unit are adjusted by using the Formula 4. Otherwise, the input is considered to be different and is added as a member of a new cluster. Therefore, the number of cluster is trained by competitive learning.

$$b_{iJ}^{new} = \frac{L * t_{Ji}^{old} * x_i}{L - 1 + \sum_{i=1}^{n} t_{Ji}^{old} * x_i} \text{ ; } t_{Ji}^{new} = t_{Ji}^{old} * x_i \quad (4)$$

A central feature of all ART systems is a pattern matching process that compares an external input with the internal memory of an active code. The major difference between ART and other unsupervised neural networks is the vigilance parameter. In the processing of ART, a degree of similarity between a new pattern and the stored pattern is defined. This similarity (ρ) is a measure to ensure whether the new pattern is properly classified. The other unsupervised learning neural networks without implementing vigilance may cause a significantly different input pattern to be forced into an inappropriate cluster. The detail of proposed fuzzy ART which is designed to perform for both continuous-valued and binary-valued input vectors is described in section 3.1.

3 Patent clustering algorithm

This research has proposed two parts in the process of patent analysis. The first part is the key phrases extraction focusing on the discrimination of patent document context. The process of extraction contains constructing the ontology of the study domain (RFID in our case study) [12], words segmentation, stopping, stemming, and term weight calculating, key phrases collection using well known

NTF-IDF method. After key phrase collection, the fuzzy adaptive resonance theory can efficiently analyze the patent document based on the frequency of key phrases extraction for providing the trained key phrases clusters. The second part of the process uses non-exhaustive overlapping partitioning clustering based on the results of fuzzy-ART to analyze the patent document clusters. The two-part process is shown in Figure 2.

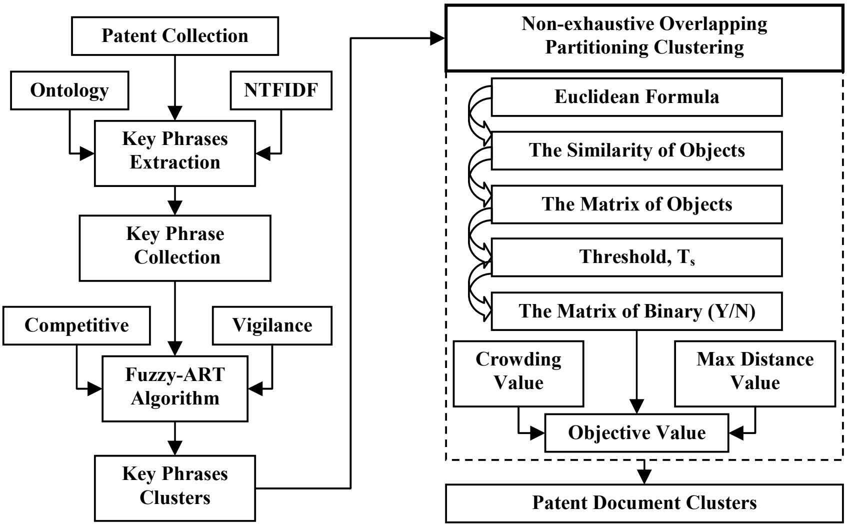


Figure 2. The two-part process of patent clustering algorithm

3.1 Fuzzy adaptive resonance theory

A fuzzy-ART model has been introduced by Carpenter et al. [13]. Figure 3 shows the structure of fuzzy-ART. And, this theory is capable of rapid stable clustering for analogy or binary input patterns. Each input is represented by an M-dimensional vector from the key phrase of patent documents, where input I = [I_1, I_2,…, I_M], and $I_{k,k=1\sim M}$ means the frequency of key phrase of patent documents in the interval [0, 1]. Moreover, each unit has a fixed amount of weights and the weights are limited to value between 0 and 1.

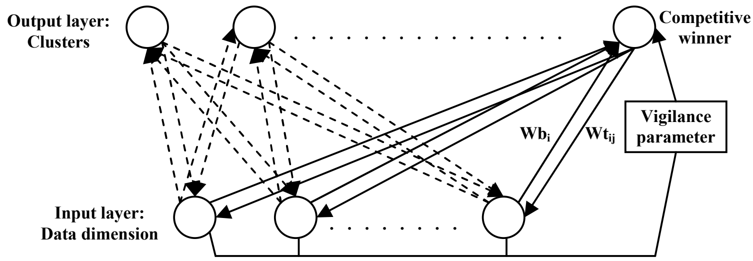


Figure 3. The structure of ART

In fuzzy-ART, the choice function (T_j) for every output node is represented $T_j = \frac{\|I \wedge w_j\|}{\alpha + \|w_j\|}$ for nodes j=1 to M, where $\wedge$ is the fuzzy AND operator, defined as $(x \wedge y) = \min(x_i, y_i)$. The best match function is $T_\theta = \max_j \{T_j\}$ and the degree of similarity with best matching exemplar for resonance test is represented as Formula 5. Finally, the update of weight is defined as Formula 6.

$$\text{Similarity} = \frac{\|I \wedge w_\theta\|}{\|I\|} \geq \rho,\ \rho \text{ is vigilance parameter} \quad (5)$$

$$w_\theta^{new} = \beta(I \wedge w_\theta^{old}) + (1-\beta) w_\theta^{old} \quad (6)$$

A fuzzy-ART model generates a new cluster if the existing clusters are not suitable for the input pattern. Competitive learning of fuzzy-ART is done as follows.

1. Set with random arrangement of weights and give all output units a chance to compute.
2. Choose the winner cluster that has the biggest value of choice function.
3. Judge the cluster resonance or reset by the match function. Resonance occurs if the value of match function is bigger than vigilance parameter.
4. Then the weight of winner cluster is updated according to input vector. Otherwise, reset occurs. And, new cluster is generated according to input vector.
5. Continue the above cycle until reset will not happen for each input vector.

The network is presented with a set of training patterns, but no target is given for each input pattern. Thus, the network organizes the training patterns into a set of clusters on its own. Moreover, fuzzy-ART belongs to competitive learning model. The architecture of a competitive learning consists of a set of hierarchically layered units in which each layer connects, via excitatory connections, with the layer immediately above it. The clusters are winner-take-all, such that the unit receiving the largest input achieves its maximum value, while all other units in the cluster are pushed to their minimum value.

3.2 Non-exhaustive overlapping partitioning clustering

This research combines the key phrase clusters of fuzzy adaptive resonance theory with non-exhaustive overlapping partitioning clustering for patent clustering. This method assumes that the distance of objects is computed by Euclidean formula. A threshold T_s for similarity is defined to determine whether two objects (x_i, x_j) belong to one cluster or not. The objects in patent analysis mean patent documents. The similarity of object x_i and x_j is expressed as Formula 7. If $Cor_{i,j}$ is larger than threshold T_s as Formula 8, this means object x_i and x_j should belong to same cluster.

$$Cor_{i,j} = 1 - \min\{d_{ij}, d_{if}\} / d_{if} \tag{7}$$

where d_{if} is the top 5% percentile of distances of all object-pairs.

$$d_{if} - \min\{d_{ij}, d_{if}\} > T_s \times d_{if} \tag{8}$$

Thus, the matrix of objects will be noted as Y, otherwise these two objects are noted as N. Finally, the Y/N matrix can be obtained from table of correlation of patent documents. Moreover, the purposed of non-exhaustive OPC algorithm is maximizing the number of objects in a cluster and maximizing the distances between each cluster center. The first objective Cv can be calculated and n_{xi} represented the number of objects in cluster center x_i. The second objective Mdv can be expressed and ndx_i represented the distance of object x_i to the nearest center. Therefore, the Cv and Mdv are defined as Formula 9. Finally, these objectives of non-exhaustive OPC algorithm are integrated by a portion $w1$ and $w2$ as CRF, shown in Formula 10.

$$Cv(x_i) = \frac{n_{xi}}{n_{\max}},\ n_{\max} = \max(n_{x1},...,n_{xi});\ Mdv(x_i) = \frac{ndx_i}{\max d},\ \max d = \max(ndx_1,...,ndx_i) \tag{9}$$

$$CRF(x_i) = w1 \times C_v(x_i) + w2 \times Mdv(x_i) \tag{10}$$

According to the CRF value of every object, the object with higher CRF value is more suitable of being cluster center. The opportunity of being cluster center can be assigned according to its CRF value, Prod(x_i) = $\frac{CRF(x_i)}{\sum_{i=1}^{n} CRF(x_i)}$. The object with higher CRF value is more possible to be randomly selected as cluster center. After the initial cluster centers is obtained, the objective value of this combination of cluster centers are calculated, as shown in Formula 11.

$$Obj = w1 \times Min\left[Mdv(x_{c1}), Mdv(x_{c2}),..., Mdv(x_{ck})\right] + w2 \times \frac{\sum_{i=1}^{k} Cv(x_i)}{k} \tag{11}$$

One of these cluster centers is replaced to form a new combination of cluster centers. The new combination will replace the former combination, if the *CRF* value is higher. Otherwise, this combination of cluster center will be kept until a new combination with higher *CRF* valued is found. This operation repeats a predefined cycles to obtain the best objective of cluster center.

4 Patent analysis – RFID case study

In our research, there are more than 300 patents collected for the RFID technology innovations and inventions, e.g., designs of antenna, reader and RFID tags, from United States Patent and Trademark Office (USPTO). The proposed algorithm which combine fuzzy adaptive resonance theory and non-exhaustive overlapping partitioning clustering algorithms can process the key phrase clusters and the patent document cluster after accomplishing the process of patent collection. In order to improve the effectiveness of key phrase extraction, this research builds the RFID ontology hierarchically by domain expert before the patent document analysis. Each node represents a specific domain key-phrase. Using the RFID ontology and the NTF-IDF methods, the key phrases are extracted as shown in Table 1. After the uploading of patent documents, this research calculates the NTF-IDF values of all key phrases to rank the top ranked phrases and matche the nodes of ontology tree.

Table 1. The sub-list of RFID key phrases extracted based on RFID ontology

Key Phrases	TF	IDF	NTF-IDF	Method
linear magnetic material	226	2.907	656.957	TF-IDF
transmitter key device	130	3.907	507.896	TF-IDF
multi meter terminal	128	3.907	500.082	TF-IDF
rfid reader writer	132	2.907	383.71	TF-IDF
control circuit	76	3.907	296.924	TF-IDF
thin film transistor	69	3.907	269.575	TF-IDF
spread spectrum radio	66	3.907	257.855	TF-IDF
antenna	209	0.737	154.026	Ontology
band	96	1.322	126.905	Ontology
reader	68	0.585	39.777	Ontology

The process of fuzzy-ART has two steps, i.e., training and testing. The first step is to use the most representative patent documents to train the weights of a competitive network and build the key phrases bank. In this research, 200 RFID patents are taken as training patents. The second step is to upload other 100 patents to match the trained weights of fuzzy-ART for testing and validating the network. The key phrases of patent documents are clustered according to their correlations by competitive learning and vigilance parameter. Therefore, the number of Key Phrase Clusters (KPC) is calculated for the non-exhaustive OPC algorithm as shown in Table 2.

Table 2. The key phrase clusters are generated using fuzzy-ART

Cluster	Phrase	Phrases in The Same Cluster
KPC 1	circularly polarized antenna	circularly polarized signal/ dual band/ different frequency bands/ ground plate/ bevels bottom/ polarized signal simultaneously/…
KPC 2	linear magnetic material	tag antenna/ soft combination tag/ rf tag/ magnetic material providing/ frequency magnetic field/ low frequency magnetic/ tag antenna comprises/…
KPC 3	electronic device	electronic protection device/ electronic tag/ magnetic electronic article/ write electronic tag/ electronic article identification/…
KPC 4	control circuit	lc circuit/ rfid circuit/ short circuited stub/ reader integrated circuit/ circuit voltage/ resonant circuit/ power storage circuit/…
KPC 5	multi meter terminal	antenna terminals/ spread spectrum radio/ radio frequency signal/ integrated radio module/ multiple rfid readers/ multi position trigger/...

After the processing of fuzzy-ART, this research analyzes the meaningful characteristics of patent documents within key phrase clusters. Based on the result of fuzzy-ART, we got the five key phrase clusters. Further, the non-exhaustive OPC algorithm sets the threshold (0.9) as similarity, the weight of crowding value (0.5), the maximal distance value (0.5), and the training cycle times (30) for learning epoch. In Table 3, 100 testing patents are clustered into five patent document clusters (PDC). The result shows the patent overlapping clusters. Using the stepwise model improvement framework, the clustering results are steadily improved through the expansion of key phrase bank by increasing the number of training patent documents.

Table 3. The overlapping patent document clusters (each cluster choose Top 4)

Cluster	Patent Number		Patent Title
PDC 1	**Center**	**7304579**	RFID reader to select code modules
		7345635	Apparatus for encapsulating radio frequency identification (RFID) antennae
		7374105	RFID tag with improved range
		7364088	Radio frequency identification tag
PDC 2	**Center**	**7317426**	Core antenna for EAS and RFID applications
		7335551	Method to fabricate a thin film non volatile memory device scalable to small sizes
		7364088	Radio frequency identification tag
		7378971	Radio frequency identification tags for digital storage discs
PDC 3	**Center**	**7364089**	Apparatus for electronically determining whether a tax for a product has been paid
		7369050	Curtain for isolating radio propagation zones
		7324061	Double inductor loop tag antenna
		7369051	Method for isolating radio propagation zones
PDC 4	**Center**	**7315283**	Dual-band circularly polarized antenna
		7333061	RFID device and method of making
		7348887	RFIDs embedded into semiconductors
		7374105	RFID tag with improved range

DC 5	**Center**	**7364088**	Radio frequency identification tag
		7324061	Double inductor loop tag antenna
		7408456	Wireless communication system
		7369049	System for isolating radio propagation zones

The most commonly used clustering approaches include k-means, k-medoids, and Self-Organizing Map (SOM). These clustering methods intend to exclusively divide n objects into k clusters exclusively. On the other hand, the proposed non-exhaustive OPC algorithm can group an object into different clusters or in none of the clusters (non-exhaustive). In this research, the methodology allows overlapping of patent documents which coincide with the patent characteristics often claiming multiple key technical inventions and novelties.

In order to evaluator the results of proposed algorithm which combines fuzzy adaptive resonance theory (fuzzy-ART) and non-exhaustive OPC, this research choices original fuzzy adaptive resonance theory and well-known K-mean as the comparison. The same 100 patents from the testing stage are clustered and the number of cluster is trained by competitive learning for proposed algorithm. But the K-mean need to assign the number of cluster by user. The result of original fuzzy-ART has 4 clusters, such as antenna, reader, tag, and application. The patents of every cluster are exclusively partitioned. The object will be assigned to one and only one cluster even when it is an outlier. We compare the results of K-mean, original fuzzy-ART, and proposed algorithm as shown in Table 4.

Table 4. Comparison of non-exhaustive OPC and fuzzy ART algorithm

Methods	**K-mean**	**Fuzzy-ART**	**Proposed**
Model	Partitioning	Competitive	Fuzzy partitioning
Input data type	Numeric	Binary, Analogy	Numeric
Cluster number	Assigned	Trained	Trained
Cluster center	Mean	None	Object
Result	Belong to one cluster	Belong to one cluster	Overlapping & Non-exhaustive
Outlier	Influenced	Influenced	Excluded

5 Conclusions

This paper has introduced a modified non-exhaustive overlapping partitioning clustering algorithm incorporating the fuzzy adaptive resonance theory. The approach is to partition *n* objects into *k* overlapping clusters non-exhaustively for patent analysis. From the experimental case study, the proposed non-exhaustive OPC algorithm can obtain better and more significant clustering result than the original fuzzy-ART method. Because the doamin ontology, clustering threshold and attribute weights can be defined by user, the algorithm can group the patents accordingly and flexibly. In the RFID patent document case, the threshold is 0.9, which means the relationship higher than 90% of all correlation is accepted into the given cluster. The proposed algorithm is influenced by the accuracy of key phrase clusters from fuzzy-ART training. Thus, the method of key phrase extraction applies the domain ontology and NTF-IDF methods as bases to build key phrase bank. The accuracy of extraction is crucial to the clustering results.

In conclusion, the proposed algorithm can maximize the distance among cluster centers and maximize the average number of objects contained in a cluster. Consequently through the case of RFID patent clustering analysis, the proposed algorithm enables R&D teams to effectively focus on their patent development strategies with a good understanding of the patent portfolios of their own and their competitors.

6 References

[1] Campbell, R. S., Patent trends as a technological forecasting tool, World Patent Information, vol. 5, no. 3, pp. 137-143, 1983.

[2] Han, J., Kamber, M., and Tung, A. K. H., Spatial clustering methods in data mining: A survey. Geographic data mining and knowledge discovery, 2001.

[3] Han, J., and Kamber, M., Data Mining: Concepts and Techniques, Morgan Kaufmann, pp. 335-391, 2001.

[4] Grilliches, Z.. Patent statistics as economic indicators: A survey. Journal of Economic Literature, vol. 28, pp. 1661-1707, 1990.

[5] Scherer, F. M., Inter-industry technology flows in the United States. Research Policy, vol. 11, pp. 227-245, 1982.

[6] Paci, R., Sassu, A., & Usai, S., International patenting and national technological specialization, Technovation, vol. 17, no. 1, pp. 25-38, 1997.

[7] Grandstrand, O., The economics and management of intellectual property: Toward intellectual capitalism, Edward Elgar Publishing, 1999.

[8] Anderberg, M., Cluster analysis for applications, New York: Academic Press, 1973.

[9] Kaufman, L., Rousseeuw, P. J., Finding Groups in Data: An Introduction to Cluster Analysis, Wiley-Interscience, 1990.

[10] Hu, H. L., Optimization in Data Mining An Overlapping Cluster Algorithm to Provide Non- exhaustive Clustering, Ph.D., Department of Information Management, National Central University, Taiwan, 2006.

[11] Grossberg, S., The link between brain, learning, attention, and consciousness. Consciousness and Cognition, vol. 8, pp. 1-44, 1999.

[12] Raza, N., Bradshaw, V., & Hague, M., Applications of RFID technology, IEEE Colloquium, No. 123, October, 1999, pp. 1-5.

[13] Carpenter, G.A., Grossberg, S., & Rosen, D.B., Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system, Neural Networks, pp. 759-771, 1991.

Performance Evaluation for an ERP System in Case of System Failures

Shin-Guang Chen[1]

Department of Industrial Engineering and Management, Tungnan University, Taipei County 222, Taiwan.

Abstract. This paper proposes an analytical method to evaluate the performance of an Enterprise Resource Planning (ERP) system in case of system failures. A system failure means the software or hardware between processes been malfunctioning. At first, an ERP system is a complex network (an ERP net) composed of various business processes. The nodes in the net denote the persons responsible for the business tasks during the processes. The arcs between nodes denote the process precedence relationships in the ERP system. When the process starts, the documents are initiated from the source node to its succeeding nodes. Finally, the documents are released in the destination node. Thus, the performance of an ERP system is related to the document flow under the net. By using the fuzzy linguistic results of the ERP examination of the users, a fuzzy linguistic performance index, defuzzified from the probability of maximal flow not less than d, is proposed to evaluate such performance, which can be used to assess the system performance either before or after the system going live.

Keywords. Enterprise resource planning, ERP net, fuzzy mathematics, performance evaluation, minimal path.

1 Introduction

In decades, adopting an ERP system has been a promised way for corporation to gain competition advantages in the world [6]. However, the performance of an ERP system is not easily identified. This is due to the volatility of the internal factors and the external environment. On the other hand, the major obstacle is the lack of an effective quantitative evaluation method been developed. So, many corporations had invested large amount of capital to the ERP systems but have no way to recognize what the performance it is or how to improve it.

Chen and Lin [1, 2] firstly proposed novel analytic methods for evaluating the real time performance of ERP systems. This is not like the conventional empirical study to find the successful key factors such as the work of De- Lone and McLean [4]. These models can be used to analyze/evaluate the performance of a general

[1] Corresponding Author E-mail: bobchen@mail.tnu.edu.tw.

ERP system. They established a theoretical basis for the performance evaluation of ERP systems. According to the approach, researchers may conduct the analysis and investigation for an ERP system. The brief of the approach is that: assuming the performance of an ERP system is closely related with the end-users, i.e., if the end-users are not familiar with the system, the performance of the underlined system will be low, but if the end-users are knowledgeable about the system, the performance of the underlined system will be appeased. Therefore, based on this knowledge, the performance of an ERP system can be analyzed and evaluated after effectively assessing the capability of the end-users about the underlined system.

The proposed approach considers that an ERP system is a complex network composited of various business processes. It can be called an ERP net. The nodes of the net are the persons responsible for the operations of the processes. The arcs are the precedence relationships between processes. The net can be activated by various business requests from source node of a process. Then, the documents are generated and sent to its succeeding nodes for further operations. Thus, the performance of an ERP system is closely related with the document flow in the net. To address the capacity of a node, it is said that a low capacity node can typically be assessed by the lack of right knowledge about the operations of the underlined system. Some researches had revealed the fact that the knowledge of end-users concerning ERP operations is an important successful key factors in the ERP systems [8]. To validate the ERP knowledge captured by the users, an objective examination/ assessment is valuable. Generally, the results of the test are given into grades such as “excellent”, “good”, “normal”, “bad” and “failed” instead of the exact scores. These linguistic grades can be advantageously handled by the fuzzy set theory [7]. This paper extends Chen and Lin’s approach [2] to cover the system failure situation. A system failure in an ERP net means the software or hardware between processes been malfunctioning. This is a usual phenomenon in practice. For example, such failures can be identified by alert lights such as “green”, “yellow” and “red” corresponding to “normal condition”, “minor damaged” and “seriously damaged” respectively. For a realistic example, assume we have a T1 network which consists of 24 channels connects two processes. If only one or two channels failed due to software or hardware reasons, since it still has 92% of bandwidth available for use, this situation can be denoted as ”yellow” alert. Whereas if 11 or 12 channels failed, the situation is considered more serious than the former and can be denoted as ”red” alert due to 50% of bandwidth available for use. Thus, based on the nodes and arcs real time conditions, a fuzzy linguistic performance index $\widetilde{R}_d$ for an ERP system can be created by defuzzifying the probability R_d of the maximal document flow not less than d. The probability calculation is based on minimum paths (MPs) methodology [1, 2]. An MP is a sequence of nodes and arcs from source to sink without cycles. The remainder of the work is described as follows: The assumption for the ERP systems discussed here is presented in Section 2. The net model for an ERP system is discussed in Section 3. An algorithm to evaluate $\widetilde{R}_d$ is proposed subsequently in Section 4. Then, the calculation of the linguistic performance index is illustrated by some numerical examples in Section 5.

2 Assumptions

Let $G = (A, N, M)$ be an ERP net where $A = \{a_i|1 \le i \le n\}$ is the set of arcs, $N = \{a_i|n+1 \le i \le n+s\}$ is the set of nodes, and $M = (m_1, m_2, \ldots, m_{n+s})$ is a vector with m_i (an integer) being the maximal capacity of component (node or arc) a_i. Such a G is assumed to satisfy the following assumptions.

1. The capacity of each component a_i is an integer-valued random variable which takes values from the set $\{0, 1, 2, \ldots, m_i\}$ according to a given distribution governed by μ_i, where μ_i is a membership function mapping from fuzzy set {excellent, good, normal, bad, failed} or {green, yellow, red} corresponding to nodes or arcs respectively to real interval [0, 1].
2. Flow in G must satisfy the flow-conservation law [5].
3. The components are statistically independent from each other.

3 The ERP net model

Suppose $mp_1, mp_2, \ldots, mp_z$ are totally the MPs from source to sink. Thus, the ERP net model can be described in terms of two vectors: the capacity vector $X = (x_1, x_2, \ldots, x_{n+s})$ and the flow vector $F = (f_1, f_2, \ldots, f_z)$ where x_i denotes the current capacity of component a_i and f_j denotes the current flow on mp_j. Then such a vector F is feasible if and only if

$$\sum_{j=1}^{z}\{f_j \mid a_i \in mp_j\} \le m_i \quad \text{for each } i = 1, 2, \ldots, n+s. \tag{1}$$

Equation (1) describes that the total flow through a_i can not exceed the maximal capacity of a_i. We denote such set of F as $U_M \equiv \{F|F \text{ is feasible under } M\}$. Similarly, F is feasible under $X = (x_1, x_2, \ldots, x_{n+s})$ if and only if

$$\sum_{j=1}^{z}\{f_j \mid a_i \in mp_j\} \le x_i \quad \text{for each } i = 1, 2, \ldots, n+s. \tag{2}$$

For clarity, let $U_X = \{F|F \text{ is feasible under } X\}$. The maximal flow under X is defined as $V(X) \equiv \max\{\sum_{j=1}^{z} f_j \mid F \in U_X\}$.

3.1 System performance evaluation

Given the level d (the required document flow), the system performance R_d is the probability that the maximal flow is not less than d, i.e., $R_d \equiv \Pr\{X|V(X) \ge d\}$. To calculate R_d, it is advantageously to find the minimal capacity vector in the set $\{X|V(X) \ge d\}$. A minimal capacity vector X is said to be a lower boundary point for d if and only if (i) $V(X) \ge d$ and (ii) $V(Y) < d$ for any other vector Y such that $Y < X$, in which $Y \le X$ if and only if $y_j \le x_j$ for each $j = 1, 2, \ldots, n+s$ and $Y < X$ if and only if $Y \leq X$ and $y_j < x_j$ for at least one j. Suppose there are totally t lower boundary points for d: $X_1, X_2, \ldots, X_t$, the probability R_d is $\Pr\{\bigcup_{i=1}^{t}\{X \mid X \geq X_i\}\}$. Let μ_R be the membership function mapping from fuzzy set {acceptable, re-certified, declined} to [0, 1]. Then, the linguistic performance index $\widetilde{R}_d$ is defuzzified from $\mu_R^{-1}(R_d)$, where μ_R^{-1} is a reverse function of μ_R.

3.2 Generation of all lower boundary points for *d*

At first, we find the flow vector $F \in U_M$ such that the total flow of F equals d. It is defined as in the following demand constraint.

$$\sum_{j=1}^{z} f_j = d \,. \tag{3}$$

Then, let $\mathbf{F}=\{F \mid F \in U_M$ and satisfies Equation (3)$\}$. We show that a lower boundary point X for d existed then there exists an $F \in \mathbf{F}$ by the following lemma [2] without proof.

Lemma 1. Let X be a lower boundary point for d, then there exists an $F \in \mathbf{F}$ such that

$$x_i = \sum_{j=1}^{z} \{f_j \mid a_i \in mp_j\} \quad \text{for each } i = 1, 2, \ldots, n+s. \tag{4}$$

Given any $F \in \mathbf{F}$, we generate a capacity vector $X_F = (x_1, x_2, \ldots, x_{n+s})$ via Equation (4). Then the set $\Omega = \{X_F \mid F \in \mathbf{F}\}$ is built. Let $\Omega_{\min} = \{X \mid X$ is a minimal vector in $\Omega\}$. Lemma 1 implies that the set Ω includes all lower boundary points for d. The following lemma [2] further proves that $\Omega_{\min}$ is the set of lower boundary points for d.

Lemma 2. $\Omega_{\min}$ is the set of lower boundary points for d.

4 Solution

4.1 Algorithm

Suppose all MPs has been pre-computed. All lower boundary points for d can be generated by Algorithm 1.

Step (1) depicts that according to the MPs, the feasible F under Equation (1) and (3) is enumerated into set $\mathbf{F}$. Then, the candidate vector set Ω for lower boundary points can be derived from $\mathbf{F}$ under Equation (4) in step (2). Finally, the set Ω_{min} of lower boundary points is filtered out by pairwise comparison from step (3).

Algorithm 1: Find all lower boundary points for d.

1. Find the feasible flow vector $F = (f_1, f_2, \ldots, f_z)$ satisfying both capacity and demand constraints.
   a) **enumerate** f_j for $1 \le j \le z$, $0 \le fj \le \min\{m_i | a_i \in mp_j\}$ do
   b) **if** f_j satisfies the following equations
   $\sum_{j=1}^{z}\{f_j \mid a_i \in mp_j\} \le m_i$ and $\sum_{j=1}^{z} f_j = d$ for $1 \le i \le n + s$,
   then $\mathbf{F} = \mathbf{F} \cup \{F\}$.
   end enumerate.
2. Generate the set $\Omega = \{X_F \,|F \in \mathbf{F}\}$.
   a) for F in $\mathbf{F}$ do
   $x_i = \sum_{j=1}^{z}\{f_j \mid a_i \in mp_j\}$ for each $i = 1, 2, \ldots, n + s$.
   b) $U_X = U_X \cup \{X_F\}$. // where $X_F = (x_1, x_2, \ldots, x_{n+s})$ may have duplicates.
   endfor.
   c) for X in U_X do //Remove the redundant vectors.
   d) if $X \notin \Omega$, then $\Omega = \Omega \cup \{X\}$.
   endfor.
3. Find the set $\Omega_{min} = \{X \mid X$ is a minimal vector in $\Omega\}$. Let $J = \{j \mid X_j \notin \Omega_{min}\}$.
   a) for $i \notin J$ and $1 \le i \le ||\Omega||$ do //where $||\Omega||$ denotes the number of elements in Ω.
   b) for $j \notin J$ and $i < j \le ||\Omega||$ do
   c) if $X_j \le X_i$, then $J = J \cup \{i\}$ and goto (3d).
   else if $X_j > X_i$, then $J = J \cup \{j\}$.
   endfor.
   d) $\Omega_{min} = \Omega_{min} \cup \{X_i\}$.
   endfor.

4.2 The generation of $\widetilde{R}_d$

From Ω_{min}, we get the set of lower boundary points. Suppose there are totally t lower boundary points for d, say $X_1, X_2, \ldots, X_t$. The probability R_d for these lower boundary points is $R_d = \Pr\{\bigcup_{i=1}^{t}\{X \mid X \ge X_i\}\}$. Let $B_i = \{X \mid X \ge X_i\}$. Then R_d is rewritten as

$$R_d = \Pr\{\bigcup_{i=1}^{t} B_i\}. \quad (5)$$

To obtain $\Pr\{B_i\}$, let's suppose the result of ERP test or system assessment for the r^{th} component is η_r and the r^{th} column of the minimal vector in B_i is l_{ir}, we get

Pr{ B_i } = $\prod_{r=1}^{n+s}\sum_{j=l_{ir}}^{m_r}\mu_r(\eta_r)(j)$ by definition. Next, we find the respective probabilities of membership function for the corresponding capacities of these components. $R_d = \Pr\{\bigcup_{i=1}^{t} B_i\}$ is then figured by applying the inclusion-exclusion rule. Finally, $\widetilde{R}_d$ is defuzzified from Equation (6), where μ_R^{-1} is a reverse membership function of $\widetilde{R}_d$.

$$\widetilde{R}_d = \mu_R^{-1}(R_d). \quad (6)$$

4.3 Complexity

The number of solutions for **F** is bounded by C_{z-1}^{z+d-1}. The number of X_F generated is then bounded by C_{z-1}^{z+d-1}. The storage space needed for $\Omega_{\min}$ is then bounded by $O(C_{z-1}^{z+d-1})$ in the worst case. In sum, the total storage space needed is $O(C_{z-1}^{z+d-1})$ in the worst case. A pairwise comparison is required for generating $\Omega_{\min}$ between C_{z-1}^{z+d-1} solutions. This takes $O((n+s)^2 C_{z-1}^{z+d-1})$ time to generate $\Omega_{\min}$. In short, the total computational time required is $O((n+s)^2 C_{z-1}^{z+d-1})$ in the worst case.

5 Numerical examples

Figure 1 shows an ERP net with one specific process. When process starts, node a_7 will initiate the document flow and send it via either node a_8 or a_9 to node a_{10} who ends the process. Table 1 gives the results of an examination for the four persons and the system alerts at a time instance during an ERP specific run. Table 2 shows the respective membership functions for the various capacities of all components under different situations. These probabilities can be obtained from statistical observations or through the work studies in the underlined cooperation.

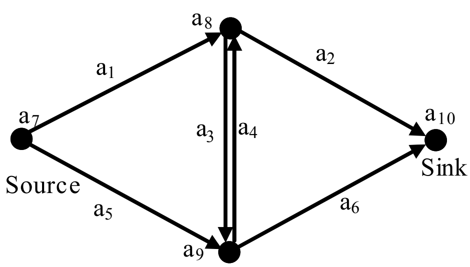


Figure 1. An one-process ERP net.

Table 1. The results for the four person's exam. and the system alerts.

Comp.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
Results	G	Y	G	R	Y	G	GD	F	GD	E
Mem. Func.	μ_1(G)	μ_2(Y)	μ_3(G)	μ_4(R)	μ_5(Y)	μ_6(G)	μ_7(GD)	μ_8(F)	μ_9(GD)	μ_{10}(E)

G:Green, Y:Yellow, R:Red, E:Excellent, GD:Good, F:Failed.

Assume that an acceptable level of the document flow is 5 documents from a_7 to a_{10} (i.e., R_5). All lower boundary points for 5 are generated as in the following calculations. Here, there are 4 MPs existed: $mp_1 = (a_7, a_1, a_8, a_2, a_{10})$, $mp_2 = (a_7, a_1,$ $a_8, a_3, a_9, a_6, a_{10})$, $mp_3 = (a_7, a_5, a_9, a_6, a_{10})$, and $mp_4 = (a_7, a_5, a_9, a_4, a_8, a_2, a_{10})$. After the calculation, the probability R_5 can be conducted in terms of 36 (some points have been filtered out from 56) lower boundary points. At first, let $B_1 = \{X|X \geq X_1\}$, $B_2 = \{X|X \geq X_2\}$, ..., $B_{36} = \{X|X \geq X_{56}\}$. From Equation (5), we get $R_5 = \Pr\{\bigcup_{i=1}^{36} B_i\}$. To calculate this, we find the respective probabilities of membership function for the corresponding capacities of components from Table 2. Then, R_5 is figured by applying the inclusion-exclusion rule as in the following calculation.

Table 2. The respective membership functions for various capacities of all components.

	Probabilities for the different capacities						
Mem. Func.	0	1	2	3	4	5	6
μ_1(G)	0.000	0.005	0.007	0.018	0.080	0.210	0.680
μ_1(Y)	0.001	0.008	0.050	0.300	0.440	0.200	0.001
μ_1(R)	0.680	0.210	0.080	0.018	0.007	0.005	0.000
μ_2(G)	0.000	0.004	0.008	0.018	0.080	0.200	0.690
μ_2(Y)	0.001	0.008	0.050	0.210	0.430	0.300	0.001
μ_2(R)	0.690	0.200	0.080	0.018	0.008	0.004	0.000
μ_3(G)	0.000	0.005	0.007	0.018	0.080	0.190	0.700
μ_3(Y)	0.001	0.008	0.050	0.211	0.410	0.300	0.020
μ_3(R)	0.700	0.190	0.080	0.018	0.007	0.005	0.000
μ_4(G)	0.000	0.005	0.007	0.018	0.080	0.180	0.710
μ_4(Y)	0.001	0.008	0.050	0.230	0.460	0.250	0.001
μ_4(R)	0.710	0.180	0.080	0.018	0.007	0.005	0.000
μ_5(G)	0.000	0.005	0.007	0.018	0.080	0.210	0.680
μ_5(Y)	0.001	0.008	0.050	0.210	0.430	0.300	0.001
μ_5(R)	0.700	0.190	0.080	0.018	0.007	0.005	0.000
μ_6(G)	0.000	0.005	0.007	0.018	0.080	0.180	0.710
μ_6(Y)	0.001	0.008	0.050	0.300	0.440	0.200	0.001
μ_6(R)	0.690	0.200	0.080	0.018	0.008	0.004	0.000
μ_7(E)	0.000	0.005	0.007	0.018	0.080	0.210	0.680
μ_7(GD)	0.002	0.003	0.005	0.010	0.130	0.650	0.200
μ_7(N)	0.001	0.008	0.050	0.300	0.440	0.200	0.001
μ_7(B)	0.005	0.200	0.650	0.120	0.020	0.004	0.001
μ_7(F)	0.680	0.210	0.080	0.018	0.007	0.005	0.000
μ_8(E)	0.000	0.004	0.008	0.018	0.080	0.200	0.690
μ_8(GD)	0.002	0.003	0.005	0.010	0.130	0.660	0.190
μ_8(N)	0.001	0.008	0.050	0.210	0.430	0.300	0.001
μ_8(B)	0.005	0.180	0.670	0.120	0.020	0.004	0.001
μ_8(F)	0.690	0.200	0.080	0.018	0.008	0.004	0.000
μ_9(E)	0.000	0.005	0.007	0.018	0.080	0.190	0.700
μ_9(GD)	0.002	0.003	0.005	0.010	0.140	0.640	0.200
μ_9(N)	0.001	0.008	0.050	0.211	0.410	0.300	0.020
μ_9(B)	0.005	0.220	0.630	0.120	0.020	0.004	0.001
μ_9(F)	0.700	0.190	0.080	0.018	0.007	0.005	0.000
μ_{10}(E)	0.000	0.005	0.007	0.018	0.080	0.180	0.710
μ_{10}(GD)	0.002	0.003	0.005	0.010	0.100	0.680	0.200
μ_{10}(N)	0.001	0.008	0.050	0.230	0.460	0.250	0.001
μ_{10}(B)	0.005	0.160	0.690	0.120	0.020	0.004	0.001
μ_{10}(F)	0.710	0.180	0.080	0.018	0.007	0.005	0.000

G:Green, Y:Yellow, R:Red, E:Excellent, GD:Good, N:Normal, B:Bad, F:Failed.

$$R_5 = \Pr\{\bigcup_{i=1}^{36} B_i\}$$
$$= \sum_{i=1}^{36} \Pr\{B_i\} - \sum_{1\le i<j\le 36} \Pr\{B_i \cap B_j\} + \ldots + (-1)^{35} \Pr\{\bigcap_{i=1}^{36} B_i\}$$
$$= 0.4311$$

In the calculation, R_5 is 0.4311. Figure 2 denotes the membership function μ_R for the linguistic performance index $\widetilde{R}_5$. We get $\widetilde{R}_5 = \mu_R^{-1}$ (0.4311) = "Re-certified". The performance that the ERP system can complete at least 5 documents is dissatisfied under such a situation. A detailed performance analysis for an ERP system can be found in the work of Chen and Lin [3].

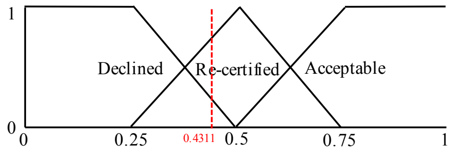


Figure 2. The μ_R membership function for $\widetilde{R}_5$.

6 Conclusion

In practical application, a clearly meaningful performance index is more advantageous than that of the traditional numbers. This paper extents the fuzzy linguistic performance index to the condition of system failures. A system failure in an ERP net means the software or hardware between processes been malfunctioning. The performance of an ERP system is related to the document flow under the net. Then, through the algorithm proposed in this paper, the probability R_d is generated. After defuzzified from R_d, a fuzzy linguistic performance index $\widetilde{R}_d$ is obtained. By the illustration of numerical example, the proposed approach is easily to be fulfilled and can be took place at a regular interval within the system execution to assess the ERP performance under the company's people floating.

7 References

[1] Shin-Guang Chen and Yi-Kuei Lin. An evaluation method for enterprise resource planning systems. Journal of the Operations Research Society of Japan, 51(4):299–309, 2008.

[2] Shin-Guang Chen and Yi-Kuei Lin. On performance evaluation of ERP systems with fuzzy mathematics. Expert Systems with Applications, 36(3):6362 – 6367, 2008.

[3] Shin-Guang Chen and Yi-Kuei Lin. Performance analysis for enterprise resource planning systems. In Proceedings of the 2008 IEEE International Conference on Industrial Engineering and Engineering Management, pages 63 – 67, Singapore, 2008.
[4] William H. DeLone and Ephraim R. McLean. The DeLone and McLean model of information systems success: A ten-year update. Journal of Management Information Systems, 19(4):9 – 30, 2003.
[5] L. R. Ford and D. R. Fulkerson. Flows in networks. NJ: Princeton University Press, 1962.
[6] Mary C. Jones and Randall Young. ERP- usage in practice: an empirical investigation. Information Resources Management Journal, 19(1):23 – 42, 2006.
[7] Bart Kosko. Fuzzy Engineering. Prentice-Hall Inc., 1997.
[8] Chian-Son Yu. Causes influencing the effectiveness of the post-implementation ERP system. Industrial Management and Data Systems, 105(1):115 – 132, 2005.

Multi-Disciplinary Design and Optimization

An Approach Based on Rough Sets Theory to Design Space Exploration of Complex Systems

Xue Zheng Chu[a], Liang Gao[a,1], Mi Xiao[a], Wei Dong Li[b], Hao Bo Qiu[a]

[a] The State Key Laboratory of Digital Manufacturing Equipment and Technology–Huazhong University of Science and Technology, China.
[b] Faculty of Engineering and Computing, Coventry University, U.K.

Abstract: In this paper, to satify the need of engineeing design, and quickly map the performance space to the regions of the design space, a promising approach combined Fuzzy c-means algorithm (FCM) and Rough Sets Theory (RST) in incomplete systems is elaborated. On the base of discretizing continuous variables by FCM, RST in incomplete systems is utilized to extract design scheme rules to assist designers with mapping performance space to regions of desgin space. Finally, a conceptual design of 50000 DWT bulk carrier is taken as a case study to prove the validity of the approach.

Keywords: Design space, performance space, Rough Sets Theory, Fuzzy c-means algorithm.

1 Instructions

In practical engineering field, complex product/system design is a significant issue, because the design of complex product/system such as automobiles, aircraft, and ship, which require designers to think from all aspects and study in multidiscipline. During the two decades, many methodologies and methods have been developed and innovated in the complex system design, for example Multidisciplinary design optimization (MDO) is an effective approach to design and optimize complex products or systems with coupled design functions and variables in a large design space from different disciplines [1]. In substance, these methodologies belong to optimization processes in which mathematical methods are used to search the best result of the object function within a model.

The traditional design optimization processes are usually blind to engineers, and optimization results are only given to engineers directly. In addition, ordinarily, these results are unacceptable, due to the limitation of the practice. Thus, designers prefer to know the whole optimization process in detail. In other words, the optimization designs are needed to be transparent and adjustable, and expert experience is also needed to be extracted to achieve the goal of intelligent design.

[1] Corresponding author. Tel.: +86-27-87557742 ; Fax: +86-27-87543074
E-mail address: gaoliang@mail.hust.edu.cn

Furthermore, in traditional optimization, a sampled point is randomly selected to start the searching process, and then the performance of this solution is evaluated. If the need is not satisfied, another sampled point will be selected, until the performance need is met. With this random sample optimization, engineers will face another problem, how to use the successful constructed products' data and experience? The design knowledge reused maybe hard to be pushed forward under the random sample situation. In general, products' performance are vital for the companies, consequently, the designers urgently expect that a method that can map the performance space (evaluation of the products' performance) to the design space (design variables and their values) is developed, in which the engineers can focus the useful regions rapidly in the whole design space at the beginning of the optimization to reduce the searching scope of the design space.

Although, the computational speed of computers are impressing improved, the expense cost of high precision computation is still too expensive, because more and more computationally intensive analysis and simulation tools, such as finite element analysis (FEA), computational fluid dynamics (CFD) and so on, are used in product design, especially in MDO. Normally, a flight test program of an aircraft may take more than one year [1]. To solve this problem, the approximation model was employed in complex system design and optimization, and has achieved many progresses [2], however, if the whole design space is reduced to some interesting regions, more progresses must be gained by using approximation models.

As discussed above, in complex system design, an approach that can map the performance space to design space directly, and help designers to make design decisions in reducing the design space, is imperative. In this paper, a promising approach integrating Fuzzy c-means algorithm (FCM) and Rough Sets Theory (RST) to in incomplete information systems is proposed, in which RST is adopted to map performance space to design space of the product design, reduce the design space, and save much product development time effectively.

2 Relevant Theories and Detailed Application Steps

To accomplish complex systems design, many experts have their own design experience, and they prefer to use them in the new system design and analysis, which can help engineers effectively avoid some unnecessary mistakes and save much optimization time. Also, companies have already stored huge data about the similar built products in the Data Base Systems, which can provide credible information to engineers for product design. Based on the expert knowledge and stored data, the proposed approach combined FCM and RST that is able to map the performance space to the design space, and help the designers to only focus on the interesting regions in the design space during the optimization process.

2.1 Discretization of Variables by Fuzzy C-Means Algorithm (FCM)

Pawlak [3] proposed a RST mathematical tool to extract rules from exited data, which has been extended and applied in many research fields [4]. The merit of RST is that, it neither needs additional information about the data nor is necessary to correct the missing or incomplete data of variables. However, RST cannot deal with the continuous variables, which is the disadvantage of this theory, so one of the crucial problems is finding a suitable method for converting useful continuous variables into discrete ones in the design space.

Fuzzy c-means clustering algorithm (FCM), first proposed by Bezdek [5], is an effective discretization method. Compared with other algorithms, it does not need lots of the user's experience except for the number of clusters assigned by the engineers at the beginning. Consequently, applying FCM to accomplish variable discretization in the whole design space is the appropriate approach. The concepts and definitions of $c, v_i, q, \varepsilon, u_{ij}, v_i$ and so on, were explained in detail in literature [5]. In this paper, application steps of FCM in the proposed approach in complex system design are summarized as below:

Step 1: Determine the variables that will be discretized in the design space;
Step 2: Confirm the sample set X of the specific variables and the *j*-th *k*-dimensional vector of each variable sampled point x_j (a design scheme);
Step 3: Assigns parameters c, q, ε for the given variables;
Step 4: Initialize the degree of membership matrix u_{ij}^0, which denote the relation between each variable point and the centroid of its cluster at the beginning;
Step 5: Update each centroid of the variable clusters with u_{ij}^0;
Step 6: Calculate $u_{ij}^{(L+1)}$, which denote the relation between each variable point and its centroid in the update process;
Step 7: If $\max\left[\left\|u_{ij}^{(L)} - u_{ij}^{(L+1)}\right\|\right] \le \varepsilon$ stop, all points of the sampled points (design schemes) are separated into c variable clusters. Otherwise go to *step 5*.

With this algorithm, numeric variables can be effectively discretized in the whole design space. In reality, lots of continuous variables have to be discretized to make preparation for utilizing RST.

2.1 Concept of RST in Incomplete Information Systems

Complex system design is a highly dynamic and interdependent process. In the design space, every design scheme (a set of sampled points of each correlative variable) might not be perfect, but has its own advantages. Because some new technology and new equipments are adopted for improving the performance of new systems, it is difficult for engineers to assess their performance. As a result, incomplete data is included in the design space. Accordingly, RST in incomplete information systems (design space) is employed for getting design scheme rules with which performance space can be mapped to design space.

The definitions and explanations of RST have been discussed in detail in [6]. In this paper, a decision table consisting of design schemes is denoted as $(U, A \cup \{d\}, V, f)$, where U is a set of design schemes (a set of sampled points of each correlative variable); A is all concerned design variables, (e.g., volume and cost) of these schemes; $\{d\}$ is a set of decision attributes/variables (the designers' comments /products' performance) of the schemes; V is the value domain of α , in which $\alpha \in A$; f is an information function $f: U \to V_\alpha$.

Kryszkiewicz first make use of RST in incomplete information systems [6], by a new thought, which can successfully deduce optical certain rules and generalized rules from incomplete and imprecise data.

In this paper, in the incomplete decision table of design space, any variable domain V_α may contain special symbol "∗" to indicate that the value of a variable is unknown or missing.

In the incomplete decision table of design space, $SIM(M)$ is defined as:

$$SIM(M) = \{(x,y) \in U \times U \mid \forall a \in M, f_a(x) = f_a(y) \text{ or } f_a(x) = * \text{ or } f_a(y) = *\} \quad (1)$$

The set of indiscernible schemes is $S_M(x) = \{y \in U \mid (x,y) \in SIM(M)\}$ (2)

Define the function as: $\partial_A(x) = \{f_d(y) \mid y \in S_M(x)\}$. $\partial_A(x)$ is the generalized decision in the incomplete decision table of design schemes, due to the effect of the incomplete data.

Let $x \in U$ and $I_A(x) \subseteq I_{\{d\}}(x)$ (i.e. $card(\partial_A(x)) = 1$), $\Delta_U(x)$ is a certain x-discernibility function iff:

$$\Delta_U(x) = \prod_{y \in Y} \sum \alpha(x,y), \text{ where } Y_U = U \setminus I_{\{d\}}(x). \quad (3)$$

Let $x \in U$, $\Delta_g(x)$ is a generalized x-discernibility function in S iff:

$$\Delta_g(x) = \prod_{y \in Y_g} \sum \alpha(x,y), \text{ where } Y_g = U \setminus \{y \in U \mid d(y) \in \partial_{AT}(x)\}. \quad (4)$$

With formula (3), optimal certain rules can be extracted, which denote one region of performance space can be mapped to one region of design space. While, base on formula (4), optimal generalized rules can be reasoned, which denote one or two regions of performance space can be mapped to one region of design space.

3 Case Study

In this section, the conceptual design of a new 50000 DWT Handymax bulk carrier is taken for example to prove the validity and necessity of the proposed approach. General data about the 16 similar cases are listed in Table 1, which build up the whole design space of the new ship. Then, RST is employed to map the performance space to interesting regions of design space.

Firstly, the continuous variables must be discretized. Considering power and speed, three semantic words "High (H), Normal (N) and Low (L)" are ordinarily used to describe them. Hence, the data of Max Power and Speed in Table 1 are redefined with the three words by FCM algorithm. Ship designers assign the number of clusters 3. As a result, these points of each dimension are separated into 3 clusters (intervals) as follows, and the results are shown in Table 2. Each of the two dimensions (Max Power and Speed) of the design space is divided into three intervals, thus, the design space is divided into regions.

Max Power: {7900, 7980, 8160, 8580}; {8930, 9300, 9360, 9480, 9600}; {10225, 10620, 10920}. Speed: {100, 105, 110}; {123, 124, 127}; {135, 148}.

Table 1. General data about the existing ships similar to 50000 DWT Handymax bulk carrier

Ship No.	Length (o.a.) (m)	Length (p.p.) (m)	Breadth (m)	Depth (m)	Draft (t)	Dead Weight (t)	Main Engine	Max Power (kW)	Speed (rpm)	Oil Consumption (t/d)	Service Speed (kn)
4	190.00	182.00	32.26	17.0	11.0	52200	Man B&W 6S50MC	8580	127	30.6	14.5
10	189.90	182.00	32.26	17.1	10.75	52000	Sulzer 6RT A48T	8160	124	27.9	14.4
23	189.99	182.00	32.26	16.67	10.75	50900	Sulzer 6RTA 48TB	8930	127	31.9	14.5
25	189.80	181.00	32.26	16.9	10.75	50000	Man B&W 6L50MC	7980	148	26.8	14.2
38	189.96	181.00	32.20	16.5	10.7	48530	Man B&W 6S50ME-C	9480	127	33.1	14.9
52	186.00	177.00	30.40	16.50	10.95	45800	Man B&W 5S50MC-C	7900	127	26.5	14.2
67	188.00	180.00	32.00	17.7	10.7	48199	Man B&W 6S50MC-C	9480	127	33.1	15.0
68	187.00	180.00	32.26	16.8	11.75	50500	MITSUBISHI 5UEC60LSII	10225	105	36.8	15.0
70	182.00	173.90	32.26	17.1	10.75	52000	Sulzer 7RTA52U	10920	135	38.2	15.2
82	190.00	181.90	32.26	17.0	11.6	54000	MITSUBISHI 6UEC60LS	10620	100	36.0	15.1
87	189.80	182.10	32.26	17.5	10.9	51000	Sulzer 6RT A48T	8160	124	27.8	14.5
98	182.00	174.90	32.26	16.67	11.9	52000	Man B&W 5L60MC	9600	123	34.6	14.5
101	190.00	181.50	31.00	17.7	10.7	46000	Sulzer 6RTA52U	9360	135	32.0	14.9
104	189.99	182.00	32.26	17.90	12.55	54000	MAN B&W 6S50MC-C	9480	127	33.5	14.3
111	188.10	180.00	32.26	16.8	10.75	52000	MITSUBISHI 5UEC60LSII	10225	105	36.9	14.7
118	179.60	172.00	32.20	19.60	12.0	47300	MITSUBISHI 5UEC60LA	9300	110	33.1	14.3

Table 2. The discretized data of max power and speed in Table 1 with FCM

Ship No.	Max Power	Speed	Ship No.	Max Power	Speed
4	L	N	70	H	H
10	L	N	82	H	L
23	N	N	87	L	N
25	L	H	98	N	N
38	N	N	101	N	H
52	L	N	104	N	N
67	N	N	111	H	L
68	H	L	118	N	L

In the engineering design, the customers' comments and experts' evaluations of products' performance must be considered carefully during the process of the new system development. In the proposed approach, based on the FCM, design space has already been divided into the some regions. According to the engineers' proposals, the performance space is mapped to regions of design space by using RST in incomplete systems. In this case, six feasible design schemes (six sets of design variables sampled points) have been confirmed. From designers' view points, five variables (valid utilization degree, maintenance cost, building cost, reliability in voyage and energy consumption cost) are thought important for these schemes. When some new subsystems are wanted to be equipped, because there is no information and experience of using the new systems in the past, except for the manufacture technical parameters that are offered by the provider, some variables of these design schemes are hard to be estimated in practice. So, the incomplete data is inevitably included in decision table. Table 3 is the decision table containing incomplete data of the six design schemes. The symbol "*" is used to denote the incomplete values of the variables.

Table 3. The incomplete decision table of the six design schemes

Scheme	V	M	C	R	E	Grade	∂_A
S_1	H	H	H	H	H	II	{II}
S_2	H	L	H	*	*	I	{I, II}
S_3	N	*	L	H	H	II	{II}
S_4	L	H	H	L	*	III	{III}
S_5	N	L	L	*	H	II	{II}
S_6	H	L	H	*	H	II	{I, II}

This decision table is expressed as: $S=(U,A\bigcup\{d\},V,f)$, where $U=\{S_1,S_2,...S_7\}$ is the set of these design schemes of the new ship; A is the set of conditional variables that consists of V (valid utilization degree), M (maintenance cost), C (building cost), R (reliability in voyage) and E (energy consumption cost), whereas the decision variable $\{d\}$ is Grade, which represents the performance grade of design schemes. The variable domains are: V_V={H (High), N (Normal), L (Low)}; V_M={H (High), L (Low)}; V_C={H (High), L (Low)}; V_R={H (High), L (Low)}; V_E={H (High), L (Low)}; V_{Grade}={I, II, III}, where I, II and III stand for the good, normal and poor performance grade of design schemes, respectively.

$S_A(S_1)=\{S_1\}$; $S_A(S_2)=\{S_2,S_6\}$; $S_A(S_3)=\{S_3,S_5\}$;
$S_A(S_4)=\{S_4\}$; $S_A(S_5)=\{S_3,S_5\}$; $S_A(S_6)=\{S_2,S_6\}$.

The optimal certain rules of Table 3 are concluded with formula (3) as follows:

Rule 1: If maintenance cost is low or valid utilization degree is high or reliability in voyage is high and energy consumption cost is low, then Grade is I;
Rule 2: If valid utilization degree is normal, then Grade is II;
Rule 3: If building cost is low, then Grade is II;
Rule 4: If maintenance cost is high and reliability in voyage is high or valid utilization degree is high, then Grade is II;
Rule 5: If valid utilization degree is low, then Grade is III;
Rule 6: If maintenance cost is high and reliability in voyage is low or energy consumption cost is low, then Grade is III.

As the result of analyzing the design schemes, Rule 1 illuminates that if the new ship design scheme is wanted to have a good performance, the maintenance cost should be low (1.6×10^4-2×10^4 dollars/week) or valid utilization degree should be high (70%-80%) or reliability in voyage should be high (92%-95%) and energy consumption cost should be low (1.82×10^5-2.1×10^5 dollars/day). Hence, the high level performance of design scheme has already been mapped to regions of the design space. These useful regions can help designers to do optimization with reduced scopes and save a lot of development time. On the contrary, Rule 5 and 6 explain which regions in design space (variables and their intervals) would result in the poor grade of design schemes. In practice, they should be avoided because of their unsatisfying performance.

If the regions focused by the optimal certain rules, cannot provided the optimum result sets, the optimal generalized rules with a wider range of restriction are advised to be adopted to help designers. Generalized decision variables are calculated with formula (2). Optimal generalized rules of design schemes in Table 3 are gotten with formula (4) as following:

Rule 1: If valid utilization degree is high, then Grade is I or II;
Rule 2: If maintenance cost is low, then Grade is I or II;
Rule 3: If valid utilization degree is normal, then Grade is II;
Rule 4: If building cost is low, then Grade is II;
Rule 5: If maintenance cost is high and valid utilization is high or reliability in voyage is high, then Grade is II;
Rule 6: If valid utilization degree is low, then Grade is III;
Rule 7: If maintenance cost is high and reliability in voyage is low, then Grade is III.

From the above results, some optimal generalized rules can be found the same as the optimal certain ones. However, some optimal generalized rules have two decision variables, so they can map two performance spaces to the same design spaces. In the light of Rule 1, in order to have a good or normal performance grade of design scheme, ship designers should amend the design scheme to be highly validly utilized (70%-80%). While, with respect to Rule 7, it expresses that if the designers want to save optimization time, the regions with the design schemes being hard maintained (more than 1.4×10^4 dollars/week) and unstable (with 50%-60% stability) in voyage should be avoided, because of its poor performance.

All the above discussions show the regions of performance space are mapped to the regions of the design space effectively by the proposed approach. It can provide another new thought for the designers to develop new complex systems rapidly, according to the past successful cases' performance.

4 Conclusion

In this paper, to overcome the disadvantages of the traditional optimization in the complex system design, a novel approach integrating FCM and RST is elaborated, in which FCM is employed to discretize continuous variables, and RST is adopted to deduce the design scheme rules to map the performance space to regions of the design space. This approach can focus the interesting regions by the performance level. The case study proves the proposed approach can reduce the searching scope in the design space, and speed the efficiency of complex system optimization and design. The future research of this work is how to construct approximation model in these interesting regions.

5 Acknowledgments

This research has been supported by the National Basic Research Program (973 Program) of China under grant No. 2004CB719405, the Natural Science Foundation of China under grant No. 50705032, the Hi-Tech Research, Development Program of China under grant No. 2007AA04Z120 and the Open Research Fund Program of the State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body.

6 References

[1] Weck OD, Agte J, Sobieszczanski-Sobieski J, Arendsen P, Morris A, Spieck M. State-of-the-Art and future trends in multidisciplinary design optimization. 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Honolulu, Hawaii, AIAA 2007-1905.

[2] Gano SE, Renaud JE, Martin JD, Simpson TW. Update strategies for kriging models used in variable fidelity optimization. Struct Multidisc Optim 2006; 32: 287–298.

[3] Pawlak Z. Rough Sets, International Journal of Computer and Information Sciences 1982;11(5): 341-356.

[4] Wong, J.T., Chung, Y.S. Rough set approach for accident chains exploration. Accident Analysis & Prevention 2007;39(3): 629-637.

[5] Bezdek JC. Fuzzy mathematics in pattern classification. Ph.D.dissertation, Cornell Univ. Ithaca, 1973.

[6] Kryszkiewicz M. Rules in incomplete information systems. Information Sciences 1999;113(3-4): 271–292.

The Set-Based Multi-Objective Satisfactory Design for the Initial Design with Uncertainties in Collaborative Engineering

Masato Inoue[a], Yoon-Eui Nahm[b] and Haruo Ishikawa[c,1]

[a]Assistant Professor, Department of Mechanical Engineering and Intelligent Systems, The University of Electro-Communications (UEC), Japan.
[b]Professor, Hanbat National University, Korea.
[c]Professor, The University of Electro-Communications (UEC), Japan.

Abstract. The advancement in computer networks and information technologies has been rapidly changing the manufacturing environment from the environment in industrial age to it in information era. The engineering philosophy has been shifted from concurrent engineering to collaborative engineering among the respective areas of expertise in the design space. Handling the uncertainties in the initial design has a great importance especially for the collaborative engineering. Due to the complexity of today's products and the incorporated overlapping activities between the lifecycle functions, the multi-objective optimized design is necessary. We have proposed a preference set-based design method that enables the flexible and robust design under various sources of uncertainties while capturing designer's preference. This paper discusses the availability of our proposal for obtaining the multi-objective satisfactory solution in the collaborative engineering by applying to an automotive front-side frame and a material design problem of multilayer porous absorber and insulator.

Keywords. Collaborative design, set-based design, initial design, multi-objective design

1 Introduction

The manufacturing industries are now experiencing fierce pressure of competition from every corner on this planet. The present market can be characterized as higher competition, shorter product lifecycle, greater product diversity, fragmented markets, variety and complexity, and smaller batch sizes to satisfy a variety of customer profiles. To date, many researchers and companies have been introducing the concurrent engineering (CE) philosophy, wishing to survive in the present

[1] Corresponding Author : Professor, Department of Mechanical Engineering and Inelligent Systems, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, JAPAN; Tel: +81 (0) 42 443 5422; Fax: +81 (0) 42 484 3327; Email: ishikawa@mce.uec.ac.jp; http://www.ds.mce.uec.ac.jp/index-e.html

market by realizing faster product development, higher quality, lower costs, improved productivity, better custom values, and so on [1]. The CE stresses the simultaneous development of products and their related manufacturing and support processes [2]. Therefore, the early focuses in the CE field were mainly on the parallelism or concurrency of entire lifecycle functions.

In order to incorporate the CE philosophy, the concept of the computer integrated manufacturing (CIM) was introduced, focusing on the integration of a number of enterprise applications, including CAD (Computer Aided Design), CAPP (Computer Aided Process Planning), CAM (Computer Aided manufacturing), PDM(Product Data Management), ERP (Enterprise Resource Management), and the like [3]. However, most research efforts also focus on special individual activities and limited integration in a single enterprise, thus resulting in the “islands of integration”.

On the other hands, the advancement in computer networks and information technologies has been rapidly changing the manufacturing environment from the environment in industrial age to it in information era. In addition, many of today’s products are so complex that no single company has all the necessary knowledge about either the product or the required processes to completely design manufacturing them in-house. Therefore, recently, the engineering philosophy has been shifted from concurrent engineering to collaborative engineering among the respective areas of expertise in the design space. This means that computer-supported collaborative engineering is required.

In the concurrent engineering, the incorporated overlapping activities between the entire lifecycle functions, *e.g.*, upstream design process and downstream manufacturing process. While there are a large number of uncertainties in describing the design, the level of uncertainty is typically high in the early phase of engineering design. Therefore, handling the uncertainties in the early phase has a great importance especially for the collaborative engineering.

Due to the complexity of today’s products and the incorporated overlapping activities between the lifecycle functions, the multi-objective optimized design is necessary. Recently, an emerging practice in concurrent engineering field is “Set-Based Concurrent Engineering” (SBCE) paradigm [4]. The SBCE is a powerful model for efficient design and management of large-scale operations. The broad set of each function corresponds to some kind of uncertainty. The previous series of the authors’ study have proposed a method of set-based design with designer’s preference (Preference Set-Based Design: PSD), that enables the flexible and robust design under various sources of uncertainties [5]. In the present study, the availability of the PSD method for obtaining the multi-objective solution in the collaborative engineering is discussed.

2 Collaborative Engineering

In the distributed and integrated design environment, a multi-functional or cross-functional team from different organizations establishes a network of linked decisions with interdependencies [6]. Then, an excellent solution from one side of design is a poor from another, making it suboptimal for the overall system design

[4]. Various sources of conflicts are avoidable, due to the different point-of-views, built-in-goals, or limited expertise.

In the present study, to overcome the conflict and the lack of the expression of interrelationship of design solutions for the different perspective design, the following ideas are considered to be required for realizing the collaboration.

(1) The leading roles of the collaboration are played by the designers and the support system for their decision making is required.
(2) The system has the functions as,
    a) The solution space shared by the designs of the different functional discipline can be explicitly represented.
    b) Through the design space, the relative location of each solution for the perspectives can be expressed.
    c) The each design space for the perspectives is flexible or robust.
    d) The designers' judgment of value or preference for the each parameter and requirement in the design space can be introduced.

2.1 Collaborative Engineering in Initial Design Process

In general, the design process starts from the conceptual design for realizing the design purpose. At the second stage, the mechanism and/or structure can be started to be considered as the following design process, namely, the initial design process. In the present study, it is proposed that from the initial design the collaborative engineering is promoted. While the early decision-making process has the greatest effect upon the lead time of development and the overall cost, there are a large number of uncertainties in describing the design.

In the early process of design, uncertainties arise from many sources of variations as follows [7, 8],

(1) Epistemic uncertainty: Designer's imprecise discretion related to ignorance, lack of knowledge, and incomplete information
(2) Inherent variation of the physical system: Aleatory uncertainty,
(3) Changes in environmental conditions,
(4) Uncertainty in the decision of other engineers,
(5) Error (numerical uncertainty): recognizable deficiency in any phase or activity of modeling and simulation.

As an example of (1), the part of automotive front-side frame as shown in Figure 1(a) was extracted from the published automotive body structure of 2.0L displacement [9], and its initial design parametric CAD model, shown in Figure 1(b), was created by defining the part sizes representing the form feature of the structure. In this design, eight design variables, shown in Figure 2, are chosen for

(a) Published front-side frame (b) Parametric CAD model

Figure 1. Front-side frame model

realizing five performance requirements, *i.e.*, bending stiffness, tie-down strength, maximum reaction force, average collapse load, and mass.

As an example of (2), a material design problem of absorption and insulator of sound, including porous material, is shown. The number of design parameters such as the parameters and production tolerance of Biot model for absorption and insulator are related to the design of performance. The Biot model is shown in Figure 3(a). If we choose five layers laminate for absorption and insulator, shown in Figure 3(b), various kinds of design parameters have inherent variation of the physical system and material properties. As an example, Table 1 shows the intervals of variation of material properties of a foam layer.

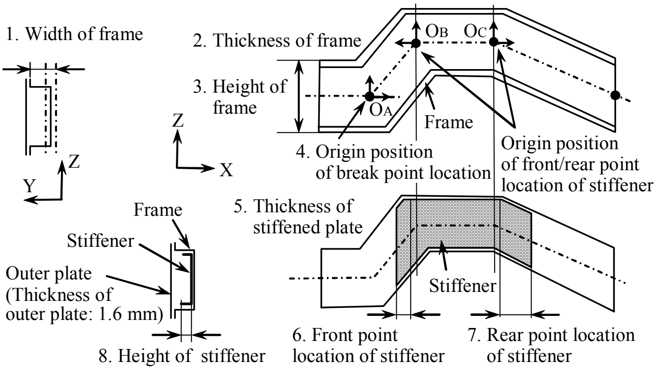
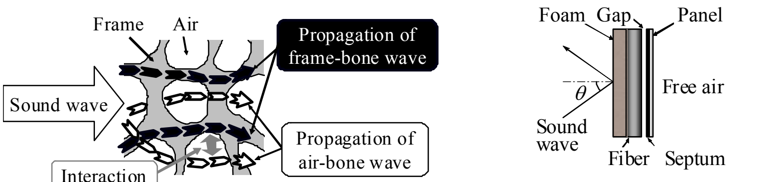


Figure 2. Design variables of front-side frame model



(a) Sound propagation of Biot model

(b) Sound absorbency and insulation material of five layers

Figure 3. Application to a material design problem

Table 1. Interval of variation of material properties of a foam layer

Design variables of foam layer	Interval of variation
Thickness d [mm]	[1.0, 6.0]
Porosity ϕ	[0.9, 1.0]
Density ρ [kg/m^3]	[40, 70]
Flow resistivity σ [$\times 10^5$ Ns/m^4]	[3.0, 8.0]
Viscous characteristic length Λ [$\times 10^{-5}$ m]	[2.0, 7.0]
Young's modulus E [$\times 10^5$ N/m^2]	[1.0, 3.0]
Loss factor η	[0.08, 0.20]
Poisson's ratio ν	[0.2, 0.35]

2.2 Design of Uncertainties

Representing uncertainties is a topic that researchers have approached from many different directions. In the engineering design community, there have been rigorous research efforts that are fuzzy set-based approaches [10], interval set-based approaches [4], probabilistic-based approaches [11], and so on. In the present study, for realizing the requirements, ((2) a, b, and c) in chapter 2, we paid attention to the interval set-based approach as set-based collaborative engineering paradigm. In addition to the set-valued assignments to design variables, the idea of the requirement, ((2) d), is introduced by capturing a preference function on the interval-set. The PSD method proposed by the authors [5] deals with the uncertainties and the preference in the collaborative engineering.

3 Preference Set-Based Design (PSD)

PSD method consists of the set representation, set propagation, set modification, and set narrowing [12].

3.1 Set Representation

To capture the designer’s preference structure on the continuous set, both an interval set and a preference function defined on this set, which is called the “preference number (PN)”, are used. The PN is used to specify the design variables and performance requirements, where any shapes of PN are allowed to model the designer’s preference structure, based on designer's knowledge, experience, or know-how. The interval set at the preference level of 0 is the allowable interval, while the interval set at the preference level of 1 is the target interval that the designers would like to meet.

3.2 Set Propagation and Modification

Set propagation method that combines the decomposed fuzzy arithmetic with the extended interval arithmetic (*i.e.*, Interval Propagation Theorem, IPT [13]) is proposed to calculate the possible performance spaces achievable by the given initial design space. Then, if all the performance variable spaces have the common spaces (*i.e.*, acceptable performance space) between the required performance spaces and the possible performance spaces, there is a feasible subspace within the initial design space.

3.3 Set Narrowing

If the overlapping regions between the possible performance spaces and the required performance spaces exist, there are feasible design subspaces within the initial design space. However, if the possible performance space is not the sub-set of the required performance space, there also exist infeasible subspaces in the initial design space that produce performances outside the performance requirement. Then, the next step is to narrow the initial design space to eliminate inferior or unacceptable design subspaces, thus resulting in feasible design subspaces. To select an optimal design subspace out of those feasible design subspaces, robust design decisions need to be made to make a product's performance insensitive to various sources of variations. The present method has been also used to define the possible design space by capturing the designer's preference structure. In addition to the design robustness, we should take into account which one is preferred by the designer. The design preference and robustness are evaluated to eliminate infeasible design subspaces.

4 Example Problems

4.1 Front Side Frame

Figure 4 and Figure 5 show the ranged set of solutions of design variables and the possibilistic distribution of performances, respectively. Figure 4 indicates that all of the ranged sets of solutions of design variables as shown in solid line are narrowed from the initial preferences of design variables as shown in dotted lines. Figure 5 indicates that all of the possibilistic distributions of performances as shown in solid line are limited within the required performances as shown in dotted lines. These results show that the multi-objective satisfactory design solutions are obtained. In this way, the proposed design method can capture the designers' preference structures and reflect the design intentions of designers in their design solutions.

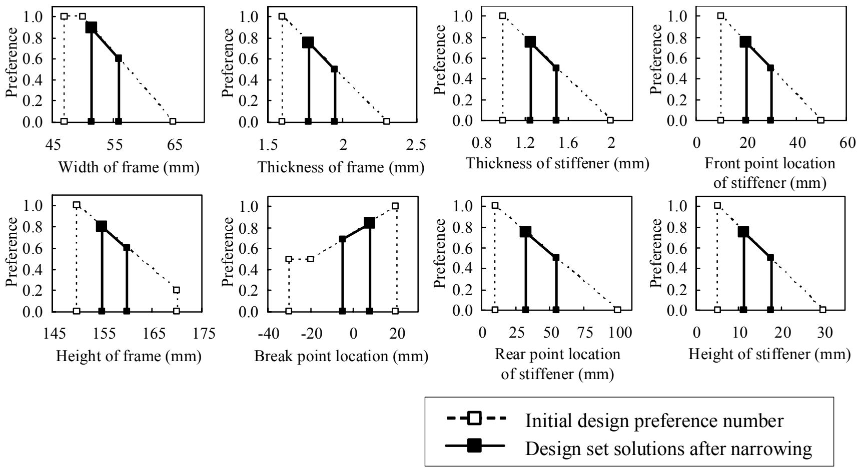


Figure 4. Preference of design variables for front-side frame

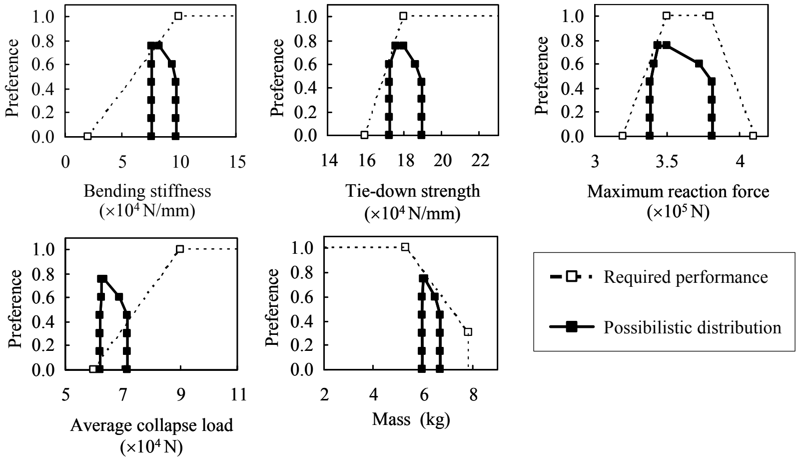


Figure 5. Preference of design performances for front-side frame

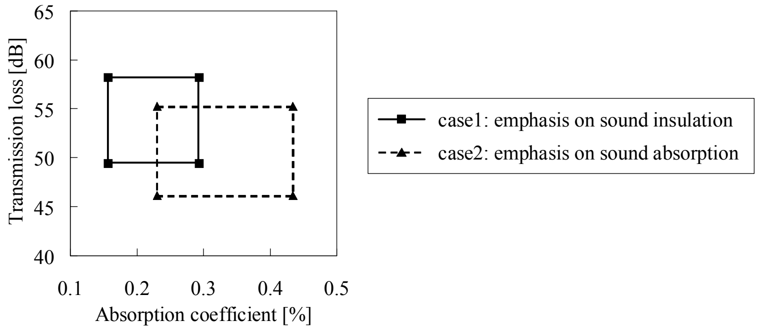


Figure 6. Comparison of possibilistic distributions between case1 and case2

4.2 Absorbency and Insulation Problem of Multilayer Porous Materials

To reflect the degree of emphasis for required performances in the design solutions, different weighting factors of required performances are set. In this paper, the case1 emphasizes the sound insulation, and the case2 emphasizes the sound absorbency. Figure 6 compares, in terms of the relation between the absorbency (absorption coefficient) and the insulation (transmission loss), the design solution of case1 and that of case2. This result indicates that the design solution of case1 is superior to that of case2 in terms of the sound insulation when the insulation performance is more highly weighted. By contrast, case2 is the superior solution in terms of the sound absorption when absorbency performance is more highly weighted. In this way, the proposed design method can obtain the multi-objective satisfactory design solutions including design parameters which have inherent variation of the physical system.

5 Conclusions

This paper proposed the preference set-based design method that enables the flexible and robust design under various sources of uncertainties while capturing designer's preference for collaborative engineering. The proposed method was applied to a real industrial multi-objective design problems (*i.e.*, automotive front-side frame and a material design for multilayer porous absorber and insulator) with uncertainties in the initial design. This presents the availability of our proposal for obtaining the multi-objective satisfactory solution in the collaborative engineering.

6 Acknowledgements

This research was supported by the Japan Society for Promotion of Science, Grant-in-Aid for Young Scientists (B) No. 21760109.

7 References

[1] Prasad B. Concurrent engineering fundamentals, Volume I: integrated product and process organization, and Volume II: integrated product development, Prentice Hall PTR, 1996.
[2] Syan C, Menon U. Concurrent engineering: concepts, implementation and practice, Chapman & Hall London, 1994.
[3] Nahm Y.-E, Ishikawa H. A new 3D-CAD system for set-based parametric design, International Journal of Advanced Manufacturing Technology 2006;29:137–150.
[4] Sobek II D.K, Ward A.C, Liker J.K. Toyota's principles of set-based concurrent engineering, Sloan Management Review 1999;40,2:67–83.
[5] Nahm Y.-E, Ishikawa H. Representing and aggregating engineering quantities for set-based concurrent engineering, Concurrent Engineering 2005;13,2:123–133.
[6] Danesh M.R, Jin Y. An agent-based decision network for concurrent engineering design, Concurrent Engineering 2001;9,1:347.
[7] Kogiso N, *et al.* Robust topology optimization for compliant mechanisms considering uncertainty of applied loads, Journal of Advanced Mechanical Design, Systems, and Manufacturing 2008;2,1:96–107.
[8] Agarwal H, S, *et al.* Uncertainty quantification using evidence theory in multidisciplinary design optimization, Reliability Engineering and System Safety 2004;85,1-3:281–294.
[9] National Crash Analysis Center. Available at: <http://www.ncac.gwu.edu>. Accessed on: Jan. 30th 2009.
[10] Suganuma S, *et al.* A fuzzy set based approach to generalized landscape theory of aggregation, New Generation Computing 2004;23,1:57–66.
[11] Chen W, *et al.* A probabilistic-based design model for achieving flexibility in design, ASME Journal of Mechanical Design 1999;121,1:77–83.
[12] Nahm Y.-E, Ishikawa H. Novel space-based design methodology for preliminary engineering design, International Journal of Advanced Manufacturing Technology 2006;28:1056–1070.
[13] Finch W.W, Ward A.C. Quantified Relations: A class of predicate logic design constraint among sets of manufacturing, Operating and Other Variations, ASME Design Engineering Technical Conference 1996; 18–22.

Minimizing Makespan for Server Testing with Limited Resource

Ping-Yu Chang[a,1], Ya-Ting Hsu[a], and Chin-An Cheng[a]

[a] Department of Industrial Engineering and Management, Mingchi University of Technology

Abstract. In recent years, the global competitive circumstance has become more and more drastic. Different strategies such as reducing cost and increase efficiency have been implemented to strengthen their competitiveness. Hence, efficient scheduling method which is proved to bring vantages to the company is usually adapted. In computer server companies, each server should be tested with different software and option cards to prevent confliction between hardware and software. However, different servers require different combinations of option cards in testing and the time of testing will also be based on the type of option card used in testing. Therefore, it will take long time to test a group of computer servers due to limited workforces and option cards. Hence, a developed scheduling program will reduce testing time of computer servers which will result in cost reduction of the company. This paper conducts research in minimizing makspan for computer server testing with limited resource. A mathematical model is developed based on Job Shop scheduling with parallel machine model to provide optimal schedule in computer server testing. Our results shows significant reduction in testing time compared to a real case and the developed model can be applicable for further research.

Keywords. Job Shop scheduling, Parallel machine scheduling, Server testing, Makespan

1 Introduction

In recent years, the global competitive circumstance has become more and more drastic. Different strategies such as reducing cost and increase efficiency have been implemented to strengthen their competitiveness. Hence, efficient scheduling method which is proved to bring vantages to the company is usually adapted. To achieve the maximum efficiency using scheduling, all the companies try their best to decrease the makespan. Reducing makespan will result in the reduction of waiting time and machine idle time under the limitation of personnel, machine, time, and resource. Therefore, an appropriate scheduling program should minimize the makesapn under the consideration of these factors.

[1] Corresponding Author Email: pchang@mail.mcut.edu.tw.

Although scheduling is guaranteed to bring vantages to company, it is usually implemented within manufacturing process. However, scheduling strategies could also be applicable in other business to bring the same benefit as it provides to manufacturing process. In computer server companies, each server should be tested with different software and option cards to prevent confliction between hardware and software. However, different servers require different combinations of option cards in testing and the time of testing will also based on the types of option card used in testing. Therefore, it will take long time to test a group of computer servers due to limited workforces and option cards. Furthermore, Figure 1 illustrates the operations of computer server testing. In Figure 1, servers are waiting for the available workforce and appropriate option cards. Also, different types of option card require different amount of time for testing where the time unit is usually hour.

Figure 1 demonstrates that a server could wait for the option card and workforce to be available for a long time since the required testing time is long and numbers of option card is limited. Hence, a developed scheduling program will reduce testing time of computer servers which will result in cost reduction of the company. Thus, this research will examine the characteristics of server testing problem and develop a quantitative model to provide optimal schedule of server testing. This research is organized as follows. Section 2 describes the literature reviews for server testing problem, Job Shop scheduling, and parallel machine scheduling. Section 3 demonstrates the quantitative model for optimizing server testing problem while Section 4 discuss the insides of the developed model and future work of this research.

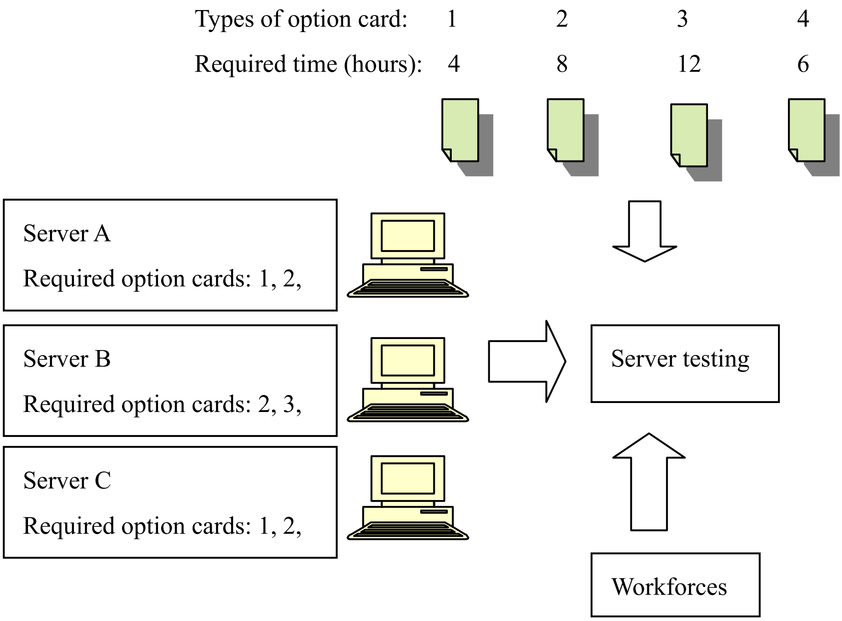


Figure 1. Operations of server

2 Literature review

Scheduling problems have been studied for various manufacturing systems such as single machine, parallel machine, flow shop, and job shop or for different objectives such as minimizing makspan, tardiness, lateness, and number of tardy jobs [8]. However, server testing scheduling problem that can be categorized into Job Shop Scheduling with parallel machine has only drawn a few attentions in academic. Therefore, only literatures within the areas of parallel machine scheduling and Job Shop scheduling will be discussed in this section.

Scheduling for parallel machine has traditionally categorized into assigning jobs to machine and sequencing the jobs that are assigned to the same machine ([3]; [8]). However, some other research focused on scheduling, job assignment, or sequencing with considerations of processing time, job arrival time, batch processing, or setup time ([8]; [7]). So [10] discussed different heuristics in assigning jobs to parallel machine with setups while Daniels et al. [4] developed optimal schedule for parallel manufacturing cells with resource flexibility. Chang et al. [2] used simulated annealing to achieve feasible schedule of batch processing on parallel machine. Their results show that an appropriate feasible schedule can be obtained within significant shorter time using simulated annealing.

The traditional job shop scheduling task consists of a finite set of jobs, each composed of a chain of ordered operations. Each operation must be performed in strict sequence and cannot begin until the previous one has been completed, that is, operations should be performed in a single uninterrupted period of time with a specific resource [11]. There are several objectives can be optimized in a job shop problem such as minimizing makespan, and earliness/tardiness. Minimizing makespan has been investigated for many year and several approaches have been developed including Mixed Integer Programming (MIP) and different heuristics [1]. However, most research focus on minimizing earliness/tardiness in the recent years ([5]; [6]; [9]). This research, however, will develop a mathematical model to minimize makespan based on job shop scheduling with parallel machine. A detailed problem description and model development will be discussed in the following section.

3 Methodology

In job shop scheduling, the main resource is machine in each workstation and jobs are assigned based their own sequence. Thus, since our main resource in this server testing problem is option card, the option card will be seen as the workstation and servers will be assigned based on their required option cards. The parameters of this server testing problem are shown as follows.

Parameters:
i: types of machine, where $i = \{1, 2, ..., I\}$,
m: numbers of machine in each type of machine, where $m = \{1, 2, ..., M\}$,
k : types of software, where $k = \{1, 2, ..., K\}$,

j: types of option card, where $j = \{1, 2, \ldots, J\}$,
r: index to select a constraint from a set of constraints, where $r = \{1, 2, \ldots, R\}$,
n: numbers of option card for each type of option card, where $n = \{1, 2, \ldots, n_j\}$,
P_j : testing time for each type of option card,
Y_{ij} =1, option card *j* is required for machine type *i*,
=0, otherwise.

The decision variable of this problem is demonstrated as follows.

Decision variables:

X_{imkjn} =1, the n^{th} card of option card *j* is used to test m^{th} machine of machine type *i* with software *k*
=0, otherwise

T_{imkjn} : the beginning time of using n^{th} card of option card *j* to test m^{th} machine of machine type *i* with software *k*

Z_r =1, select a constraint in a set of constraints, other constraints will become redundant,
=0, otherwise.

In this server testing problem, we have *i* types of machine and each type of machine has *m* numbers of machine. All the machine should be tested through *k* different types of software with different combinations of *j* different types of option card. However, the combinations of option card will be based on the types of machine and will not be affected by the types of software. Hence, our main limitation in this server testing problem is the numbers of option card since the testing time is also based on the type of option card. Furthermore, the assumption of this problem is discussed as follows.

Assumptions:
1. All the servers arrive testing station at the same time.
2. Since testing time is long and based on the types of option card, numbers of workforce become uncritical in testing procedure. Therefore, numbers of workforce can be assumed to be enough for the operation and can be neglected from the model.
3. The time for testing is constant.

Based on the parameters and assumptions, the mathematical model of this server testing problem can be demonstrated as follows.

minimize $$\{(T_{imkjn} + p_j)x_{imkjn}\}_{\max} \quad (1)$$

subject to

$$\sum_{n=1}^{n_j} X_{imkjn} = Y_{ij}, \quad \text{for all } i, m, k, j \quad (2)$$

$$MZ_1 + M(1 - X_{imkjn}) + T_{imkjn}X_{imkjn} - T_{(imk)'jn}X_{(imk)'jn} \geq P_j X_{(imk)'jn} \quad (3)$$

$$i \neq i' \vee m \neq m' \vee k \neq k' \quad \text{for all } j, n$$

$$M(1 - Z_1) + M\left(1 - X_{(imk)'jn}\right) + T_{(imk)'jn} X_{(imk)'jn} - T_{imkjn} X_{imkjn} \geq P_j X_{imkjn} \tag{4}$$
$$i \neq i' \vee m \neq m' \vee k \neq k' \quad \text{for all } j, n$$

$$MZ_2 + M(1 - X_{imkjn}) + T_{imkjn} X_{imkjn} - T_{imkj'n} X_{imkj'n} \geq P_{j'} X_{imkj'n} \tag{5}$$
$$j \neq j' \quad , \quad \text{for all } i, m, k$$

$$M(1 - Z_2) + M(1 - X_{imkj'n}) + T_{imkj'n} X_{imkj'n} - T_{imkjn} X_{imkjn} \geq P_j X_{imkjn} \tag{6}$$
$$j \neq j' \quad , \quad \text{for all } i, m, k$$

$$MZ_3 + M(1 - X_{imkjn}) + T_{imkjn} X_{imkjn} - T_{imk'jn} X_{imk'jn} \geq P_j X_{imk'jn} \tag{7}$$
$$k \neq k' \quad , \quad \text{for all } i, m$$

$$M(1 - Z_3) + M(1 - X_{imk'jn}) + T_{imk'jn} X_{imk'jn} - T_{imkjn} X_{imkjn} \geq P_j X_{imkjn} \tag{8}$$
$$k \neq k' \quad , \quad \text{for all } i, m$$

$$Y_{ij}, \ X_{imkjn}, \ Z_r \ \in \{0,1\} \tag{9}$$

Constraint set (1) is the objective function which minimizes the makespan of the server testing. Constraint set (2) ensures that each server will only be tested with one option card and one software at one time. Constraint sets (3) and (4) ensure that each option card will be assigned to test a server at one time while constraint sets (5) and (6) implies that only one option card can be used in testing at one time. Constraint sets (7) and (8) ensures that a single server can be tested with one specific software at one time and constraint set (9) is the binary constraint for the decision variables.

4 Conclusion and future research

A mathematical model is developed in this research to demonstrate the scheduling problem of server testing. Through the model verification and validation, the developed model is proved to be useful for providing optimal server testing scheduling with minimum makespan. Although this research has demonstrated the usefulness of the mathematical model, some directions could be conducted for further research. For example, the testing time is set to be constant in this research. However, the testing time could be varied due to the conditions of server, option card, and software. Also, the arrival of servers is assumed to be the same which is not practical for the real case and can be investigated further in the future research.

5 References

[1] Blazewicz J, Domschke W, Pesch E, The job-shop scheduling problem: Conventional and new solution techniques. European Journal of Operational Research, 93(1):1–33, 1996.

[2] Chang PY, Darmadaron P, Melouk S, Minimizing Makespan on Parallel Batch Processing Machines, International Journal of Production Research, 42(19), 4211-4220, 2004.
[3] Cheng TCE, Sin CCS, A state-of-the-art review of parallel-machine scheduling research. European Journal of Operational Research, 47, 271-292,1990.
[4] Daniels RL, Hoopes BJ, Mazzola JB, Scheduling parallel manufacturing cells with resource flexibility, Management Science, 42, 1260-1276, 1996.
[5] Mark S. Fox and Stephen F. Smith. ISIS: A knowledge-based system for factory scheduling. Expert Systems, 1(1):25–49, 1984.
[6] Hoogeveen JA, van de Velde SL, A branch-and-bound algorithm for single-machine earliness-tardiness scheduling with idle time. INFORMS Journal on Computing, 8: 402–412, 1996.
[7] Karabati S, Kouvelis P, Flow-line scheduling problem with controllable processing times, IIE Transactions, 29, 1-15, 1997.
[8] Pinedo M, Scheduling: Theory, Algorithms, and Systems. Prentice-Hall, Englewood Cliff, NJ, USA, 1995.
[9] Sadeh N, Fox MS, Variable and value ordering heuristics for activity-based job-shop scheduling. In Proceedings of the Fourth International Conference on Expert Systems in Production and Operations Management, 1990.
[10] So KC, Some heuristics for scheduling jobs on parallel machines with setups, Management Science, 36, 467-475, 1990.
[11] Tan DS, Gergle D, Mandryk R, Inkpen D, Kellar M, Hawkey D, Czerwinski M, Using job-shop scheduling tasks for evaluating collocated collaboration, Pers Ubiquit Comput, DOI 10.1007/s00779-007-0154-3, 2007.

Exchange of Heterogeneous Feature Data in Concurrent Engineering and Collaborative Design Environments

Zhiyong Huang[1], Fazhi He, Xiaoxia Li, Xiantao Cai and Huajun Liu

School of Computer Science and Technology, Wuhan University, P. R. China

Abstract. Exchange of heterogeneous product data is one of enable techniques in concurrent engineering and collaborative design environments. Beyond the traditional geometry-based data exchange, the feature-based data exchange has a lot of advantages, such as preserving the design intent. Firstly, the existing methods of feature-based data exchange are discussed. Secondly, our latest research on feature-based data exchange is presented. Finally, the comparison of our methods with former typical methods is analyzed with cased study.

Keywords. Concurrent engineering design, Collaborative design, Feature-based data exchange

1 Introduction

Concurrent engineering involves different design teams, who work together and use heterogeneous CAD tools to develop complex and large-scale products. There are many kinds of classification related to the collaboration and integration technique in concurrent engineering design environment.

(1) Online collaboration and offline collaboration. The offline methods exchange the data files, such as STEP files. The online method is complicated and is impacted by the software architecture. The middleware and distributed component is an ideal method. This architecture directly supports the integration of the distributed application in the network [1]. If CAD systems are not based on the middleware and distributed component, there are two methods to construct the online collaboration. The first method is to encapsulate standalone legacy CAD applications into distributed components, which can be transparently called by the middleware [2]. The second method is based on the network programming (such as Win Socket API as plug-in) to implement the collaboration [3, 4, 5, 6].

(2) Online collaboration: transparency collaboration VS. awareness collaboration. As discussed in (1), we can adapt a standalone legacy CAD application into collaborative environment either using middleware encapsulation or socket plug-in. This kind of developing roadmap was named as "transparency

[1] Corresponding Author Email : cacg@whu.edu.cn

collaboration" in CSCW research areas [7]. van den Berg published a good report on this kind of "transparency collaboration" on 3D CAD [8]. On the other hand, we can also develop a collaborative CAD application from scratch or just using a very fundamental geometric computation API (such as ACIS) [9, 10]. The second kind of developing roadmap was named as "awareness collaboration".

(3) Interaction centric collaboration and data centric collaboration [11]. Interaction centric collaboration supports the real-time collaborative design, such as CSCW based CAD system [5, 6]. Data centric integration focus on collaborative data sharing or knowledge integration. Date exchange can be regarded as offline and coarse-grained collaborative design. Knowledge collaboration can be regarded as the extension of intelligent CAD to collaborative one [11].

(4) Low-level collaboration and high-level collaboration. According to the inner data structures and information types of the CAD system, the collaboration can be classified to low-level integration and high-level integration. The high-level integration exchanges the parametric features, functions, high-level commands and APIs among the CAD systems [2, 3, 4, 5, 6, 8, 9, 10]. The low-level integration focuses on exchanging the B_rep model (geometric data) among the CAD systems.

(5) Homogeneous collaboration and heterogeneous collaboration. According to the collaborative environment, the integration can be classified to homogeneous integration and heterogeneous integration. In the homogeneous environment, the systems, commands and data of every CAD nodes are homogeneous in the collaborative design, which means that every cooperative designer should setup the same kind of CAD system. Conversely, the heterogeneous environment has to integrate the different CAD systems for collaborative design [6, 7, 12, 13].

(6) Horizontal integration and vertical integration [11]. According to the phase of the product in its whole life, the collaboration can be classified to horizontal integration and vertical integration. In design phase, horizontal integration means the integration among the CAD systems. The vertical integration means the integration among the CAD systems, CAE systems and CAM systems.

There are three major research motives behind our paper. (1) Collaboration is one of the key techniques to deal with the economic globalization. As the rapid development of the network and information technology, the collaborative work has been expanded from inner enterprise to the outer enterprise. Different enterprises adopt their familiar and heterogeneous CAD systems to design products. For different enterprises work collaboratively, the collaboration and integration among the heterogeneous CAD systems are indispensable. (2) Feature-based data exchange includes design semantics, such as design constraints, design parameter and design history, while geometric-based data exchange includes only the B_rep model which is the evaluation result of parametric feature-based CAD systems. Feature-based data exchange is becoming more and more important in concurrent engineering design environment. (3) Some companies (such as Proficiency Inc and so on) have developed data exchange tools. However, for the reason of commercial secret, the key implementation techniques are not available.

Related work can be found in [14, 15, 16, 17, 18, 19, 20, 21, 22].

2 A New Approach for Feature-based Data Exchange

2.1 The Modeling Process to Create the Parametric Model

As shown in Figure 1, feature based parametric CAD systems have two basic functions. First category user (the creator of the parametric feature model) adapt the interactive modeling of CAD system to create a parametric feature model, the model records the user's designing semantics(including constraint, parameter, feature and so on), and includes a geometric instance according to the feature model. Second category user (the user of the parametric feature model) edits and modifies the parameters of a created parametric feature model, and the CAD systems revalue it and recreate a new feature model by the new parameters.

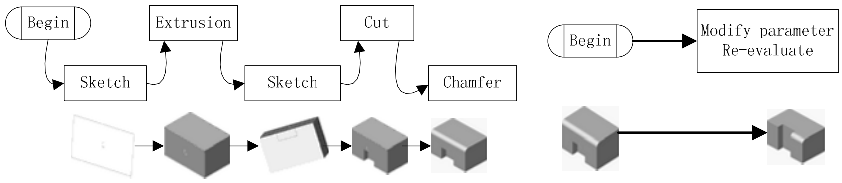


Figure 1. First category user and second category user

It should be noted that the users discussed above are the meta-user. In the real life a designer maybe is a combination of above meta-users.

The aim of feature based data exchange among the heterogeneous CAD systems is that the model created by the first category user in the source CAD system can be edited and modified by the second category user in the target CAD system. In the ideal case, the exchange is transparent, the user don't know the heterogeneity and just like to user the same CAD system. But the problem is that: among the homogeneous CAD systems, the parametric feature model can be exchanged directly; but among the CAD heterogeneous systems, one system can't directly read the parametric feature model of another system. So the feature based data exchange among the heterogeneous CAD systems exchanges the creating process of the model but not the model-itself, that means "simulating the first category user to create the parametric feature model" in the target CAD system. So we need to recover the "creating process of the parametric feature model" in the source CAD system and transfer the process to the target CAD system. Let's take the model in the Figure 1 as an example, the whole creating process can be described as the Figure 2, the whole process includes operations to creating the model and information related to the operating commands including interactive selecting information and parameters. The interactive selecting information means the feature's reference element, basic element and orientation element when the feature is being created. The inner information is produced automatically by the

CAD system when the feature is created, it includes constraint, implicit parameters, featuregeometry, physical attributes and so on.

When the first category user is creating the parametric feature model, he can get two results. The first kind of result is the final “parametric feature model”, as shown in the real-line frame of Figure 2; The second kind of result is a macro file, most of CAD systems supply the “macro recording” function, if user needs, he can use the function to record the macro command sequence of the model creating process, as shown in the broke- line frame of the Figure 2.

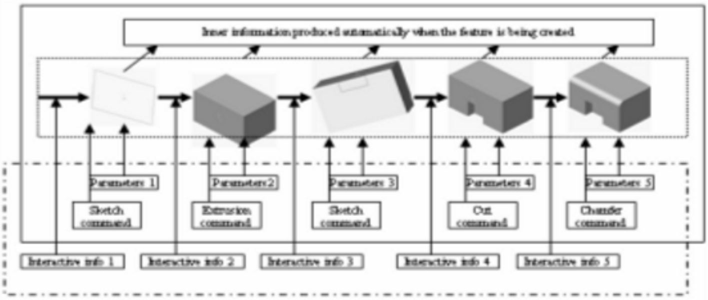


Figure 2. Creating process of parametric feature model

The integration of the above two kind of results is a whole “creating process of the parametric feature model”. Comparing the two kinds of results, the data exchange based on the “parametric feature model” is more meaningful, because most of first category user always only saves the first kind of result and the enterprise need the first kind of result for the information integration. Because the first kind of result is not enough, we must recover the whole creating process of the model in the source system based on the “parametric feature model”, and use the computation, conversion and interaction to get the information which is not matching between the source system and target system.

We call the information recorded in the “parametric feature model”as “first-order information” and call the absent information as “second-order information”.

The first-order information means the information saved in the data base after the model being successfully created, the information includes: the parameters of every features, constraint, model geometry, relation of the elements, data structure and so on.

The second-order information includes the information that the information needed by the target system but not in the data base (the interactive information, the temporary information and the inconsistent information among the heterogeneous CAD systems)

In the UPR, all the information is the first-order information or deduced from the first-order information. So if the information of a feature needed by the target CAD system can’t be deduced form the first-order information, the parametric feature can not be created in the target CAD system and it is replaced by the geometric expression. Comparatively speaking, in our approach, we try to retrieve the second-order information to satisfy the need of other CAD system, that ensure the parametric features can be created more validly in the heterogeneous CAD systems.

2.2 Our approach to recovery the modeling process of feature model

Our approach to recovery the modeling process of feature model is based on the feature model level. It includes two parts: feature retrieval in the source system.and feature model reconstruction in the target system. The key idea in our approach is two steps retrieval strategy in retrieval aspects. Therefore, there are two feature retrieving steps and one model reconstructing step in our approach: First-Order Info Retrieval, Second-Order Info Retrieval and Model Reconstruction, as shown in Figure 3.

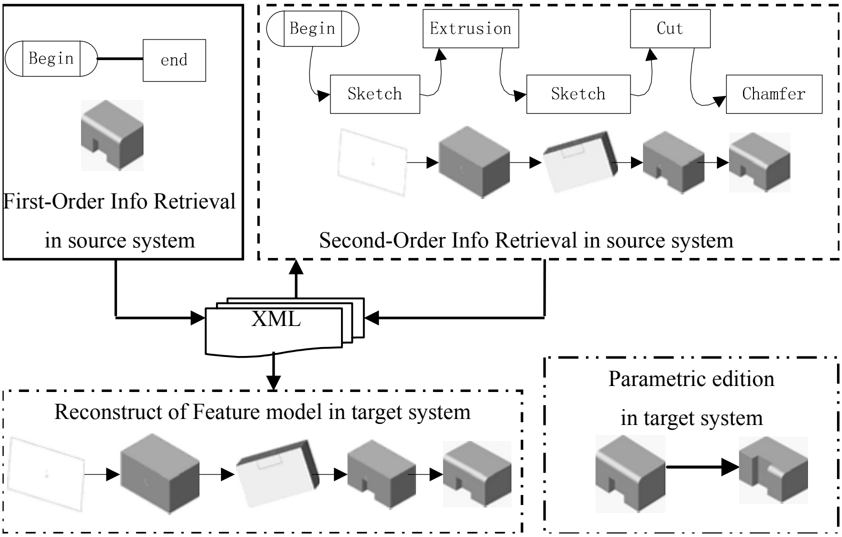


Figure 3. Information retrieval and model reconstruction

First-Order Info Retrieval. Most of parametrical CAD systems provide a feature tree for history based design model. The model information is saved in the data base in accordance with the feature tree. The First-Order Info Retrieval retrieves the most of the model information from the data base of the source CAD system, but there are some kinds of important information are not in the data base.

Second-Order Info Retrieval. Some information, such as the temporary information during the creating process of the model, the lost information and the information is absent in the source system but needed by the target system (the information is inconsistent between the source system and the target system).

Model Reconstruction: The whole design information is recorded in a XML formatted file and transferred to the target system on the network. The Model

Reconstruction read the XML file and acquire necessary parameter list to construct corresponding feature model in the target system.

The details of our method can be found in [12, 13, 23].

2.3 Prototype system and experiments

This approach has been validated by the exchanging of gear model among some popular heterogeneous CAD systems (Solidworks, Pro/E, UG and CATIA) as shown in Figure 4.

The tested part model in our experiment is apparently complex than the published examples in other typical researches, such as UPR method and Macro-command method. The reasons are analyzed in section 3.

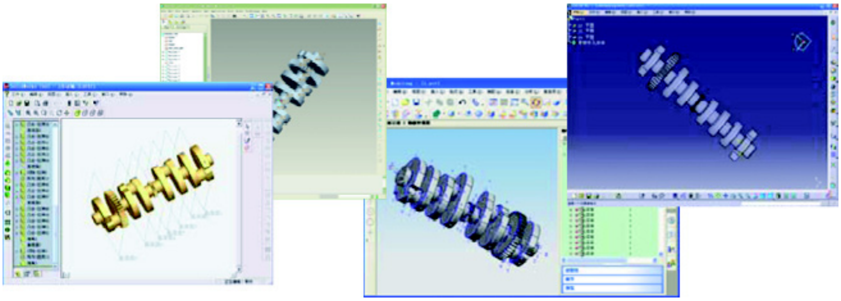


Figure 4. Experiments among Solidworks, Pro/E, UG and CATIA

3 Comparison with other approaches

Compared with the UPR [14, 15, 16, 17], our ultimate solution has following advantages. (1) The UPR just retrieves the first-order information, and all the information in the UPR is transformed from the first-order information. But in our approach, it retrieves both the first order and second order information, and the information transformed from them is more enough. (2) In the UPR, there are two cases for the feature to be replaced by the geometric presentation in the target system: one is because the information is not enough to recreate the feature; the other is because the feature is singular. But in our approach, because of the Second-Order Info Retrieval, the information is enough and the feature is replaced by geometric presentation only when it is a singular feature. (3) In the UPR, if the information is not exists in the UPR, it maybe back to the source system to get it. And in the network environment, that method is inefficient. But in our approach all the information has been transferred to the target system in the XML firstly.

On the other hand, compared with the macro based approach [18, 19], our method has following advantages. (1) The macro based data exchange is based on the neutral file, but our approach is based on the information union. (2) The macro commands supplied by CAD systems are limited and many modeling operations can’t be done by the macro commands; (3) The information coming form the macro files is not enough to reconstruct the parametric feature model (discussed in the paper above). But the approach in this paper is based on the APIs, which are more powerful and the two-step information retrieval ensure the information is enough. (4) The macro based data exchange can’t deal with the singular feature, but in our approach the singular feature can be replaced by its geometry.

4 Conclusions and prospects

This paper presents a feature data exchange approach among the heterogeneous CAD systems based on the process recovery. The key of our approach is to recovery the creating process of a feature model. The future work includes at least two aspects. One is how to improve the accuracy and efficiency. The second one is how to extend our approach to the assembly models.

Acknowledgement

This research has been partially supported by the National Science Foundation of China (NSFC Grant No.60673027), National High Technology Project (863 Project) of China (Grant No. 2007AA04Z149).

References

[1] URBAN SD; DIETRICH SW; SAXENA A. Sundermier A. Interconnection of Distributed Components: An Overview of Current Middleware Solutions. Journal of Computing & Information Science in Engineering. vol. 1, no.1, p. 23-31, 2001.
[2] LIAO B; HE F; CHEN J; MA Y. Using a Middleware Agent to Bridge Standalone CAD Systems for Distributed and Collaborative Applications. In: Shen W, et al, editor. LNCS Volume 3865 / 2006, Springer, 2006; 304-312.
[3] HE F; WANG S; SUN G. Cooperative Design System for Mechanical Drawings. Proceedings of 1999 International Conference on CAD&CG, Shanghai, China, 1999; 303-307
[4] HE F; GAO S. Concurrency control and Consistency Maintenance: A Case Study in Collaborative Engineering Graphics Design System. Chinese Journal of Computer-Aided Design & Computer Graphics. vol. 14, no. 10, p. 928-932, 2002. (in Chinese)
[5] HE F; GAO S. Telepointers and cooperative awareness in semi-transparency and semi-awareness CSCW systems. Chinese Journal of Computer Research and Development. vol. 39, no. 9, p.1009-1105, 2002. (in Chinese)
[6] HE F; HAN S. A Method and Tool for Human-Human Interaction and Instant Collaboration in CSCW Based CAD. Computers in Industry, vol. 57, no. 8/9, p.740-751, 2006
[7] LI M; GAO S; WANG CL. Real-Time Collaborative Design with Heterogeneous CAD Systems Based on Neutral Modeling Commands. Journal of Computing and Information Science in Engineering. vol. 7, no. 2, p. 113-125, 2007

[8] van den Berg E. Collaborative Modeling Systems. Technical Report, Delft University, 1999
[9] JING S, HE F, CAI X, LIU H. Collaborative Naming For Replicated Collaborative Solid Modeling System. Proceedings of 2008 ASME DETC/CIE, New York, USA, 2008; ASME DETC-49233
[10] JING S; HE F; HAN S; CAI X; LIU H. A Method for Topological Entity Correspondence in a Replicated Collaborative CAD System. Computers in Industry. to appearing, article in press, available online 17 March 2009.
[11] FUH JYH; LI WD. Advances in collaborative CAD: the-state-of-the art. Computer-Aided Design. vol. 37, no. 5, p. 571-581, 2005.
[12] CAI X; HE F; LI X; CHEN X. A Direct Feature Retrieval and Reuse Method for Heterogeneous Collaborative CAD Systems. Proceedings of Third International Conference on Pervasive Computing and Applications (ICPCA2008), Alexandria, Egypt, 2008; 718-723
[13] LI X; HE F; CAI X; CHEN Y; LIU H. Using Procedure Recovery Approach to Exchange Feature-based Data among Heterogeneous CAD Systems. Accept to CSCWD 2009.
[14] RAPPOPORT A. An architecuture for universal CAD data exchange. Proceedings of 2003 ACM Symposium on Solid and Physical Modeling, Seattle, WA, USA, 2003; 266 - 269
[15] SPITZ S; RAPPOPORT A. Integrated feature-based and geometric CAD data exchange. Proceedings of 2004 ACM Symposium on Solid and Physical Modeling, Genova, Italy, 2004; 183-190
[16] RAPPOPORT A; SPITZ S; ETZION M. One-dimensional selections for feature-based data exchange. Proceedings of 2005 ACM Symposium on Solid and Physical Modeling, MIT, Massachusetts, USA, 2005; 125-134
[17] RAPPOPORT A; SPITZ S; ETZION M. Two-dimensional selections for feature-based data exchange. Proceedings of 2006 Geometric Modeling and Processing (GMP 2006), Pittsburgh, USA, 2006; 325-342
[18] CHOI G; MUN D; HAN S. Exchange of CAD Part Models Based on the Macro-Parametric Approach. International Journal of CAD/CAM. vol.2, no. 1, p. 13–21, 2002.
[19] MUN D; HAN S; KIM J; Oh Y. A set of standard modeling commands for the history-based parametric approach. Computer-Aided Design. vol. 35, no. 13, p. 1171-1179, 2003.
[20] PRATT MJ; KIM J. Experience in the Exchange of Procedural Shape Models using ISO 10303(STEP). Proceedings of 2006 ACM Symposium on Solid and Physical Modeling, Cardiff University, Wales, UK, 2006; 229-238
[21] Pratt MJ; ANDERSON B; RANGER T. Towards the standardized exchange of parameterized feature-based CAD models. Computer-Aided Design. vol. 37, no. 12, p. 1251-1265, 2005.
[22] KIM J, PRATT MJ, IYER RG, Sriram RD. Standardized data exchange of CAD models with design intent. Computer-Aided Design. vol. 40, no. 7, p. 760-777, 2008.
[23] HE F; CAI X. A Feature Data Exchange Approach among Heterogeneous CAD Systems Based On Procedure Recovery. P.R.China Patent, Application number 200810197046.1, Accepted Date: 2008-09-28

An Ergonomic Assembly Workstation Design Using Axiomatic Design Theory

Xiaoyong Wang[a,b], Dunbing Tang[a,1], Peihuang Lou[a]

[a] College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics. Nanjing 210016, P.R. China
[b] Department of Mechanical Engineering, Nanjing Institute of Industry Technology, Nanjing, 210046, P.R. China

Abstract: To improve the productivity and protect the worker's health, a practical manual assembly workstation is designed using axiomatic design theory with the ergonomics requirements. In the ergonomic workstation design, the function criteria are divided into tolerance type, evaluation type and ergonomic type. To make the information axiom fit to the workstation design, the computation method of the information content is improved. The independence axiom is used to analyze the workstation design process. Through the comparison of the information content of several design solutions, a best alternative is selected. Finally, a simulation has been performed to confirm the ergonomic effectiveness of the selected solution.

Keywords: Axiomatic Design; Assembly Workstation Design; Ergonomics.

1 Introduction

A reasonable industrial workstation design affects not only the productivity but also the health of workers. But in the most of industrial workstation design, the primary concern has usually been the improvement of the performance of the equipment alone. Little consideration is given towards matching the abilities of the operator with the task requirements [1]. Consequently, many industrial workstations are poorly designed, resulting in lower worker productivity and unnecessary injury at the workplace. In recent years the ergonomics has been applied into the workplace and workstation design. Bullinger described the ergonomic advantages of the assembly system, the individual planning steps involved and the actual realization of the system [2]. Lin and Chan evaluated the effect of ergonomic workstation design on musculoskeletal risk factors (MRFs) and musculoskeletal symptoms (MSSs) reduction among female semiconductor fabrication room worker [3]. Using ergonomic simulation software, Santos

[1] Corresponding Author E-mail: d.tang@nuaa.edu.cn

performed ergonomic analysis on some workstations of non-repetitive manufacturing processes [4]. Irad and Joseph presented a new structural methodology based on virtual manufacturing for the ergonomic workstation design which could reduce the number of examined design solutions and obtained estimation for the best design configuration with respect to multi-objective requirements [5]. The investigations of Landau and Petters revealed poor ergonomic design, unsatisfactory logistics and quality problems at assembly and inspection workplaces in a large corporation supplying fenders to the automotive industry and then they designed ergonomically optimized assembly and inspection workplaces compatible with production flow [6].

In above literatures, ergonomics and computer simulation technology have been widely applied into workstation design. But the design approach of workstation in the literatures mostly is based on trial-and-error approach. The current trial and error approach to ergonomic workstation design is inefficient and often leads to cost overrun and missed schedules. Workstation design is a system design process. Although ergonomics is a design-oriented discipline, ergonomists are just responsible for human associated design function and do not design the whole systems. But the workstation, as a system, must satisfy multiple functional requirements including ergonomics, production process, time and cost. Otherwise, the iterative design will be not avoided, owing to some functions unsatisfied. Therefore, a new design methodology is needed to fuse ergonomics into workstation system design. Suh's axiomatic design theory overcomes the shortcomings of conventional workstation ergonomics design and has been applied in the ergonomic design in recent years. For example, Shin presented a design process sequence that uses axiomatic design in order to satisfy all the regulations and designed an automobile seat that could satisfy all the regulations [7]. But this ergonomic research used the first axiom (the Independence Axiom) of Axiomatic Design, and the second axiom (the Information Axiom) is neglected. Helander and Li presented a new conceptualization of the Information Axiom in ergonomics by considering selection of a table based on the information content of specific table design parameters [8]. But after deep investigation, it is found that direct application of the Information Axiom is not appropriate for ergonomic design.

In this paper, we present a design process of a manual assembly workstation using the axiomatic design theory, and both the Independence Axiom and the Information Axiom are used for the assembly workstation design with the consideration of ergonomics. In our research, it is noted that the information content computation method in the information axiom is improved to be suitable for ergonomic design.

2 Principles of axiomatic design

Design is interplay between what a designer wants to achieve and how the designer achieves it. It is a creative process but must become a principle based process. In order to establish a scientific foundation of design, Suh [9, 10] has created and popularized axiomatic design theory. Axiomatic design has two design axioms: the independence axiom and the information axiom. They are formally stated as:

Axiom 1: The Independence Axiom: Maintain the independence of the FRs (functional requirements).

Axiom 2: The Information Axiom: Minimize the information content of the design.

Information is defined in terms of the information content, I_i, which is related in its simplest form to the probability of satisfying the given FRs. I_i determines that the design with the highest probability of success is the best design. Information content I_i for a given FR_i is defined as Equation 4.

$$I_i = \log_2 \frac{1}{P_i} = -\log_2 P_i \qquad i = 1,2,\cdots,n \tag{4}$$

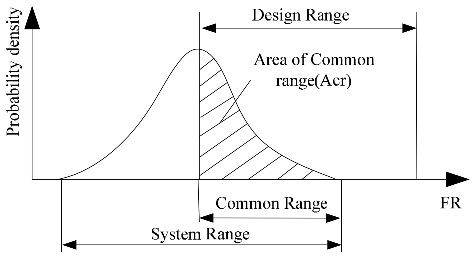


Figure 1. Definition of design range, system range and common range for the calculation of information content.

In any design situation, the probability of success is given by what designer wishes to achieve in terms of tolerance (i.e. design range) and what the system is capable of delivering (i.e. system range). As shown in Figure 1, the overlap region between the design range and system range is called common range (cr) and it is the only region where the FR is satisfied. The success probability, P_i, which shows the design's probability achieving the specified FR, can be defined as Equation 5.

$$P_i = \frac{A_{cr_i}}{A_{sr_i}} \qquad i = 1,2,\ldots,n \tag{5}$$

Common area, A_{cri}, is the area under common range, and system area, A_{sri}, is the area under system range. The probability P_i of achieving FR_i in the design range can be expressed as (if FR_i is a continuous random variable)

$$P_i = \int_{dr^l}^{dr^u} p(FR_i)dFR_i \qquad i = 1,2,\ldots,n \tag{6}$$

Where $p(FR_i)$ is the system pdf (probability density function) for FR_i. Equation 6 gives the probability of success by integrating the system pdf over the entire design range (i.e., from the lower bound of design range, dr^l, to the upper bound of the design range, dr^u). In Figure 1, the area of the common range (A_{cr}) is equal to the probability of success P_i. Therefore, the information content is as follows,

$$I_i = \log_2\left(\frac{1}{A_{cr}}\right) \tag{7}$$

In case of having FRs, information content of overall system, I_{sys}, is defined in Equation 8.

$$I_{sys} = \sum_{i=1}^{n} I_i \qquad i = 1,2,\cdots,n \tag{8}$$

3 Improvement of the information axiom in workstation design

The information axiom has been applied to the evaluation of product design schemes. In the inforamtion axiom, system range (sr) is the capability of the current system, given in terms of tolerances; common range (cr) refers to the amount of overlap between the design range (dr) and the system capability, and design range is the range specified by the designer. If the system range falls into the design range completely, the designer requirement is well satisfied. The designer doesn't mind the degree that the system range satisfies the design range. In the workstation design, the function criteria include not only the tolerance type, but also the evaluation type and the ergonomic type. For the evaluation type criteria, the decision maker cares about the degree that the attribute value satisfies the object. For example, for the benefit criteria, decision-maker not only requires the attribute value falling into the object range, corresponding to design range in the product design, but also expects that the attribute value is larger as possible. The information axiom must take into account the criteria types and the information content should reflect the system ability and decision maker satisfaction degree together. For the ergonomic type criteria, Helander and Li [8] provided an improvement of the Information Axiom in ergonomics by considering selection of a table based on the information content of the adjustable table height ranges. In the context of ergonomics design, the probability (p) of achieving FR, i.e. probability of satisfying human users with regard to a particular FR, can be calculated using a criterion of accommodating the desired range of a specific design variable. In such a case, the information content for the design can be defined as follows:

I=log2 (Desired Range/Common Range) (9)

But in the Equation 9, the tolerance type and evaluation type criteria are not considered. To integrate all type criteria, a new computation formula of the information content is presented, as shown in Equation 10.

$$I_i = -(\alpha_i \log_2 P_i + \beta_i \log_2 S_i), i = 1,2,\cdots,n \tag{10}$$

Where S_i is the satisfaction degree of decision maker on the attribute i. α_i is the weight of P_i and β_i is the weight of S_i. They are subjected to $\alpha_i \geq 0$, $\beta_i \geq 0$, and $\alpha_i + \beta_i = 1$.

For the tolerance type criteria, $\alpha_i=1$, $\beta_i=0$. For the ergonomic type criteria, $\alpha_i=0$, $\beta_i=1$ and $S_i=\log_2$(Desired Range/Common Range). For the evaluation type, the criteria can be classified into the benefit type and the cost type. The system probability density function for FR_i, $p(FR_i)$ and the satisfaction function, $s(FR_i)$ are shown in Figure 2.

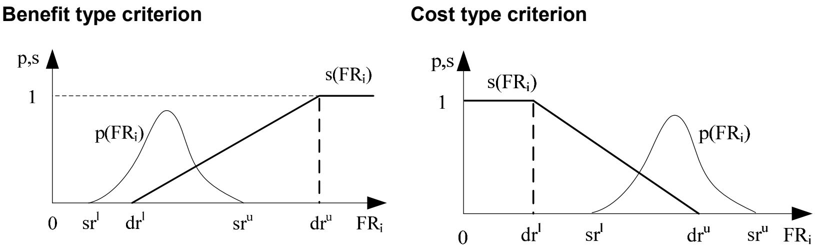


Figure 2. The satisfaction function of evaluation type criteria

S_i is defined as Equation 11. P_i is also defined as Equation 6.

$$S_i = \int_{dr^l}^{dr^u} p(FR_i)\, s(FR_i)\, dFR_i \qquad i = 1,2,\ldots,n \tag{11}$$

Where dr^l and dr^u denote the lower bound and the upper bound of design rang, sr^l and sr^u denote the lower bound and the upper bound of system range.

4 A manual assembly workstation design based on AD principles

4.1 Description of the tasks

There is a Chinese enterprise (named Jiuzhu company) focusing on the production of electrical controlled folding barrier. Because the present assembly workshop is very inefficient and the working condition is very poor which is complained by the workers, this company wants to design a barrier assembly line. Our project team visited the enterprise four times to take notes and measurements of the production line and the workplace. We also recorded the video of the operations of whole assembly process. The recorded data assisted us to build and design a new assembly line. Through the data analysis, we found that the cross-bar assembly workcell was the bottleneck of the assembly line. Hereby we use the cross-bar assembly worksation as a case of our research. The process of assembly operations is as follows: At first, six plastic bushings are installed in the holes of the two long poles respectively, and then a casing and a bar are assembled between the two poles. In this worksation, there is no assembly work-table. The workers sit on a small stool. In such kind of work condition, the worker is easily tired. The assembly efficiency is slow and the health of workers is easy to be harmed. To improve the productivity and the work situation, we designed the cross-bar assembly workstation using axiomatic design theory with the ergonomics requirements.

4.2 The workstation design based on independence axiom.

The assembly workstation design based on axiomatic design can be divided into three steps. Step 1: choose FRs in the functional domain; Step 2: map FRs in the physical domain; Step 3: decompose FRs in the functional domain—zigzagging between FR and DP domains. The repetition of the three steps can detail the design process until the bottom DPs are implemented. The decomposition of assembly workstation design is shown in Figure 3.

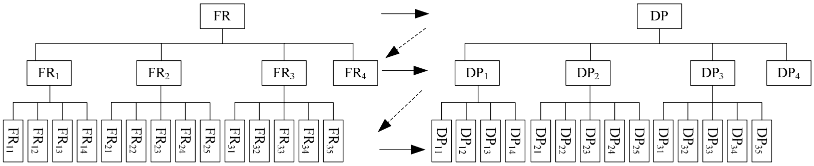


Figure 3. The decomposition of assembly workstation design.

At the first level, according to the requirement of the enterprise, the highest FR can be described as below.

FR：Improve the efficiency of the cross-bar assembly and the worker can operate comfortably, which avoids potential harm to workers after the long term working.

The best solution to this requirement is to design an assembly workstation which is convenient for worker operation and meets the ergonomics demand. The DP may be stated as below:

DP：Use an ergonomic assembly workstation.

Above DP cannot be implemented without further clarification. AD principles recommend returning to the functional domain for decomposing the FR into lower FRs set. The assembly workstation must achieve the assembly process demand. Because there is not heavy manual labor in the cross-bar assembly process, the ergonomics in the station design just contains anthropometric design to make workers of different height keeping a good work posture. The least investment and the most benefit are the target of the enterprises. So we shall reduce costs of the equipments and improve efficiency in the workstation design. The decomposition of FR and DP is described in Table 1.

Because DP1, DP2 and DP3 can't be implemented, FR1, FR2 and FR3 shall be decomposed further. The decomposition of FR1 and DP1 is described in Table 2, the decomposition of FR2 and DP2 in Table 3, and the decomposition of FR3 and DP3 in Table 4.

Table 1. The decomposition of FR and DP.

FR	DP
FR1: Equipment meets the assembly process demand.	DP1：Select equipments according to operating process demand.
FR2: Provide a good work posture for operators.	DP2: Use a height adjustable workstation.

FR3: Investment as little as possible.	DP3: Reduce the cost of equipments.
FR4: The efficiency is improved 20%~30%. (The origin Takt time is 48[s].)	DP4: Takt time as little as possible.

Table 2. The decomposition of FR1 and DP1.

FR1	DP1
FR11: The equipment to contain long poles.	DP11: A material shelf.
FR12: The equipment to contain small parts.	DP12: Plastic boxes.
FR13: The equipments by which assembly operations are implemented.	DP13: Workbench and chair.
FR14: The equipment to contain assembled products.	DP14: An assembled product shelf.

Table 3. The decomposition of FR2 and DP2.

FR2	DP2
FR21：Long poles can be gotten by workers with different arm length.	DP21: Removable long pole shelf. (trolley)
FR22：Assembled product can be put on the shelf by workers with different arm length.	DP22: Removable assembled product shelf. (trolley)
FR23: The worker's feet can be supported on the floor.	DP23: Adjustable chair height.
FR24: Workbench at elbow height.	DP24: Adjustable table height.
FR25: Small parts can be gotten by workers with different arm length.	DP25: Adjustable shelf height.

Table 4. The decomposition of FR3 and DP3.

FR3	DP3
FR31: The intended cost of trolley for containing long poles.	DP31: The realistic cost of trolley for containing long poles.
FR32: The intended cost of trolley for containing assembled product.	DP32: The realistic cost of trolley for containing assembled product.
FR33: The intended cost of chair.	DP33: The realistic cost of chair.
FR34: The intended cost of table.	DP34: The realistic cost of table.
FR35: The intended cost of plastic boxes.	DP35: The realistic cost of plastic boxes.

So far, the initial design of the assembly workstation is completed. A complete design matrix for workstation is shown in Figure 4. The design matrix is a lower triangular matrix and the workstation design is a decoupled design.

FR		DP1				DP2					DP3					DP4
		DP11	DP12	DP13	DP14	DP21	DP22	DP23	DP24	DP25	DP31	DP32	DP33	DP34	DP35	
FR1	FR11	X														
	FR12		X													
	FR13			X												
	FR14				X											
FR2	FR21	X				X										
	FR22				X		X									
	FR23			X				X								
	FR24			X				X	X							
	FR25		X					X	X	X						
FR3	FR31	X				X					X					
	FR32				X		X					X				
	FR33			X				X					X			
	FR34			X					X					X		
	FR35		X												X	
FR4		X	X	X	X	X	X	X	X	X						X

Figure 4. A complete design matrix for the assembly workstation design.

4.3 The computation of the information content based on information axiom.

According to above analysis based on independence axiom, we get three alternatives (A1, A2 and A3). The differences of them are the adjustable range and cost of the equipments belonging to different venders. The values of FRs and DPs are shown in Table 5 and Table 6. The information content is calculated by Equations 6, 10 and 11. Because the FR1, FR21 and FR22 have been fulfilled when equipment type are chosen, their information content has been equal to zero. The data about them is not included in below tables. The unit of length is millimeter, the unit of cost is Yuan and the unit of time is second.

Table 5. The values of FRs

FR23 [mm]	FR24 [mm]	FR25 [mm]	FR31 [Yuan]	FR32 [Yuan]	FR33 [Yuan]	FR34 [Yuan]	FR35 [Yuan]	FR4 [s]
40-60	70-90	110-140	1000-1300	900-1000	300-500	1800-2200	800-1000	37-40

Table 6. The values of DPs

Solution \ DP	DP23 [mm]	DP24 [mm]	DP25 [mm]	DP31 [Yuan]	DP32 [Yuan]	DP33 [Yuan]	DP34 [Yuan]	DP35 [Yuan]	DP4 [s]
A1	30-60	65-85	100-140	1000	850	400	2000	800	38-40
A2	40-55	80-100	110-130	1100	920	350	1850	850	38-40
A3	30-55	70-100	100-140	1260	950	380	1900	880	38-40

Table 7. Information content

Solution \ I	I_{23}	I_{24}	I_{25}	I_{31}	I_{32}	I_{33}	I_{34}	I_{35}	I_4	I_{sys}
A1	0	0.42	0	0	0	1	1	0	0.29	2.71

A2	0.42	1	1	0.58	0.32	0.42	0.19	0.42	0.29	4.64
A3	0.42	0	0	2.9	1	0.74	0.42	0.74	0.29	6.51

The system information of A1 is the least in the three alternatives. Therefore A1 is the best solution.

5 Ergonomics simulation

We designed the virtual assembly workstation using the CAD module of eM-Workplace. We used the standard eM-workplace mannequin as it is adaptable to real worker morphology. We analyzed the ergonomics characteristic of the designed assembly workstation using the Ovako Working Posture Analyzing System (OWAS) in the virtual environment, as shown in Figure 5. The “OWAS” gives information about the physical stress recorded in correspondence of each working posture of shoulders, arms and legs and in relation to the weights handled during the operations. The 5th, 50th and 95th percentiles mannequin were used to verify the height adjustability ranges for chair, table and material shelf and the positions of trolleys of the selected alternative solution. According to design matrix for the assembly workstation design, the adjusting sequence should follow the order of chair, table and material shelf. Otherwise, the iterative adjustment will occur. The OWAS analysis results show that all of the working postures of 5th–95th percentile operators belongs to category 1 (The body posture has no adverse effects on the muscular system of the worker; the stress level is optimum, no corrective interventions are required.). The OWAS has not revealed any particular posture problem, and the ergonomics effectiveness of the selected alternative solution has been confirmed.

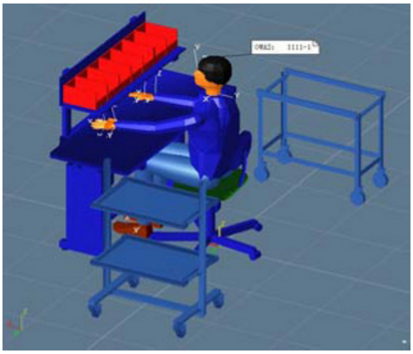


Figure 5. Simulation model of the final designed assembly workstation

6 Conclusion

The assembly workstation design is a complex system design process which involves multiple disciplines such as mechanical design, industry engineering, decision making and ergonomics. Any one discipline can not be competent for the work. In this paper, axiomatic design theory is applied to the ergonomic workstation design. To fuses multiple disciplines in the workstation design, the imformation axiomatic is improved. Function criteria are divided into tolerance type, evaluation type and ergonomic type and the computation method of information content are improved to adapt the different types of criteria. The independence axiom is applied to decompose the assambly workstion design process. According to the imformation content of the alternatives, a best solution is selected. Finally, the eM-workplace software is applied to perform ergonomic analysis, the result confirms the ergonomic validity of the selected solution.

7 Acknowledgement

This work is supported by National Natural Science Foundation of China (NSFC) under Grant No. 50505017 and No.50775111, the Open Foundation of Jiangsu Province Precise Manufacturing Engineering R&D Centre, China (No. ZK06-01-05 and ZK08-01-02), R&D Foundation of Nanjing Institute of Industry Technology, China (NO.YK08-01-04) and Qing Lan Project, China.

8 References

[1] BIMAN D.; ARIJIT K. S. Industrial workstation design: A systematic ergonomics approach. Applied Ergonomics. vol 27, no. 3. p. 157-163, 1996.

[2] BULLINGER H. J.; RALLY P. J.; SCHIchipfer J. Some aspects of ergonomics in assembly planning. International Journal of Industrial Ergonomics, vol 20, no. 5, p. 389-397, 1997.

[3] LIN R. T.; Chan C. C. Effectiveness of workstation design on reducing musculoskeletal risk factors and symptoms among semiconductor fabrication room workers. International Journal of Industrial Ergonomics, vol. 37,no. 1, p. 35-42, 2007.

[4] SANTOS J.; SARRIEGI J. M.; SERRANO N.,; TORRES J. M. Using ergonomic software in non-repetitive manufacturing processes: A case study. International Journal of Industrial Ergonomics, vol. 37, no. 3, p.267-275, 2007..

[5] IRAD B. G.; BUKCHIN J. The ergonomic design of workstations using virtual manufacturing and response surface methodology. IIE Transactions, vol. 34, no. 4, p375-391, 2002.

[6] LANDAU K.; PETERS H. Ergonomic demands in automotive component inspection tasks. Occupational Ergonomics,vol. 6 , no. 2, p. 95-105, 2006.

[7] SHIN M.K.; KIM Y.I.; KANG B.S.; PARK G.J. Design of an automobile seat with regulations using axiomatic design. Proceedings of the Institution of Mechanical Engineers -- Part D -- Journal of Automobile Engineering, vol. 220, no. 3, p269-279, 2006.

[8] HELANDER M. G. and LI L. Axiomatc design in ergonomics and an extension of the information axiom Journal of engneering design,vol. 13, no. 4, p.321–339, 2002.

[9] SUH N. P. The Principles of Design. New York: Oxford University Press, 1990.

[10] SUH N. P. Axiomatic Design, Advances and Applications. New York: Oxford University Press, 2001.

Heterogeneous Material-Based Biomedical Product Development

W.D. Li[a,1], L. Gao[b], D.B. Tang[c] and K. Popplewell[a]

[a] Coventry University, UK.
[b] Huazhong University of Science and Technology, China.
[3] Nanjing University of Aeronautics and Aerospace, China

Abstract: With the rapid advance in material engineering, product design methodologies are experiencing dramatic paradigm shifts. The new trend is that designers cooperate with material scientists closely to tailor materials to achieve desirable properties, and the developed products can therefore exhibit superior performance than those using other available materials. A major characteristic of biomedical products is heterogeneous in terms of material composition and gradation, and heterogeneous materials (HE) which can mimic such the function have been increasingly used in biomedical product developed. In this paper, design methodologies and a software prototype system have been developed in the following three aspects: (i) A HE material design methodology for complex biomedical products; (ii) Designer-centric features to represent and visualize HE material-based biomedical product models; and (iii) Layer-based virtual manufacturing simulation to verify the fabrication feasibility and efficiency of HE material-based biomedical products. A case study is presented in the paper to demonstrate the effectiveness of the developed methodologies and system.

Keywords: Heterogeneous material, biomedical product, virtual manufacturing.

1 Introduction

The rapid advances in material engineering, information technology and manufacturing have brought a number of opportunities for developing innovative biomedical products. Accordingly, design methodologies and tools have experienced dramatic paradigm shifts. The traditional product design process is to select suitable materials from databases according to bio-functional, mechanical, thermal, processing and other criteria to meet product specifications and regulations. The new trend is that, based on digital design and manufacturing simulation toolkits, designers cooperate with material scientists closely to design new materials to generate products with superior performance, more sophisticated functionalities and better structures than those using traditional materials [1]. The

[1] Corresponding Author E-mails: weidong.li@coventry.ac.uk

scenario is depicted in Figure 1, in which the design team specifies the overall structure as well as the chemical constituents of biomaterials, micro-/meso-/macro-structures and manufacturing paths of a biomedical product iteratively. Furthermore, rapid and virtual (digital) manufacturing facilities are dynamically and efficiently used to verify the manufacturability, assemblability and profitability of new materials-based products.

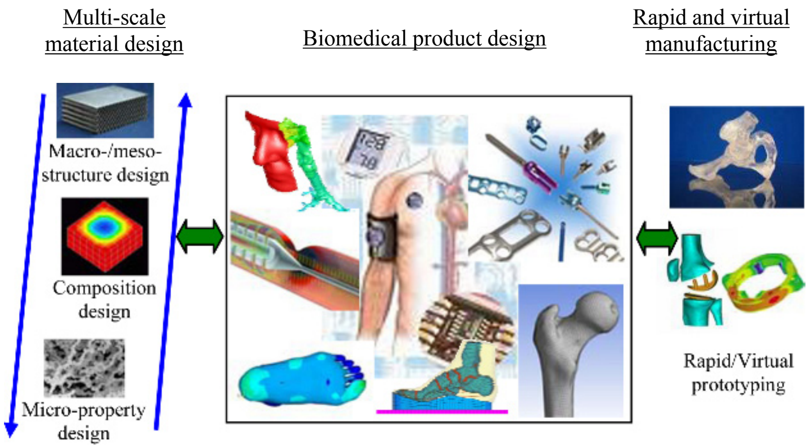
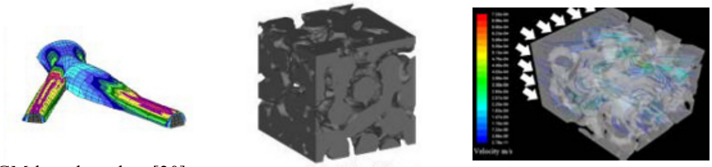


Figure 1. New design paradigm with multi-scale materials and rapid/virtual manufacturing

From the perspective of new materials, a successful implementation of the above scenario is the application of heterogeneous materials (HE materials) to biomedical product development. HE materials, which have shown broad potentials in various applications, including biomedical, automotive, aerospace, energy, civil, nuclear and naval engineering, are presently in the forefront of material research. The heterogeneities are either from varied material compositions/distributions (material heterogeneity), specific mesoscopic structure design (structural heterogeneity), or both. A typical HE material is a Functionally Graded Material (FGM) shown in Figure 2. Different from a conventional homogenous material, FGM contains spatial and gradual variations in composition/distribution from two or more materials for stress concentration avoidance and better local control of thermal, structural, electrical or functional properties. By adopting heterogeneities, the advantages of various materials can be fully exploited and conflicting design requirements can be alleviated. Meanwhile, functional gradation is an important characteristic of biomaterials, and FGM can mimic such gradient. For instance, heterogeneous Titanium (Ti)/Hydroxyapatite (HAp) is under investigation to develop artificial teeth with the composition of pure Ti at one end achieving excellent mechanical properties, increasing the concentration of HAp towards 100% at the other end to be more appetitive for the lower part implanted inside the jaw bone. FGMs-based tissue scaffolds with properly arranged porosity, pore size/shape and interconnectivity, such as a HAp-ceramic graded pore structure, has been recently researched to create physiologic tissues to restore, maintain or improve the function of human tissues and the relevant functionalities [2-3, 14].

However, the extraordinary characteristics of the material heterogeneities have not been fully leveraged in product development including biomedical products. One of the major barriers is that the current Computer-Aided Design/Engineering/Manufacturing (CAD/CAE/CAM) systems have been developed mainly based on homogenous materials, and the latest research results in HE materials have not been effectively updated. The authors have identified, through active discussions and research with manufacturing and software companies and the investigation of the state-of-the-art research, the following challenges should be addressed as a matter of urgency:



(a) A FGM-based product [20]

(b) Tissue scaffolds with meso-structures to enable body fluid flow [18]

Figure 2. Typical HE materials and structures

HE material modeling methodology: To fully take advantage of the knowledge-based digital technologies (i.e., CAD/CAE/CAM), it needs to digitalize the representation and modeling of typical HE materials applicable to biomedical products, develop HE material-based biomedical CAD/CAE/CAM systems, and establish seamless workflows and information exchange protocols between the systems.

HE material design methodology: The incorporation of HE materials in biomedical product design brings forth a complicated design space characterized with a highly interdisciplinary nature and intractable constraints/requirements. To identify design functional modules and parameters, it is imperative to develop a systematic design methodology of HE material design for biomedical products.

Flexible manufacturing simulation of HE material-based biomedical products: Considering the high cost incurred in HE material-based manufacturing processes and rapid prototyping, it is desired to adopt virtual manufacturing simulation to refine design to achieve optimization in terms of manufacturability and assemblability. The virtual manufacturing are able to provide a flexible, efficient and economical means to verify, prioritize and optimize various options of HE material-based design.

In this paper, preliminary results from a pilot research to satisfy the above requirements are presented. A software prototype system with the following three functions has been developed: (i) A HE material design methodology for complex biomedical products; (ii) Designer-centric features to represent and visualize HE material-based biomedical product models; and (iii) Layer-based virtual manufacturing simulation to verify the fabrication feasibility and efficiency of HE material-based biomedical products. In order to demonstrate the effectiveness of the developed methodologies and system, a case study is presented in the paper.

2 Related Works

The research on biomedical products and applications is very active worldwide. A number of cutting edge methodologies and technologies across a wide spectrum of engineering disciplines have been developed, including healthcare/biomedical design methodologies [4], medical imaging processing and simulation [5-6], assembly and automation technologies [7], surface measurement and processing [8-9], rapid and virtual manufacturing [10-15], etc.

Manufacturing technologies of HE materials are essential because design needs to be not only theoretically correct but also realizable in manufacturing. Considering the characteristics of HE material-based products, one of the major challenges is to develop affordable and customized free-form manufacturing technologies. In UK, the relevant research and funded projects in HE materials focus on the exploration of suitable rapid manufacturing technologies applicable for biomedical product development, including: the ultrasonic technology for manufacturing functionally graded cellular microstructures from the Heriot-Watt University [12], the laser-based manufacturing technology from the Birmingham University funded through EPSRC GR/S02365/01 (A Feasibility Study of the Manufacture of Functionally Graded Alloys and Composites using Laser Fabrication, 2003-2006), the ink-jet 3D printing technology from the Manchester University funded through EPSRC EP/C01328X/1 (Manufacture of Bio-Functional Components by Freeform Fabrication, 2005-2007), the powder fusion technology from the Rapid Manufacturing Research Group funded by EPSRC IMRC [13], and the platform establishment of the powder ink-jet, electro-hydrodynamic jetting, filament free-forming and dry-powder dispensing technologies from the University College London funded through EPSRC EP/E045839/1 (Intercollegiate Platform on Powder-based Synthesis and Modeling, 2008-2011) [14].

Industries witness extensive deployment of CAD/CAE/CAM systems to facilitate various product development. However, no or limited functionalities have been developed in the systems to support the modeling, design, optimization and free-form fabrication of HE materials, due to the current digital systems' limitations in terms of design methodology and process, infrastructure, data structure, etc. To tackle this problem, new modeling methodologies of HE materials to facilitate product design are in active research. The National Science Foundation (NSF) of USA has continuously funded the research in recent years (www.nsf.gov). A research team from the Massachusetts Institute of Technology developed parametric and feature-based HE materials modeling with local composition control to support 3D printing [16]. The advent of a B-Spline-based method from a team of the University of Michigan is able to represent complex HE material-based features and products [17]. Teams from the Carnegie Mellon

University and the Drexel University investigated new HE materials modeling methods for scaffold structures and porous artefacts in tissue engineering, which have varying surface, internal and volumetric properties and geometrics [3, 18]. A team from the University of Hong Kong has carried out research of introducing the Axiomatic Design to guide HE materials and product design. Product components, constituent compositions, periodic meso-structures of HE materials and inclusions are linked with product requirements to model design in a structural way to deduce good solutions [19]. A team from the Georgia Institute of Technology has conducted a series of research in the integration of system design and multi-functional/scale materials including HE materials [20].

However, research gaps still exist, especially in the following aspects:

HE material modelling: The underlying data schemes of biomedical models generated from CT (Computed Tomography) and MRI (Magnetic Renounce Imaging) are cell (voxel)-based or mesh-based. Models are usually costly and problematic in further computing and processing. It is imperative to synthesise the existing research efforts, and develop parametric features with extensiblity and flexibilities for complex heterogeneities.

HE material design: The previous decision-making processes in HE material design are intuitive and ad hoc. It needs to introduce a set of design governing principles for typical HE materials to guide conceptual and detailed product design.

Manufacturing evaluation: The physical manufacturing processes are relatively time-consuming and uneconomical, in particular for complex HE materials. Research on the efficient verification of fabricating HE materials and the relevant products is rare. It is beneficial to industries if a set of virtual HE materials manufacturing tools are applied to minimize the lead time and expense of the iterative processes of material and product design.

3 System Framework and Design Methodologies

In this research, design methodologies and a software prototype system have been developed to support HE material-based biomedical product development. The system consists of three components, which are shown in Figure 3.

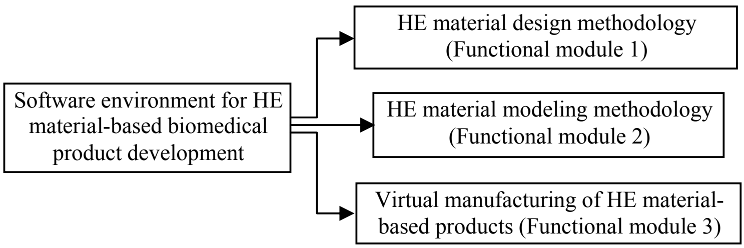


Figure 3. System Framework and functional modules

3.1 HE Material Design Methodology

HE material design is comprised of the selection of primary materials and the composition design and optimization of the materials. One property of a material can be displayed as a ranked list or bar-chart. However, it is rarely that the performance of a product just depends upon one property. The relationships between the governing principles and criteria need to be understood to achieve the best trade off. Based on a series of key data and appropriate mathematical approximation models, two or three criteria can be put into a two or three dimensional space for a comparative study. The Ashby's approach [1] is to use material property charts for material selection. For instance, in some cases, it needs to have a higher modulus while the density needs to be relatively lower. A Modulus-Density chart can be plotted using the CES software based on the Ashby's approach (Figure 4). In the plot, data for member of a particular family of materials are usually cluster together using an envelop for better illustration. A line of $Modulus / Density$ is drawn to filter infeasible materials. Once the primary materials are selected, the materials are formed as a composite and the finite element analysis is used to verify the structural integrity. The detailed procedure is explained in a case study of Section 4.

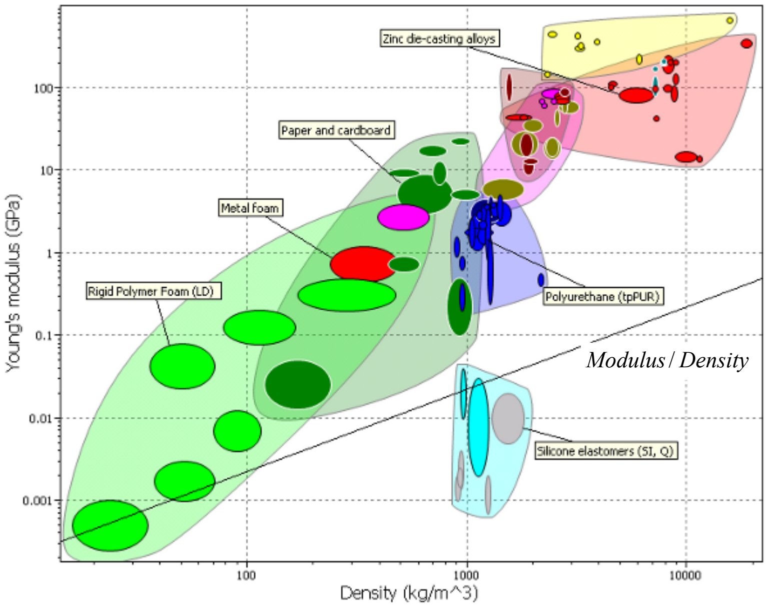


Figure 4. Young's Modulus vs. Density plotted using the CES software [1]

3.2 HE Material Modelling and Virtual Manufacturing

Three relationships of HE material design have been developed in an open source modeling software – Open Cascade (www.opencascade.org), i.e., linear heterogeneity, radial heterogeneity and NURBS-based heterogeneity (shown in Figure 5). The relationships can be used to meet the requirements of the majority of HE material design. The software system is able to be interoperated with commercial CAD/CAE/CAM systems via neutral file exchanges (e.g., ANSYS Workbench). Meanwhile, a virtual rapid prototype machine consisting of multiple machine tools, multiple material nozzles, manufacturing planning algorithms and section slicing algorithms has been developed and integrated into the HE material design system (some examples are shown in Figure 6).

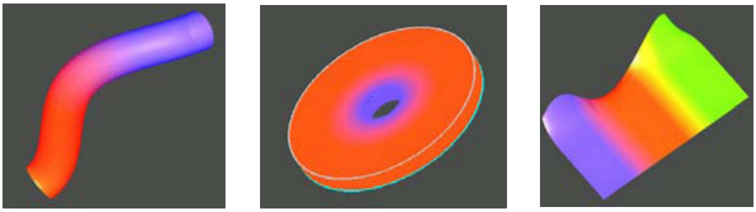


Figure 5. Modelling cases for various HE materials (two or more materials with transition phases)

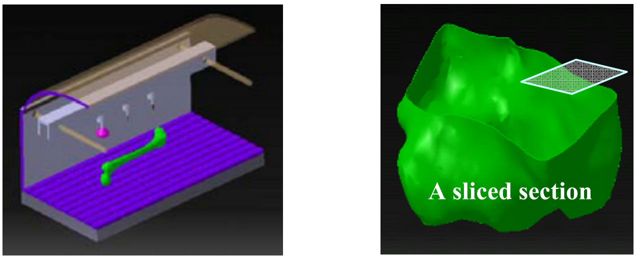


Figure 6. A virtual prototyping machine and a slicing algorithm to drive the machine

4 A Case Study

Osteoporosis is a common and costly disease. Skeletal cancer and serious bone fractures also require expensive treatment. Normally, worn bone parts need to be replaced with bone implants. In this case study, a HE material-based bone implant has been designed using the Ashby's approach. Finite element analysis (ANSYS Workbench) has been introduced to verify the strength and stiffness of the bone. In the end, the bone has been virtually manufactured.

In Figure 7 (a), a bone CAD model (stl format) can be generated through data processing of micro CT images (e.g., Mimics and Simpleware software). In order to use the Ashby's approach [1] to determine the primary material, the model has been simplified as Figure 7(b) and a mathematical model has been deduced as follows.

(a) A CAD model of a femur bone (b) A mathematical model of the bone

Figure 7. A bone model and its simplified mathematical model

The parameters of the bone implant are: density - ρ, young's modulus - E, length - L, maximum load - F, and radius - r. The objective is to minimize the mass:

$$m = r^2 L\rho\pi \tag{1}$$

subject to the constraint that is supports the load without buckling. The elastic buckling load F' is:

$$F' = \frac{\pi^2 EI}{L^2} = \frac{\pi^3 E r^4}{4L^2} \tag{2}$$

$F' < F$ then,

$$m \geq (\frac{4F}{\pi})^{1/2}(L)^2[\frac{\rho}{E^{1/2}}] \tag{3}$$

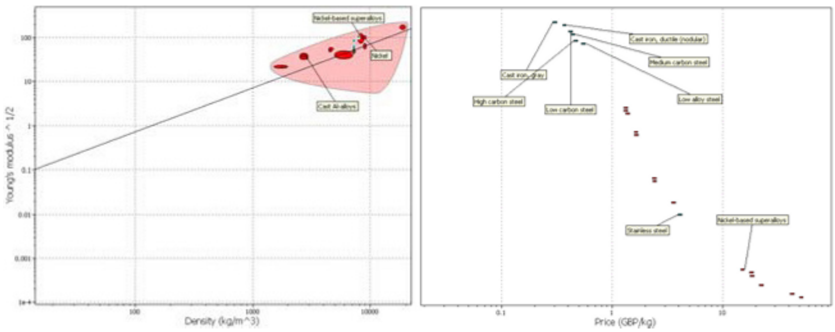
In order to minimize m, it needs to select the materials with the greatest value of:

$$\frac{E^{1/2}}{\rho} \tag{4}$$

In addition, two more conditions are added: (i) all the selected materials need to be metal or alloys; and (ii) the selected materials are ranked in terms of cost. The plots of these two parameters have been generated in the CES software and shown in Figure 8.

To avoid erosion, stainless steel has been eventually selected as the primary material. HAp to improve the biocompatibility has been then coated on the stainless steel as a linear gradual variation HE.

The designed bone implant has been designed in the Opencascade-based design system and passed to ANSYS Workbench for finite element analysis. It can be found that the strength meets the safety requirement. The manufacturing simulation module has been invoked to verify the manufacturability. The intermediate result of the finite element analysis is shown in Figure 9.



(a) Plot of $\frac{E^{1/2}}{\rho}$ (b) The final qualified materials ranked in terms of price

Figure 8. The selection of primary material using the CES software

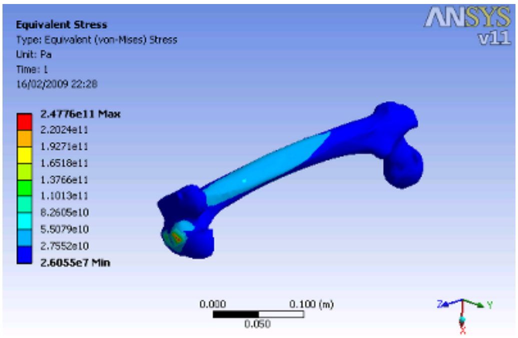


Figure 9. The finite element analysis result of the HE material-based bone implant

5 Conclusions

In this paper, design methodologies and a software prototype system have been developed for HE material-based biomedical product development. The following three functions are covered: (i) A HE material design methodology for complex biomedical products; (ii) Designer-centric features to represent and visualize HE material-based biomedical product models; and (iii) Layer-based virtual manufacturing simulation to verify the fabrication feasibility and efficiency of HE material-based biomedical products. To explain the effectiveness of the developed methodologies and system, a case study is presented in the paper.

6 References

[1] Ashby, M.F., 2005, *Materials Selection in Mechanical Design (3rd Edition)*, BH Publisher.

[2] Pompe, W., Worch, H., Epple, M., Friess, W., Gelinsky, M., Greil, P., Hampel, U., Scharnweber, D. and Schulte, K., 2004, Functional graded materials for biomedical applications. *Materials and Engineering A*, 362(1-2): 40-60.
[3] Sun, W., Starly, B., Nam, J. and Darling, A., 2005, Bio-CAD modelling and its applications in computer-aided tissue engineering. *Computer-Aided Design*, 37(11): 1097-1114.
[4] Buckle, P., Clarkson, P.J., Coleman, R., Ward, J. and Anderson, J., 2006, Patient safety, system design and ergonomics. *Applied Ergonomics*, 37(4): 491-500.
[5] Long, Q., Merrifield, R., Xu, X.Y., Kilner, P., Firmin, D.N. and Yang, G.Z., 2008, Subject-specific computational simulation of left ventricular flow based on magnetic resonance imaging. *IMechE, Part H*, 222(4): 475-485.
[6] Simpson, D.J., Brown, C.J., Yettram, A.L., Procter, P. and Andrew, G.J., 2008, Finite element analysis of intramedullary devices: the effect of the gap between the implant and the bone. *IMechE, Part H*, 222(3): 333-345.
[7] Williams, D.J., Ratchev, S. and Hirani, H., 2006, The application of assembly and automation technologies to healthcare products. *Annals of the CIRP*, 55(2): 617-642.
[8] Ramsden, J.J., Allen, D.M., Stephenson, D.J., Alcock, J.R., Peggs, G.N., Fuller, G. and Goch, G., 2007, The design and manufacturing of biomedical surfaces. *Annals of the CIRP*, 56(2): 687-711.
[9] Blunt, L.A., Bills, P.J., Jiang, X.Q. and Chakrabarty, G., 2008, Improvement in the assessment of wear of total knee replacements using coordinate-measuring machine techniques. *IMechE, Part H*, 222(3): 309-318.
[10] Lim, T., Ritchie, J.M., Dewar, R.G., Corney, J.R., Wilkinson, P., Calis, M., Desmulliez, M. and Fang, J.-J., 2007, Factors affecting user performance in haptic assembly. *Virtual Reality*, 11(4): 241-252.
[11] Goodridge, R.D., Wood, D.J., Ohtsuki, C. and Dalgarno, K.W., 2007, Biological evaluation of an apatite–mullite glass-ceramic produced via selective laser sintering. *Acta Biomaterialia*, 3(2): 221-231.
[12] Torres-Sanchez, C. and Corney, J.R., 2008, Effects of ultrasound on polymeric foam porosity. *Ultrasonics Sonochemistry*, 15(4): 408-415.
[13] Savalani, M.M., Hao, L., Zhang, Y., Tanner, K.E. and Harris, R.A., 2007, Fabrication of porous bioactive structures using the selective laser sintering technique. *IMechE, Part H*, 221(8): 873-886.
[14] Gupta, A., Seifalian, A.M., Ahmad, Z., Edirisinghe, M.J., Winslet, M.C., 2007, Novel electrohydrodynamic printing of nanocomposite biopolymer scaffolds. *Journal of Bioactive and Compatible Polymers*, 22(3): 265-280.
[15] Sachlos, E., Reis, N., Ainsley, C., Derby, B. and Czernuszka, J.T., 2003, Novel collagen scaffolds with predefined internal morphology made by solid freeform fabrication. *Biomaterials*, 24: 1487-1497.
[16] Liu, H., Maekawa, T., Patrikalakis, N.M., Sachs, E.M. and Cho, W., 2004, Methods for feature-based design of heterogeneous solids. *Computer-Aided Design*, 36, pp. 1141-1159.
[17] Qian, X.P. and Dutta, D., 2003, Physics-based Modelling for Heterogeneous Objects. *Transactions of ASME: Journal of Mechanical Design*, 125, pp. 416-427.
[18] Starly, B. and Sun, W., 2007, *Biomimetic Design and Fabrication of Tissue Engineered Scaffolds*, VDM Verlag, ISBN: 978-8364-2464-6.
[19] Chen, K.Z. and Feng, X.A., 2004, CAD modelling for the components made of multi heterogeneous materials and smart materials. *Computer-Aided Design*, 36, pp. 51-63.
[20] Messer, M., Panchal, J.H., Allen, J.K., McDowell, D.L., and Mistree, F., 2007, A function-based approach for integrated design of material and product concepts. *ASME 2007 DETC Conference*, Las Vegas, USA, DETC2007/DAC-35743.

Research on Variable Parameter Set in Complex Multi-Domain Physical System and Its Repeatedly Simulating Arithmetic

Renwang Li[a,1], YiZhong Wu[b], Liang Gao[b], Zhansi Jiang[b], Zefei Zhu[a]

[a] Department of Industrial Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, P.R.China
[b] School of Mechanical Science and Engineering, Huazhong University of Science and Technolog, Wuhan 430074, P.R.China

Abstract: As a unified modeling simulation language for complex multi-domain physical system, now Modelica is getting more and more attention, and its compiler and solver method emerges in endlessly. However, up to now little solver method consider the efficiency of repeatedly simulation after modifying variable parameters. Here, a method focuses on improving the efficiency of repeatedly simulation for complex model is set forth. Firstly, elucidate some related definitions such as coupling block, data dependence, coupling block dependent graph, variable parameter set and so on. Secondly, illustrate the hierarchy expression of coupled block dependence graph, variable sub-graph and minimum solution tree. Finally, taking a concrete object as an example, some related conclusions are obtained.

Keywords: Variable parameter set, complex Multi-domain physical system, arithmetic.

1 Introduction

With the development of modeling and numerical technology, mathematical modeling and simulation is playing a more and more important role in today's design of new product, and has become an important measure to test and analyze product technology performance [1]. However, with the continual enhancement of product complexity, as well as more and more attention focused on whole product performance, now popular method for simulation of sub-system in different domain by oneself is limited, and apparently modeling and simulation techonology is needed urgently for multi-domain complex physical system. Fortunately, now it seems that integration and data exchange between different domain-models can be carried out wonderfully based on unified modeling language. Modelica[2,3],

---

[1] Corresponding Author E-mail: lirenwang2008@googlemail.com. The work is supported by the National Natural Science Foundation (Grant No. 50775084, 50675208), Zhejiang Province National Natural Science Foundation for Outstanding Youthful Research Team (Grant No. R6080403) and Zhejiang Province Qianjiang Plan Project (QJD0702002), P.R.China

appeared in 1997, is one of the most popular multi-domain physical systems modeling language. Modelica describes physical rule and phenomenon of different domain sub-system in the form of mathematical equation, and carry out modeling construction and multi-domain integration based on language inner component link mechanism according to topological structure of physical system.

Moreover, up to now there are many literatures about simulating complex physical system, however little of them are focused on repeated simulation under the circumstance of variable parameters.

2 Basic Definition

Definition 1 coupling block
DAG (directied acycline gragh) G' is obtained from Directed G through coagulating strongly connected component[4]. As shown in Figure 1, vertex of G' is a strong coupled equation subset, called copling blockt.

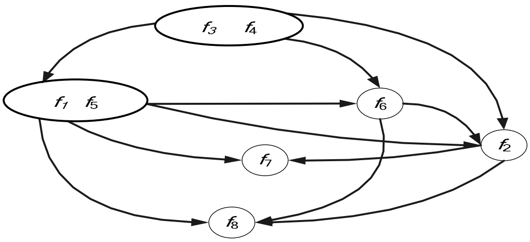


Figure 1. Coupled block dependence graph

Definition 2 Data dependence
Supposed B1and B2 are coupling block, B2 depends B1 if there is variable x in B2, and x's solution comes from B2.[5]

Definition 3 Coupled block dependence graph
If vertex set B is a coupled block of $G=(B,E)$ dealt with independently, $(s,k)\in E$ is a kind of data dependence relationship between vertex s and k, G called Coupled block dependence graph (BDG).

Definition 4 Variable parameter set
During model parameter testimation or parameter optimization, usually set different values repeatedly only for a group of parameters, and then obtain related answer or variable numerical solution. Here, this group of parameters called variable parameter set.

3 Variable parameter set

In Modelica syntax, parameters of main model usually are transferred to component through variant mechanism of component. If there exist parameter dependence, component's parameter depends on main model's parameter, and then no modification is allowed here. Moreover, when changing main model's parameter, related component's parameter also changed. In order to change equation solution area when modifying parameters, firstly it is necessary to remove parameter dependence relationship, i.e. replace dependent parameters with independent paramenters; secondly locate vertex of changed parameters in relation graph, search this vertex's spread area, and find out sub-graph needed to be dealed with again.

3.1 Hierarchy expression of coupled block dependence graph

In order to simplify solution dependence relationship in block dependence graph, it is useful to remove unnecessary dependence border though hierarchy expression of coupled block dependence graph [5,6]. Apparently, no dependence relationship between vertexs in the same layer, after describing coupled block dependence graph in the form of hierarchy expression.

For example, hierarchy graph shown in Figure 3, can be obtained from coupled block dependence graph in Figure 2.

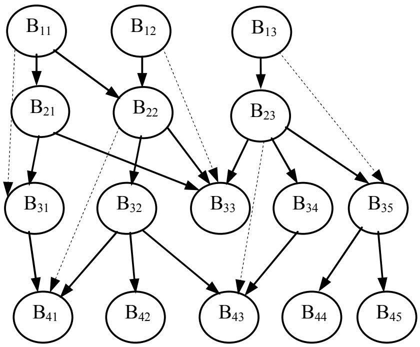


Figure 2. Original coupled block dependence graph

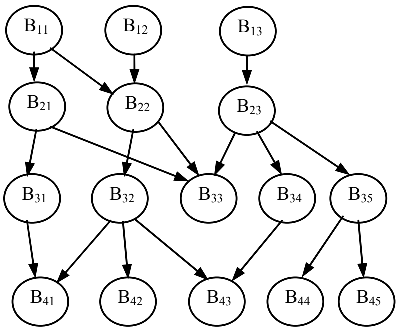


Figure 3. Hierarchy expression of original

3.2 Variable sub-graph

Variable sub-graph refers to a new sub-graph through transferring a group of independent variable parameters in coupled block dependence graph. The method of creating variable sub-graph is to find out vertex set of top layer in the hierarchy expression of coupled block for each paramenter in variable parameter set, and then variable sub-graph is the merge set of those vertex set's spread area.

For example, coupled block dependence graph in Figure 4 corresponds to variable sub-graph of variable parameter set Pv={p1, p2, p3}, i.e. thick line shown in Figure 5.

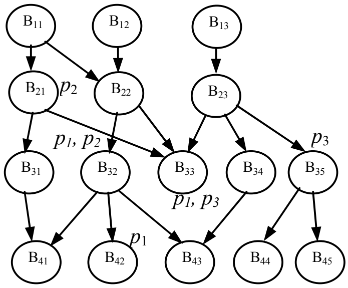


Figure 4. Variable parameter locations in hierarchy expression of coupled block

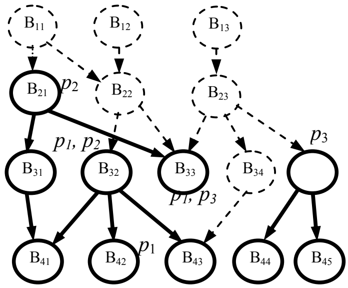


Figure 5. Variable sub-graph corresponding to variable parameter set Pv={p1, p2, p3}

3.3 Minimum solution tree

In Modelica, it is already possible to get the solution of variable sub-graph in coupled block dependence graph. After having obtained the first simulation solution, follow-up solution can be gotten through simply dealing with its variable sub-graph corresponding to variable parameter set. However, here it is necessary to create minimum solution tree corresponding to variable sub-graph, in order to avoid repeatedly dealing with those vertexes in the same layer.

For example, Figure 6 is the minimum solution tree of variable sub-graph corresponding to Figure 5, and Figure 7 is the serial solution relatated to Figure 6.

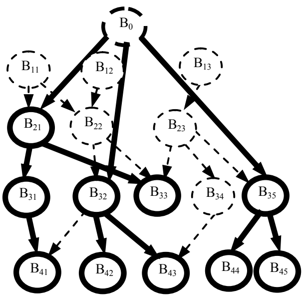


Figure 6. Minimum solution tree

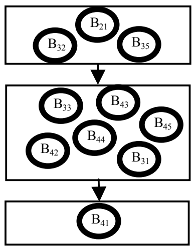


Figure 7. Serial solution related to Figure 6

4 Arithmetic realizations and its instance

All related arithmetic realizations here are based on MWorks [6], i.e. multi-domain physical system modeling and simulation platform. The following is some simple introduction of key data structure related to arithmetic.

Here, taking the simplified battle-plan F14 model [8] as an example.

1) Equation tree

Leafage node is operation number or function parameter, which can be constant, parameter or variable; other node is operator or function symbol.

Usually equation expression read from Modelica model file is described in the form of infix, however this kind of expression is inconvenient to elucidate the tree structure. Therefore, it is necessary to change the infix mode of equation expression into postfix. And then, construct Binary Tree through making use of stack operation according to postfix expression.

For example, change the equation $\cos(a)+5*(b-c)=d/e$ into $a\cos 5bc-*+de/=$, "=" is root node.

2) Bipartite Graph

Data structure of Bipartite Graph [6] as shown in Figure 8. Here graph, border and vertex separately correspond to clustering bigraph, edge and node.

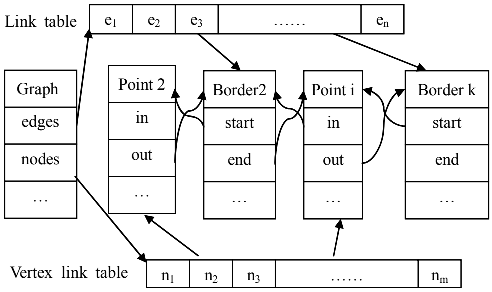


Figure 8. Data structure of Bipartite Graph

The construction process of restriction expression graph for equation system can be described as follows.

Step1, create an instance of Bipartite Graph

Step2, construct vertex for each equantion in equation system in turn, and obtain inclusive variable from equantion object. If vertex corresponding to variable already created, obtain this vertex, and then make up an edge between eqation vertex and variable vertex. Otherwise, create a new vertex for variable, and then make up an edge between eqation vertex and new variable vertex.

Steps 3, repeat Step 2 until all vertexes for each equation is created.

3) Coupled block dependence graph

The data structure of coupled block dependence graph is similar to the above, but the data structures of vertexes are all coupled block.

The Coupled block dependence graph of Controler_F14 model equation system is shown as Figure 9.

4) Minimum solution tree

The data structure of minimum solution tree conserves the data-pointer of root node, and the related operation of realization. The corresponding minimum solution of K_f in the simplified battle-plan F14 model [7] is as shown in Figure 10.

5) Repeated simulation

Based on the above, necessary minimum solution block can be easily carried out through making use of variable parameter set based on minimum solution tree of all corresponding variable parameter set. Apparently, solution series can be obtained through taking use of breadth-first search of minimum solution tree, and no restriction of order for dealing with the nodes in the same layer of minimum solution tree.

Figure 9. The Coupled block dependence graph of Controler_F14 model equation system

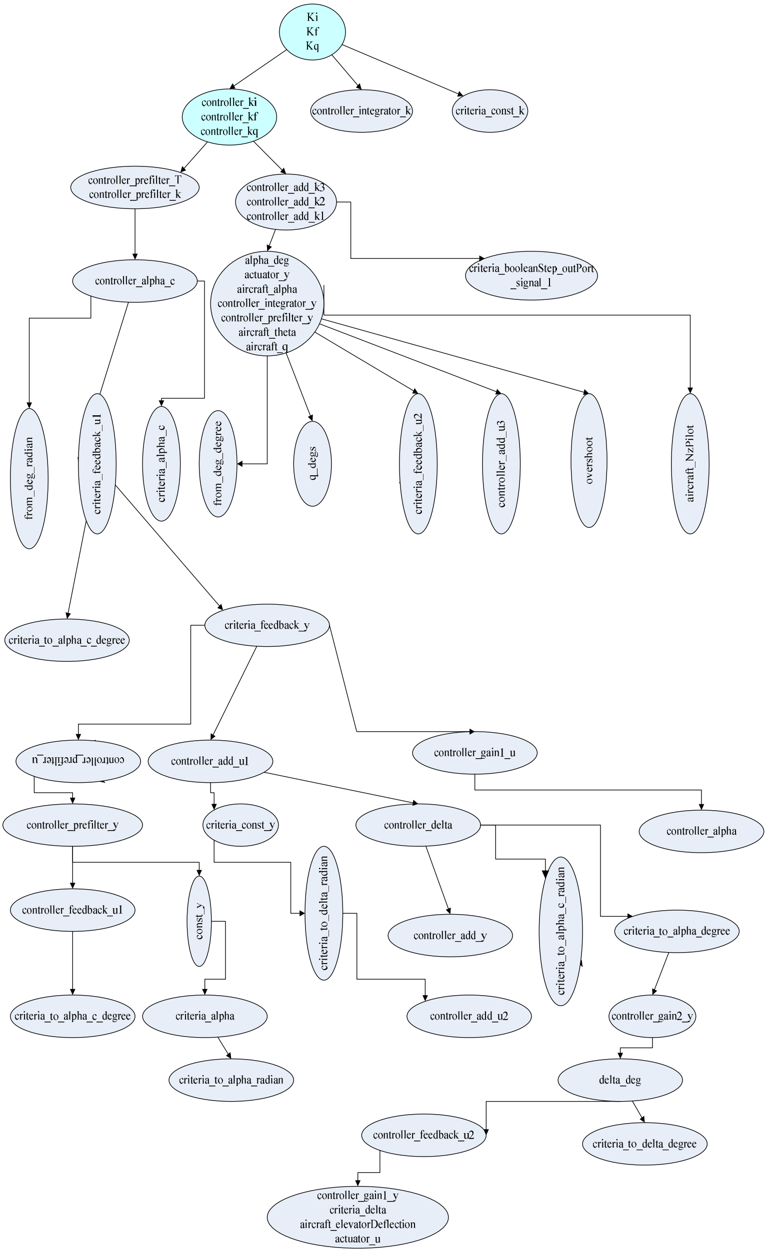


Figure 10. The corresponding minimum solution of K_f

In the process of simulating based on Dymola platform, modifying any one or several parameter took almost the same time for repeatedly simulation. However, In the process of simulating based on MWorks platform, the first simulation takes the most time, successively repeated simulation usually takes less and less time.

5 Conclusion

1) A kind of method about simulating complex physical system is set forth, which is based on scale decomposition of equation system and Modelica syntax. This method can efficiently avoid dealing with all coupled block when repeating simulation.
2) When repeatedly simulating different parameters of variable parameter set, only minimum solution tree according to hierarchy expression of coupled block dependence graph is needed, after making use of parameter dependence relationship between coupled BDG, and setting up minimum solution tree of variable parameter set.
3) As to other ordinary simulation paltform, the work here is useful for reference, especially for improving follow-up design optimization and model experiment based on simulation analysis.

6 References

[1] Andersson M. Object-Oriented Modeling and Simulation of Hybrid Systems: [PhD thesis]. Sweden: Lund Institute of Technology, 1994
[2] Fritzson P, Engelson V. Modelica-A unified object-oriented language for system modeling and simulation, Proc. of the 12th European Conf. on Object-Oriented Programming. Brussels: Springer-Verlag, 1998, pp. 67-90
[3] Fritzson P. Principles of object-oriented modeling and simulation with Modelica 2.1. New York: IEEE Press, 2003
[4] Steward, D. On an Approach to Techniques for the Analysis of the Structure of Large systems of Equations, SIAM rev. Vol5. No. 4 1962, pp322-342.
[5] Bunus P, Fritzson P. Semi-automatic fault localization and behavior verification for physical system simulation models. Proc. of the 18th IEEE International Conference on Automated Software Engineering, Montreal, Canada, 2003, pp.236-241.
[6] Ding, J.W., Chen, L.P., Zhou, F.L. et al. An analyzer for declarative equation based models, Proc. of the 5th International Modelica Conference, Vienna, Austria, 2006, pp.349-357.
[7] Dynasim AB. User's Manual Dymola 6 Additions, 2006
[8] Powell M J D. Problems Relate to Unconstrained Optimization. Murray, W. ed. Academic Press, 1972.

Two Stage Ant Coordination Mechanisms for Sequencing Problem in Mixed Model Assembly Lines

Qiong Zhu, Jie Zhang[1]

CIM Research Institute, School of Mechanical Manufacturing, Shanghai Jiao Tong University, Shanghai, 200240, China

Abstract: The minimization of material usage rates variation in the mixed model assembly lines (MMAL) is one of the major optimal problems to effectively execute Just-in-Time (JIT) production. In this paper we present a novel approach with two stage ant coordination mechanism (2S-ACM) to optimize the sequence of mixed product models. In the 2S-ACM approach, the elements in MMAL sequencing problem are transferred to different functional agents, and these agents simulate the ant foraging behaviour to coordination with each other and find the near optimal solutions. The computational results show that the 2S-ACM approach outperforms the other two stage heuristic method and one stage heuristic method in minimizing material usage rates variation and general sacrifice some CPU time in an accepted range. Therefore, the 2S-ACM approach can be used as one of effective approaches for solving sequencing problem in MMAL

Keywords: mixed model assembly lines, sequencing, ant coordination mechanism

1 Introduction

Current market trends such as variety of consumer demands and shorter product life cycles have resulted in a competitive business environment, in which the enterprises should have the ability of rapid responding to customer demand and keeping stable production.

The Just in Time (JIT) approach is one important tool that supporting the stable production. The JIT approach requires that the materials pass through are very evenly matched and balanced, such that manufacturing wastes are eliminated and the flow of materials as they get manufactured through the line is smooth. Accordingly, JIT approach aims at keeping the material flow as constant as possible and increasing the stability of production procedure.

The MMA) production system is another important tool that rapid processing customer demands. MMAL production is a practice of assembling a variety of different product models similar in product characteristics on the same line without changeovers. In the MMAL production system, customer orders can be scheduled

[1] Corresponding Author E-mail: zhangjie@sjtu.edu.cn

and produced without holding large inventories or incurring large shortages. Currently, MMAL production system can be found in many industrial environments because of their capability of giving faster response to market changes.

This paper deals with the JIT based MMAL sequencing problem with objective of minimizing the variation of material usage rate. The objective of this research is to develop an efficient scheduling policy that should be simple and robust.

The remainder of this paper is organized as follows. This paper is organized as follows: section 2 reviews the related research background. Section 3 presents the two stage model formulation. Section 4 proposes a dynamic schedule policy to optimize the sequence of product models. In section 5, the experiment for performance measures and case study is made. Section 6 is conclusions.

2 Literature review

As part of the famous “Toyota production system”, MMAL sequencing problem with JIT objective received wide attention from both theoretical and practical point of view and is still discussed up to now. JIT objective focus on guarantee a steady flow of material on the assembly line. Here, materials may refer to part level material or product level material, or both levels material.

The part level JIT objective aims at minimizing the variation of material usage rates so that a JIT supply of material is facilitated. Toyota Corporation (Mondern, 1993) transferred part level MMAL sequencing problem to a mathematical model with a pure quadratic distance function that minimizes the total variation of each model’s practical consumption rate of parts from that of the ideal value, and developed a Goal Chasing Algorithms (GCA) to handle this problem [1]. GCA selects a model that minimizes the one stage material usage rates variation at each stage. Here, a “stage” represents a position in the order of a sequence, and the “one (two) stage” variation refers to the changed value associated with a progression of one (two) position in the sequence. Miltenburg and Sinnamon (1989) presented a multi level part model in sequencing mixed model assembly line. This multi part level model considers the variation at the product, sub-assembly, component, and raw material levels. Heuristics were developed based on direct enumeration of one and two stage variations at all levels [2]. Bautista et al. (1996) presented several goal chasing method improvements including a transformed approach that used model terms in the minimizing the one stage variation, an approach that further considers the two stage variation and a rate-preserving approach to prevent the impoverishment of the remainder sequence. They also presented a bounded dynamic programming procedure to obtain either optimal or heuristic solutions to the problem of minimizing variation in parts usage rate [3]. Kubiak et al. (1997) developed a dynamic programming approach to obtain optimal solutions for the part level model in sequencing mixed model assembly lines. It was shown that the part level problem is NP-hard in strong sense [4].

Some researchers focus on JIT objective as well as other related objectives. Ding, Zhu et al. (2006) presented a joint objective of multiple objectives that include minimizing variation of part level usage rate, variation of production rate (product level variation), and station-time usage rate variation. Potential interaction among the various objectives is discussed simultaneity [5]. Rahimi (2007) considered three objectives simultaneously: minimizing total utility work, product level variation and total setup cost and developed a Multi-Objective Scatter Search algorithm to solve the model. And the computational results show that the proposed algorithm outperforms the existing genetic algorithm, especially for the large sized problems [6].

Other researchers investigated the relationship between part level JIT objective and product level JIT objective. Kubiak (1993) addressed that whether all models requiring the same number of mixed parts is a sufficient condition for distinguishing of the product level and part level problem, and the product and part level problem is discussed under the assumption that all models require the same number of mixed parts [7]. Zhu and Ding (2000) show that additional cases lead to equivalence of part level and product, the author consider a transformation of product level variation to an "un-weighted" part level variation, where the variation function weights all variation by the same constant factor. The authors considered the simplification and transformation process of the two stage part level variation as a heuristic solution approach [8].

It was observed the schedule problem in MMAL for the purpose of achieving part level JIT objective is becoming an extremely important task. Boyson, et al. (2009) consider about the part level JIT objective and discuss experimental results obtained from many different approaches listed in literatures, and conclude that the part usage variation is certainly an interesting base model of high theoretical value and might be directly used in computer scheduling problems [9]. The experimental results obtained from literatures [3, 4, 8] indicate that the mixed-model sequencing methods of minimizing the two stage variation are generally better than the ones of minimizing the one stage variation. However, the presented two stage variation optimization is based on heuristic rules, the high performed artificial intelligence approaches are never applied in this field. This note further the research in directly considering the two stage part level variation as a heuristic solution approach, and in understanding the effect of minimizing the part level variation. An approach of 2S-ACM approach is proposed on minimizing two stage material usage rates variation. The objective is minimizing the variation of material usage rates so as to smooth the requirement of parts. The performance in terms of optimal results and CPU time will be compared with one stage results.

3 Model formulation

The following notations are defined for the sequencing problem.

Index:

$n \in \{1,2,\cdots\cdots,N\}$ is the index of stages;

$m \in \{1,2,\cdots\cdots M\}$ is the index of types of product model;

$p \in \{1,2,\cdots\cdots P\}$ is the index of parts used in MMAL;

d_m demand amount for product type m ;

b_{mp} the required amount of part p for product m ;

a_p the ideal usage rate of part , $a_p = \sum_{m=1}^{M} b_{mp} d_m / N$;

θ_{np} the cumulative consumption of part p from 1 through n stages.

x_{nm} the Boolean value identifies whether position n is mapping with model m , if yes, 1, otherwise 0 instead.

To reduce the material usage rates variation, that is, to sequence the product models so that the cumulative usage rate of every part p at any stage n is as close as possible to the ideal material usage rate at stage n, the objective can be expressed as follows:

$$\min \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{p=1}^{P} (na_p - \theta_{n,p})^2 x_{nm} \tag{1}$$

This part level sequencing problem was shown to be NP-hard by Kubiak (1993), direct one stage approaches to (1) adopts the transition rule that minimizes one stage variation $\sum_{m=1}^{M} \sum_{p=1}^{P} (na_p - \theta_{np})^2 x_{nm}$ at each stage n , those include the GCA approach, and the approach by Miltenburg and Sinnamon (1995) [10]. In the two stages aspect, Miltenburg and Sinnamon (1989) presented a two stage method based on directly enumeration method [2]. Zhu and Ding (2000) developed a transformed two stage approach [8]. The approach defines two stage transition rule as formula (2). The approach perform a greedy search process, that is to say, at stage n , the two stage material usage rates variation is optimized by selecting a product model pair (a,b) that minimize formula (2):

$$2A_{m_1} + B_{m_2} + Q_{m_1 m_2} \tag{2}$$

Where : $A_{m_1} = \sum_{p=1}^{p} b_{m_1 p} (\theta_{np} - (n+0.5)a_p - b_{m_1 p} / 2)$,

$B_{m_2} = \sum_{p=1}^{P} b_{m_2 p} (\theta_{np} - (n+1)a_p - b_{m_2 p} / 2)$,

$Q_{m_1, m_2} = \sum_{p=1}^{P} b_{m_1 p} b_{m_2 p}$

The transition rule (formula (2)) can also be employed as heuristic value to guide the optimization search process in other heuristic approaches. This paper adopted this two stage transition rule. Following these general conditions, the remainder of this paper will present a special two stage ant coordination mechanism approach to solve instances of MMAL sequencing problem.

4 A special 2S-ACM approach

The artificial intelligence approach that simulates the ant colony foraging behavior was first proposed by Dorigo and Gambardella in 1997 and has been applied to investigate combinational optimization problems [11]. Ant foraging behavior describes that a colony of ants has a natural ability to find the shortest path between a food source and their nest. If there were two closed paths that a colony of ants can take, where one would be longer than the other, initially, nearly half amount of the ants will take the short path and the rest will take the long path. During their moves from and to the food source, they deposit a chemical substance (i.e., pheromone). Ant instinct to follow each other by sensing the pheromone attracts them to move to the path that has more pheromone. Since pheromone evaporates quicker in the long path than the short one, the pheromone amount in the short path will be more condensed, which will attract more ants to that path over time.

Later, some other approaches that based on ant colony foraging behavior are produced, for example, the ant coordination mechanism (Zhu et al 2008) for a dynamic job shop scheduling problem. Since this publication is in Chinese, a brief introduction is given here. In the ant coordination mechanism, the elements in a job shop are classified into different types of ant agents, including resource ants, central coordination ant and order ants. In the dynamic scheduling process, resource ants release the real-time available resource information, and order ants build individual solutions step by step based on the released resource information. During this procedure, the central coordination ant sorts different solutions generating by different order ants and determines the quantity of pheromone left between resource ants, and the pheromone is the other information to guide the order ants. In this way, dynamic information can be good solutions emerge from the coordination between the ants. To our knowledge, the literature reporting the use of artificial intelligence approaches on MMAL sequencing problem which minimizing two stage material usage rates variation is scarce.

In the proposed 2S-ACM approach for the two stage mixed model sequencing problem, several different types of ants coordinate with each other to find solutions. Figure 1 presents a 2S-ACM sequencing model with 4 types of product model. In this research, three types of ant are proposed including product *model ant*, sequence *search ants* and *central coordination ant*.

- *Model ants* represents different product models, they take charge of real-time updating and dissemination of relevant information about product models, in the format (*m, Rm, n, Am, at*), where *m* is the type index of product models, *Rm* checks the remained quantity of product models, *n* is the current stage number, *Am* defines the first part of heuristic value in formula

(2), *{Bj}*present the set of possible value defined by the second part of heuristic value in formula (2). Model ants interact with and central coordination ant through the information channel.

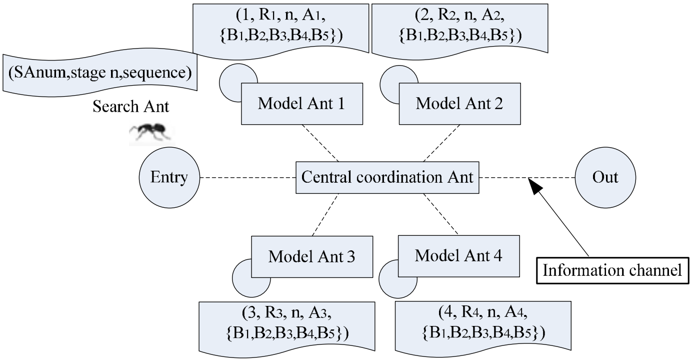


Figure 1. The 2S-ACM sequencing model (*M=4*)

- *Search ants* figure on establishing a sequence solution in Figure 1. In the searching procedure search ants takes its local memory in the format (*SAnum*, *n*, *sequence*), where the *SAnum* is the serials number of search ant, *n* displays the current stage, *sequence* records the list of selected models. Search ants interact with central coordination ant through the information channel.
- *Central coordination ant* plays a role as an information processing centre. As illustrated in Figure 1, central coordination ant exchanges information with all the elements though the information channel.

Figure 2 presents the outline of 2S-ACM approach. The 2S-ACM approach is started by creating a new search ant colony with a predefined number of search ants. Each ant in the colony renews its local memory (*SAnum*, *n*, *sequence*) to keep its traversal information. In the second phase, search ants enter the sequencing model (Figure 1) from the entry point and exit though out point. At each stage, search ant reports its local memory information to central coordination ant, and the central coordination ant release the current stage information to model ants, then model ants update the so called Kanban information (*m, Rm, n, Am, {Bj}*). In this way, the search ant can circulate between different model ants to acquire the current stage product model by applying a local decision policy named *two stage transition rule*. The rule takes refer to the pheromone density information and two stage heuristic information displayed by model ants to compute two stage transition probability and help search ants select product models. The third phase is triggered after all search ants in the new colony have constructed their own sequences. In this phase central coordination ant employs a *local search rule* to find the best solution from afore generated search ant colonies. In the forth phase, central coordination ant applies *pheromone update rule* to change the pheromone onto the path between different model ants. The pheromone intensity would be decreased due to the *evaporation rule* to reflect the natural phenomenon so as to forget bad decisions ever made. The pheromone intensity would be increased on the basis of *pheromone award rule* to build an autocatalytic feedback process for awarding the search ant colony with good sequence. Until the termination is satisfied, repeat afore procedure. Ants coordinate with each other to find a near optimal sequence in the end.

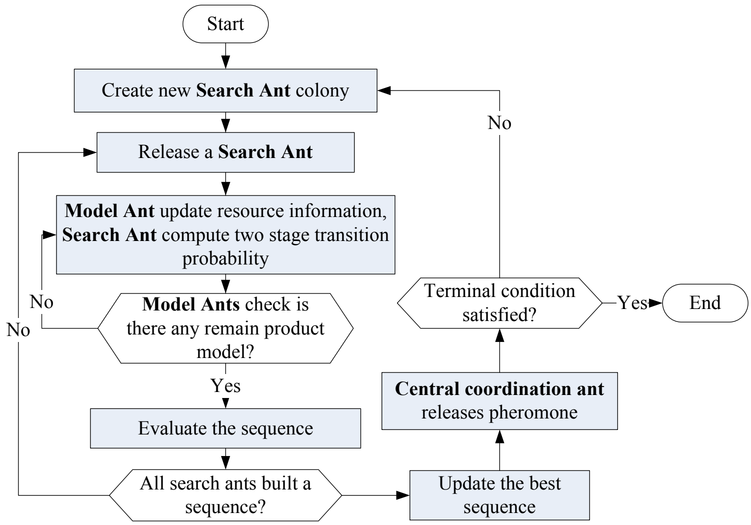


Figure 2. Outline of 2S-ACM

4.1 Two stage transition rule

The designated ants in 2S-ACM approach are individuals with artificial intelligence. When the search ants select product model pairs, they simulate the ant movement in the foraging behavior. Dorigo and Gambardella (1997) summarized a famous transition rule that directly guide the movement [11]. They take TSP (Tour Salesmen Problem) as an example to introduce the rule. An ant at city w chooses the next city which possessing the maximal probability. The probability is computed according to formula (3). Where, $\tau(w,u)$ is the amount of pheromone trail on edge (w,u). $\eta(w,u)$ is the inverse of heuristic information (such as the distance value) between city w and u. α is a parameter that allows a user to control the relative importance of pheromone trail, β is a parameter that determines the relative importance of heuristic information. $candidate(w)$ is the set of available cities to be selected by ant in city w.

$$\Pr o_{wv} = \begin{cases} \dfrac{[\tau(w,v)]^{\alpha} \cdot [\eta(w,v)]^{\beta}}{\sum_{u \in candidate(w)} ([\tau(w,u)]^{\alpha} \cdot [\eta(w,u)]^{\beta})}, & v \in candidate(w) \\ 0, & others \end{cases} \quad (3)$$

The two stage transition rule follows the principle of formula (3). Suppose w is the last stage product model in the $sequence$ of search ant's local memory. As mentioned in section 2, formula (2) defines the transformed two stage heuristic value which effects selecting a pair of product models (i^*, j^*). Thus, the two stage heuristic information is defined as formula (4). The two stage pheromone considers about the pheromone intensity between model ants w, i, j. The two stage pheromone computing formula is defined as (5). Then we developed the two stage transition rule as the formula (6). Other parameters such as α and β play the same role as in (3), $candidate(w)$ is the available models set.

$$\eta_{(i,j)} = 1/(2A_i + B_j + Q_{ij}) \quad (4)$$

$$\tau[w,(i,j)] = \tau(w,i^*) + \tau(w,j^*) \quad (5)$$

$$\Pr o_{w,(i^*,j^*)} = \begin{cases} \dfrac{\tau[w,(i^*,j^*)]^{\alpha} \cdot [\eta_{(i^*,j^*)}]^{\beta}}{\sum_{(i,j) \in candidate(w)} (\tau[w,(i,j)]^{\alpha} \cdot [\eta_{(i,j)}]^{\beta})} & (i^*,j^*) \in candidate(w) \\ 0, \quad others & w \in (1,2,\ldots\ldots M) \end{cases} \quad (6)$$

In this way, the search ants transit though different model ants follow a two stage rule.

4.2 Local search rule

The coordination ant implements local search rule to find the sequence with minimal material usage rates variation as soon as the search ant colony construct their sequences. In this paper, Taboo Search (TS) is adopted by the central coordination ant. Formula (1) computing material usage rate variation is used to evaluate the sequences obtained by search ants.

TS is a memory based approach that can avoid being trapped in local optimum by allowing non-improving moves and can prevent cycling back to previous visited solution. To generate new "best so far" solution, the TS is modified and used. First, a new set of feasible solutions is generated by ants to form a neighborhood; second, the best one of these solutions is selected to compare with the "best so far" solution. If the best candidate is the better one, a replacement operation is performed; otherwise, the best candidate is tabooed and recorded in the taboo list. Contrary to available TS, the modified TS is easier for the unnecessary stopping criteria is canceled.

4.3 Pheromone update rule

Since the local search is completed, the central coordination ant cooperating with model ants to perform the pheromone update rule. The updating of the pheromone intensity between model ants contains two steps, the first step is to perform evaporation rule to reflect the natural phenomenon so as to decrease the effects of the path on the sequence search procedure. The second step increase the pheromone intensity between model ants on the basis of pheromone award rule to build an autocatalytic feedback process for awarding the search ant colony with good sequence.

4.3.1 Evaporation rule

A portion of pheromone on all paths between model ants is evaporated according to the equation (7). Where, $\rho(0 \le \rho \le 1)$ is the pheromone evaporating coefficient.

$$\tau(i,j) = (1-\rho)\cdot\tau(i,j) \quad (i,j)\in(1,2,\cdots,M) \tag{7}$$

4.3.2 Pheromone award rule

The central coordination ant controls pheromone awarding procedure between model ants. Pheromone award rule is applied after the evaporation rule is completed. It firstly provides basic refreshment to the pheromone trails in graph models based on the ordinal number of result solutions. The award progress is performed as equation (8). Where, Q is the basic pheromone increase coefficient, V_{SAnum} is the material usage variation value of sequence that constructed by the search ant with number $SAnum$.

$$\tau(i,j) = \tau(i,j) + Q(V_{SAnum})^{-1} \quad (i,j)\in(1,2,\cdots,M) \tag{8}$$

Second, an elitist strategy is performed by the central coordination ant to award the "best so far" sequence obtained in the local search procedure. Elitist strategy has long been considered as an effective tool to preserve the best information. The principle of elitist strategy is to ensure that the information carried by the best performing members of the colony can be emphasized. In this note, the central coordination ant award the "best so far" sequence follows equation (9). Where, Q_{best} is the extra pheromone increase coefficient.

$$\tau(i,j) = \tau(i,j) + Q_{best} \quad (i,j)\in \text{best so far sequence} \tag{9}$$

5 Experiments

Computational experiments are conducted to evaluate the effectiveness of the 2S-ACM approach on balanced material consumption variation. Specifically, the experiment is to determine the computational advantage of 2S-ACM approach compared with transferred two stage heuristic, and GCA approach, the GCA approach chooses the model with the smallest one stage variation. The computational experiment was coded in C# and run on a Pentium 2GHz PC.

At the beginning, some information about parameter analysis of 2S-ACM approach is briefly introduced. The analysis process contains two phases: separated test and combinational test. Firstly, the separated test aims to determine the tendency for the values of parameters and then select the candidate values for each parameter. Secondly, combinatorial test is done to decide the best set of values suitable for different sized problems. After completing the tests, the best sets of parameters obtained in computational experiments are set as the default value of parameters. The parameters for test problems are set as: $c = 10, \alpha = 0.5, \beta = 1, \rho = 0.2, Q = 1000, Q_{best} = 100, NC = 50$.

Several problem combinations are randomly generated and solved by three sequencing methods. For each combination, a set of product structures is randomly generated by requiring a random quantity ranging from 0 to 9 for each part. Ten demand patterns are generated for each combination with the demand of each model following a uniform distribution within [0, 40]. The results of the problems are given in Table 1.

From Table 1, it can be seen that the 2S-ACM approach obtains the better results GCA or Transferred two stage method in optimal performance except, in a few case, round off errors as well as tie breaking caused the average solution values to be slightly different. The two stage methods performed better than the one stage GCA in material usage rate variation. The two stage method using direct enumeration turns out to be very time consuming for large problems. The CPU time requirement of the 2S-ACM is in the same magnitude as that of the other two methods, and these CPU times are relatively small, thus it is in an acceptable range in this paper.

Table 1. Material usage variation and CPU time (in milliseconds) of the tested problems

Problem	GCA		Transferred two stage method		2S-ACM	
	Variation	CPU time	Variation	CPU time	Variation	CPU time
1	**12927**	0.377	15428	0.82	41645	0.307
2	4877	0.282	3781	0.582	**2873**	0.142
3	3693	0.302	3121	0.577	**2576**	0.115
4	**1362**	0.243	1476	0.516	1796	0.064
5	986	0.228	**930**	0.48	986	0.454

6 Conclusions

For a manufacturer that applies JIT production, properly using and sequencing mixed models to achieve evenly material usage rate can result in a smoother production environment at all levels of production activities. In this paper, a 2S-ACM was developed to tackle the MMAL sequencing problem. The procedure minimizes the two stage material usage rate variation. Computational comparison showed that the artificial intelligent approach generally has an advantage over the heuristic rules in optimal performance of minimizing material usage rate variation. The 2S-ACM approach can be a useful tool to obtain near optimal solutions, which reduce the material usage rates variation in MMAL,.

Acknowledgements

This research is supported by National High Technology Research and Development Program of China (863 Program, No.2007AA04Z109), National Science foundation of China (No.50875172) and State Key Laboratory of Digital manufacturing equipment and technology of Huazhong University of Science and Technology, Wuhan, China.

7 References

[1] MONDEN, Y. Toyota Production System. Norcross: Industrial and Management press, 1993.

[2] MILTENBURG, J.; SINNAMON, G. Scheduling Mixed-model Multi-level Just-in-time Production Systems. International Journal of Production Research, vol. 27, pp. 1487-1509, 1989.

[3] BAUTISTA, J.; COMPANYS, R.; COROMINAS, A. Heuristics and Exact Algorithms for Solving the Monden Problem. European Journal of Operational Research, vol. 88, pp. 101-113, 1996.

[4] KUBIAK, W.; STEINER, G.; YEOMAN, J.S. Optimal Level Schedules for Mixed-Model Multi-level Just-in-time Assembly Systems. Annals of Operations Research, vol. 69, pp. 241-259. 1997.

[5] DING, F.Y.; ZHU, J.; et al. Comparing Two Weighted Approaches for Sequencing Mixed-model Assembly Lines with Multiple Objectives. International Journal of Production Economics, vol. 102, no. 1, pp. 108-131, 2006.

[6] RAHIMI-VAHED, A. R.; RABBANI, M.; et al. A Multi-objective Scatter Search for a Mixed-model Assembly Line Sequencing Problem. Advanced Engineering Informatics, vol. 21, no. 1, pp. 85-99, 2007.

[7] KUBIAK, W. Minimizing Variation of Production Rates in Just-in-time Systems: A Survey. European Journal of Operational Research, vol. 66, pp. 259-271, 1993.

[8] ZHU, J.; DING, F.Y. A Transformed Two-stage Method for Reducing the Part-Usage Variation and a Comparison of the Product-level and Part-level Solutions in Sequencing Mixed-model Assembly Lines. European Journal of Operational Research, vol. 127, no. 1, pp. 203-216, 2000.

[9] BOYSEN, N.; FLIEDNER, M.; SCOLL, A. The Product Rate Variation Problem and Its Relevance in Real World Mixed-model Assembly Lines." European Journal of Operational Research, vol.197, no. 2, pp. 818-824, 2009.

[10] MILTENBURG, J.; SINNAMON, G. Revisiting the Mixed-model Multi-level Just-in-time Scheduling Problem. International Journal of Production Research, vol. 33, pp. 2049－2052, 1995.

[11] DORIGO, M.; GAMBARDELLA, L. M. Ant Colonies for the Travelling Salesman Problem. Biosystems, vol. 43, no. 2, pp.73-81, 1997.

[12] ZHU, Q.; CHEN, X. F.; ZHANG, J. Multi-agent based Dynamic Job Shop Scheduling Method. Shanghai Jiaotong Daxue Xuebao, vol. 42, no. 7, pp. 1046-1050, 2008.

Product Design

A Study of Design by Customers: Areas of Application

Risdiyono and Pisut Koomsap[1]

Industrial and Manufacturing Engineering, Asian Institute of Technology, Thailand.

Abstract. In order to serve specific individual customers' personal needs which could not be fulfilled by mass production nor by mass customization , a Design by Customer concept has been recently introduced by researchers. Even though some applications seem to be visible from manufacturing point of view, the customer requirements may not be the same. A research to investigate the application area of design by customer concept from customer view is presented in this paper. The Kano model of customer satisfaction was used to gather and to analyze voice of customers. Customer satisfaction coefficients were generated based on the questionnaire. In addition, the importance level of design by customer system's capabilities were also investigated. The study concluded that the design by customer concept is considered to be applicable to some specific products like gifts and souvenirs but not in general. It also predicted that the concept will be successfully applied when the uniqueness is considered as an important feature of the products. However, deeper investigation is needed to get some more supports.

Keywords. Design by customer, collaborative design, customer satisfaction, Kano method.

1 Introduction

It is arguable that the only way for businesses to win in the marketplace is to adopt a customer driven strategy which delivers products or services to meet or exceed customer expectations [1]. Therefore, identifying customers, segmenting them, prioritizing them, knowing their main expectations and needs are all key marketing activities. This has forced producers to change their product design concept from manufacturer-oriented to customer-oriented [2].

Although has included the voice of customer in product development, traditional mass-produced product providers offer "take all or leave it" option for the customer since they use voice of majority as a great consideration. As a result, this approach cannot accommodate individual or specific customer preferences. This leads to fundamental changes from mass production to mass customization concept [3]. A popular way of product customization is by configuration design,

[1] Corresponding Author : Industrial and Manufacturing Engineering, School of Engineering and Technology, Asian Institute of Technology, Pathumthani 12120, Thailand, Tel: +66-2-5245678; Fax: +66-2-5245697; E-mail: pisut@ait.ac.th; http://www.ait.ac.th

where customers can choose different components and assemble them together to form a product. Product customization faces a problem when the diversity of customer needs is quite wide. Exponential increase of variety will occur and causes high cost and long lead time. This leads to the new concept of design by customer in which customers are no longer only searching for goods which satisfy the requirements such as quality and function, but they can also involve in making their own design [4]. This fundamental change of product design concept is shown in Figure 1.

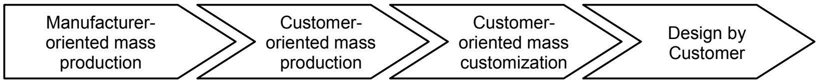


Figure 1. Fundamental changes of product design concepts

Even though some applications seem to be visible from manufacturing point of view, the customer requirements may not be the same. Hence, a research to investigate the application area of design by customer concept from customer view is needed. Customer perseptions about products' attractiveness and their feeling on their own capability to design may lead to the type of products which give possibility for this concept to be applied on.

This paper mainly aims to investigate the area in which design by customer concept will succeed when it is applied and to explore its important features to consider. However, in order to build a same perception about the meaning of 'design' in design by customer concept and to observe the recent design environments, a brief literature review is presented.

2 Literature Review

Noble and Kumar [5] considered that design can be classified into three categories i.e. Utilitarian Design, Kinesthetic Design and Visual Design.

Utilitarian design focuses on the practical benefits a product may provide. This approach attempts to achieve functional differentiation through making products that simply work better in very tangible ways, including effectiveness, reliability, durability, safety and to other competitive advantages relative to other offerings like multi functionality and modular product architecture.

A kinesthetic design emphasizes how a user physically interacts with the product. One interesting aspect of this strategy is the ability to potentially enhance both functional differentiation and emotional value. There are several tactics a firm can pursue to enhance the kinesthetic of their goods. Ergonomics is probably the best-understood concept in this group.

Visual Design is probably the closest element aligned with what design means to most observers. Visual design is driven by form, color, size and the desire to communicate value to the consumer without necessarily interacting with the product. Visual design is primarily focused on the generation of emotional value.

Therefore, this design category is what the word ‘design’ means in the concept of design by customer.

Design is considered as a collaborative process and outcome rather than as a single profession in this work. For most, the process of designing includes other members of their profession and members of other professions [6]. Designers from different disciplines need to communicate and interact with one another. It is assumed that problems are decomposed into sub-problems and that each team member focuses on a particular portion of the sub-problem (Figure 2). One of the key roles of an integrated modeling system is to mediate the information-flow between the participants [7].

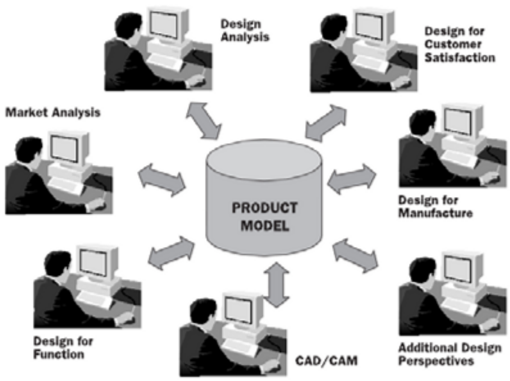


Figure 2. Multi-discipline extended design team [7]

Internet-based collaborative product design and virtual design spaces are significantly got more attention from many researchers. Shyamsundar and Gadh [8] for example, proposed a new geometric representation in order to overcome the problem with the limited bandwidth of the Internet transfer of detailed CAD based assembly models so that the designers will be able to perform real-time geometric modification, assembly constraints specification and concurrent design of different components/sub-assemblies. This will enable designers to collaborate effectively during design. This issue is becoming more popular as a collaborative environment often includes bringing people together through the Internet or corporate Intranet for face-to-face communication even though those people are located at different places [9,10].

The implication of collaborative design in computer technology (including CAD/CAM) is also becoming an important issue and interesting area of research. For designing a promising product, there is always a need for collaboration among the design, marketing, finance and procurement departments, and the top management. Global manufacturing makes it difficult to frequently gather all the departments in a meeting room to discuss, because of geographical constraints. In order to address this issue, recently, a number of software tools and research works have arisen to provide collaborative solutions [11]. Pohl et al. [12] suggested collaborative and distributive as one principle in computer-aided design systems for the 21st century. The necessity of collaborative CAD system comes from an established fact that the use of computers in design and manufacturing is the key to substantial productivity gain. Future manufacturing organizations will definitely be information oriented, knowledge driven and most of their daily operations will be automated around the global information network that connects everyone together [10]. The development of the next generation of collaborative CAD systems supported by digital multimedia communication tool will undoubtedly bring enormous advantages in design, drafting and machining. The Internet, incorporating computers and multimedia, has tremendous potential for remote integration and collaboration in CAD/CAM applications [13].

Since not all members of collaborative design team have CAD/CAM software and due to the limitation of data transfer technology of internet, Liu et al. [9] utilized the combination of VRML, JavaScript and HTML to make a web-based environment collaborative design system that allows the manufacturer and their clients to work collaboratively on the product design and to improve the design in less time with less cost. By using this web-based collaborative environment, engineers and designers can participate in design reviews; marketers can see the appearance of a new product on store shelves; and salesperson can give their feedbacks on possible consumer acceptance of new products. There is, however, still a lot of research on internet collaborative design, especially in the area of large-scale project management and security [14]. In addition, the concept of customer-oriented virtual factory system (Figure 3) has also been under investigation [15].

Based on the foregoing, the feasibility to involve customers in the design process (design by customers) is quite clear. The supports from internet based collaborative design as well as CAD/CAM, rapid manufacturing and information technology are considered as main factors should be ready before applying the concept. On the other hand, selecting type of products and their requirements based on the voice of customers is also an important thing to focus on. Therefore, this study is an important measure to analyze this new concept from customer view.

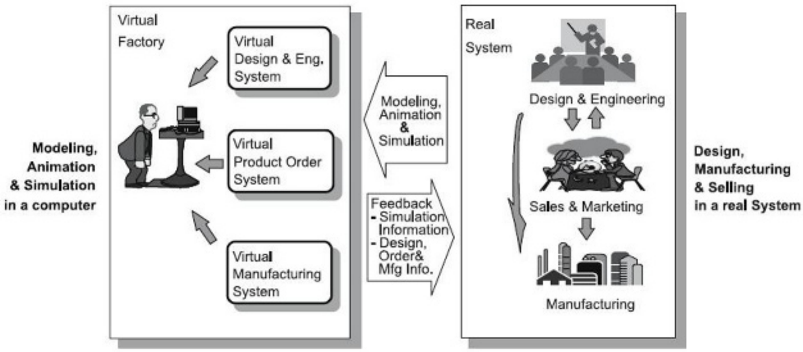


Figure 3. The concept of virtual factory [15]

3 Methodology

In order to achieve the objective of this study, the Kano model of customer satisfaction was used to gather voice of customers (Kano questionnaire) and to analyze the data (Kano evaluation table). This method is recommended by many researchers when manufacturer want to get customer responses about the product requirements or when a new feature will be introduced. It captures the relationship between product performance and customer satisfaction. The Kano model is constructed through customer surveys, where a customer questionnaire contains a set of question pairs for each and every product attribute. The question pair includes a functional form question, which captures the customers' response if a product has a certain attribute, and a dysfunctional form question, which captures the customers' response if the product does not have that attribute. The questionnaire is deployed to a number of customers, and each answer pair is aligned with the Kano evaluation table, revealing an individual customer's perception of a product attribute. The final classification of a product attribute is made based on a statistical analysis of the survey results of all respondents.

The survey incorporated 127 respondents, mostly students of Asian Institute of Technology Thailand, coming from Indonesia, Thailand, Vietnam, Bhutan, Pakistan and Bangladesh. Customer satisfaction coefficients were generated based on the questionnaire. There are 5 pairs of questions constructed in order to investigate customer satisfaction to 3 design criteria i.e. uniqueness, modification ability and design by customer option and also to 4 kinds of products including general products, gifts, souvenirs and jewelry. In addition, the importance level of design by customer system's capabilities including the option to modify the shape, to change the color, to modify the size, to add identities, to use own design and to order using online system (internet) are also investigated.

4 Results and Discussion

The Kano model of customer satisfaction is a useful tool to classify and prioritize customer needs based on how they affect customer's satisfaction. It captures the relationship between product performance and customer satisfaction. In practice, four types of product attributes are identified:

(1) Must-be (basic) attributes are expected by the customers and they lead to extreme customer dissatisfaction if they are absent or poorly satisfied,
(2) One-dimensional (performance) attributes are those for which better fulfillment leads to linear increment of customer satisfaction,
(3) Attractive (excitement) attributes are usually unexpected by the customers and can result in great satisfaction if they are available, and
(4) Indifferent attributes are those that the customer is not interested in the level of their performance.

By using Kano method, customer satisfaction coefficient (CS coefficient) can be calculated from qualitative value. The CS coefficient states whether satisfaction can be increased by meeting a product requirement, or whether fulfilling this product requirement merely prevents the customer from being dissatisfied. Different market segments usually have different needs and expectations so sometimes it is not clear whether a certain product feature can be assigned to the various categories. It is especially important to know the average impact of a product requirement on the satisfaction of all the customers. The CS coefficient is an indicator of how strongly a product feature may influence satisfaction or, in case of its ”non-fulfillment” customer dissatisfaction. CS coefficient can be calculated by:

- $$\text{Satisfaction} = \frac{A+O}{A+O+M+I} \qquad (1)$$
- $$\text{Dissatisfaction} = \frac{O+M}{(A+O+M+I)\times(-1)} \qquad (2)$$

where A = attractive, O = one dimensional, M = must be, I = indifferent

The resume of Kano evaluation table is presented in table 1 while the customer satisfaction coefficient (CS coefficient) is shown in figure 4. By using simple classification of table 1, uniqueness of general products is categorized as indiferent design criteria in design by customer system (I=40%). However, the value of attractiveness is quite high (A=34.29%) compared to the value of one dimension criteria (O=17.14%). Hence, when CS coefficient is calculated (figure 4), the level of satisfaction (if this feature is available) is higher than the level of dissatisfaction (if this feature is absent). Based on the CS coefficient plotted in the diagram, it is reasonable to clasify the uniqueness as attractive design criteria.

When the uniqueness of a specific product i.e. gift is investigated, it is found that both Kano evaluation table and CS coefficient diagram classify it as an atrractive design criteria with a very high level of satisfaction. This finding leads to a logical thought that analyzing the design criteria of a specific product is preferable than of a general one since customers will be very articulate in giving their opinion about a specific product they are familiar with. Uniqueness is a significant selling point for some products but not for all kind of products. Unique gifts and souvenirs, for example, are very attractive in respondents’ opinion, but for some other products which are intended for their functionality (for example: machinery), unique design may have less attractiveness than their performance.

In general, modifying the product is still considered as indifferent design criteria by both Kano evaluation table and CS coefficient diagram. But, when specific products (souvenir and jewelry) are investigated, it is found that own design of souvenir is categorized as attractive while designing jewelry by customer is considered as indifferent. This verifies the previous finding that analyzing a specific product is preferable than the general ones. In addition, it can be assumed that for jewelry products, customers prefer to have a famous designer design rather than their own design. It is arguable considering the fact that jewelry is an esteem need in which a good brand has higher social value and is preferable than a unique design.

Table 1. Resume of Kano Evaluation Table

Product Requirement	A	O	M	I	R	Q	Total	Category
Uniqueness (general)	34.29%	17.14%	5.71%	**40.00%**	0.00%	2.86%	100%	I
Modification (general)	23.53%	5.88%	5.88%	**61.76%**	0.00%	2.94%	100%	I
Uniqueness (gift)	**61.76%**	2.94%	0.00%	26.47%	2.94%	5.88%	100%	A
Design by Cust. (jewelry)	18.18%	3.03%	6.06%	**69.70%**	0.00%	3.03%	100%	I
Design by Cust. (souvenir)	**44.12%**	8.82%	2.94%	41.18%	0.00%	2.94%	100%	A

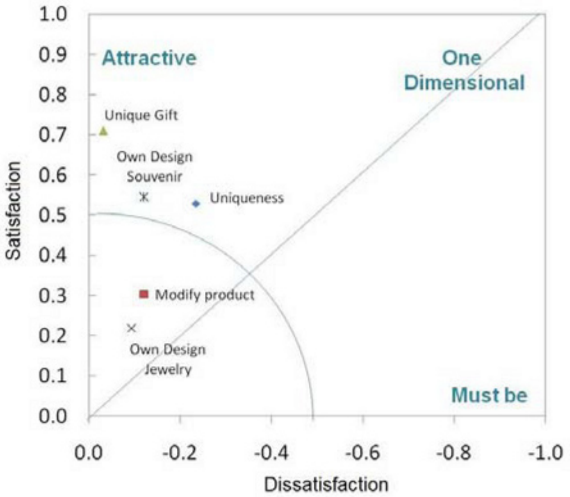


Figure 4. Customer satisfaction coefficient (CS coefficient)

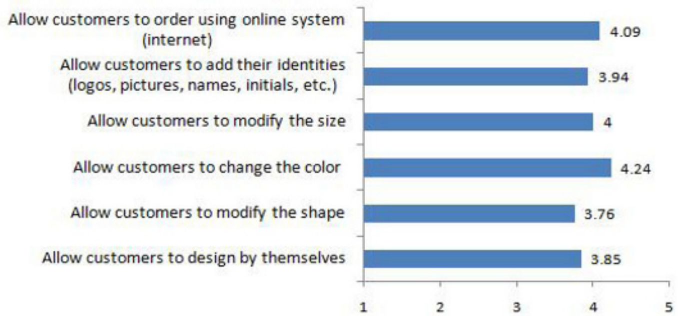


Figure 5. Importance level of design by customer system's capabilities

Figure 5 shows the importance level of design by customer system's capabilities including the option to modify the shape, to change the color, to modify the size, to add identities, to use own design and to order using online system (internet). The figure shows that all features are considered important. The system's capability that allows customer to modify the color and shape of product as well as to order by online system seems to be the most important features to consider. These are not big problems since, as discussed in literature review, the future manufacturing organizations will definitely be information oriented, knowledge driven and most of their daily operations will be automated around the global information network that connects everyone together.

5 Conclusion

The study concludes that the design by customer concept is considered to be applicable to some specific products like gifts and souvenirs but not in general. It also predicts that the concept will be successfully applied when the uniqueness is considered as an important feature of the products. The system's capabilities that allow customer to modify the color and shape of product as well as to order by online system seem to be the most important features to consider. However, deeper investigation is needed to get some more supports. Furthermore, establishing a fundamental framework to support the real world applications of design by customer concept will be the near future work.

6 References

[1] Wallace TF. Customer-Driven Strategy. Winning through Operational Excellence, Oliver Wight Publications, Inc., 1992.

[2] Anderson DM. Design for manufacturability. CIM Press, Lafayette, California, 1990.

[3] Tseng MM, Jiao J, Merchant ME. Design for mass customization. CIRP Annals - Manufacturing Technology 1996; 45(1): 153-156.

[4] Risdiyono, Koomsap P. Toward a Design by Customers for Automotive Accessories. In: Proceedings of the Third National Conference on Automotive Engineering, Bangkok, 2007.

[5] Noble CH, Kumar M. Using product design strategically to create deeper consumer connections. Business Horizons 2008; 51(5): 441– 450.

[6] Pahng GDF, Bae S, Wallace D. Web-based collaborative design modeling and decision support. In: Proceedings of DETC ASME Design Engineering Technical Conferences, Atlanta, Georgia, 1998.

[7] Kvan T. Collaborative design: what is it?. Automation in Construction 2000; 9(4): 409-415.

[8] Shyamsundar N, Gadh R. Internet-based collaborative product design with assembly features and virtual design spaces. Computer-Aided Design 2001; 33: 637-651.

[9] Liu PC-H, Teng SG, Tsai H-C. A collaborative product design environment for small to medium size companies. In: Proceedings of the 2002 IE Research Conference, 2002.

[10] Tay FEH, Roy A. CyberCAD: a collaborative approach in 3D-CAD technology in a multimedia-supported environment. Computers in Industry 2003; 52(2): 127–145.

[11] Fuh JYH, Li WD. Advances in collaborative CAD: the-state-of-the art. Computer-Aided Design 2005; 37(5): 571–581.

[12] Pohl J, Chapman A, Pohl KJ. Computer-Aided Design Systems for the 21st Century: Some Design Guidelines. In: Proceedings of the fifth International Conference on Design and Decision, 2000.
[13] Wang HF, Zhang YL. CAD/CAM integrated system in collaborative development environment. Robotics and Computer Integrated Manufacturing 2002; 18: 135–145.
[14] Li WD, Lu WF, Fuh JYH, Wong YS. Collaborative computer-aided design–research and development status. Computer-Aided Design 2005; 37: 931–940.
[15] Yang SM, Ahn B, Seo K. Development of a prototype customer-oriented virtual factory system. International Journal of Advanced Manufacturing Technology 2006; 28(9-10): 1031–1037.

Dual Lines Extraction for Identifying Single Line Drawing from Paper-Based Overtraced Freehand Sketch

Natthavika Chansri and Pisut Koomsap[1]

Industrial and Manufacturing Engineering, Asian Institute of Technology, Thailand.

Abstract. Freehand sketch on a piece of paper is a quick rough drawing for portraying ideas, and commonly used during a conversation to elaborate an explanation. Overlapping lines provide additional information to make the rough sketch clearer, but the overtraced sketch should be simplified first to a single line drawing to be useful for downstream processes (e.g., input for generating 2D tool paths or for constructing 3D model). Presented in this paper is a new approach for identifying a single line drawing from a paper-based overtraced freehand sketch. Key activities in this approach are thick line sketch creation, boundary extraction, contour expansion and shrink, and line drawing creation. The current stage of its implementation is also reported in this paper.

Keywords. Paper-based overtraced freehand sketches, single line drawing.

1 Introduction

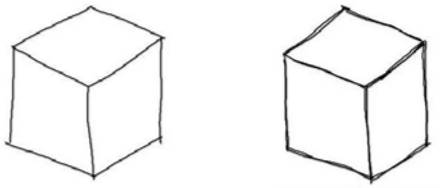
Despite of living in today's digital world, pencil and paper remains a popular tool when it comes to express ideas. According to experiment conducted by Won [1], sketchers felt more comfortable to describe ideas naturally on a paper than on computer. A simple quick freehand sketch can help elaborate and make ideas easier to be understood [2, 3]. Typically, a designer starts with a single line sketch, also known as non-overtraced strokes. Unless these single lines unambiguously represent ideas, additional lines are drawn repeatedly over the existing lines to provide additional information to make the sketch clearer. This is known as multistroke or overtraced stroke [3, 4]. Both types of sketches are illustrated in Figure 1.

Although it is simple and convenient, the extended application of paper-based sketch is limited compared with online sketch because only available information

[1] Corresponding Author : Industrial and Manufacturing Engineering, School of Engineering and Technology, Asian of Institute of Technology, Pathumthani 12120, Thailand, Tel: +66-2-5245678: Fax: +66-2-5245697; Email: pisut@ait.ac.th; http://www.ait.ac.th

is a batch of data points representing lines [5-7]. For an online sketch, in addition to line information, stroke information is also available. By knowing their starting points and endpoints, all strokes can be identified. As a result, an online sketch is more often used with other applications such as 2D model identification [4, 8-10], 2D face identification [11-14] and 3D model reconstruction [10, 12, 15-21].

Presented in this paper is an attempt to identify a single line drawing from a paper-based overtraced freehand sketch in order to extend its applications. This is a step toward our main research on direct fabrication of a physical prototype from an overtraced freehand sketch that commands for rapid prototyping will be generated directly from an overtraced sketch without reconstruction a 3D CAD model. This direct interface will allow designer-customer realize and experience their ideas rapidly. The next section is the literature review of the types of sketch commonly found in practice, and of line drawing identification. Proposed approach and its current stage of implementation are presented in sections 3 and 4 respectively. The conclusion and future work are addressed in the last section.



(a) (b)

Figure 1. The types of sketches (a) a single stroke sketch and (b) an overtraced stroke sketch

2 Related Work

Two basic media used for portraying ideas are digital-based sketch and paper-based sketch. As aforementioned, the online nature of digital sketch allows data points belonging to strokes be recognized separately. This simplifies line interpretation. Groups of data to be considered are determined from strokes of interest. The data on each stroke are arranged in order but they become unordered for multiple overlapping strokes. The process for interpreting line is similar for both, but thinning process is required additionally for unordered points before identifying line. Curvature analysis [22], conic section equation [8], angular and distance test [17], linearity test [9] and principal component analysis [23] have been applied to determine type of lines whether they are straight lines, elliptic arcs, parabolic arcs or corners.

On the other hand, a batch of points is only available information for paper-based sketch. Few researchers developed methods for identifying line drawing from non-overtraced freehand sketch. Marti et al. [5] proposed a method for extraction and interpretation of solid lines and hidden lines. For line extraction, isolated points were filtered out from a binarized image. This image was then gone through thinning process. Points on solid lines and hidden lines were detected by weighted K curvature method and graph-based method respectively. Those points were then merged and line labeling was applied for feature interpretation from those junctions and lines. Later on Farrugia et al. [6] proposed a framework for identifying line drawing, and demonstrated on symmetrical parts. The binarized image was thinned using skeletonization algorithm before obtaining straight lines and curves using Hough transform. Its corners were determined from the image by using chain code. Moreover, chain code and T-junction have been also applied to identify boundary and internal line drawings in thinned image and to detect corners respectively [7].

Our prior attempt to identify line drawing from overtraced freehand sketch was developed based on the assumption that point density is higher at corners than on edges [24]. The algorithm filled overtraced lines to form thick line drawing before filtering out the low density areas. The remaining high density areas were represented by their centroids that were paired and checked to form lines. There, however, were two main drawbacks in this approach: it was not guaranteed that all corners would always have higher point densities than edges; and some imaginary lines created during pairing were not marked on the original image, although the representative points were found.

3 Methodology

Illustrated in Figure 2 is our new approach for identifying line drawing from paper-based overtraced freehand sketch. Similar to our previous approach, the overtraced lines are combined to form a thick solid line sketch but instead of seeking for corners to identify lines, the new approach tries to represent the thick line sketch with its center lines. This can be done by extracting the boundary of the thick line sketch which always gives a set of closed contours, representing exterior and interior. By expanding internal contours and shrinking an external contour simultaneously, these contours will meet their neighbors, but all meeting points may be obtained at different times depending upon the uniformity of line thickness on the sketch. The contours will all meet their neighbors simultaneously in case of uniform thick lines. In case of non-uniform thick lines, the contours around the narrowest sections will meet first. The obtained meeting points will be used to form a single line drawing of a sketch.

The input to this approach is a binary image of a sketch. For both scanned binary and grayscale images, image processing technique such as thresholding method is required to remove noise that normally exists before inputting. The steps for implementing this approach are presented in Figure 3.

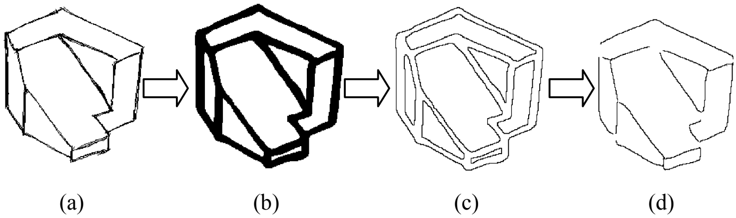


Figure 2. The proposed line drawing identification (a) an overtraced sketch, (b) a thick line sketch, (c) a dual line sketch, and (d) a line drawing

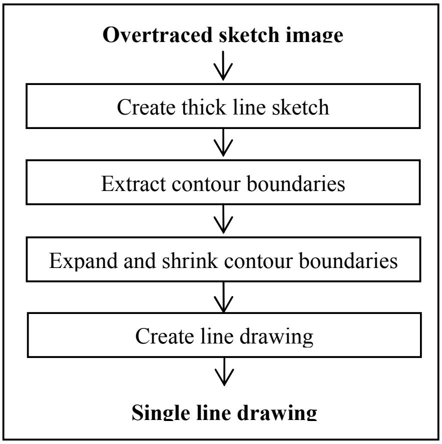


Figure 3. The overview of proposed line drawing identification

3.1 Thick Line Sketch Creation

A sliding window is applied for creating a thick line sketch from an overtraced sketch as shown in Figure 4. A square window with its size about the largest pixel gap in the sketch is slid along the image from left to right and top to bottom. At each position, the number of data points (black pixels) inside the window is determined, and compared with the minimum allowable number of data points inside the window, obtained from the window width. If the number of data points is greater than or equal to the threshold value, the center point of the window is drawn on a new image. The drawing is skipped, otherwise. By allowing overlapping between windows, these center points will form a thick line sketch. The obtained thick line sketch is sent to the next process for extracting contours.

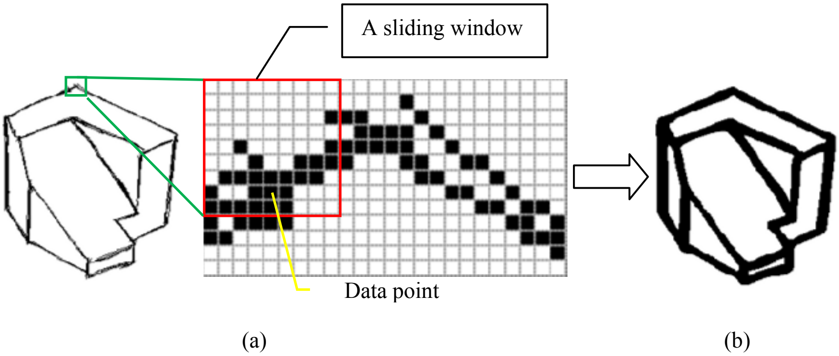


Figure 4. Creation o f a thick line sketch from an overtraced sketch
(a) a sliding window on an overtraced sketch image
(b) the thick line sketch image

3.2 Contour Boundary Extraction

Morphological boundary extraction, in which logic operations are applied to erode data points on an image with a predetermined structuring element, is applied to extract one pixel wide closed contours on the thick line sketch. A typical 3×3 pixel window is slid on the sketch image from left to right and top to bottom. The center point of the window is drawn on a new image, if it is a point on the sketch and surrounded by at most seven neighboring points. This operation results in a dual line sketch as shown in Figure 2c. A contour tracing algorithm proposed by Kuagoolkijgarn et al. [25] is then applied on the new image for obtaining ordered sequences of boundary pixels representing all closed contours.

3.3 Contour Boundary Expansion and Shrink

By expanding internal contours and shrinking an external contour simultaneously, these contours will meet their neighbors. The positions that they meet are used later for creating a single line drawing. For implementation, every two neighboring contours are paired as shown in Figure 5. The two contours are resized by one pixel wide every iteration. If the contours are odd pixels apart, they will be at the same position. If they are even pixels apart, they will not be at the same position but will be adjacent. The positions of either one of them are recorded. However, the relationships between any pairs of contours are not available. Therefore, their neighbors must be determined first, and this can be done by determining the shortest distance between the two contours.

For each point on the first contour, its shortest distance to the second contour is determined from the minimum value of its direct distances to all points on the second contour.

Let P and Q be a set of M and N data points on a pair of contours respectively. $P = \{(x_i, y_i) \mid i = 1, 2, \ldots, M\}$ and $Q = \{(x_j, y_j) \mid j = 1, 2, \ldots, N\}$.

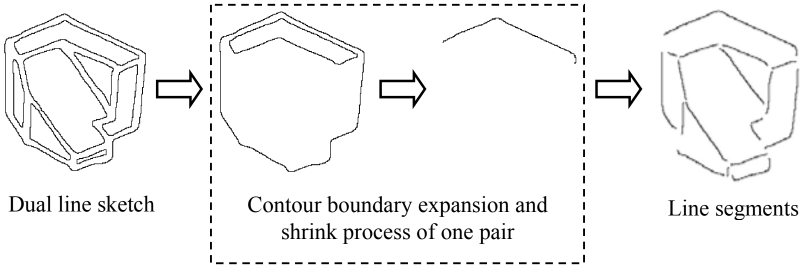


Figure 5. Contour boundary expansion and shrink

For any point i on contour P, its shortest distance to contour Q is given as

$$d_{i.} = \min\left(d_{ij}\right)$$

where $d_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}$ and the set of shortest distances from all points on contour P to contour Q is $D = \left\{d_{i.}, i = 1:M\right\}$.

The two contours are said to be neighbor, if there exists at least one pair of points whose shortest distance is shorter than or equal to the maximum allowable distance between neighboring contours (L_a). The set of qualified shortest distances is represented as $L = \left\{L_k \mid L_k \le L_a, \forall k\right\}$, and $L \subseteq D$.

After a pair of contours is identified as neighbor, they are resized. The meeting points start to appear from the pairs of points that have the minimum shortest distance, and only the pairs of points whose shortest distances are shorter than the maximum allowable distance qualify. They will all appear after the qualified pairs with longest distance meet. The wide span from the first iteration the meeting points appearing to the last iteration may result in a line segment with more than one pixel wide. Therefore to achieve a one pixel wide line segment, the span should be controlled within one standard deviation of qualified pairs. A stopping condition for expansion and shrink can then be described as

$$\textit{Stopping condition} = \left\lceil 0.5\left(\bar{L} + S\right)\right\rceil$$

where $\bar{L}$ and S are the average and standard deviation of the qualified shortest distances.

3.4 Line Drawing Creation

The sets of meeting points obtained from the previous step are used in this step to form one pixel wide line segments. When its data points are clearly recognized, parametric curve fitting, i.e. linear equation and conic section equation, can be applied for identifying each individual line segment as being done in online non-overtraced sketch. The discrete obtained line segments are then connected. This is done by calculating shortest distance from an end point to be connected to another end point that belongs to different line segment. The pair with the shortest distance is linked.

4 Implementation of Line Drawing Identification

This proposed approach is currently under the implementation. Its program has been developed on LabVIEW software. Figure 6 shows the current user interface screen of the program. The input of this program is a thick line sketch that is obtained from another program developed for our previous approach. It is displayed on the left side of the screen. The middle of the screen shows the result of contour extraction, and displayed on the right side of the screen is all the meeting points obtained from expansion and shrink operation. In this particular example, the integration of these meeting points gives a single line drawing.

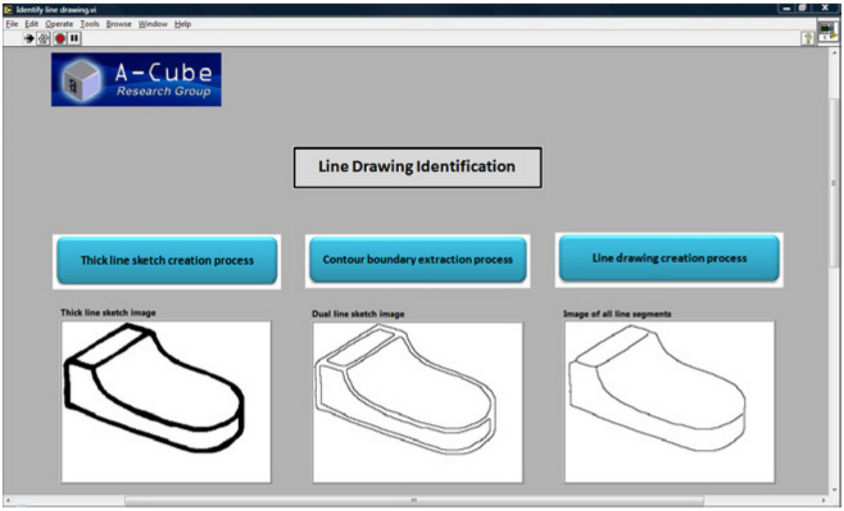


Figure 6. The user interface screen of a program

5 Conclusion and Future Work

This paper presents a new approach for identifying line drawing from paper-based overtraced freehand sketch. Instead of seeking for corners to identify lines, the new approach tries to represent the thick line sketch with their center lines. Shrink and expansion are performed on all pairs of neighboring closed contours, extracted from the sketch, to determine meeting points which lead to a discrete one pixel wide line segments. The program has been developed to this stage. Curve fitting and connecting will be incorporated in this approach to achieve a one pixel wide line drawing.

6 References

[1] Won PH. The comparison between visual thinking using computer and conventional media in the concept generation stages of design. Automation in Construction 2001; 10: 319-325.

[2] Plimmer B, Apperley M (2002) Computer aided sketching to capture preliminary design. In: Proceedings of the 3rd Australasian User Interfaces Conference (AUIC2002), Melbourne, 2002; 7: 9-12.

[3] Lim S, Qin SF, Prieto P, Wright D, Shackleton J. A study of sketching behavior to support free-form surface modeling from online sketching. Design Studies 2004; 25: 393-413.

[4] Sezgin TW, Davis R. (2004) Handling overtraced strokes in hand-drawn sketches. In: Proceedings of the AAAI Spring Symposium Series: Making Pen-Based Interaction Intelligent and Natural, Washington DC, 2004; 141-144.

[5] Marti E, Regions J, Lopez-Krahe J, Villanueva JJ. Line drawing interpretation as polyhedral objects to man-machine interaction in CAD systems. Pattern Recognition and Image Analysis 1995; 1-12.

[6] Farrugia PJ, Borg JC, Camilleri KP, Spiteri C, Bartolo A. A cameraphone-based approach for the generation of 3D models from paper sketches. EUROGRAPHICS Workshop on Sketch-Based Interfaces and Modeling 2004; 33-42.

[7] Haron H, Shamsuddin SM, Mohamed D. A New corner Detection Algorithm for Chain Code Representation. International Journal of Computer Mathematics 2005; 82(2): 941-950.

[8] Shpitalni M, Lipson H. Classification of sketch strokes and corner detection using conic sections and adaptive clustering. Journal of Mechanical Design 1995; 119: 131-135.

[9] Ku DC, Qin SF, Wright DK. (2006) Interpretation of ovetracing freehand sketching for geometric shapes. In: Proceedings of WSCG' 2006 Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Plzen, Czech Republic, 2006; 263-270.

[10] Shesh A, Chen B. SMARTPAPER: an interactive and user friendly sketching system. In: Proceedings of Eurographics 2004, 2004.

[11] Shpitalni M, Lipson H. Identification of faces in a 2D line drawing projection of a wireframe object. IEEE Transactions on Pattern Analysis and Machine Intelligence 1996; 18(10): 1000-1012.

[12] Oh BS, Kim CH. Progressive reconstruction of 3D objects form a single freehand line drawing. Computers & Graphics 2003; 27: 581-592.

[13] Liu JZ, Lee YT. A graph-based method for face identification from a single 2D line drawing. IEEE Transactions on Pattern Analysis and Machine Intelligence 2001; 23(10): 1106-1119.

[14] Liu JZ, Lee YT, Cham WK. Identifying faces in a 2D line drawing representing a manifold object. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002; 24(12): 1579-1593.

[15] Lipson H, Shpitalni M. A new interfaced for conceptual design based on object reconstruction form a single freehand sketch. Annals of the CIRP 1995; 44: 133-136.

[16] Lipson H, Shpitalni M. Optimization based reconstruction of a 3D object from a single freehand line drawing. Computer Aided Design 1996; 28(8): 651-663.

[17] Matsuda K, Sugishita S, Xu Z, Kondo K, Sato H. Freehand sketch system for 3D geometric modeling. IEEE 1997; 55-62.

[18] Company P, Contero M, Conesa J, Piquer A. An optimization-based reconstruction engine for 3D modeling by sketching. Computers & Graphics 2004; 28: 955-979.

[19] Marsy M, Kang D, Lipson H. A freehand sketching interface for progressive construction of 3D objects. Computers & Graphics 2005; 29: 563-575.

[20] Kara L, Stahovich T. An image-based trainable symbol recognizer for hand drawn sketches. Computers and Graphics 2005; 29(4): 501-517.

[21] Kara L, D'Eramo C, Shimada K. Pen-based styling design of 3D geometry using concept sketches and template models. In: Proceedings of ACM Solid and Physical Modeling Conference (SPM' 06), 2006.

[22] Jenkins DL, Martin RR. Applying constraints to enforce users' intentions in freehand 2D sketches. Intelligent System Engineering 1992; 31-49.

[23] Pusch R, Samavati F, Nasri A, Wyvill B. Improving the sketch-based interface: forming curves from many small strokes. In: Proceedings of Computer Graphics International (CGI 2007), 2007.

[24] Chansri N, Koomsap P. (2008) Identifying single line drawing from paper-based overtraced freehand sketch. In: Proceedings of 9th Asia Pacific Industrial Engineering and Management Systems Conference, Bali, 2008.

[25] Kuagoolkijgarn P, Koomsap P, Chansri N. A new algorithm for tracing nests of interconnected contours. Working paper to be submitted to Pattern Recognition.

A Formal Representation of Technical Systems

Baiquan Yan and Yong Zeng[1]

Institure for Information Systems Engineering – Concordia University, Canada

Abstract. The concept of technical systems is the foundation of many existing design theories and methodologies, such as systematic design methodology and TRIZ. A formal representation of technical systems is proposed to support the more logical application of the existing methodologies to various practical problems. This formal representation is based on axiomatic theory of design modelling. It shows that all the components of a technical system can be effectively represented in a formal system.

Keywords. Technical systems, formal representation, axiomatic theory of design modelling, systematic design methodology

1 Introduction

The knowledge about technical system (TS) and technical process (TP) is an area of design science [1]. It deals with the object of designing: technical system to be designed. In [1], Hubka and Eder have described a transformation system, in which TS, TP, and therefore their relations play a major role. Of all the knowledge about TS and TP, the representation of TS is the foundation of other aspects. An intuitive explanation about the importance of the representation is immediate. First, the representation tells us how to use the system; secondly, the representation is a tool in aiding design activity and the development of new design methods. Hubka once asked in [1], in which manner can technical systems be modeled and represented in all their states of existence, so that the resulting models support different functions of designing? This question can be interpreted as: what should be represented and what are the requirements for a good representation? First of all, the representation should include all elements included in technical system. Secondly, and more importantly, the representation should correspond to different abstract levels in order to support different function of designing.

Considering that a product design is divided into conceptual design, embodiment design, and detail design, we can correspondingly classify representation into three abstract levels, namely, detailed level, layout level, and topological level. In detailed level, the representation is a detailed description about a product, such as dimensions, surface finish as in mechanical engineering.

[1] Corresponding Author Email: zeng@ciise.concordia.ca

In layout level, the representation of a system has the information about its subsystems and their relationships. In topological level, the representation is just a concept that is described with symbols or text. Obviously, the first two levels are domain dependent; the third one is very generic, applicable to any technical system. The research presented in this paper deals with the third levels, aiming to provide a representation fit for all technical systems.

Technical system (TS) is a finite set of related technical objects. The theory about TS plays an important role in helping to develop new design theories. The term technical system was introduced to emphasize the most important characteristic of technical objects [1]: any technical object belongs to a system. Form the designer's point of view, technical system is the object that designers intend to generate. The whole knowledge about technical system forms Theory of Technical System (TTS). In understanding technical systems, the following aspects have been studied [2].

1) the purpose of technical systems
2) the structure of technical systems
3) the properties of technical systems
4) the states of technical systems
5) the working principles of technical systems
6) the input and output of technical systems
7) the environment of technical systems
8) the representation of technical systems
9) the classification of technical systems

All technical systems are designed for some purposes: automobiles are designed for transporting people or goods, machine tools for making other products, nuclear power stations for generating electricity, etc. The structure of a technical system is the spatial relationships between the elements. In order for a system to work well, all constituents of the system must be correctly integrated or assembled. Properties are the characteristics that the system holds and that differentiate this system from other systems. There are two types of system properties [3]: pragmatic and physical properties. Pragmatic properties are goal-oriented, such as suitability for manufacture and transportation, heat-insulating ability, stability, and correction resistance. Physical properties include characteristics and constants of substance and fields. A state of the system is the total set of all constants, fixed parameters, and measured parameter values of all properties of the subsystem at a given time. State decides the system's behaviors; the working principles ascribe how a system works; and environment is where the system works. Because the number of technical systems is infinite, many classification methods have been presented, among which the widely accepted one is the classification according to the principles of action [3]. So, we have mechanical systems, electrical systems, hydraulic systems, pneumatic systems, etc. This classification, however, creates confusion in many cases, because many modern systems are hybrids.

A technical system is a complex product that may be made up of other simple products. The representation of technical system has two main goals: first, from engineering point of view, a good representation of TS is the prerequisite for manufacturing and using it; secondly, from science point of view, how to describe a technical system is the foundation for any design theory. Nobel Laureate Simon in his famous book *The sciences of the Artificial* once wrote "to use it (a complex system), to achieve the simplification, we must find the right representation[4]." By reviewing the existing representation approaches such as graphic and set-based ones, we have observed that they are able to achieve the first goal, but may fail to achieve the second one. In other words, they are not the right representation by Simon's standard. Based on this observation, this paper presents a formal representation of technical systems by using the axiomatic theory of design modelling. Such a formal representation intends to achieve the two aforementioned goals.

Technical process is any artificial single action or consequences of procedures to perform an activity with assistance of a technical system or a natural object [3]. It is a set of transformations of input elements into products: respecting constraints, requiring resources, and fulfilling some desired purpose [2]. Whether natural or technical, all process must follow the natural laws. The only difference between natural process and technical process is that the latter is controllable. Technical process can be accelerated, slowed down, or stopped according to different purposes. All technical processes must be achieved through corresponding technical systems. It is interesting to notice that there is no one-one relationships between technical systems and technical processes. In many cases, one TP can be achieved with different TS's; and one system may perform different processes. For example, the combustion process can be regarded as both natural and technical process. If we want to use this process for our special purpose, we can design different technical systems, such as an engine or an oven.

2 Axiomatic Theory of Design Modeling

Axiomatic theory of design modelling (ATDM) is a logical tool for representing and reasoning about object structures [5]. It provides a formal approach that allows for the development of design theories following logical steps based on mathematical concepts and axioms. The primitive concepts of universe, object, and relation are used in the axiomatic theory of design modelling, based on which two axioms are defined.

[Axiom1] Everything in the universe is an object.

[Axiom2] There are relations between objects.

Structure operation is developed in the axiomatic theory of design modeling to model the structure of complex objects. Structure operation, denoted by the symbol $\oplus$, is defined by the union ($\cup$) of an object and the interaction ($\otimes$) of the object with itself.

$$\oplus O = O \cup (O \otimes O), \tag{1}$$

where $\oplus O$ is the structure of an object O. Both the union and interaction are specific relations between objects. Since the object O may include other objects, Eq. (1) indeed implies a recursive representation of an object. The following simple derivations clearly demonstrate this recursive nature. Suppose that an object O is a compound object composed of subobjects O_1 and O_2, we have:

$$\begin{aligned}
O &= O_1 \cup O_2 \\
\oplus O &= \oplus(O_1 \cup O_2) \\
&= [O_1 \cup O_2] \cup [(O_1 \cup O_2) \otimes (O_1 \cup O_2)] \\
&= O_1 \cup O_2 \cup (O_1 \otimes O_1) \cup (O_2 \otimes O_2) \cup (O_1 \otimes O_2) \cup (O_2 \otimes O_1) \\
&= \oplus O_1 \cup \oplus O_2 \cup O_1 \otimes O_2) \cup (O_2 \otimes O_1).
\end{aligned} \tag{2}$$

This shows the recursive nature of operator $\oplus$.

3 Formal Representation of Technical Systems

3.1 Requirements for a Good Representation

In Section 1, we have briefly discussed all aspects of a technical system. In discussing the demands for a representation, Hubka has argued that models (representations) should serve all purpose for designing, such as communication, information, experiment, calculation, thinking aid; support methodical and systematic procedure, and the applications of computers; guarantee uniqueness of interpretation and efficient reading; and consider the efficiency and economy of modeling and representation processes [1]. Very similarly, we observe: a representation must include all the information about the system to be studied including purpose, behavior, structure, relations between system members, properties, states, input and output, types, and environment; and moreover, a good representation approach must reflect the nature of design process and the logic of design, and therefore to help the development of design theories. As Zeng in [6, 7] has pointed out that the design process is of recursive nature and the design activities follow recursive logic. Product design usually needs a series of processes or stages. Accordingly, the representation should be able to be applied to all those stages. Concretely, a representation should meet the following requirements:

(1) Completeness: all information about the system must be included in the representation.
(2) Clarity: a clear and simple representations help understand the system; meanwhile clarity must imply uniqueness of interpretation.
(3) Independence: the representation does not depend on any technical fields.
(4) Flexibility: the representation should be able to describe systems in different design stages or in different states.
(5) Adaptability: the representation is open to updating the system

These five requirements show different priorities: the first three suggest basic requirements, by meeting which a representation approach can achieve the first goal, i.e. to help engineers understand the system; the last two present advanced requirements, by meeting which the representation can achieve the second goal, i.e. to lay a foundation for developing new design theories.

3.2 Existing Representations

Graphic representation is the most intuitive and simplest method in describing a technical system. A commonly used graphic method is hierarchical tree representation. It clearly shows the hierarchic characteristic of complex system, as well as the components in the system. But the relationships between components are not clear. Moreover, it is difficult for graphic representation to satisfy the flexibility and adaptability requirements, about which a detailed discussion will be given later in this paper.

Another well-accepted method is Set-theoretic Representation. In the literature, this method was used by Yoshikawa in the General Design Theory [8], Braha *et al* in their Formal Design Theory [9], and Salustri and Venter in their Axiomatic Theory of Engineering Design Information [10]. Simply speaking, by using set theory, a system can be described in the form $S=\{C_1, C_2, \ldots, C_n\}$, where C_i is a component of *S*. C_i can be a compound object that consists of other sub-components, or is an atomic object that cannot be decomposed into any sub-objects. However, like graphical representation, set representation only tells the constituent components included in a system, but has difficulty in representing the aggregation process, in which a system can be built up from sub-systems and atomic objects.

Both representations-graphic representation and set-theoretic representation satisfy the first three requirements, and therefore the first goal of representing a system, to help people understand, manufacture, and use the system, can be achieved with all of them. But, they cannot achieve the last two requirements because both of them cannot clearly express relations between different elements. Therefore, we propose our new formal representation based on the Axiomatic Theory of Design Modeling in Section 3.2. In section 4, we will exam this new representation against the five requirements by using a case study. In the last section, we conclude the present research and propose future direction along this line.

3.3 Formal Representation

3.3.1 Description of technical systems

In terms of ATDM, everything in the universe is an object, and any object is related to other objects; a technical system can be regarded as an object that contains a number of other objects, which are subsystems in the conventional system language. Using the structure operator, as was given in Eq. (1), defined in ADTM, we can write a system as:

$$\oplus S = S \cup (S \otimes S). \tag{3}$$

Without loss of generality, suppose that *S* is composed of n subsystems, some of which may be only atomic objects. We have:

$$S = \bigcup_{i=1}^{n} S_i \tag{4}$$

Therefore, the system *S* can be represented as:

$$\oplus S = \oplus(\bigcup_{i=1}^{n} S_i) = \bigcup_{i=1}^{n} S_i \cup \bigcup_{i,j=1}^{n} (S_i \otimes S_j) = \bigcup_{i=1}^{n} (\oplus S_i) \cup \bigcup_{\substack{i,j=1 \\ i \neq j}}^{n} (S_i \otimes S_j) \quad (5)$$

Equation (5) shows that the structure of a technical system can be represented with its subsystems and the relationships between them.

Also, it should be noted that physically $S_i \otimes S_j$ is not necessarily the same as $S_j \otimes S_i$. Let R be a relation that contains all possible relations encountered in all technical systems. Let $r_{ij} = S_i \otimes S_j$. Then we can define r_{ij} as,

$$r_{ij} = \begin{cases} r \subset R & \text{if } S_i \text{ is diretly related to } S_j, \text{or } i = j \\ \Phi & \text{otherwise} \end{cases} \quad (6)$$

3.3.2 Environment of a technical system

In the previous subsection, we focus only on the description of systems themselves. However, environment plays an important role for the system to work well. By applying ATDM, we define:

[Definition] The environment of a system is everything except the system itself.

We can graphically describe this in Figure 1. Any system works under a certain environment. They interact with each other.

Figure 1. Technical system and its environment.

3.3.3 Relations in a technical system

Since relation is an indispensable part of any systems, we need to further discuss the relations. A relation is the way how two objects are related. In a technical system, this relation ascribes how two objects are structurally and functionally connected. The system exerts a requirement for the type of relation, and in turn the nature of the relation influences the performance of a system. There are a huge number of relation types. It is a challenging task to classify them, and yet it is crucial for describing and constructing a system. In this paper, we classify relations into several types as follows.

- Force relations: two objects are related by force. For example, two mating gears are related by force relations.
- Geometrical relations: two objects are related mainly from geometry point of view, for example an O-ring and a shaft in mechanical engineering.
- Movement relations: one object moves with respect to another object.
- Information relations: two components have information exchange. These relations are most often seen in the field of information engineering.
- Energy or field relations: two objects are not physically contacted, but related through a field. For example, in a motor, the rotor and the magnets.

- Fixation relation: two objects are permanently fixed together.

In the process of designing a system, the identification of relations and the design of structures are performed recursively. The details will be discussed in another paper.

It is noted that this list is far from exhaustive. In fact, the classification of relations in a very abstract level together with formal representation is a crucial part for creating design concepts, and is a future research topic in design science.

3.3.4 Completeness of the formal representation

As discussed in the previous sections, a representation of a technical system must include the following information about the system: purpose, behavior, structure, relations between system members, properties, states, input and output, types, and environment. Completeness requires that a representation embrace all these aspects. Based on the discussion in [Hubka], we give the definitions as follows:

[Purpose]: the purpose of a technical system is what we design for or what the system is used for.

[Structure]: structure is the set of elements in a system and the set of relations that connect these elements.

[Behavior]: a behavior is a set of successively attained states of a system. It includes predicable and unpredictable behaviors. The predicable is the desired ones; the unpredictable, in most cases, are not desired.

[Environment]: the environment of a system is everything except the system.

[Input]: the input represents the external relationship from environment to system.

[Output]: the output represents the external relationship from system to environment

[Property]: a property is any characteristic or quality possessed by an arbitrary object.

[State]: the state is the total of the measures of all properties of a system at a given time.

Now, we will exam the formal representation against all these aspects. A typical technical-environment system is shown in **Figure 1**. Let Ω stand for the whole system, T, technical system, E, environment, and then we have:

$$\oplus\Omega = \oplus(T \cup E) = \oplus E \cup \oplus T \cup (T \otimes E) \cup (E \otimes T). \tag{7}$$

[Note]: E is the environment that can be categorized into three parts [11]: natural environment, built environment, and human environment. The relation $E \otimes T$ is from environment to the system, so it stands for the input of the system. The relation $T \otimes E$ is from the system to environment, so it stands for the output of the system. Now, we consider $\oplus T$. Suppose that T is composed of n elements T_1, $T_2, \ldots, T_k$, we have:

$$\oplus T = \oplus\left(\bigcup_{i=1}^{n} T_i\right) = \bigcup_{i=1}^{n} (\oplus T_i) \cup \bigcup_{\substack{i,j=1 \\ i \neq j}}^{n} (T_i \otimes T_j). \tag{8}$$

Obviously, $\bigcup_{i=1}^{n}(\oplus T_i)$ is the elements of the system, and $\bigcup_{\substack{i,j=1 \\ i\neq j}}^{n}(T_i \otimes T_j)$ stands for all relations between any two elements; therefore, $\bigcup_{i=1}^{n}(\oplus T_i)$ and $\bigcup_{\substack{i,j=1 \\ i\neq j}}^{n}(T_i \otimes T_j)$ form the structure of the system. As Hubka has pointed out in [2] that structure decides the behaviors. Therefore, Eq. (8) includes behavior information. In a similar way, we can prove that Eq. (7) includes information about properties and therefore states of a system.

The discussion above shows that the formal representation satisfies the completeness requirement. In addition, the form and logic underlying this representation are simple and independent of any specific system, which makes this representation satisfy the requirements for clarity and independence. In section 4, we will demonstrate the flexibility and adaptability using a case study.

4 Examples

It is shown in section 3 that the formal representation can be applied to describing technical systems as well as constructing systems. It also shows that this representation is complete, clear, and independent of any specific domain. In this section, we will demonstrate that this representation is not only able to clearly describe the structure of a technical system in different design stages but also able to represent any adaptations of a system.

4.1 A brief review of some fundamental concepts

Structure of Design Problem: A design problem is implied in a product system and composed of three parts: the environment in which the designed product is expected to work, the requirements on product structure, and the requirements on performances of the designed product.
Design process: The design of a product is a recursive and evolution process [6, 12], during which the structure and relationships are continuously changing. Figure 2 graphically explains this process.

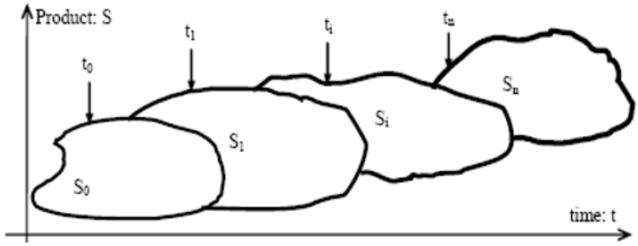


Figure 2. Evolution of product in the design process.

In the mean time, the requirement also is changing. As a result, the state of design is also changing with time. Similarly, we have

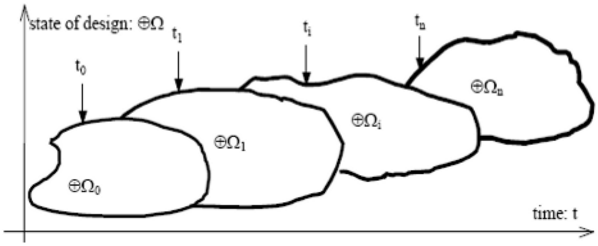


Figure 3. Evolution of the design process: refined.

Any product design, especially complex technical system designs, undergo this evolution process indicated in Figure 2, 3. During the design process, the product description must be evaluated against the prescribed design requirements to determine if the designed product satisfies the requirements. And, as product descriptions include all the results generated in the dynamic design process, the representation scheme must imply different levels of product descriptions [13].

4.2 Case study

4.2.1 Description of design problem

Design a hydraulic system used for moving a worktable left to right, and return after reaching at desired position. To make the design problem clearly understood, one possible final system to be designed is illustrated as in Figure 4.

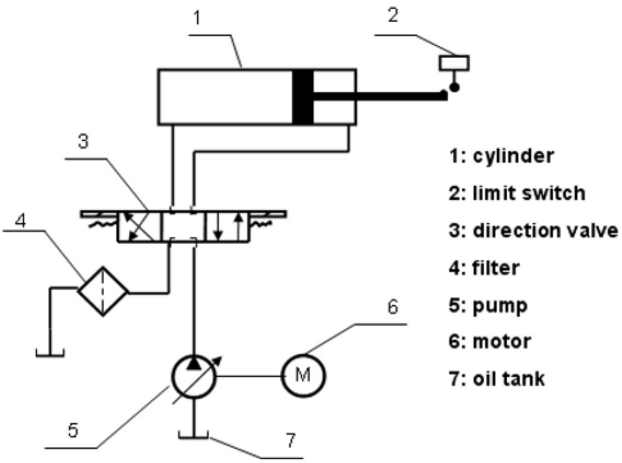


Figure 4. A hydraulic system.

4.2.2 Description of design process:

[Definition]: primitive products are those products that cannot be decomposed into other products, or that need not to be decomposed with respect to the system. For example, when we design a hydraulic system, we do not need to know the detail structure of a valve, so in this case, this valve is a primitive product.

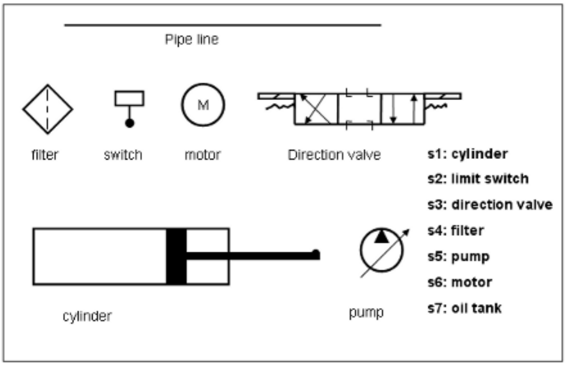


Figure 5. Primitive products.

Figure 5 lists a set of possible primitive products; surely there are other design solutions and primitive products. How to generate a design concept is beyond the discussion of this present paper, please refer to [14], or other papers. This example just demonstrates how to represent system structure in different design process.

At the very beginning, the system $S = S_0 = \Phi$. Φ is an empty set.

Step 1: $S = S_0 \cup s_1$, we obtain a partial system as shown in Figure 6. and,

$$\oplus S = \oplus(S_0 \cup s_1) = \oplus s_1 \quad (9)$$

Figure 6. Partial system 1 $S = S_0 \cup s_1$.

Step 2: $S = S_0 \cup s_1 \cup s_2$ **we obtain a partial system as shown in** Figure 7, **and,**

$$\oplus S = \oplus(S_0 \cup s_1 \cup s_2) = \oplus s_1 \cup \oplus s_2 \cup (s_1 \otimes s_2) \cup (s_2 \otimes s_1) \quad (10)$$

In this special case, we can write $(s_1 \otimes s_2) = (s_2 \otimes s_1) = r_{12}$, so, Eq. (11) can be written as:

$$\oplus S = \oplus(S_0 \cup s_1 \cup s_2) = \oplus s_1 \cup \oplus s_2 \cup r_{21} \quad (11)$$

Figure 7. Partial system 2, $S = S_0 \cup s_1 \cup s_2$.

Step 3: $S = S_0 \cup s_1 \cup s_2 \cup s_3$, we obtain the third partial system shown in Figure 8. And the representation is

$$\oplus S = \oplus(S_0 \cup s_1 \cup s_2 \cup s_3) = \oplus s_1 \cup \oplus s_2 \cup \oplus s_3 \cup r_{21} \cup_{13} \quad (12)$$

where $(s_3 \otimes s_1) = (s_1 \otimes s_3) = r_{13}$ and $(s_3 \otimes s_2) = (s_2 \otimes s_3) = 0$

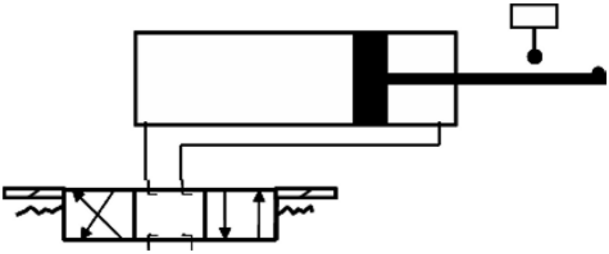


Figure 8. Partial system 3 $S = S_0 \cup s_1 \cup s_2 \cup s_3$.

In a similar way, we can step by step complete the design of the whole system, which is shown in Figure 5. And the formal representation is as in Eq. (14).

$$\begin{aligned} \oplus S &= \oplus(S_0 \cup s_1 \cup s_2 \cup s_3 \cup s_4 \cup s_5 \cup s_6 \cup s_7) \\ &= \oplus s_1 \cup \oplus s_2 \cup \oplus s_3 \cup \oplus s_4 \cup \oplus s_5 \cup \oplus s_6 \cup \oplus s_7 \\ &\quad \cup r_{12} \cup r_{13} \cup r_{34} \cup r_{35} \cup r_{56} \cup r_{57} \end{aligned} \quad (13)$$

Where, $(s_3 \otimes s_4) = (s_4 \otimes s_3) = r_{34}$ $(s_3 \otimes s_5) = (s_5 \otimes s_3) = r_{35}$
$(s_5 \otimes s_6) = (s_6 \otimes s_5) = r_{56}$, $(s_5 \otimes s_7) = (s_7 \otimes s_5) = r_{57}$

This design example clearly shows that the formal representation explained in Section 3 can describe technical systems in different design stages. From Figure 2 to Figure 8, we can see this representation indeed reveals the recursive and evolution nature of design process. Consequently, the formal representation scheme imply different levels of product descriptions, and therefore satisfy the last two requirements: flexibility and updatability.

5 Concluding Remarks

This paper presents a formal representation for describing technical systems by using axiomatic theory of design modeling. Based on Hubka's observations [1] and the two goals of representing a system, five requirements for a good representation are also proposed. Both theoretical derivation and case study show that this formal representation satisfies the five requirements. It encompasses all necessary information about a technical system, reflects the recursive nature of design process, and presents a platform of representing systems in different design stages. As a result, this representation not only applies to describing existing technical systems, but also provides a foundation for developing theories in the field of design science.

A unique characteristic of this representation lies in the application of the two axioms about objects and of structure operator ⊕. Based on Axiom 1, relations also are objects. A natural conclusion follows: the elements of a system and the relations between them are mathematically in the same abstract level. This enables the structure operation in different design stages, and therefore the formal representation thus derived can be applied to the description of technical systems during the design process.

This paper also tries to classify all possible relations existing in technical system, but this classification is not a comprehensive one. The lack of a clear classification of relations combining with a formal representation of system still remains a bottle neck in computer-aided concept design. A lot of work is needed along this line.

6 Acknowledgments

This work is partially supported by NSERC (Grant number RGPIN 298255).

7 References

[1] Hubka, V. and Eder, W., Design Science. 1996, London: Springer-Verlag.
[2] Hubka, V. and Eder, W., Theory of Technical Systems. 1984, Berlin: Springer-Verlag.
[3] Savransky, S., Engineering of Creativity. 2000, Boca Raton, Florida: CRC Press.
[4] Simon, H., The sciences of the Artificial. Third ed. 1996, Cambridge, Massachusetts: The MIT Press.
[5] Zeng, Y. Axiomatic theory of design modeling. Transaction of SDPS: Journal of Integrated Design and Process Science 2002; 6(3): pp. 1-28.
[6] Zeng, Y. and Cheng, G. On the logic of design. Design Studies 1991; 12(3): pp. 137-141.
[7] Zeng, Y. and Gu, P. General Design Governing Equation. in NSFC Grantee's Conference on Design and Manufacturing Engineering. 1999. Vancouver, BC, Canada.
[8] Yoshikawa, H., General Design Theory and a CAD System. 1981: North-Holland Publishing Company.
[9] Braha, D. and Maimon, O., A Mathematical Theory of Design: Foundations, Algorithms and Applications. 1998: Kluwer Academic Publishers, The Netherlands.
[10] Salustri, F. and Venter, R. An Axiomatic Theory of Engineering Design Information. Engineering with Computers 1992; 8(4): pp. 197-211.
[11] Morris, H., et al. Information Integration Framework For Product Life-Cycle Management of Diverse Data. Journal of Computing and Information Science in Engineering 2004; 4(4): pp. 352-358.
[12] Zeng, Y. and Gu, P. A science-based approach to product design theory Part I: Formulation and formalization of design process. Robotics and Computer-Integrated Manufacturing 1999; 15(4): pp. 331-339.
[13] Zeng, Y. and Gu, P. A science-based approach to product design theory Part II: formulation of design requirements and products. Robotics and Computer-Integrated Manufacturing 1999; 15: pp. 341-352.
[14] Chen, Z. Y. and Zeng, Y. Classification of Product Requirements Based on Product Environment. Concurrent Engineering 2006; 14(3): pp. 219-230.

Design Knowledge Assets Management with Visual Design Progress and Evaluation

Gundong Francis Pahng[a,1] and Mathew Wall[b]

[a] Zionex, Inc., Korea.
[b] Oculus Technologies Corporation, USA.

Abstract. Product design and development involves knowledge-intensive activities, tasks, and communications. As the market becomes more demand or consumer driven, the use of information technologies, computer-based tools and methods, and networking capabilities is essential to expedite the progress and obtain the desired results on time. Although these enabling technologies and solutions have improved the quality and speed of product development processes, there remain many inefficiencies due to ineffective use and sharing of design knowledge and errors in communication. This paper identifies the typical issues in design knowledge sharing in collaborative design environments and introduces a computer-based solution to enhance collaboration and knowledge sharing. The solution is focused on the timely use of design knowledge assets, including design documents, specifications, test results, reports, reviews, memos, meeting minutes, and the like.

Keywords. Design knowledge management, design collaboration, product development process innovation, progress management.

1 Introduction

This paper identifies opportunities to enhance the product development process by enabling effective management of design knowledge assets and facilitating communication relevant to evaluations and decisions on alternative designs. A new software solution was introduced to facilitate the capture of project assets, decisions, and communication.

2 Product design and development process

Every product design and development project has phases. Although general phases are often described, detailed activities and tasks depend on the product type. Products can be categorized generally as innovative, market-driven, platform and customized [1].

[1] Corresponding Author E-mail: francis_pahng@zionex.com

During the product design and development process, many organizational and/or functional stakeholders must communicate and interact. Since each stakeholder has preferences and/or requirements on the expected values, functions and performances of the product, extensive communications and decision making amongst the stakeholders are essential for a project to be successful. To integrate the decision-making process among heterogeneous and distributed stakeholders, the use of information technologies and networking capabilities has become a critical capability [2].

2.1 Design Alternatives and Associated Knowledge Assets

During the life of a project, many design alternatives may be proposed, analyzed, and compared to one another. Based on individual and/or group decision criteria decisions are made for a certain design to pursue further. Each alternative typically results in branches of development and the generation of assets for design, analysis, and evaluation. Some of the branches are deemed infeasible and die. Others produce useful concepts that are incorporated into other branches. Figure 1 illustrates the project phases augmented with branches of concept development.

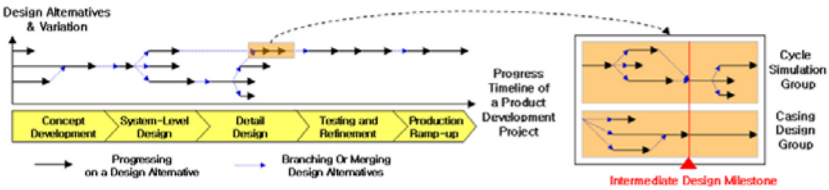


Figure 1. Hierarchical Design Alternatives and Variations

As depicted in Figure 1, design alternatives are often considered in parallel, branched into sub alternatives, and merged. Design alternatives and design decisions may also be hierarchic. The significance of each design alternative and associated analysis and decisions may change over time. It is important to capture, manage, share and reuse this design knowledge asset as well as to manage the history and progress of design alternatives. A single participant might work on a single branch, or multiple participants may share a single branch. There must be intermediate or formal milestones where the evaluations and decisions made locally are aggregated to higher levels and communicated to sibling levels, especially when there are dependencies. Such information sharing ensures that each stakeholder can monitor performance and catch miscommunications before they become bottlenecks.

2.2 Communications and Knowledge Sharing

Many systems are disconnected from individuals' work environment. For example, analysis results from CAE simulation may be stored locally or in a fileserver and the engineer reports the result to others via separate documents or emails. While a designer works on parts and assemblies using a CAD tool, he/she may summarize the design candidates and their comparisons in a document and sends it to colleagues for review. Upon a group design review a certain design alternative is selected for the further progress, and relevant reports and files may be submitted in a project management system for deliverables of a milestone. As a result, informal notes, documents, and engineering memos may not be transparent to colleagues or team managers.

Communication paths are often broken. In some cases, it is not clear who is responsible. In other cases, there is simply no communication channel for one team member to report an issue to another team member. In other cases, there is no mechanism for recording the communication so that the issues can be tracked and prioritized. If a system could assist designers to manage personal work environment and formal/informal work deliverables and simultaneously convey the individual evaluations and decisions seamlessly to the upper-level of communication, it could create a collaborative work and communication environment where each individual's working progress and knowledge can be shared and monitored. Furthermore, such a system could help product development teams to identify and analyze the value of individual and/or team's knowledge assets so that they can better manage and enhance their design knowledge [3].

3 Design Progress and Knowledge Assets Management

At each stage in a project, there exist many paths of activity, branches of development, and interactions amongst participants. Although many projects share similarities, each is unique. Even when the product is evolutionary rather than revolutionary, it is often difficult to predict the lifespan of design alternatives and the life cycle of communication and deliverables generated during the development process.

3.1 Types of Design Progress

The progress of product design and development has a lifecycle, in which design concepts are initially conceived, modified, branched, merged, and released as described in Table 1. In each step of this lifecycle decision making activities and tasks are conducted and design/knowledge assets will be utilized or generated.

Table 1. Progress Types in a Generic Product Design Project

Category	Key Characteristics
Initiation	- Initial design alternative(s) is defined. - Key evaluation measures are determined and defined.
Revision	- A design is modified in order to improve the evaluation measures, but the overall concept of the design remains unchanged.
Branching	- New design concepts are identified. - Further analysis and validation of the concepts are decided. - Tasks are planned for exploration of distinct design concepts.
Consolidation	- A design concept is merged into another concept.
Release	- A design reaches a level of definition to be shared or a milestone.

3.2 Carrots and Sticks

An ideal system provides value to every project participant. For example, a project manager needs accurate estimates and actuals on which to base a project plan. These data are typically collected by requesting estimates and periodic status reports from team members. However, estimates are often inaccurate and status often changes, so the overhead of collecting and maintaining accurate estimates and status frequently becomes a burden. The manager does not have enough information to ensure accuracy, and the team members do not have enough time to update the information.

Contrast this with a project in which estimates are drawn from historical performance, and status is drawn directly from the status of deliverables in a repository. As work progresses, both managers and participants can see the status on any component of the project. Meetings can focus on “what needs to be done?” rather than “Where are we?” Estimates can be based on the actual time to complete deliverables. This is straightforward for evolutionary work, and a growing organizational knowledge, backed up by data in the repository, makes it easier to forecast on revolutionary components of a project that have no direct historical precedent. Estimates are based upon fact rather than perception.

The first step to realizing such a system is to get team members to commit their work to a repository. The repository should contain all of the project assets, not just summaries. There are many incentives for individual participants:

- keep track of changes and go back in time to a previous configuration
- maintain multiple designs/concepts/releases and resolve conflicts
- track multiple projects and understand/communicate changes
- re-use with confidence

There are also incentives for project managers and system engineers:

- gather actuals and deliverables for planning/scheduling
- see where activity is and is not happening
- understand sources of conflict and expose bottlenecks
- facilitate audits
- determine and analyze root causes
- track data transfers to/from vendors and customers

Requirements, performance objectives, and tasks are first entered and provide the sketch that outlines the vision for the project. As the project progresses, assets are added, branches are made, and alternatives are explored, thus reinforcing parts of the sketch and erasing others. Tags and milestones, made explicit in the repository, are the ink that realizes the project goals.

3.3 Communication and design decision making

As the number of stakeholders and participants involved in a project increases, communication and effective decision making become a challenge. Product design and development involves a series of decision-making processes: identification of options, development of expectations on the outcomes of each choice, and formulation of a system of values for ordinally ranking the outcomes and thereby obtaining the preferred choices [4]. To facilitate these decision-making processes in multi-disciplinary design projects with multiple stakeholders in various viewpoints and communication levels, a probabilistic interpretation of design's acceptance is well suited [5]. In the probabilistic evaluation scheme, decision making can be associated with the three factors shown in Figure 2.

Both the visibility of performance indicators and the frequency with which they are updated are crucial to the success of a project. Indicators must be visible to those who can influence them, are being evaluated by them, and are evaluating them. They should be updated as close to real-time as possible. Weekly or monthly meetings might catch problems, but daily, hourly, or change-based updates to indicators can expose problems before they have a chance to become project-threatening. Requirements inevitably change over the life of a project, so the system for recording, quantifying, and displaying the metrics must be flexible. Finally, the metrics must measure things that matter.

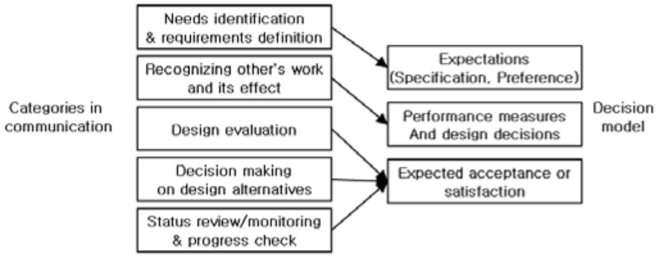


Figure 2. Association between Communication Types and Decision Models

This implies that the capturing and managing the decision model factors within the individual work environment and seamlessly projecting them onto the group and company level work environment could facilitate the communication in a project. Thus, the product design and development process can be greatly enhanced if a tool can provide an effective way to visualize and manage these decision model factors as well as keep track of the history of designs and decisions made while participants work on each task.

4 Revision Management

Most projects use a variety of tools to facility communication, process and capture. For example, a simple project might use telephone and email for communication, a wiki for collaborative document editing, and a shared file system in which to store assets. Each of these tools serves a specific purpose, and each has a significantly different user interface and experience. However, the tools are disconnected, and without end-user discipline, project knowledge often ends up being difficult to find, complicated to re-use, and increasingly fragmented.

This study used a Subversion repository as the primary storage mechanism for all project assets. The intent was to keep all project assets in the repository, with links to domain-specific tools and version control, then provide various context-appropriate interfaces to view and modify the repository contents.

4.1 History Matters

Why not use a database or shared file system? History is the key benefit of using a repository rather than a shared file system. A repository provides not only a hierarchy for where to locate an asset, but also a history for identifying when an asset was in a specific state. This combination of both where and when is incredibly powerful when applied to every project asset.

4.2 Single Point of Entry

Even if project data are located in different databases, file formats, or physical locations, providing a single entry point makes the data accessible to participants. In many cases project data are kept directly in the repository, even if their 'native' environment has its own notions of version control. In other cases, the repository contains versioned links to various states of project data in other locations.

4.3 Centralized vs Distributed

There are two major types of repositories: centralized and distributed. In the centralized model, each participant may work in a sandbox, but a connection to the repository is required for committing changes or getting changes from other participants. In the distributed model, each participant keeps a complete clone of the repository, thus enabling 'offline' commits and other operations. Merging happens when the participant connects with other participants or with the 'reference' repository. This study used a centralized model.

4.4 Per-File vs Aggregate Versioning and

Some revision management systems track changes to file contents. Others systems track changes to not only the contents of files but also directories. Most product development efforts involve a variety of files and file types, and the history of what files were created/destroyed and how they are related is just as important as the history of what the files contain.

4.5 Ease-of-Access

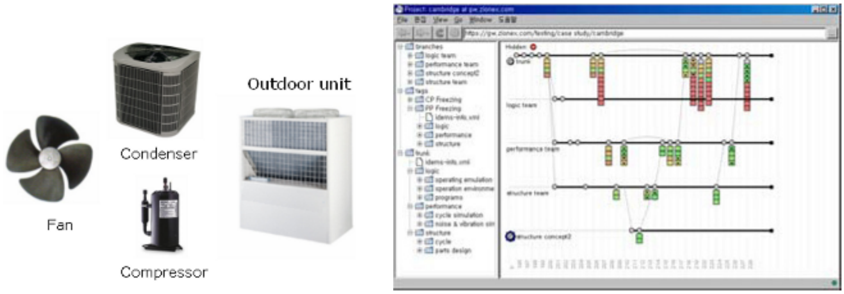
Experience has shown that the system must be accessible from a variety of contexts. A PDM system from a CAD vendor is inappropriate for housing a design document when it requires the owner of the design document to install/start the CAD software just to update the document. A purely web-based interface to a document storage system does not provide enough context for many product development activities. So an open, extensible, stable interface is necessary.

4.6 Subversion

Subversion was designed with these characteristics in mind [6]. Oculus' IDEMS Revision Manager software provides a graphic user interface to Subversion repositories. The graphic display of branches and changes correlates directly to the product development structure described in Section 2 [7].

5 A Case Study

This section describes a simplification of an ongoing pilot study that uses the Revision Manager in a product design project. Air conditioners are a typical electronics product which requires expertise from multiple disciplines to design. The example case of this paper is the design of a commercial air conditioner's outdoor unit as shown in Figure 3(a).

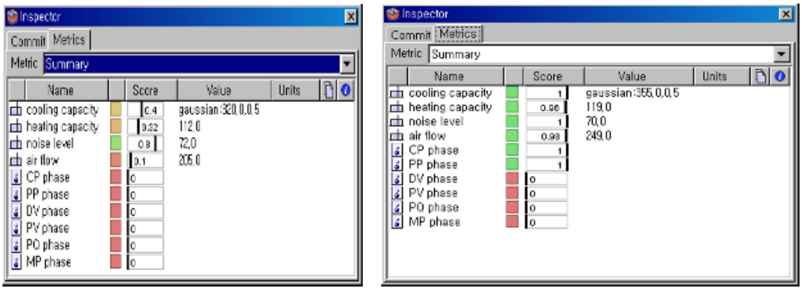


(a) Product Structure (b) Project View

Figure 3. Product Structure and Project View of a Commercia A/C's Outdoor Unit Design

Figure 3(b) shows the main window for viewing the design project. As shown in the hierarchy, project assets are divided into three folders corresponding to three teams in the project. The horizontal lines show how each team is working based on changes made to the design. The *Trunk* indicates the integrated design from which all design teams are branched out. From these branches each team works on its own parts of the design project. It is also common that a certain team may consider an alternative design in parallel with the original design.

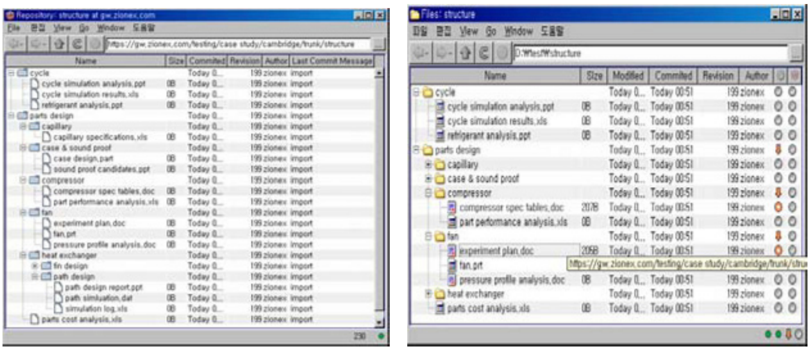
Figure 4 shows the performance metrics of the design project. The first four metrics are the key design criteria of the outdoor air conditioner unit. The next six metrics track formal design milestones (i.e., CP and PP for concept and production planning, respectively). By defining metrics in this fashion the manger can visually monitor how well the design criteria are satisfied and how the project is progressing in the formal project process. The key performance metrics can be modified by the design engineers or the manager via entering values. However, if possible, it is better to integrate these metrics with files and documents that design engineers actually work with. For instance, Figure 4(a) indicates the initial status of the four key performance metrics in the early stage of the project whereas Figure 4(b) indicates that all four metrics are satisfactory after the PP phase.



(a) (b)

Figure 4. Design Key Performance Metrics

Figure 5(a) and 5(b) show the design knowledge assets managed in the Revision Manager. While Figure 5(a) shows the assets managed in the repository, Figure 5(b) shows the locally managed assets that are checked out from the repository and always synchronized with the server. If a design engineer works on a certain sets of files and documents, the changes can be easily updated onto the server. If those files and documents are associated with performance metrics in Figure 4(a), the changes in metrics are also updated onto the repository and readily available for other team member and the manager to monitor.



(a) KA in the repository (b) KA in the local directory

Figure 5. Design Key Performance Metrics and Knowledge Assets (KA)

5 Concluding Remarks

Product design and development is a complex process requiring the extensive use of knowledge assets and effective communications amongst stakeholders. In this study, the lifecycle of progress in a project is viewed in five categories. With the perspective that the communication in projects is in general associated with decision making factors, capturing those factors generated from design knowledge assets of individual work environment and seamlessly sharing them with other stakeholders will make the communication more effective and thus expedite the process.

A computer-based solution to facilitate this concept and on-going projects were introduced. This solution reduces communication errors in a product design and development project by effectively capturing and sharing key factors of design decisions across the different levels of work environments as well as efficiently managing design knowledge assets.

This study used IDEMS Revision Manager from Oculus Technologies corporation (www.oculustech.com), groupware from Zionex (www.zionex.com), and Subversion from Tigris (subversion.tigris.org).

6 References

[1] Ulrich, K. and Eppinger, S.D. Product Design and Development, 1995 (McGraw-Hill).

[2] Nasr, E.A and Kamrani, A.K. Computer Based Design and Manufacturing, 2007 (Springer).

[3] Burton-Jones, A. *Knowledge Capitalism: Business, Work, and Learning in the New Economy*, 1999 (Oxford University Press).

[4] Hazelrigg, G.A. *Systems Engineering: An Approach to Information-Based Design*, 1996 (Prentice Hall).

[5] Wallace, D.R. *A probability specification-based design model : applications to search and environmental computer-aided design*, 1995 (Doctoral Thesis, MIT)

[6] http://subversion.tigris.org

[7] https://support.oculustech.com/revmgr/intro

Product Redesign Using TRIZ and Contradictive Information from the Taguchi Method

Ying-Ting Shen and Shana Smith[1]

Department of Mechanical Engineering, National Taiwan University

Abstract. In industry, products often need to be redesigned, after they on the market for some time, to adapt to new market trends. Compared with designing a new product from scratch, product redesign can reduce design cycle time and resources. The concept of product redesign is not extraordinary, but many prior studies only focus on particular products. A generic redesign method is still needed. To deal with the general majority of products, we present a comprehensive redesign methodology using TRIZ and the Taguchi method. Product redesign often requires tackling problems related to improving the functions of a product. Introducing or adding new functions might cause some indistinct contradictions between the new and existing functions of the product. Often, such contradictions might not be found until the final design stage is reached. With the proposed approach, the Taguchi method is used to find contradictions in the early design stage. Taguchi parameters are replaced by product components. In the past, designers eliminated function interactions by eliminating interactive parameters or changing the levels of interactive parameters. With the proposed approach, after finding interactive components, rather than removing the interactive components, TRIZ is used to eliminate the contradictions by keeping and improving the existing functions of the interactive components. After eliminating all contradictions, a conceptual design solution is created for the new product.

Keywords. Product redesign, Taguchi, Interactions, TRIZ

1 Introduction

Presently, the quality of human life is constantly improving and market demands are much different than in the past. To be competitive in modern industry, companies need to quickly release new products to meet rapidly changing market demands. Products often need to be equipped with more functions to satisfy customers' need. Therefore, to meet customers' demands, manufacturers need to actively release new products. Usually, a new design is a derivative work, with changes made to make it more suitable for a new application.

Some researchers have started to pay attention to product redesign. Salhieh proposed a method to obtain customer required functions for existing products

[1] Corresponding Author, No.1, Sec. 4, Roosevelt Road, Taipei, Taiwan, R.O.C; E-mail: ssmith@ntu.edu.tw

in the same series. Common product components can then be found and redesigned for reducing manufacturing and maintenance costs [1]. Li et al. developed a problem solving algorithm for product redesign based on model-based diagnosis [2]. In addition to the methods above, product family formation [3, 4], value-oriented life cycle costing [5], and QFD [6] have also been used to redesign existing products. Furthermore, adaptable design, as a concept, has been used to adapt products to various customer demands and to take into account changing performance requirements, based on physical, cultural, and environmental needs [7].

Many consumers prefer to buy the latest products and to throw old products away, even when the old products still function well. Some consumers might purchase many similar products because they need the various indistinct functions which the multiple products offer. As a result, valuable natural resources can be wasted and excessive environmental pollution can be released. Therefore, product lifecycle design should hold a more important position in overall environmental protection efforts. For instance, designers need to consider eco-friendly manufacturing processes when upgrading an existing product design or to determine when and which components of a product should be upgraded through the results of market and technology trend surveys [8].

To reduce the product waste, manufacturers need to produce multi-functional products which can replace multiple products. Currently, most designers first add fundamental functions to a product and then improve the existing functions gradually. For instance, many people prefer a portable digital camera. However, the use and requirements of a portable digital camera are similar to those of a cell phone. The similarities include portability, an LCD screen, electronic components, small size, and light weight. Therefore, cell phone manufacturers have been actively developing various camera phones to cater to the multiple but similar demands of consumers. Nevertheless, currently, a camera phone’s functions for taking a picture apparently cannot compare with those of a standalone digital camera.

If the deficiencies of a new product lead to customer dissatisfaction, the product will be quickly phased out. On the other hand, for the given example, if a portable camera phone can be redesigned to posses the complete functions of a digital camera, the phone will have a longer lifecycle. The same concept can also help shorten design cycles. For the example, the complete and acceptable functions of a digital camera can be considered the “target functions” and the mobile phone can be considered the “original product”. After the designers find the design interactions, they need to eliminate the contradictions. Then, the complete conceptual design can be completed.

Unfortunately, finding and eliminating contradictions can be a tedious and time consuming task. Currently, designers usually spend about 60% of their time searching for the right design information and parameters, which they rate as the most frustrating activities for engineers [9]. To help solve contradiction problems, some prior methods have been used to help create new designs [10]:

(1) Inspiration method. Designers who depend upon inspiration might be able to improve a design, but designers who depend upon the inspirational method can never be sure that they will be able to find reasonable solutions to a design

problem. In addition, inspirational solution is a very subjective method for most problems.

(2) Trial and error method. Although the trial and error approach is considered to be a conservative method, designers who depend on the method can spend a lot of time and resources on exploring design options. In addition, the outcomes might not be complete with inexperienced designers.

(3) Brainstorming method. The brainstorming approach is one of the most common methods used in industry. During 30-40 minutes of discussion, a leader shepherds a group to speak out their ideas without hesitation. The method provides opportunities for designers to hear other people's ideas, but the designers might also be influenced by others.

To help designers find necessary and correct design information in a short period of time, we introduce an approach based upon the Taguchi method. In the proposed approach, the Taguchi response is determined by customer demands. Through Taguchi analysis, the interactive information between "target functions" and "original product functions" is revealed. As a result, designers can focus the "original product" redesign efforts on the specific components which caused the contradictions.

With prior methods, designers eliminated interactions by eliminating the interactive parameters or by changing levels of the interactive parameters. Therefore, after designers changed the parameters, they might not have obtained optimal solutions, but they could create more robust solutions [11]. However, in product redesign, all of the "target functions" and the components in the "original product" are important and non-replaceable because of their particular functions.

With the proposed approach, rather than replace the interactive components with other components, we redesign the interactive components. In proposed approach, components are used as the parameters in a Taguchi analysis, to determine which components cause interactions. Designers eliminate the contradictions by keeping and improving existing functions of the interactive components, using a more systematic method, TRIZ. Target functions which meet market demands and trends are used as the objectives for redesign of the "original product". With the proposed approach, if the target functions can be easily integrated into the original product, product redesign time can be significantly shortened.

2 Taguchi and TRIZ

2.1 Taguchi method

Shortly after WW II, the Electronic Communication Laboratory in Japan was established to improve the quality of communications in Japan. At the Laboratory, Genichi Taguchi developed the Taguchi method, which uses statistical methods to improve product quality and reduce manufacturing costs [11, 12]. Taguchi method is a well known tool for parameter analysis and engineering quality improvement.

2.1.1 Orthogonal array

The Taguchi method depends upon an inner orthogonal array and an outer orthogonal array. The former provides a particular combination of control factors called a trial. The latter provides a combination of noise factors. In the proposed approach, the orthogonal arrays are used to provide information concerning interactions between factors.

2.1.2 Interactions

Interactions indicate an outcome with additive effects between control factors. Interactions can be classified into two categories, weak interactions and strong interactions, as shown in Figure 1. Usually, designers can accept weak interactions, but they cannot accept strong interactions.

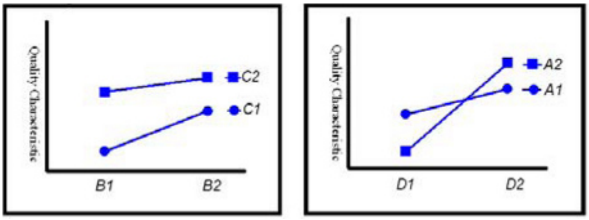


Figure 1. Weak interactions and strong interactions

2.2 TRIZ

TRIZ is a tool which was developed by Genrich Altshuller in 1940s to solve most engineering problems in an innovative way [13]. To develop TRIZ, Altshuller gathered approximately one hundred people by screening over 200,000 patents, from which he extracted 40,000 innovative patents. They concluded three innovative methods in inventions. The first method is to extract innovative principles for solving engineering problems by using a contradiction matrix. The second method is called the separation principle, which is used to separate certain effects. The third method is used to highlight the deficiencies of a design using Substance-Field analysis.

Figure 2 shows the resulting conceptual schema of TRIZ. Originally, designers are faced with a given design problem. The designers transfer the design problem into a TRIZ problem and then obtain a TRIZ solution. With the TRIZ solution, designers can solve their original design problem. Although TRIZ was developed in the 1940s, it is still very useful.

Most problems and knowledge can be categorized into 2 parts: existing problems or knowledge and new problems or knowledge. Table 1 shows four common combinations in engineering design. Using new techniques to solve new problems might cost a lot in time and resources and the results might be uncertain. As a result, TRIZ focuses on providing new concepts to solve new problems by applying existing knowledge to the new problem domain. Such problems are called inventive problems.

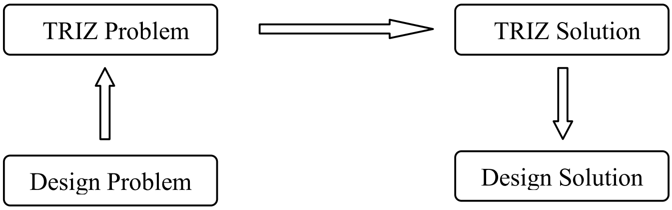


Figure 2. Conceptual schema of TRIZ

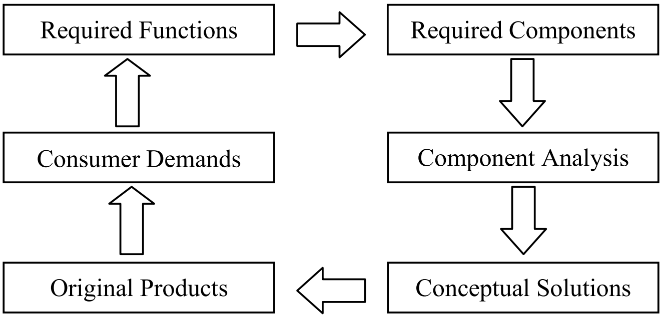


Figure 3. Relationship between customer demands and redesign objectives

Table 1. Knowledge and problem [14]

New Knowledge	New knowledge applied to known problems Example：New plastics provide strong and lightweight products.	New knowledge applied to new problems Example：Various uses of lasers.
Existing Knowledge	Existing knowledge applied to known problems Example：All tasks with generally known solutions.	Existing knowledge does not provide satisfactory solutions. Inventive Problems
	Known Problem	**New Problem**

3 Redesign Methodology

Understanding market demands and market trends is very critical in product redesign. Figure 3 shows the relationship between customer demands and redesign objectives. If designers clearly realize current market trends, they can quickly redesign a product by understanding the distinct relationships between components and their functions. Therefore, the most important task in our proposed approach is to build a communication platform between designers and customers.

3.1 Product Redesign Flowchart

After obtaining explicit design objectives, designers can define evaluation functions for evaluating components. Figure 4 shows a process flowchart for the proposed redesign methodology. The details of each step are described in the next section.

3.2 Product Redesign Details

3.2.1 Step 1: Specify the Products

The major goal of this paper is to improve and redesign existing products based upon market trends. During the redesign process, two or more existing products are chosen. Products with the "target functions" need to belong to the same product cluster as the original product, based upon certain similarities. For example, in Table 2, cell phone, digital camera, PDA, and GPS are in the same product cluster, based on their functions and attributes.

Products are then decomposed into their basic components, as shown in Figure 5. For example, a cell phone can be decomposed into a screen, shell, etc. Customers' demands for desired functions are then analyzed to find corresponding TRIZ parameters. Figure 5 shows product components and their corresponding functions (input by customers) and TRIZ parameters (input by design engineers).

3.2.2 Step 2: Build the component factor table

In Figure 5, product *A* has a screen (component 1_A). Product *B* has a screen (component 1_B), as well. The two screens might have different sizes and operation methods. However, the two screens can be considered the same component with, perhaps, different levels. Table 3 shows an example of a component factor table which we used as the Taguchi factor table for the example. Component *1* has three levels, which are component 1_A (used by product *A*), component 1_B (used by product *B*), and component 1_C (used by product *C*).

3.2.3 Step 3: Select the orthogonal array

After determining components and their levels, a suitable orthogonal array can be created. Subsequently, interactions that appear can be considered by designers and, then, entered into particular columns of the orthogonal array based upon the triangular table [13].

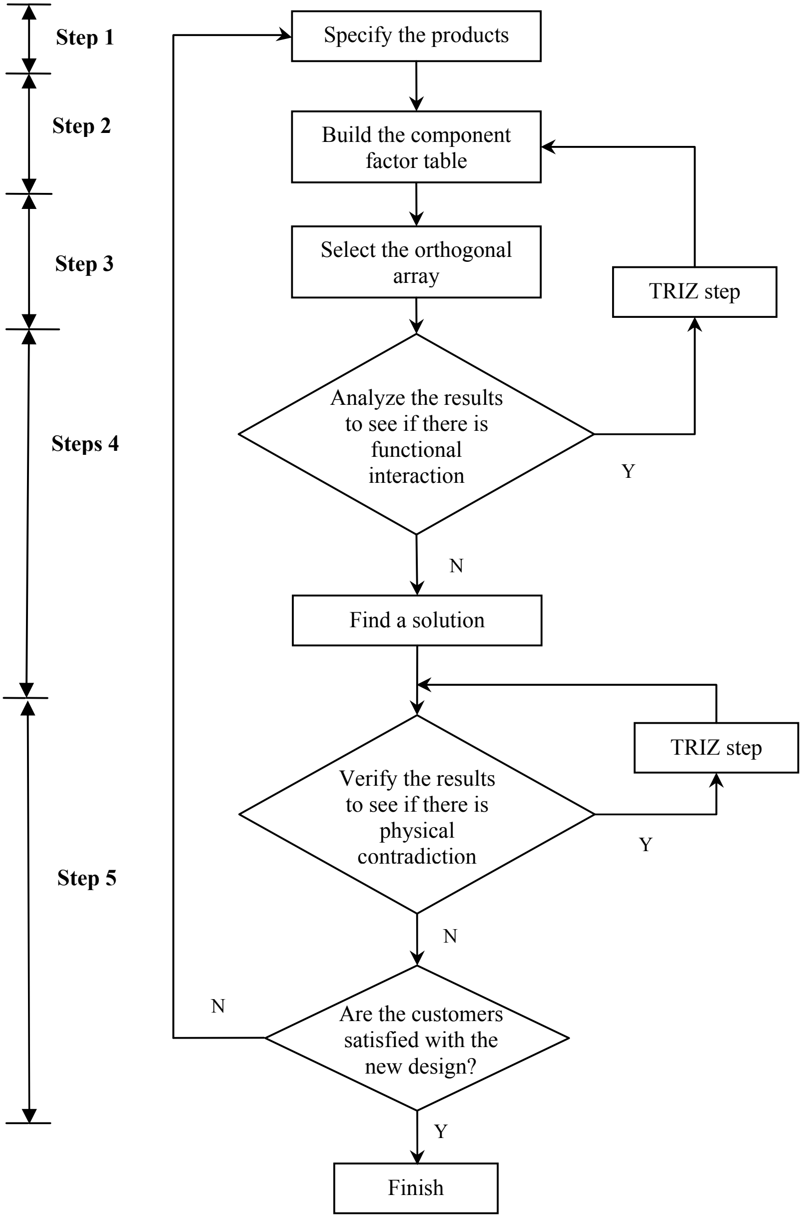


Figure 4. Product redesign flowchart

Table 2. Example of Product Clusters

Product / Product Cluster	Product 1	Product 2	Product 3	Product 4
Consumer electronic products				
Office products				
Machining products				

3.2.4 Steps 4: Analyze the results

Based upon the impacts of a component in the new product, for our example, we consider seven grades, as shown in Table 4. The scores can be set based upon customer demands and the designers' judgments. Each level of a component can be assigned a score from Table 4. Summing up the scores for each trial gives the response values for the new design.

3.2.5 Step 5: Verifying the results

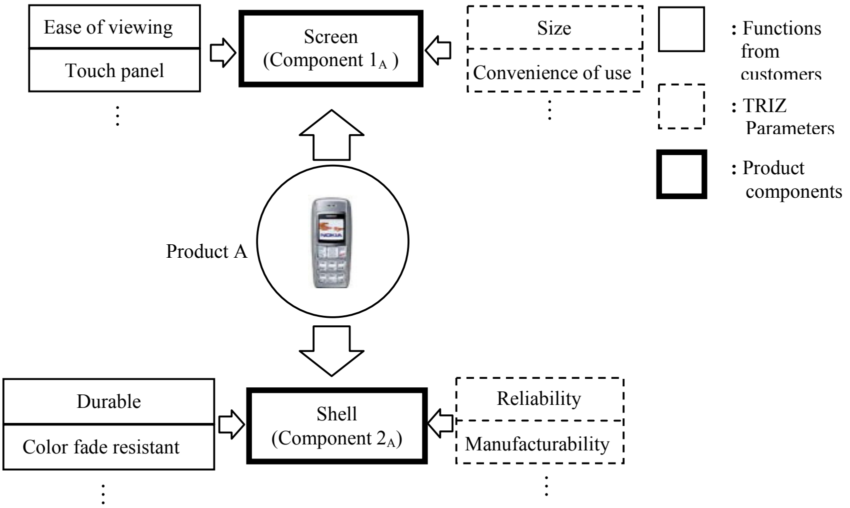
After Taguchi analysis, contradictions between components may be discovered. In general, one or more TRIZ parameters are associated with each component. The designers need to check the associated engineering parameters to determine which parameters caused the contradictions. Thus, the designers need to find all contradicting parameters between each pair of components, based upon customer demands, as shown in Table 5. For example, if customers want a larger screen but still want a lightweight cell phone, the size parameter of the screen contradicts with the lightweight parameter of the cell phone.

To deal with the contradictions, we use a TRIZ contradiction matrix to find innovative design principles. If we cannot find a suitable principle in the matrix, TRIZ provides two other processes which we use to solve the problem: separation principles and substance-to-field analysis.

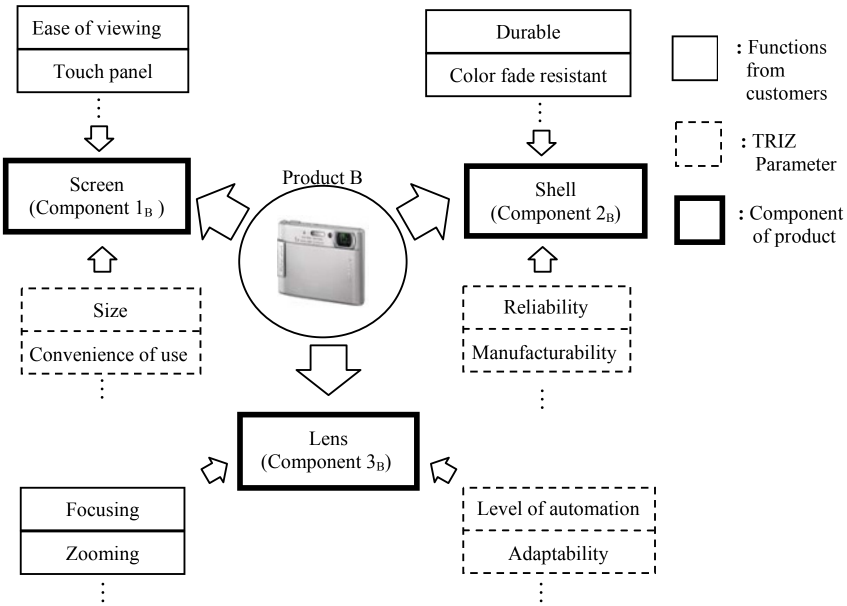
3.2.6 TRIZ Step

TRIZ is a multi-selection tool for a design. Designers can obtain more conceptual solutions by using TRIZ for redesigning components. After the TRIZ step, designers can check to see if any contradictions still exist between the original components and the new designs. If there are still interactions that need to be eliminated, designers can repeat the TRIZ step again.

In addition to functional contradictions, some physical contradictions might be found in the solutions, when designing a new product. Thus, designers can use the TRIZ method again to resolve the physical contradictions.



(a) Cell phone



(b) Digital camera

Figure 5. Product decomposition

Table 3. Component factor table

Levels / Factors	Camera	Mobile phone	GPS
Screen	1	2	3
Shell	1	2	3
Lens	1	-	-
Press-Button	1	2	3

Table 4. Scaling grade

Grade	Impacts to the new product	Score
1	Very important	+5
2	Somewhat important	+3
3	Slightly important	+1
4	Neutral	0
5	Slightly harmful	-1
6	Somewhat harmful	-3
7	Very harmful	-5

Table 5. Example of All Possible Contradictive Parameter Combinations

Contradicting components		Interaction (Y/N)
Component *1*	Component *2*	
TRIZ parameters		
Parameter 1	Parameter 1	N
Parameter 1	Parameter 2	Y
Parameter 1	Parameter 3	N
Parameter 2	Parameter 1	N
Parameter 2	Parameter 2	Y
Parameter 2	Parameter 3	Y

4 Conclusions and Discussions

Product redesign or reuse is becoming increasingly important. Analyzing existing products can help understand customer demands and current market trends. In this paper, we present a systematic product redesign method. The proposed method can be used to analyze existing products with complex relationships between components and parameters. The methodology integrates the parameter analysis ability of the Taguchi method and the innovative design principles of TRIZ. With the proposed methodology, designers only need to clearly define the redesign objectives and focus on improving particular components. The proposed method can maintain the desired functions of existing components while eliminating design contradictions. Future work will focus on the commonalities between products and on modularizing components to further facilitate the redesign process. Furthermore, methods for using TRIZ to effectively explore the relationships among engineering parameters will also be studied.

5 References

[1] Shieh MD, Yan W, Chen CH, Soliciting customer requirements for product redesign based on picture sorts and ART2 neural network. Expert Systems with Applications 34 (2008) 194–204

[2] Salhieh SM, A Methodology to redesign heterogeneous product portfolios as homogeneous product families. Computer-Aided Design 39 (2007) 1065–1074

[3] Li ZS, Kou FH, Cheng XC, Wang T, Model-Based Product Redesign. International Journal of Computer Science and Network Security, 2006, VOL.6 No.1A,

[4] Thevenot HJ, Nanda J, Simpson TW, Redesign Product Families using Heuristics Shared Ontology Component Information. Information Reuse and Integration, 2006 IEEE, 330-335.

[5] Alizon F, Shooter SB, Simpson TW, Improving an existing product family based on commonality/diversity, modularity, and cost. Design Studies Vol. 28, Issue 4, July 2007, 387-409.

[6] Janz D, Sihn W, Product Redesign Using Value-Oriented Life Cycle Costing. CIRP Annals - Manufacturing Technology Vol. 54, Issue 1, 2005, 9-12.

[7] Bovea MD, Wang B, Redesign methodology for developing environmentally conscious products. International Journal of Production Research, Vol. 45, Issue 18 & 19 September 2007, 4057-4072.

[8] Mary EK, Janis PT, Dan I, Karl RP, Jelesko J, Asli S, Jaeil P, Design for adaptability (DFAD)—a new concept for achieving sustainable design. Robotics and Computer-Integrated Manufacturing, 2007, 23: 727–734.

[9] Huppes G, Ishikawa M, Upgrade planning for upgradeable product design. Quantified co-Efficiency. Springer Netherlands, 2007, 261–281.

[10] Natraj I, Yagnanarayanan K, Kuiyang L, Subramaniam J, Karthik R. A reconfigurable 3D engineering sharp search system part I: sharp representation. Proceedings of ASME DETC' 02 CIE Conference.

[11] Zhu YZ, A study on Integrating the TRIZ Method with Functional Analysis. Thesis. National Cheng Kung University, 2003

[12] Phadke MS, Quality Engineering Using Robust Design, Prentice-Hall. 1989

[13] Ranjit KR, Design of Experiments Using the Taguchi Approach. pp.247, 2001

[14] Glenn M, Theory of Inventive Problem Solving (TRIZ). Available at: <http://www.mazur.net/triz/>. Accessed on: Jan. 15th 2008

[15] What is an Inventive Problem. Available at: <http://www.ideationtriz.com/source/12_Inventive_Problem.htm>. Accessed on: Aug. 20th 2008

Radio Frequency Identification (RFID)

Toward Full Coverage UHF RFID Services - An Implementation in Ubiquitous Exhibition Service

Tung-Hung Lu[a,1,] , Li-Dien Fu[b]

[a] Researcher, Industrial Technology Research Institute, Taiwan
[b] Engineer, Industrial Technology Research Institute, Taiwan

Abstract. RFID systems are widely introduced in supply chain management, logistics and medical services. Lately, applying the technology onto exhibition industry, which raises its own challenges, is in a beginning stage. Existing systems for exhibition, such as barcodes or HF RFID access control systems, are facing problems with short interaction distance and disconnected to database, which lead to inconvenient in usage and lack of extension services. This study proposes an innovation in UHF RFID system, with longer sensing distance and real-time connection to database, providing a near-ubiquitous service for exhibition industry. A real case is implement in Taipei World Trade Center. Valuable findings and implication are learned though the study.

Keywords. RFID, Ubiqouitous, MICE Industry

1 Introduction

Many researches [1, 3, 9-12] already showed the economic impacts on international convention and exhibition. According to ICCA's annual report, "There were 6,500 international conventions and exhibitions which took place in 2007. Taipei is a remarkable newcomer in the top 20 as it took fortieth place in 2006" [13]. The meetings, incentives, conventions, and exhibitions (MICE) industry, as a new service industry, has become a value creation focus in Taiwan. Exhibitions provide unique environments for customers, prospects and industry colleagues to meet face to face. Due to the globalized competition, a trend of continuously expending international exhibitions is expected. For example, there were 1200 vendors in the Taipei International Electronic Autumn Show 2007, and the number will be increased to 2000 vendors on 2008 according to expectation from Taiwan Electrical & Electronic Manufactures' Association (TEEMA).

[1] Corresponding Author : Researcher, Identification and Security Technology Center, Industrial Technology Research Institut.; Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan , R.O.C. Tel: +886-3-5913408; Fax:+886-3-5826464; Email: tonyly@itri.org.tw

However, few applications deployed on providing interactive information service to visitors are found. Moreover, legacy technologies deployed for access control may cause inconvenient situation for visitors such as badges with 2D barcodes and lacks of considerate service. These shortages could result in bad images to foreigners and reduce the willing to participant the same convention or exhibition next year. Hence, the goal of this research is to develop a serious of RFID based interactive applications to provide visitors a better experience during the convention or exhibition.

2 Related Works

Exhibition using cutting edge technology to demonstrate the technology achievement of a country already become a trend. The 2005 world exposition held in Aichi, Japan is a good example of applying newest technology in exhibition. The exposition applied RFID technology to provide access control, registration service, and exhibits introduction for visitors [5,7,14]. The goal of these applications is to enhance the visitor's experience. RFID technology provide several unique features such as tags can be detected at a longer distance, multi tags can be read at the same time, and tags can be read/write in milliseconds. These features make RFID technology a popular front end for triggering a certain information service.

There are two types of configuration for deploying RFID based applications. The user-with-reader configuration and the user-with-tag configuration; both types have its weakness and advantage. Mobile RFID readers allow users decide when to retrieve the data from tags. On the one hand, Gråsjö et al. [6] applied RFID tags close to objects, with a mobile RFID reader held by user, and a content database deployed at backend. Gråsjö 's research established a way to obtain object information on demand. For another example, Blache et al. [2] developed an indoor navigation system for a structured location-management infrastructure maintaining information about all fixed, movable and mobile features of a limited space, and RFID technology was used to enable the position detection and moving prediction of tagged objects.

On the other hand, the service system without disturbing users' original behaviours usually attaches RFID tags on users. The eXspot project [7] employed RFID to support, record, and extend exhibit-based, informal science learning. Fleck et al. [5] established a conspicuous, portable, freestanding unit called Pi-station in their interactive museums research. The Pi-station equips with several identification technologies including 2D barcode reader, HF RFID reader, and camera to provide exhibition content on the fly. These two projects are enabled by contactless feature of RFID technology. However, the sensing distance of the reader was not discussed in their research. In our point of view, the sensing distance would determine the way users interact with the system. Generally speaking, services with a longer sensing distance could be more proactive and friendly, while shorter ones could be more private.

In indoor navigation and location-based application research area, most of researches mentioned using PDAs or other handheld devices to access backend service. However, in real world exhibition, excepting most visitors own a RFID reader compatible handheld device is not practical. In this research, a full coverage UHF RFID system will be deployed in commercial environment. The simplicity of using the service for visitors is the major concern. Thus, RFID tags are attached on exhibition badges. In order to provide an attentive service, the long distance sensing is crucial. Hence, UHF RFID system is the first choice of the research. More details of the system components will be specified in the next section.

3 The Ubiquitous Exhibition Services

3.1 Differentiations of Ubiquitous Exhibition Services

With the evolution of MICE service, there exists different ways to acquire visitors information. Some typical methods are shown in table 1.

Table 1. Existing method to acquire visitors information

Method	Advantage	Disadvantage	Available Information
Human contact	-most friendly -almost no operational training needed -good privacy ensuring	-inefficient -potential human mistake	-visitors basic Information
Barcode	-cheap & popular - some operational training -good privacy ensuring	-very short sensing distance -requiring line-of-sight -visitor bothering	-visitors basic Information -visitor summary report from exhibition host
HF RFID	-matured RF technology -good privacy ensuring	-very short sensing distance -no better than barcode but more expensive -visitor bothering	-visitors basic Information -visitor summary report from exhibition host

Generally speaking, traditional method can retrieve only visitors' basic information such as name, contact plus a visitor summary report from exhibition host. With barcodes or HF RFIDs, exhibitors can save some key-in time and reduce human mistake while inserting visitors' information, however, those technologies require visitors to hand out their badges for the short sensing distance. This is very inconvenient for visitors, and could disturb the process of visiting an exhibition.

In order to ease the inconvenience and promote a better service level for exhibition, this study proposed a ubiquitous exhibition service. The objective of ubiquity is to provide a more flexible service trigger method, which doesn't bother the visiting behaviours of buyers. By using UHF RFID technology, the sensing distance could be extended to about 1.5 meters with a passive RFID tag attached on visiting badge, which results in a more convenient scenario that visitors don't need to be interrupted during their visiting. With the feature of longer sensing distance of UHF RFID, the moving history of visitors with RFID badges are tracked silently through the exhibition. So visitors don't need to hand out their badges any more as well as exhibitors don't need to ask visitors for their badges or contacts.

As mentioned in section 2, a longer sension distance provides a more convinent service trigger method but raises some privacy problems. Table 2 shows a comparison between short distance service model (the traditional ones) and proposed long distance service model.

Table 2. Existing method to acquire visitors information

Method	Advantage	Disadvantage	Available Information
Short distance	-good privacy ensuring -some operational training -cheap & popular -good privacy ensuring	-inefficient -potential human mistake -disturbing	-visitors' basic information -visitor summary report from exhibition host
Long Distance (UHF RFID)	-smooth experience through the exhibition -detecting valuable behaviours of visitors -no operational training needed	- expensive reader - privacy issues - unexpected radio wave collision	-visitors basic Information -visitor summary report from exhibition host -moving history of visitors

The 2 major challenges faced by long distance model are privacy issues and unexpected radio wave collision, however, detecting behaviours of visitors is far more valuable than collecting visitors' basic information or the summary report of the exhibition. For example, mining the visitors' behaviour data base from a series exhibition of the same topic through years, could create a huge value while apply the mining result into customer relationship management systems (CRMs), and potential leads could be easy identified by analyzing visitors behaviour histories. In contract to value for exhibitors, visitors can have their own value-added. For example, by collection one's visit-log, he or she can have an auto-generated business report of visited booth, product or vendor comparison. This could save lot of time for B2B visitors and provide for organized procurement information automatically for them as customers. With the potential value, following challenges of mass deployment of UHF RFID seems worth to be overcome.

3.2 Challenges of Ubiquitous

3.2.1 Challenge of Mass-IO-threading

As an embedded network device, the client application needs at least 2 threads to work with our micro reader, one for commanding, and the other for listening. This model works fine while the amount of readers is under a small size, for example 5 or 10. Conversely, if a huge amount, for example, 200 readers, is needed, the total tread amount will increase to double of that. In this example, there will be 400 threads with heavy RFID data IO running at the same time. Further, each thread might insert up to 50 row data into database 5 times a second. A maximum estimation of total database insertion action frequency in this example would be:

200 (readers) * 5 (times) * 50 (rows) = 50,000 (insertion / per sec)

Though the real case could be not so tough, a prevention mechanism for IO-threading burst effect must be employed, or the potential risk of database dead-locking could break the whole system.

3.2.2 Challenge of Unexpected Radio Fequency Collision

The operation of RFID systems often involves which numerous tags or readers in a elimited area at the same time. Not only tags could collide and cancel each other out, but also readers colide to each other and might cause very bad reading performance [4,8]. While expecting a full coveraging ubiquitous service in a exhibition hall, the collision effect should be considered seriously for the reason that hundreds of reader and over tousands of tags exists and functions at the same time.

To sum up, for a ubiquitous service model, the mass IO-threading problems and RF collision problems must be well considered. Next section describes the proposed system architecture to overcome these problems.

4 System Analysis, Architecture & Implimentation

4.1 System Analysis & Design

In this study, Object Oriented (OO) system analysis and development methodology is adopted. Figure 1 shows the top level use case for the whole picture. Generally, while visitors approach an UHF RFID reader, the reader automatically collects visitors data. The coordinater then coordinate all readers in the exhibition hall, and filter mass data into backend perstance server. After that, information kiosk and web server could obtain proper information for both visitors and web users, for example, location of friends for visitor or booth performance report for exhibitor manager through the web.

The most complex part of the use case coordinating readers. While in the real case, the reader amount could be up to hundreds, the total data flow should be very huge. Thus, it is important to design a algorithm for front-end redundant data filter as well as a back-end transaction manager. Besides, mass I-O problem could be expected, so that a well-structured multi-threading framework is needed to be considered carefully.

Figure 2 shows a top level class diagram for the system. The OO objects are designed in 3 layers: model layer, DAO layer and manager layer. The model layer consists of entity models which are designed in POJO (Plan Object java Object) format. The POJO format could be considered as a more light-weighted EJB object.

The POJO focus purely on entity attributes, so that make it is more suitable for mapping the object into relational data base scheme, or called O-R Mapping (Object-Relational Mapping) process.

In the DAO layer, basic CRUD (Create, Retrieve, Update and Delete) behaviours are designed in base DAO object. In advance, further data accessing behaviours are defined in the DAO interface objects. The implementation of each interface differs by cases. In this study, Hibernate framework is selected to be the implementation framework for its high performance, scalability and flexibility. In addition, Hibernate framework provides a matured O-R mapping feature. For these advantages, it makes Hibernate a perfect framework for DAO layer implementation.

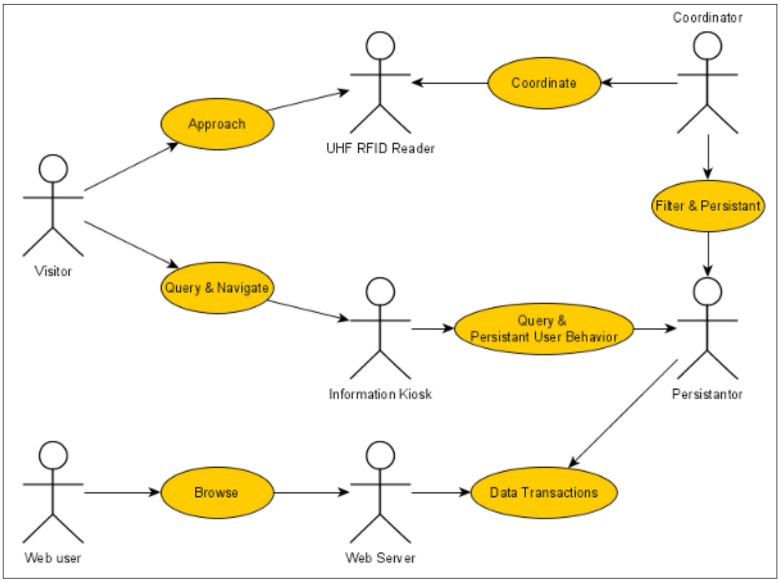


Figure 1: Use Case Diagram

In the manager layer, interfaces with different behaviour groups are designed for reporting and user managing. Each of their implementations is assigned with corresponding DAO interface.

Each class in above layers, including interfaces and implementations, are glued via Spring framework. With the injection technology provided by Spring framework, each component of the system can be plugged-n-played by case. With those system framework designs, the advantage of OO could be fully utilized.

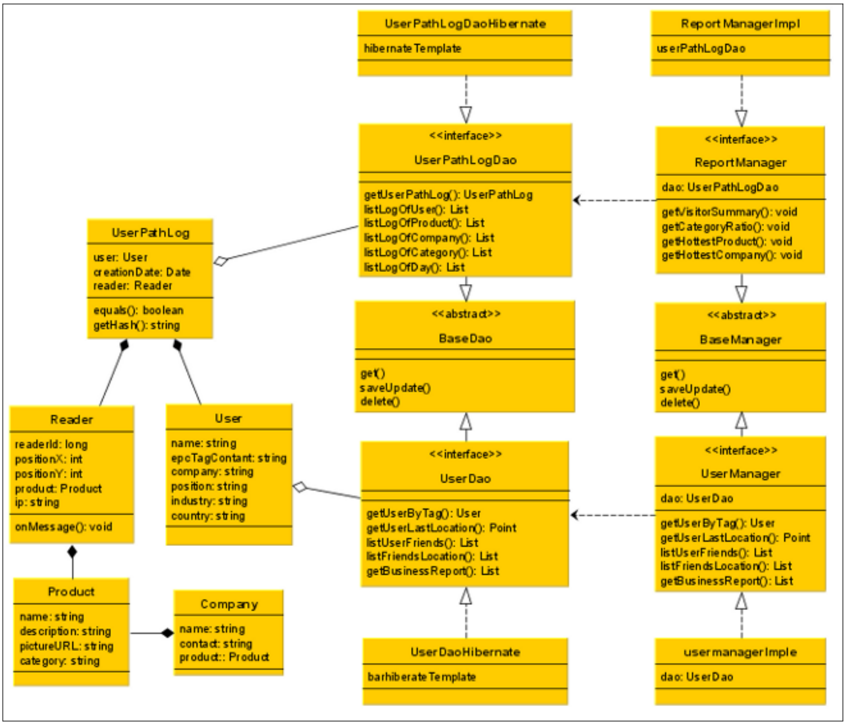


Figure 2: Class Diagram

4.2 System Architecture

To overcome the mass IO-threading problem, the system is designed with following nodes:

- *Reader Coordinator*: coordinating all threads including commanding threads and data receiving threads.
- *Message Queue Server*: queuing tag data while IO of database is overloading and ensuring no losing data during the transaction.
- *Database / Datamart*: storing visitors' behaviour histories for separated exhibition.
- *Data Warehouse*: storing visitors' behaviour histories across exhibitions and for further data mining usage.
- Application Server: providing front-end application service such as an web-container for running web-based applications.
- UHF RFID Readers: TCP/IP based reader, collecting visitors' moving histories and behaviours.
- Information Kiosks: providing personalized exhibition information such as personal visiting reports for visitors.

- Web-based User Clients: providing on-line analysis of visitor behaviours, leads indentifying and booth ROI reporting for exhibitors.

By employing this architecture, the heaviest par of loading – massive IO between readers and database– is coordinated by the reader coordinator. The coordinator summarizes massive data collected from all readers, and sent organized data sets unto a message queue. By queuing those data sets, all data can be smoothly inserted into database in order, even if a sudden burst of data comes in a very short time. The conceptual architecture is shown in figure 3.

In addition to the IO-threading problem, the RF collision problems is considered to be overcome via sliding readers' polling time into groups. This function is implemented in reader coordinator.

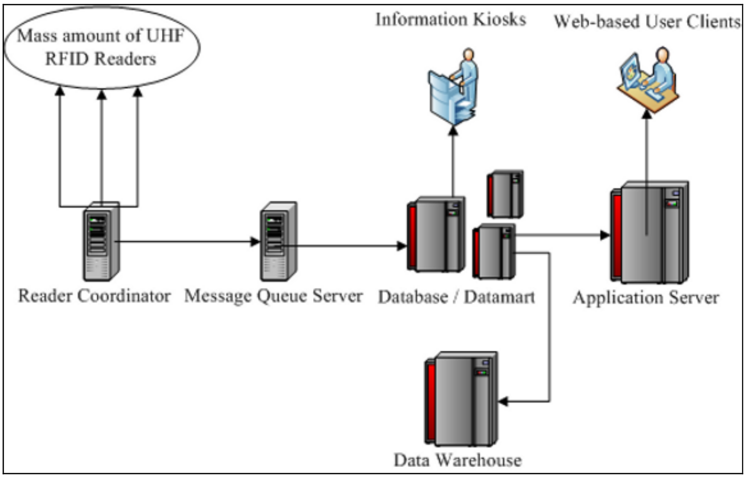


Figure 3. System Architecture

4.3 Implementation

A field-trial deployment is conducted by Industrial Technology Research Institute (ITRI), and cooperated with Taiwan Electrical & Electronic Manufactures' Association (TEEMA) and Taiwan External Trade Development Council (TAITRA). The scale of pilot running is as follows:

- 70 micro UHF RFID readers
- 50 booths
- 10 information kiosks with RFID readers embeded

The pilot system was deployed at Taipei World Trade Center Nangang Exhibition Hall (Nangang Hall) on October 2008. Nangang Hall is opened on April 2008, and will be the major exhibition center of Taiwan in the future. There are two exhibition floors in the hall, both with 22,680 m^2 of area. In addition, a standard booth space is set to 3x3 m^2.

The result of pilot shows that the full coverage configuration of UHF RFID is powerful to acquiring visitors' behavior and moving histories. Besides, the system is also able to be an indoor RTLS (real-time-location-service) foundation framework. The massive IO-threading problem is well handled by the designed reader coordinator and message queue combination. Though the threading challenge is overcome, the polling-time sliding method doesn't effectively ease RF collision problem. This result suggests a future work direction of designing more flexible and refined model for reducing RF collision problem.

5 Implication & Conclusion

This study carries out a full covarage UHF RFID deployment implementation, which is a near ubiquitous exhibition service system pilot in Taipei World Trade Center Nangang Exhibition Hall. The pilot demostrated the value of tracking visitors' behaviour histories, and is useful for exhibitors to identify their potencial leads as well as for visitors to experience a more convinient exhibition.

In addition, while a full coverage deployment is conduct, the anti-collision strategy becomes imperative for coordinating each RFID readers. This study suggests that sliding readers' polling time into groups could ease some collision effect among readers, however, not only time segment, but there are also more factors involves with the RF collision, for example, the RF channel, the environment reflections and so on. For further improvement, FDMA (Frequency division multiple access) and TDMA (Time division multiple access) methods, even SDMA (Space division multiple access) and CDMA (Code division multiple access) should be considered to be combined.

In 2009, the idea will be implemented in a world class fair in Taipei. The challenges of RF collision and costing down become a major working topic in the near future.

6 Acknowledgment

Authors would like to appreciate Taiwan Electrical & Electronic Manufactures' Association (TEEMA) and Taiwan External Trade Development Council (TAITRA) for providing professional experience in hosting international exhibition and a field trial deployment space during the system pilot.

7 References

[1] Bauer, T.G., Kowloon, H.H., Lambert, J. and Hutchison, J., Government Intervention in the Australasian Meetings, Incentives, Conventions and Exhibitions Industry (MICE), Journal of Convention & Exhibition Management, 2001, 3:65-87.

[2] Blache, F., Chraiet, N., Daroux, O., Evennou, F., Flury, T., Privat, G., and Viboud, J.P., Position-Based Interaction for Indoor Ambient Intelligence Environments, Lecture Notes in Computer Science, Springer, 2003, 2875:192–207.

[3] Colin, B., Yahia, B., Mohamed, E., and Achmad W., Revamping the value chain in exhibition facilities: the case of the Dubai exhibition industry, Journal of Facilities, 2007, 25:419-436.

[4] Dong-Her Shih , Po-Ling Sun , David C. Yen , Shi-Ming Huang, Taxonomy and survey of RFID anti-collision protocols, Computer Communications, 2006, 29:2150-2166.

[5] Fleck, M., Frid, M., Kindberg, T., O'Brien-Strain E., Rajani, R, and Spasojevic, M., From informing to remembering: Ubiquitous systems in interactive museums, Pervasive Computing, 2002, 1:13-21.

[6] Gråsjö, U., Jansson, E., Lind, K., Snis, U.L., Pareto, L., and Svensson, L., Experience Design: Understanding the Role of Technology for Virtual Experiences, Proceedings of the 29th Information Systems Research Seminar, 2006.

[7] Hsi, S., Semper, R., Brunette, W., Rhea, A., & Boriello, G., eXspot: A wireless RFID transceiver for recording and extending museum visits,2004, Proceedings of UbiComp, 2004.

[8] Jiexiao Yu, Kaihua Liu, Ge Yan , A Novel RFID Anti-collision Algorithm SDMA, 2008, International Conference on Wireless Communications, Networking and Mobile Computing, 1-4.

[9] Kim, S.S., and Chung, K.Y., Convention industry in South Korea: an economic impact analysis, Journal of Tourism Management, 2003,24:533-541.

[10] Kozak, N., The Expectations of Exhibitors in Tourism, Hospitality, and the Travel Industry A Case Study on East Mediterranean Tourism and Travel Exhibition, Journal of Convention & Event Tourism, 2006, 7:99-116.

[11] Lee, M.J. and Back, K.J., A review of economic value drivers in convention and meeting management research, Journal of Contemporary Hospitality Management, 2005, 17:409-420.

[12] Rogerson, C.M., Conference and exhibition tourism in the developing world: The South African experience, Journal of Urban Forum, 2005, 16: 176-195.

[13] Top 20 country and city rankings 2007, Available at : < http://www.iccaworld.com/ newsarchives/archivedetails.cfm?ID=1759 >. Accessed on: Jan. 12th 2009.

[14] Use of Hybrid Communicator at EXPO 2005 Is Announced, Japan Association for the 2005 World Exposition, Available at : < http://www.expo2005.or.jp/en/newstopics/ 2004/041213_hybrid.html>, Accessed on: May 19th 2008.

[15] Wu, J.L., Huang, X.Q., E-commerce and exhibition industry go hand in hand – the ideal combination in new century, Journal of Kybernetes, 2004, 33:406-413.

Building a RFID Anti-Collision Environment for Conference and Exhibition Industry

Min-Hsien Weng[1], Chih-Wei Chao, Kuo-Shu Luo, Li-Dien Fu and Tung-Hung Lu

Researcher, Industrial Technical Research Institute, Taiwan.

Abstract: Conference and exhibition industry is a new industry with high growing power. Radio frequency identificaion (RFID) technology has been available for decades and applied in conference and exhibition industry recently. However, the system's performance is expected to be critically reduce and unstabe due to the radio frequency collision. This paper developes an algorithm and a hierarchical process for quickly deploying the system and reducing the interference between RFID readers. With the hierarchical process, a simulation experiment is conducted to test the proposed anti-collision algorithm.

Keywords: RFID, conference and exhibition industry, anti-collision

1 Introduction

Conference and exhibition industry is a new industry with high growing power. [6].During the last ten years, the exhibition spaces in Asia growth doubled because countries in Asia were thirsted for boosting their economics via international trading and exhibitions. But new challenges from traditional markets in Western Europe and North America, including emerging MICE (meetings, incentives, conferences, and exhibitions) destinations in Central and Eastern Europe, is diminishing Asia's strength in the international MICE market [9].

In order to provide a better visitor experience, some advanced technologies are introduced into the MICE industry, such as RFID (Radio Frequency Identification), Wi-Fi, 3G et al. Especially, RFID is a suitable application for exhibition and conversion [4,7]. In a scenario of employing RFID technology in conference or exhibition, all registrars or visitors should have a unique badge embedded with an RFID tag. RFID readers should be arranged in front of booths' promoting products, so as to collect and record nearby visitors' information. It's convenient for visitors or buyers to trace and track and find potential leads during exhibition period to establish business relationship. For the whole system, there might be a great

[1] Corresponding Author : Engineer, Identification and Security Technology Center, Industrial Technology Research Institut.; Tel: +886-9-20080295; Fax:+886-3-5826464; Email: samweng@itri.org.tw;

number of RFID readers deployed in a predetermined space, performing highly reliable visitor locating function. Unfortunately, the system's performance is expected to be critically cut down with the readers' collision. In order to solve the collision problem for employing RFID system into exhibition, this research focuses on developing an agorithm and a hierarchical process to quickly implement the system and reduce the interference between readers.

The remaining of this paper is organized as follows. Section 2 is the literature review of countermeasure of readers collision problem. Section 3 depicts the agorithm and hierarchical process to slove the RFID deployment problem. Section 4 sets up a simulation to verify the algorithm's performance. The final section concludes the results.

2 Countermeasure of surveyed readers collision problem

Readers to reader collisions occur when multiple readers have an overlap reading area and tags. Generally, there are four procedures to reduce RFID collisions: space division multiple access (SDMA), frequency division multiple access (FDMA), time division multiple access (TDMA) and code division multiple access (CDMA) [3].

According to the researches [2,5,8], TDMA is often applied to solve reader to reader collision's problems. This procedure is widespread implemented in the field of digital mobile radio system. It relates to techniques that allot readers in different time slots thus avoiding simultaneous transmissions. Colorwave is the first one to apply TDMA's methodology to solve the reader to reader collision [1]. Each reader chooses a random timeslot to read. If a reader is affected by another reader, it can select a new timeslot and send a message to its neighbors to inform the new timeslot (color). If any neighbor has the same color, it changes to new timeslot and inform its neighbors and this continues. The timeslots will increase to assign to the conflict readers until the successful transmission percentage is accepted by the system administrator.

FDMA refers to techniques in which several transmissions on various carrier frequencies are simultaneously available to the communication participants [3]. The reader to reader collision happens when two readers are using the same frequency. The total available channels for RFID application are between 922 MHz to 928MHz in Taiwan. One disadvantage of the FDMA procedure is that the reader collision also happens if two readers use close frequencies such as 922.25 MHz and 922.75 MHz, and different brands' RFID readers show different conflict situation. It should take several experiments to determine how far frequency interval can reduce the collision between readers.

Because of conference and exhibition's characteristic, an exhibition's lead time only two days. The RFID system should be implement in very short time including setup the hardware, the readers, sensor network and completes the overall test. In order to quickly deploy the RFID system with mass readers and avoid readers' collision, this paper provides a hierarchical process to comfortably assign the timeslot and frequency's list based on TDMA and FDMA procedure.

3 Anti-Collision algorithm

3.1 Reader collision problems in Dense RFID readers exhibition environment

According the size and owner's requirement of a booth, single or multiple readers to each booth is deployed in an exhibition. In this paper, it is assumed that each reader has the same reading and interfering range. In situations that two or more readers work closely, the RF collision problem occurs. Figure 1 shows two common interference situations among readers. The solid circle is the reader in the exhibition hall. The circle shows the read range of a reader. The dotted circle represents the interfering zone of a reader. Reader collision problems result from those readers whose interfering zones have overlaps and are operated at the same time. As Figure 1 shows, reader 1 will be in collision with reader 2 and reader 3. Others also have their own collision problem.

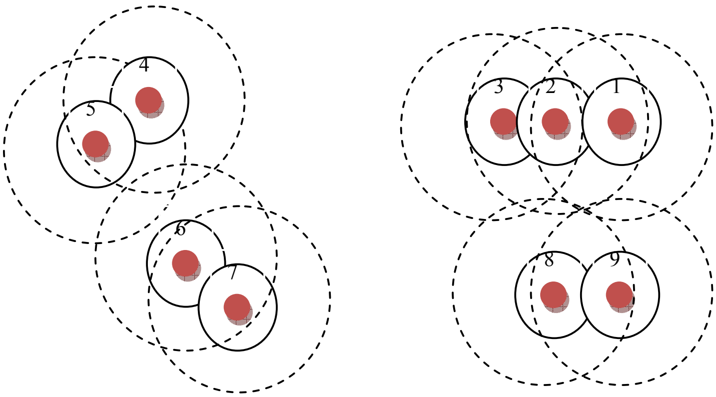


Figure 1. Reader Allocation Layout in an Exhibition

3.2 Exhibition Hall Block Network

The Length of the exhibition hall is defined as L and the width of the exhibition is W.The exhibition is divided into m blocks. The unit block size will be $\frac{L \times W}{m}$. The Block set B = {b1, b2, …, bm} is the set of unit Block in a exhibition hall. The reader set R = {r1, r2, …, rn} is the set of all readers in an exhibition hall . Each reader has the same available frequency channel F = {f1, f2, …, fo}. The readers in block bi are grouped together. Readers in a block work with the same frequency channel. Only one reader in a block is allowed to read tags for a timeslot defined as ts. The reader amount set K = {k1, k2, …, km} is the set of reader amount of each block in the exhibition hall. For example, there are ki readers in Block Bi, the cycle time is ki*ts. Each reader in Block Bi is given a polling sequence set S = {1, 2, …, ki}. For example, Figure 2 is the division of the exhibition hall shown in Figure 1.

Reader 4 and Reader 5 are grouped together with same block number 1. Therefore, the reader amount in block 1 k1 is 2. The polling sequence of readers in block 1 is first to open and close reader 4 for 400ms, and then to open and close reader 5 for 400ms cyclically. The polling cycle time of Block 1 is 800ms. All readers in this exhibition hall are assigned block number shown in Figure 2.

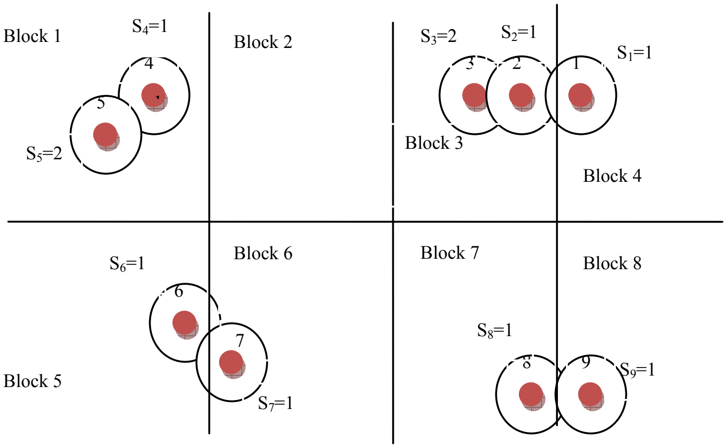


Figure 2. Block Division and Polling Sequence Assignment

3.3 Heuristic Distributed Frequency Selection (HDFS)

Step 1: Initialization

In order to avoid readers' collision, readers in adjacent blocks should be operated with different frequency channels. Ni denotes adjacent block set of block Bi.

Step 2: Adjacent Block Frequency Assignment

The assignment sequence is randomly generated. Ifthe fi is chosen from the available frequency channel F of B_i, remove f_i from the available frequency channels F of N_i. Then, the intersection of available frequency channel set T_i in N_i is checked with exception for those blocks that had been assigned. If T_i in N_i is null, another channel is reassigned to Block B_i until T_i is not null. For example, the frequency channel set of each reader is $F = \{f_1, f_2, f_3, f_4\}$. The neighbour block set of Block 1 is N_1 = {Block 2, Block 5, Block 6}. In Block 1, f_1 is assigned as the operating frequency channel. Then, remove f_1 from the frequency channel of N_1 shown in Figure 3. The intersection of available frequency channels will be $T_1 = \{f_2, f_3, f_4\}$.

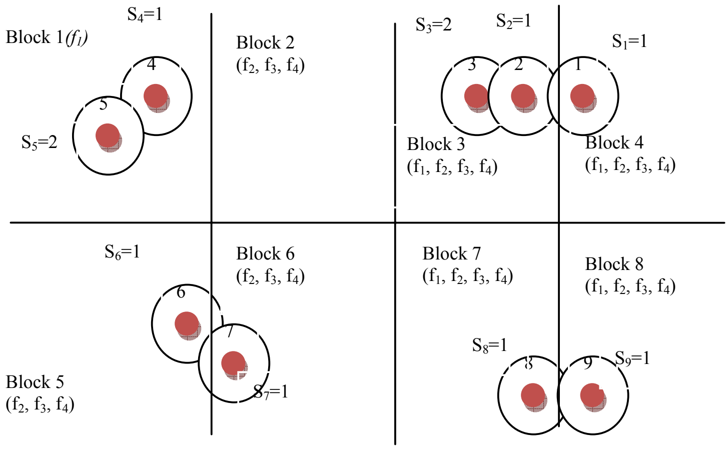


Figure 3. Available frequency channel of all blocks in exhibition

Step 3 : Kick resolution

The resolution in Step 2 is deployed into the exhibition. Then,the performance of this resolution is evaluated by calculating the average transmission rate of all readers in the exhibition.

4 Simulation Experiments

This section presents a simulation experiment to show how to implement the agorithm and hierarchical process in a conference and exhibition scenario and evaluate the reduction of readers' collision. The experiment used 9 UHF RFID readers in 11 frequency channels between 922MHz to 928MHz and adjusted the readers' transmission power in 27dBm with a 2dBi antenna. With these configurations, the readers' reading range is 1.2 meters and interference range is 15 meters .

In the first step, The experiment's area is partitioned into blocks. The blocks' width and length are equal to the reader's interference range, 15 meters. Then 9 readers are randomly put in the experiment's area. After assigning readers into blocks, each reader could get its frequency channel according to the algorism. The second step is determining polling sequence for each block. Readers in the same block will be listed in a reading group. Each of them are assigned an own timeslot of reading for 400 millisecond. After its reading slot, the reader stops polling until the group completes a full rotation and polls another 400 millisecond. Figure 4 presents the experiment's environment and the readers' parameters.

The RFID system's architecture shows in Figure 5. All readers are connected to a central backend server by Ethernet. The server controls each reader and receives the tags' content before logs them into database.

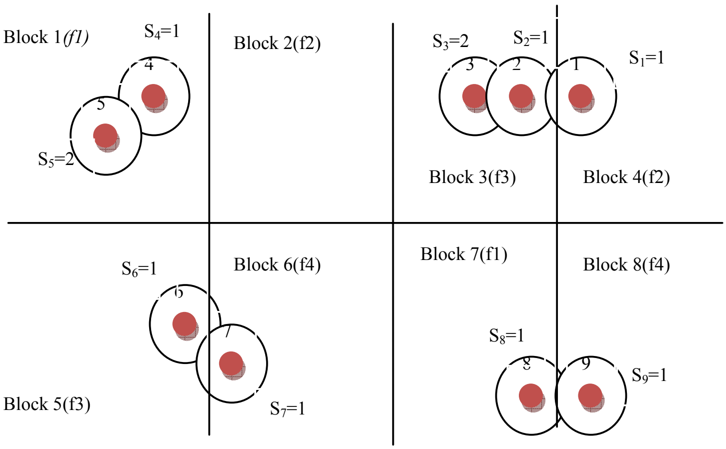


Figure 4. Experiment environment

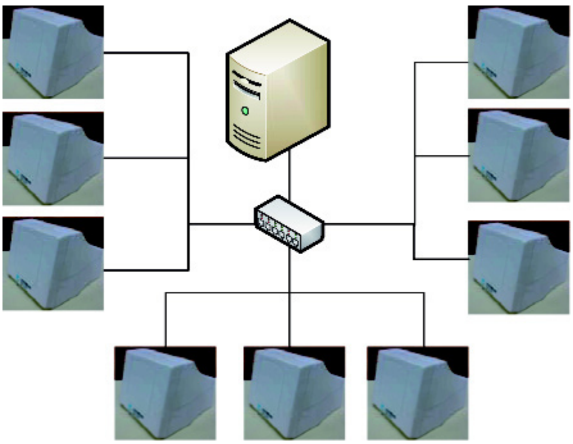


Figure 5. System's architecture

A successful hit is counted if a reader sends signals and receives the response with unique identification number from a tag in the read range. The reader's reading rate are defined as followings.

$$\text{Reading Rate of a Reader} = \frac{\text{Successful Hit counts}}{\text{Total time}}$$

Then the readers are tested in a chamber that can absorb radio waves without interference of other RF. The reading rates of readers in this chamber are the basis of reading performance of readers. As a result, the reading efficiency of a dense reader exhibition is as followings.

$$\text{Reading Efficiency of a Dense Reader exhibition} = \frac{\text{Average Reading Rate of all readers in a dense Reader exhibition}}{\text{Average Reading Rate of all readers in the none interference chamber}}$$

The experiments were conducted and studied the effects of proposed anti-collision algorithm on 9 readers for 30 runs in a small exhibition room. The reading efficiency of readers in different is shown in Figure 6. Following are the results:

(1). The efficiency of a RFID collided exhibition is about 40%, which means 60% of transmissions in the exhibition are wasted due to readers' collisions.

(2). The efficiency of a RFID anti-collision exhibition is about 80 % and outperforms the one of the none-interference chamber in some runs. As a result, the proposed RF anti-collision algorithm is successful in reducing the possibilities of RF collisions problems.

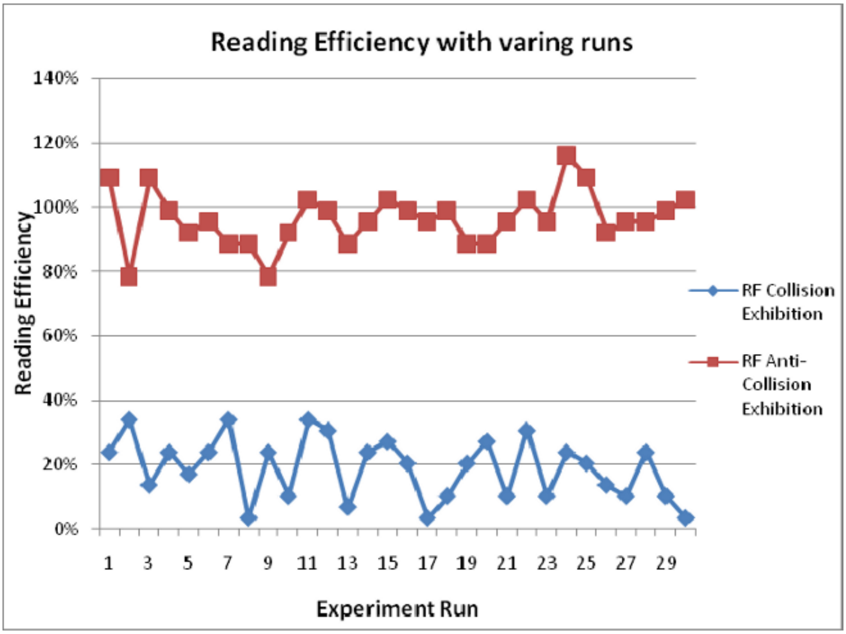


Figure 6. Experiment's result

5 Conclusion

This paper presents a deployment methodology to prevent the RF collision in a small exhibition room with a heuristic distributed frequency selection algorithm. The experiment results shows, the algorithm can provide an effective collision reduction control mechanism.

This paper develops an algorithm and a hierarchical process for quickly deploying the RFID exhibition system and reducing the interference between RFID readers. With the hierarchical process, a simulation experiment is conducted to test the proposed anti-collision algorithm. According to our HDFS progress present in this paper, it could be a good way to resolve RF anti-collision issues.

Future research will be focused on two broader areas of activity. One is the study of reflection and collision differences between chamber and exhibition environment. The other is how to analytics the behaviours from visitors' routing path and collection data for further value adding.

6 Citations and References

[1] Waldrop J, Engels DW, Sarma SE, Colorwave: An anticollision algorithm for the reader collision problem. IEEE International Conference on Communications 2003:2:1206-1210.
[2] Birari SM, Iyer S, PULSE: A MAC protocol for RFID networks, Lecture Notes in Computer Science 2005:3238:1036-1046.
[3] Shih DH, Sun PL, Huang SM, Taxonomy and survey of RFID anti-collision protocols, Computer Communications 2006:29:2150-2166.
[4] Myung J, Lee W, Adaptive splitting protocols for RFID tag collision arbitration, International Symposium on Mobile Ad Hoc Networking & Computing 2006:202-213.
[5] Maina JY, Mickle MH, Lovell MR, Schaefer LA, Application of CDMA for anti-collision and increased read efficiency of multiple RFID tags. Journal of Manufacturing Systems 2007:26:27-43.
[6] Wu Jie, Wang Dong, Sheng Huanye, Building a context-aware smart exhibition environment, International Conference on Wireless Communications, Networking and Mobile Computing 2007:2177-2180.
[7] Zhou S, Luo Z, Wong E, Tan CJ, Luo J, Interconnected RFID reader collision model and its application in reader anti-collision, IEEE International Conference on RFID 2007:212-219.
[8] Yu J, Kiu K, Yan G, A novel RFID anti-collision algorithm, International Conference on Wireless Communications, Networking and Mobile Computing 2008:4:1-4.
[9] Lara RA, Har C, Reassessing the Need for the Development of Regional Standards for the MICE Sector for the ASEAN and Asia Pacific Region, Journal of Convention & Event Tourism 2008:9:3:161-181.

Networking Dual-Pair-Tele-Paths for Logistic and Parking Structures with RFID Applications

Li-Yen Hsu[1]

Associate Professor, Department of Aviation Management, China Institute of Technology, Taiwan ROC.

Abstract. Parking and logistic structures, facilities that are significantly dominated by vehicles, are essential for urban or airport development. To enhance safety and sustain development, dual-surveillance based supervisory control and data acquisition (SCADA) networks with radio-frequency identification (RFID) applications are modeled to be spatially integrated within structures' circulation patterns. Network information tansmission is studied to establish mutually-independent hamiltonian path (MIHP) performance to cope with radio interference and multipath effects. With sequential order, integrative connectivity, and reliability, SCADA networks can also promote maintenance, environmental control, safety and security, and a customer service-oriented image, which are due to the networks' effective dual sensing, systematic fault-tolerance, and integrative management in area perspectives.

Keywords. Air freight, supervisory control and data acquisition (SCADA), telecommunication, transportation

1 Introduction

Throughout the last decade, radio frequency identification (RFID) technology has emerged for application in many areas of transportation, such as parking management and logistic processing. Parking and logistic structures are facilities that are significantly dominated by vehicles and are essential for urban or airport development. However, they are sensitive to disasters because there is relatively little intrinsic compartment protection in such a building type. Moreover, the incorporation of both security and privacy is becoming increasingly important in the application of RFID, especially on its active modes. Therefore, a supervisory control and data acquisition (SCADA) network that is integrated into the circulation and distributed intelligence is suggested to better serve customers and manage both environments and the system itself.

The examination of casino surveillance systems, where more than eight cameras may be required for a game table, helps us to recognize that the line-of-sight environment can be dynamic. One must be concerned about unexpected

[1] Corresponding Author Email : liyenhsu@ms19.hinet.net

faults. Hence, dual surveillance is recommended for a critical/dynamic environment [4]. Although the traditional approach to traffic surveillance employs the use of only one or two cameras which are installed in the middle of a section of road, such an approach cannot provide left- or right-side views of vehicles in the left- or right-hand traffic system. Moreover, because traffic is dynamic, monitored targets can easily be hidden by other vehicles.

Contactless communication must address interference (Figure 1(a)) issues, noises, and ill-integrated images which are caused by multipath effects or other physical factors (Figure 1(b, c)). In order to sufficiently serve customers, manage the surrounding environment and the system itself, security-oriented SCADA networks which are placed along the path have introduced dual-surveillance capabilities. For example, these capabilities include monitoring the space or entities with two sensors/sensor nodes that are located at different positions in the vicinity for assisting space-related judgment, especially for detecting abnormal vehicular conditions or moving objects from both sides, and providing tolerance if one sensor or one sensor node is faulty.

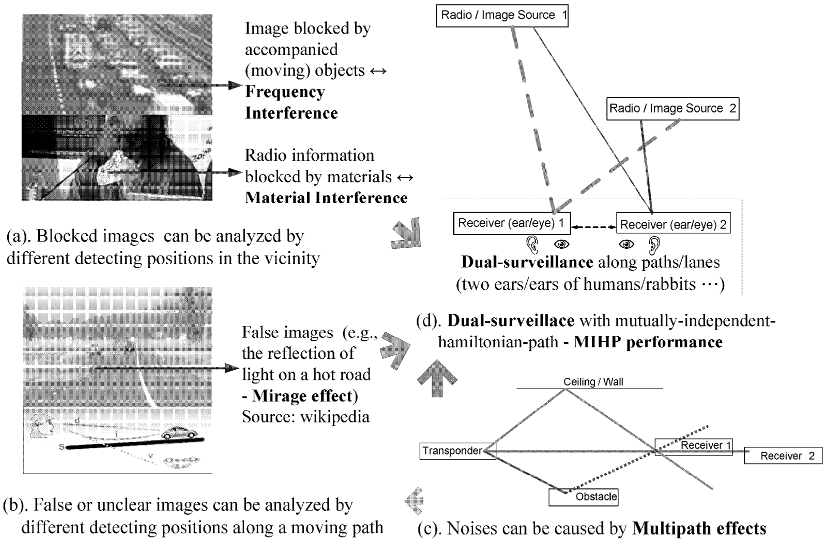


Figure 1. Illustrative case study on RFID demand of parallel surveillance and tracking.

Dual-surveillance based SCADA network prototypes can provide better functionality, just as human vision is better with two eyes instead of one. Such a network also can be capable of executing efficient and effective checking or maintenance of a SCADA network in a systematically sequential-order environment and allows for more fault-tolerance for real-time traffic management. Moreover, such networks are intended to have the performance of mutually independent hamiltonian paths (MIHP), by which various environment impact and ill-integrated images or noises can be analyzed.

In parking structure circulation paths, two rows of parking stalls are typically aligned with a vehicle path that is approximately 18 meters wide. Its center coincides with the helix (Figure 2). Six nodes are assumed along the helix at certain intervals (for example, a structure bay has two six-node rows), and the main detection direction (forward or backward) can be changed alternatively. The spider web network SW(m,n) is considered, especially SW(6,n) because it has already been found to perform using dual-pair, mutually independent hamiltonian paths. It can benefit countering multipath effects such as those present in radio frequency identification (RFID) applications. Moreover, it can offer a platform for the area-based SCADA networks, which can benefit area-oriented radio/wireless information processing and emergency response [3, 6]. The degree of SW(m,n) is optimal in terms of offering hierarchical order for maintenance or operation [1, 3].

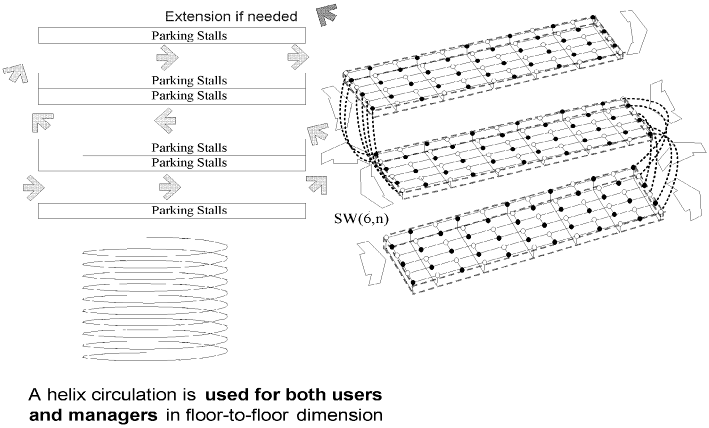


Figure 2. Integrate a dual-surveillance network with the circulation.

2 Mathematical Preliminaries

Communication networks are usually illustrated by graphs in which nodes represent processors and edges represent links between processors. Let $G=(V,E)$ be a graph if V is a finite set and E is a subset of $\{(a,b) \mid (a,b)$ is an unordered pair of $V\}$. A path is delimited by $(x_0,x_1,x_2,\ldots x_{n-1})$. A path is called a hamiltonian path if its nodes are distinct and span V. A cycle is a path of at least three nodes such that the first node is the same as the last node. A cycle is called a hamiltonian cycle or hamiltonian if its nodes are distinct except for the first node and the last node, and if they span V. A bipartite graph $G = (V,E)$ is a graph such that $V=A\cup B$ and E is a subset of $\{(a,b) \mid a\in A$ and $b\in B\}$; if G-F remains hamiltonian for any $F=\{a,b\}$ with $a\in A$ and $b\in B$, then G is 1_p-hamiltonian. A graph G is 1-edge hamiltonian if G-e is hamiltonian for any $e\in E$; moreover, if there is a hamiltonian path between any pair of nodes $\{c,d\}$ with $c\in A$ and $d\in B$, then the bipartite graph G is hamiltonian laceable.

The spider web network, SW(m,n), is the graph with the node set $\{(i,j) \mid 0\le i<m, 0\le j<n\}$, where m and n are positive integers with m,n being even, such that (i,j) and (k,l) are adjacent if they satisfy one of the following conditions: (1) $i=k$ and $j=l\pm 1$;

(2) $j=l$ and $k=i+1$(**mod** m) if $i+j$ is odd or $j=n-1$; (3) $j=l$, $k=i-1$(**mod** m) if $i+j$ is even or $j=0$. SW(m,n) is proved 1-edge hamiltonian if $n\geq 4$; 1_p-hamiltonian if $n\geq 4$ [8], Figure 2. Thus, the fault-tolerance engaged in is systematically based. Moreover, SW(m,n) are hamiltonian laceable if $n\geq 4$ [9].

The number of links connecting a node is called the degree; a network regularly has fewer degrees and is generally economic [10]. Two hamiltonian paths $P_1=(u_1, u_2,\ldots,u_{n(G)})$ and $P_2=(v_1,v_2,\ldots,v_{n(G)})$ of G from u to v are independent if $u=u_1=v_1$, $v=u_{n(G)}= v_{n(G)}$, and $u_i\neq v_i$, for every $1<i<n(G)$. A set of hamiltonian paths, $\{P_1,P_2,\ldots,P_k\}$, of G from u to v are mutually independent if any two distinct paths in the set are independent from u to v. It is found that SW(m,n), n≥4 even, has the performance of at least two mutually independent hamiltonian paths between any pair of bipartite nodes.

3 MIHP performance on Spider Web Networks

3.1 Basic Methodology Concepts on MIHP in SW(m,n)

The patterned SW(m,n) can be extended by two basic means: linear extension (Figure 3) and helix extension (Figure 4). A vertical linear extension path can be defined as L_i = (i,j), (i,j+1), …, (i,j+k), k being an even integer, i= 0,1,…,m-1, is along each column, and its adding nodes and edges are located at the same elevation of each column ; i.e., the location of the second index or j index is consistent. The helix path, such as H_i= (i,j), ($[i+1]_m$,j), ($[i+1]_m$,j+1), ($[i+2]_m$,j+1), ($[i+2]_m$,j), …, ($[i+m-1]_m$,j), ($[i+m-1]_m$,j+1), (i,j+1), (i,j+2), passes every column's two nodes in a patterned cyclic path for extension; however, every column's two nodes may be adaptively changed to different elevation. The hamiltonian path has two end nodes (x_a, y_a), (x_b,y_b), let $y_a\geq y_b$. In these research results, the linear extension is only approached in the part above y_a and below y_b (in terms of pipe-shaped SW configuration and y_a is not lower than y_b); the helix extension is only approached in the part between y_a and y_b.

By configuring basic patterns of dual-pair MIHP for the specifically targeted SW(6,n) with certain relationship of end nodes, we can extend basic patterns through trying processes to see whether they can keep the feature of dual-pair MIHP. After checking all extended basic patterns that can support dual-pair MIHP, the SW(6,n), the mainly targeted spider-web network type for the application on parking structures, for having the feature of dual-pair MIHP can be acquired.

We first consider rules for checking whether a linear extension supports MIHP. Based on the definition of mutually independent hamiltonian paths, we can assume u_i, v_i are the sequential order numbers for a node i (patterned basic node) of a pair of hamiltonian-path patterns, P_1 and P_2 (basic patterns of dual-pair MIHP) respectively, for SW(6,n), where n is a specific number.

3.2 Checking Rules for MIHP of SW(m,n)

Let a and b be the linear extension of above y_a or below y_b respectively, where a or b can be any even positive integer. These two extensions can be geometrically coded. The upper extension, a, is from w to x, where w is in the relatively higher position than x. The lower extension, b, is from y to z, where y is in a relatively higher position than z. SW(6,n) has six columns; i.e., x_a and x_b can be 0, 1, 2, 3, 4, 5. For presentation consideration, these column indices can be represented by codes A, B, C, D, E, D, F respectively.

Hence, in column F, for example, the upper extension can be presented as Fw-Fx or Fx-Fw, depending on the sequence direction that is affected by one of two end-nodes being assigned as the start. Let Fw, Fx or similar codes represent the extensive end-node cluster, and let Fw-Fx, Fx-Fw, or similar codes, represent the extending block (Figure 3, 5, 6). The number of extending blocks before a certain basic node i of two alternative extended patterns, P_{1e} and P_{2e}, which are extended from P_1 and $P_{2,}$ respectively, may be different. Let them be α_{i1}, α_{i2}, β_{i1}, β_{i2}; α_{i1} is coded for the number of upper-extending blocks before node i in P_{1e}, α_{i2} is coded for the number of upper-extending blocks before node i in P_{2e}, and similarly β_{i1} or β_{i2} is coded for the number of lower extending blocks before node i in P_{1e} or P_{2e}.

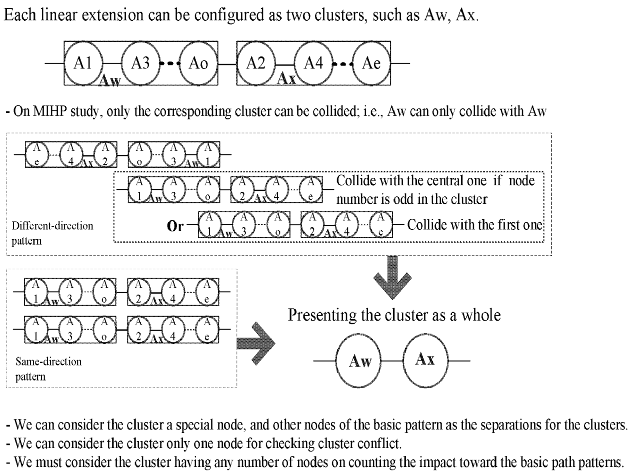


Figure 3. Node presentation for linear extension with column codes.

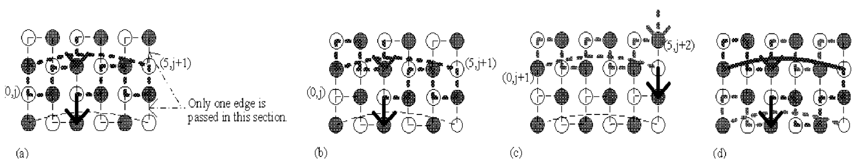


Figure 4. Helix extension. (a). typical helix extension (with start/end node on the even-number indexed column), (b). adapted helix extension (adaptively change nodes with corresponding nodes two rows below/above), (c). typical helix extension (with start/end node on the odd-number indexed column), (d). adapted helix extension for (c) on comparing (a) with lead/lag.

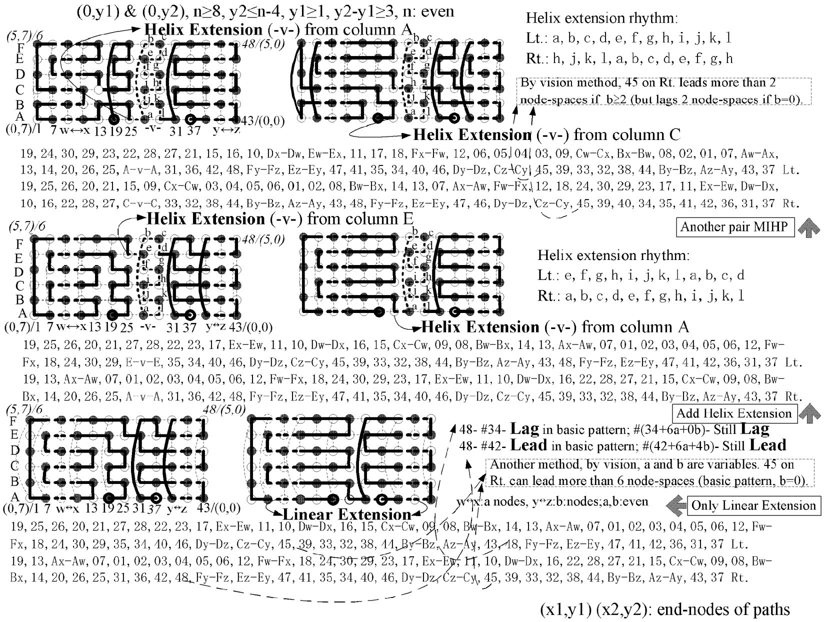


Figure 5. Concept diagram for proving dual-pair MIHP in SW(6,n).

In SW(6,n), $0 \leq \alpha_{i1}$, α_{i2}, β_{i1}, $\beta_{i2} \leq 6$. The number of nodes in each upper extending block, i.e., variable a, is naturally the same; similarly, for lower extending blocks.

Three rules can be developed from basic logical or mathematical inference for assessing dual-pair MIHP extended from P_1, P_2 to P_{1e}, P_{2e} in linear extension as the following:

(a). The extensive end-node cluster can only collide with the same coded end-node cluster in another extended pattern alternative at the same sequential position. The extensive end-node cluster will not collide with the patterned basic nodes and other extensive end-node clusters, it only can collide with the extensive end-node cluster of the same attribute; i.e., Fw can only collide with Fw of another extended pattern alternative in the same sequential position.

(b). It is not necessary to have both the upper and lower extensions; in other words, a or b can be zero. After checking rule (a), check every basic node, generally named i here, whether it can satisfy the following condition. Initial basic patterns of dual-pair MIHP boundary conditions can be extended and checked every two rows of nodes; moreover, once if $u_i > v_i$ (i.e., the sequence of node i in P_1 leads that in P_2) and $(u_i+\alpha_{i1}*a+\beta_{i1}*b)>(v_i+\alpha_{i2}*a+\beta_{i2}*b)$ (i.e., the sequence of node i in P_{1e} still leads that in P_{2e}), or if $u_i < v_i$ (the sequence of node i in P_1 lags behind that in P_2) and $(u_i+\alpha_{i1}*a+\beta_{i1}*b)<(v_i+\alpha_{i2}*a+\beta_{i2}*b)$ [2-5] (i.e., the sequence of node i in P_{1e} still lags behind that in P_{2e}), then dual-pair MIHP exists in P_{1e} and P_{2e} if helix extension does not need to be concerned temporarily, where a or b can be infinity or a number within a certain scope, depending on trying concerns; however, infinity is generally targeted. Moreover, α_{i1}, α_{i2}, β_{i1}, and β_{i2} values should cope with the geometrical relationship; hence, even $u_i<v_i$ and $(u_i+\alpha_{i1}*a+\beta_{i1}*b)>(v_i+\alpha_{i2}*a+\beta_{i2}*b)$ the MIHP still may exist if the lead or lag relationship can be consistently kept after initial conditions.

On horizontal extension [7]; i.e., considering on SW(m,n), not all cases with MIHP performance can be horizontally extended from SW(6,n) or even SW(4,n) to SW(m,n), m≥4 and n≥4 being any even integers. Generally, those cases with MIHP performance, which can be horizontally extended, generally have similar appearances or some similar sequences in a part of their configurations. Such a consideration on appearance for MIHP study is rational because extensive factors with same attributes, whose appearances rationally have some common features, may be crossed out, i.e., expelling the necessity to count some uncontrollable factors, to helpfully keep the original lead/lag relationship or no-collision relationship after at least a certain scope of m or n. Such as: if $u_i<v_i$ then $(u_i+\alpha_{i1}*a+\beta_{i1}*b+\sum_0^{s=s1-1}\gamma_{i1s}*H_r+\sum_0^{t=t1-1}\lambda_{i1t}*H_l) < (v_i+\alpha_{i2}*a+\beta_{i2}*b+\sum_0^{s=s2-1}\gamma_{i2s}*H_r+\sum_0^{t=t2-1}\lambda_{i2t}*H_l)$.

Or if if $u_i>v_i$ then $(u_i+\alpha_{i1}*a+\beta_{i1}*b+\sum_0^{s=s1-1}\gamma_{i1s}*H_r+\sum_0^{t=t1-1}\lambda_{i1t}*H_l) > (v_i+\alpha_{i2}*a+\beta_{i2}*b+\sum_0^{s=s2-1}\gamma_{i2s}*H_r+\sum_0^{t=t2-1}\lambda_{i2t}*H_l)$ [2].

At least, the left part cannot equal to the right part in a certain scope of m or n. We can check the following two conditions. (1). Although the lead/lag relationship will change, finally the lead/lag relationship will be kept and will not collide before the change. (2). The separation of their sequence is a number which cannot be equaled by a constantly increasing means; e.g., the separation is six, and the constantly increasing means is the multiple of four, six cannot be equal to four or its multiples.

On the previous formula, $\sum_0^{s=s1-1}\gamma_{i1s}$ represents the total number of nodes after the location-right insertion on the edges which affects the sequence of node i in path 1. On γ_{i1s} or γ_{i2s}, the code γ represents the embedding with right insertion which locates at the right side of both end nodes of the pipe-shape presentation of the case, with i representing node i, with 1 or 2 representing path pattern 1 or path pattern 2, with s as an index for counting the number of affected path-edges in embedding, and with s1 as the real number of affected path-edges in embedding path 1. The approach $\sum_0^{s=s2-1}\gamma_{i2s}$ in path 2 is similar to that in path 1. H_r is the number of horizontal extension in the location-right part; infinity is targeted.

Similarly, $\sum_0^{t=t1-1}\lambda_{i1t}$ represents the total number of nodes after the location-left insertion on the edges which affects the sequence of node i in path 1; similar to γ_{i1s} or γ_{i2s} basically, yet the code λ represents the left insertion which locates at the left side of the end node on the column without the origin (0,0) of the pipe-shape presentation of the network case, with t as an index for counting the number of affecting cut edges in insertion, and with t1 as the real number of affected path-edges in embedding path 1. The approach $\sum_0^{t=t2-1}\lambda_{i2t}$ in path 2 is similar to that in path 1. H_l is the number of horizontal extension in the location-left part; infinity is targeted.

Location-left insertion may be considered after the location-right insertion or may not need to be considered due to the following reasons. (1). Both end nodes may be arranged at the same column (assumed with the origin (0,0)). (2). Based on the study result, generally the horizontal extension happens after m≥6 and the general presentation in the research can provide more column bays for location-right insertion. The possibility of linear extension in $\gamma_{i1s,}$ $\gamma_{i2s,}$ $\lambda_{i1s,}$ or λ_{i2s} need be considered; i.e., the node number of linear extension surely can be infinity. The possibility of location-right insertion or location-left insertion is also infinity targeted.

(c). If a helix extension exists, the linear extension for dual-pair MIHP can be checked independently. In other words, the sequence of nodes in the central part is ignored temporarily, and we apply rule (b) only to either the upper or lower extension each time (we may see the network in an upside-down way).

The typical helix extension connects two rows of nodes. In SW(6,n) or SW(4,n), 12 or 8 nodes are linked as a path unit in a cyclic, regular, and extensive approach. The extension path of path 1, P_{1e} and the extension path of path 1, P_{2e} may have only helix extension.

The typical helix extension essentially will separate the whole network to three parts (Figure 5, 6). The upper part or the lower part can be independent for applying the rules to the previous sub-section because the end nodes of the circulation extension are located on specific coordinates with only one edge connecting the upper part and the lower part. In other words, if the helix extensions for both P_{1e} and P_{2e} happen at the same two rows, P_{1e} and P_{2e} can have the same number of nodes in either the upper (extension) part or the lower (extension) part, which consequently can be considered independently, as mentioned before. Then, both the start and the end of the helix extensions for P_{1e} and P_{2e} can have the same sequence in the whole path, but all the corresponding nodes are different because original basic paths are mutually independent and helix extension is essentially cyclic.

Depending on whether the column index of end nodes is even or odd, the helix extension may be classified into two basic types. In these two types, the edges connecting the upper part and the lower part are located at different row separations. However, for clear presentation, their presentations are basically the same in this article, i.e., such as F-v-F, where -v- represents a helix extension, F or 5 is the index of the start/end column. Consequently, we can easily know that the start/end column index of the typical helix extension is odd (or even).

If the typical helix extensions for both P_{1e} and P_{2e} happen at different two rows, P_{1e} and P_{2e}, for example, can naturally have different number of nodes in either the upper (extension) part or the lower (extension) part. However, lead/lag with an assured number can be counted. Moreover, at most after one round of helix circulation, a same group of two rows' nodes of any helix extension for P_{1e} or P_{2e} can be found but with a lead/lag relationship each other. Once such MIHP performance in a group has been assured in this perspective, this performance can be assured for any possible vertical helix extension in the same mode.

If two typical helix extensions have different mode – from even or odd numbered column, one of them may be presented with inverse v or ^ when they are figured; however, another concept, the adapted helix extension may be used instead. Typical helix extension pattern can be adapted if some nodes in extension are regularly changed by corresponding nodes with two-node separations, such as (0,j), (1,j) are changed and are corresponded by (0,j+2), (1,j+2) respectively (Figure 4); "~v~" can be used for representing adapted helix extension. An adapted helix extension still can recursively, sequentially passes two rows of nodes for vertically extending the basic path pattern. Moreover, once such MIHP performance in a group has been assured in this perspective, this performance can be assured for any possible vertical helix extension in the same mode. Because the embedded node numbers are same, and the embedding location can naturally have corresponded relationship, hence, whether the horizontal helix extension can be easily checked and inferred from the rhythm of embedding.

One of typical example MIHP case is shown on Figure 6.

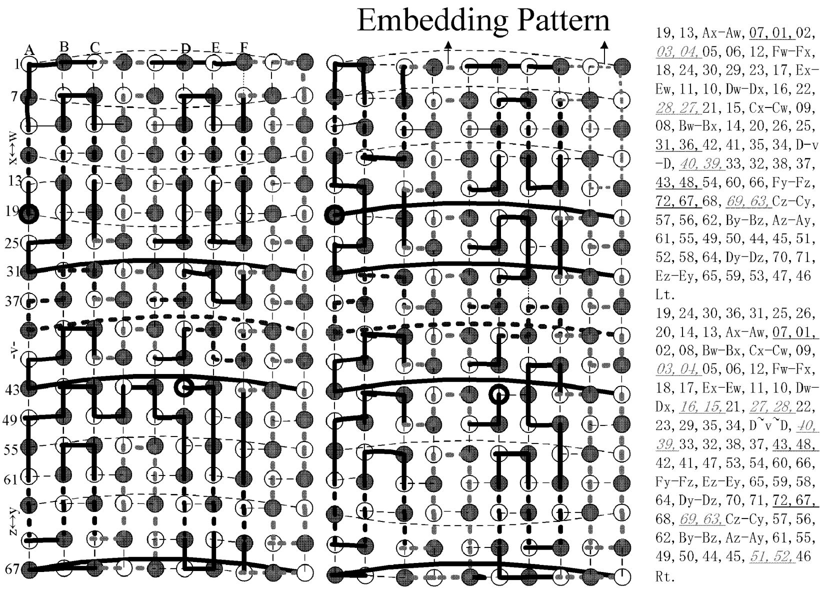


Figure 6. An example pattern of SW(m,n) having MIHP performance.

4 Significance

4.1 Reliable Parallelism Supporting Network Performance

False detections may happen for many reasons such as multipath effects, node or link (transmission) faults, and a combination of the two. After configuring adaptable dual-surveillance as a basic detection-availability platform, systematic fault-tolerance, connectivity, and management efficiency enhances detection-availability. Furthermore, a diagnostic performance MIHP can be established to analyze time-series related adversary conditions using independent, alternative, time-series recordings for data mining or diagnosing problems. This is similar to physicians employing independent alternatives and time-series records to diagnose a disease.

It must be assumed that rivals are sometimes privy to ciphered information. In order to protect infrastructures, confidential information must be transmitted without alerting such individuals. Therefore, depending only on ciphered words may not always be the best approach for transmitting security-related information. A security network also requires routine maintenance, auditing, and accounting, aspects for which radio-frequency identification can naturally be considered. Such tasks can be operated or managed by different authoritative hierarchies and initiated on random occasions. To maximize efficiency and effectiveness, such routine tasks can be operated in an orderly way through either the entire or the defined part of the network. This article proposes a network in which the sequence of such orderly operation can be both flexible and logically adjusted in a hamiltonian way, because the network has hamiltonian laceability.

Such routine work may be initiated from different nodes, so that special meanings can be assigned in the operation from both active and passive sides. In other words, the routine operation can be accompanied with special meanings, including any necessary authorization or authentication from its processing sequence and operation scope. The execution of MIHP-related authentication/authorization can also be initiated via RFID, and active RFID can be better reauthorized to deal with privacy issues.

4.2 Adaptability to Freight Distribution

Senior citizens represent one of the fastest-growing segments of the population. Therefore, it is important for parking structure design to provide demand-responsive environments in order to best accommodate elderly or disabled individuals [1]. Hence, the aforementioned, highly reliable SCADA networks can offer continuous and thoughtful protection for parking structure users. It is vital to make pedestrian travel safe and accessible when planning future urban infrastructure. Addressing such significant aspects of safety and welfare can benefit the economic development of various locations in the community [1, 3].

It is also important to consider the ideal story height for vans which transport the elderly or disabled individuals. The ideal height (i.e., the level of service is A / LOS-A standard) of a parking garage should be greater than nine feet. Moreover, airport parking structures are typically designed with the LOS-A standard [1]. Such an environment can offer another benefit, since parking spaces would be able to potentially provide an alternative place for freight distribution [1]. It should be mentioned that there is a trend to use smaller trucks with low floors that are quieter in congested cities [1]. Smaller truck capacity can be coordinated with a high-frequency strategy, since timing is a very important issue for contemporary freight services. Small trucks that better facilitate embarkation and disembarkation are being developed for urban distribution.

Moreover, accommodations for undersized trucks are especially important to consider for air freight, which generally uses smaller containers (or ULDs) and is currently facing a relatively high growth rate. Air cargo containers are primarily designed for use in the lower deck of an airplane, and they cannot be more than sixty-four inches high [1]. Therefore, the use of small trucks would allow air freight to be handled in our proposed parking spaces. In this study, we considered a feasible pattern for freight distribution.

4.3 Sustainability

The act of strategically keeping freight in transportation modes instead of warehousing or inventory stocking often helps to lower the total cost of operation. Hence, from the perspective of place marketing, it is a wise strategy to provide coordinated and flexible alternative spaces for freight handling. Besides, we must also consider the sustainable perspective. For example, if we fail to account for sufficient freight-handling space as an infrastructure element, the logistics of expanding cities may force the use of street space for such a task. Therefore, the result will impact the quality of and available space for pedestrian areas. Urban sprawl, which has been called land-use cancer, may also result [1, 3].

In the event of a disaster, this kind of highly reliable network system can help to prevent potential water damage because of its intricate fire protection design. However, the radio's multipath effect is of concern. Therefore, the proposed networks are better equipped to accommodate the RFID applications and make better use of the closed space, since there is a higher probability of reflection or transmission factors which involve metal or other incompatible materials –i.e., material interference. Hence, the proposed parking space can serve as an adaptable platform. It can also be used for cross-docking, which is based on the efficient integration of information, facility (including concerns about both fire protection and water damage), and total transportation systems. Moreover, since basements are often used for parking areas in East Asia, the ability to provide the proposed communication networks can allow for the use of more reliable security measures that use land resources.

The flexible and reliable nature of the proposed networks can provide more adaptable means for parking structures, such as logistic stations, museums, offices, or parking spaces for transit vehicles or rental cars Therefore, the proposed networks can contribute to sustainable urban development.

5 Conclusion

A supervisory control and data acquisition (SCADA) network with distributed intelligence is beneficial for integrally managing parking or logistic structures. Mathematical spider-web networks, SW(m,n), can have conformity, adaptability as interference-free oriented cellular telecommunication networks and as dual-surveillance based (SCADA) networks along paths. The degree of SW(m,n) is optimal in offering hierarchical order for maintenance or operation. For any pair of bipartite nodes, spider-web networks can offer at least two mutually independent hamiltonian paths (MIHP). With reliability, flexible connectivity, noise-reducibility, and detection-availability, spider-web networks are prototyped as SCADA networks for protecting logistic/parking structures.

Similar to physicians' employing independent alternatives and time-series recordings to correctly diagnose the disease, abnormally changing signals and multipath effects of radio/wireless communication can be analyzed via MIHP performance, to improve information reception and to prevent false alarms via independent time-series recordings. Classified authentication/authorization can be dynamically inherently implied via the specified sequence of routine operation.

A specific spider web network, SW(6,n), can well fit the network of cellular communication and the (interior) path configuration of parking structures; it also has been found having the performance of dual-pair MIHP. MIHP performance can help to better promote radio frequency identification (RFID) applications because they are essentially wireless communication based. Such performance also can help to well use parking spaces located at basements, where probabilities of metal interference are relatively high.

Wayfinding assistance, energy conservation, and other environmentally sustainable operations can be designed within parking structures' SCADA networks. The system's sequential order can also allow for the management of parking structures while supporting a company's customer-conscious image, since it can provide services for the elderly and the handicapped. The proposed networks can reduce the probability of using damaging fire protection equipment. Consequently, parking structures can provide a sustainable means for other possible uses, such as express-delivery stations, offices, or parking for transit vehicles. Hence, the proposed networks can be extremely beneficial to sustainable urban development.

6 References

[1] HSU, L.-Y. Networking Dual-Surveillance/Dual-Tele-Paths to Sustain Parking Structures. IEEE Conference on Measuring Technology and Mechatronics Automation, Printing House, Zhangjiajie, Hunan, 3: 530-535, 2009.
[2] HSU, L.-Y. Shaping Parallelism in Pervasive Spider-Web Communication Networks. Proceedings of PDPTA (Parallel and Distributed Processing Techniques and Applications) '09, CSREA, Bogart, GA, accepted, 2009.
[3] HSU, L.-Y. Networking Dual-Surveillance/Dual-Pair-Tele-Paths for Critical Urban Areas. Journal of Southeast University - English Edition, Southeast University, Nanjing, 2008.
[4] HSU, L.-Y. Developing Critical/Dynamical Area Security with Dual-Surveillance/Radial-Ring Configured Networks. IEEE Conference on Technologies for Homeland Security, Printing House, Boston, 240-245, 2008.
[5] HSU, L.-Y. Developing Dual-Pair Parallel Processing for Pervasive Communication with Spider-Web Networks. Proceedings of PDPTA (Parallel and Distributed Processing Techniques and Applications) '08, CSREA, Bogart, GA, 200-206, 2008.
[6] HSU, L.-Y.; KAO, S.-S. Developing Dual-Surveillance Based Vehicle Security Management with Generalized Honeycomb Tori. Proceedings of SAM (Security and Management) '06, CSREA, Bogart, GA, 314-320, 2006.
[7] HSU, L.-Y.; KAO, S.-S. Extending Security/Sustainability Through Pervasive Spider Web Networks in Urbanism. Proceedings of SAM (Security and Management) '07, CSREA, Bogart, GA, 230-235, 2007.
[8] KAO S.-S.; HSU, L.-H. Spider Web Networks: A Family of Optimal, Fault Tolerant, Hamiltonian Bipartite Graphs. Applied Mathematics & Computation, 160(1): 269-282, 2005.
[9] KAO, S.-S.; HSU, L.-H. Hamiltonian Laceability of Spider Web Networks. Chung Yuan Journal (Taiwan), 33(1): 1-10, 2005.
[10] STOJMENOVIC, I. Honeycomb Networks: Topological Properties & Communication Algorithms. IEEE Transactions on Parallel & Distributed Systems, 8: 1036-1042, 1997.

Applying RFID to Picking Operation in Warehouses

Kai-Ying Chen[a,1], Yu-Feng Hwang[a] and Mu-Chen Chen[b]

[a] Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei, Taiwan, R.O.C.
[b] Institute of Traffic and Transportation, National Chiao Tung University, Taipei, Taiwan, R.O.C.

Abstract: In the past decades, warehouse management system has been generally considered as a bottleneck in the whole organization of business. According to the literature, the cost of order picking activity is the major expense of operating in a typical warehouse due to the introduction of new operating programs such as Just-In-Time, cycle time reduction and quick response. For these reasons, we propose a simulated RFID-based order picking system to improve efficiency of operating in the warehouse, especially deal with the important rush orders. Furthermore, a location sensing prototype system with the application of RFID and back-propagation algorithm for locating objects is also presented in this research. The main advantage of applying RFID sensing system is to enhance efficiency of operating via novel dispatching way based on instantaneous message advised from optimize route when the real-time information for a picker is received.

Keywords: RFID, back-propagation, picking operation, warehouse,

1 Introduction

Over the last few decades, there has been a dramatic increase in the number of applications of Radio Frequency Identification (RFID) from supply chain management to pet tracking. RFID, in simple terms, is a device of automatically identified and wireless used primarily for tracking products and transforming information. It reshapes the traditional supply chain and speeds up the operations, such as decreases significantly service time by tagging the pallets, cases, or items. Besides, RFID has the potential of helping workers correctly put right things at the right place at the right time, thus maximizing sales and profits. RFID provides the technology to identify uniquely each container, pallet, case and item being manufactured, shipped and sold, thus providing the building blocks for increasing visibility throughout the supply chain.

Supply chain or logistic network, the integration of key business process from end users to original suppliers, provides products, services, and information that

[1] Corresponding Author E-mail: kychen@ntut.edu.tw

can add values for customers and other stakeholders [1]. Supply Chain Management (SCM) was developed in the 1980s, and whose primary function was to enable companies to exchange useful information about market fluctuations and production capabilities. The warehouse management system (WMS) plays a more critical role in SCM than other business functions such as production planning, order process or billing when any member in supply chain hopes to improve their performance. Besides the influence of the supply chain, the warehouse management has many advantages such as ensuring the productivity, providing additional services, reducing transport costs, balancing required and delivered quantities, and using the market position. The functions in warehouse management can be categorized into shipping, order picking, storage, and receiving. Order picking is the most cost activity in a typical warehouse as a matter of fact. Hwang [2] proposed the cost of order picking in supply chain could be estimated as much as 55% of the total warehouse operating expense.

To facilitate warehouse management and maximize efficiency, a quickly evolving technology called RFID has become essential to track inventory units (case, pallet, item, etc) throughout the supply chain. To study benefits created from applying RFID in a warehouse, researchers have described that RFID-based WMS can reduce general operation time in warehouse processes, as well as updating inventory information. In [3], Kim's study illustrated how a third-party logistics (3PL) provider aspires to be a leader in the newly introduced 3PL industry in South Korea via innovative information technology like RFID. The RFID-based WMS indeed increased the accuracy of inventory management and cut down work time to one-third by replacing the bar code system. It reveals that RFID can lead to real-time control of warehouse operation, easy communication with the other parts of the supply chain, and high levels of automation.

In order to maximize the efficient and decrease non-value-adding activities, as well as providing a corresponding solution of solving rush orders, we propose an RFID-based order picking system to achieve these targets. In short, this research is designed to achieve the following features: (1) Development of a complete process for order picking system with the help of RFID technology. (2) Development of a new location sensing prototype system for certain purpose to instantaneously realize where each pallet truck locates. (3) Development of a Re-optimization method to enable the dispatcher to interact with and modify the generated routes while some special circumstances are happened to such as receiving rush orders.

2 RFID Technology and RFID Based Sensing System

RFID system uses radio waves to automatically identify objects or people. A typical RFID system is often consists of three principle components: (1) transponder, commercially called tag. (2) A reader with an antenna. (3) A host system. A reader in all RFID system is responsible to communicate with the transponders. Reader in all system can be reduced to two fundamental functional blocks: the control system and the HF interface, consisting of a transmitter and receiver. The reader's interface activates the transponder and supplies it with power by generating high frequency transmission power. The major goal of the reader's control unit is to communicate with the application software and the execution of commands from the application software. Transponders (or tags) which are affixed to the item that is to be tracked or identified. Transponders are either passive or active devices. If transponders are able to transmit radio frequency signals to the environment without needing power of some sort from an outside source, they are called active transponder. That is to say, if transponders need energy from the reader to work, then they are called passive transponders or semi-passive transponders.

For the most part, various sensing systems with RFID always use received radio signal strength as a sensor measurement for making an estimation of distance. For instance, Hightower et al. [4] designed the Spot-on which utilized ideas from the ad-hoc networking, and the Spot-on tags used radio signal attenuation to estimate inter-tag distance. Ni et al. [5] presented LANDMARC (Location Identification based on Dynamic Active RFID Calibration), an indoor location sensing prototype system based on RFID. Instead of disposing many RFID readers, they installed many low-cost reference tags. People or objects wore a tracking tag and traveled around to test the robust of LANDMARC. Each reader detected and measured the signal strengths of all tracking tags, as well as the reference tags. Then, the relationship between the reference tags and the tracking tags were computed by the Euclidean distance (E value). The location of each tracking tag was estimated by the k-nearest neighbor algorithm, which was approach to the location of a reference tag with minimum E value. Similar to Spot-on and LANDMARC, the Microsoft Research RADAR [6] used signal strength information to aware or track objects in a building, but the radio signal measurement are collected via an 802.11 wireless networking device, not via RFID. RADAR captured diverse features at which sensors placed, to represent and compare according to scene analysis. The disadvantages of RADAR are that shifts the perceived characteristics of the scenes may be reconstructed when environments had been changed.

Thiesse and Fleisch [7] implemented an RFID-based RTLS in a semiconductor lab. They proposed a simulation model that acquires the features of the real manufacturing process and resulted in a set of dispatching rules. Considering the electromagnetic reflections generated by complicated environment, they not only chose active RFID to be the foundation for the desired solution, but also introduced ultrasound sensors. Yun and Kim [8] pointed out that a perfect location system has not presented yet. In the other hand, as a result of the received radio signals were so sensitive that the diversity of propagation media might be present. They suggested a dual layer particle filter to surmount this problem. This layer structure was split up into two parts; the lower layer specified whether a tag was in the block level computed by a triangulation technology or the support vector machine (SVM), and the upper layer used the conventional particle filter within the pre-computed or classified block to precisely estimate the tag's location. An increasing number of recent publications and empirical studies have been done in this field to accomplish asset tracking technologies, consequently, Patil, et al. [9] showed an automatic location sensing system prototype named Bluebot that employs a robot, with an attached RFID reader. Bluebot scanned periodically surroundings and detected associating items with its own location determined with the Wi-Fi positioning system.

3 Back Propagation Network

The back-propagation is an optimization procedure based on gradient descent that adjusts weights to reduce the system error. During the learning phase, back-propagation network is partitioned into two stages: forward pass and backward pass. In the forward pass, input patterns are applied to the sensory nodes of the network and propagate its effect layer by layer through the network until a set of output patterns are generated as the response of the network. Then these output patterns are respectively compared with a group of desired or target outputs and error values are determined. The error values are used to adjust the synaptic weights in the backward pass so as to make the actual response of the network move closer to the realistic answers. Figure 1 presents the architectural layout of a multilayer perceptron. A back-propagation algorithm consists of a set of nodes that constitute the input layer, one or more hidden layers, and an output layer of computational units. That is the reason why this algorithm is also commonly referred to as multilayer perceptrons (MLP).

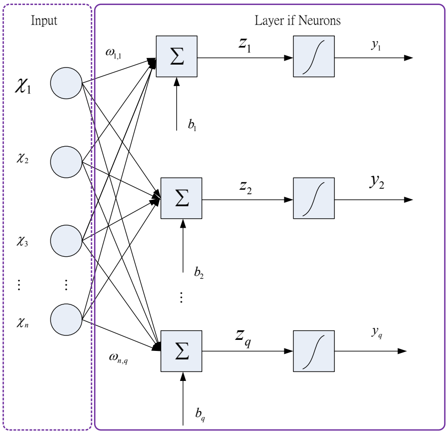


Figure 1 Architecture graph of three-layer back-propagation network.

The back-propagation is an optimization procedure based on gradient descent that adjusts weights to reduce the system error. During the learning phase, back-propagation network is partitioned into two stages: forward pass and backward pass. In the forward pass, input patterns are applied to the sensory nodes of the network and propagate its effect layer by layer through the network until a set of output patterns are generated as the response of the network. Then these output patterns are respectively compared with a group of desired or target outputs and error values are determined. The error values are used to adjust the synaptic weights in the backward pass so as to make the actual response of the network move closer to the realistic answers.

4 RFID Based Sensing System

The location tracking experiment is performed in MiBi laboratory, National Taipei University of Technology, which has many obstacles such as machines and computers in the laboratory. We place all the needed requirements onto a 1.7m × 0.8m desk and modify the distance between tags and the antenna to ensure that the RSSI value of each tag can be transmitted to the antenna. More specifically, the experimental campaign consist of a number of passive RFID tags placed as an array, and one or two antennas fixed in front of these tags. The tracking object is put into the array, and there are above 21 positions at which the tracking object can be stood. The layout of RFID array is shown in Figure 2. This system belongs to scene analysis location-sensing technique which uses features of varied spot to sum up about the location of objects in the scene. Therefore, every experiment we made that has to record all RSSI value of every object's position. These RSSI values are treated as input vectors and separated into train and test sets. The output value is the speculative position.

We conduct a series of experiments to evaluate performance of the positioning of the locating system. As a performance metric we choose the error position estimation, e, defined as the difference between the true positions of the RFID tag (x_0, y_0) we want to locate and the one computed by back-propagation algorithm (x, y), represented by:

$$e=\sqrt{(x-x_0)^2+(y-y_0)^2} \tag{1}$$

Figure 2 Layout of RFID Tag array

In order to find the set of network parameters (weights) that provide the best classification performance, we design an experiment which assesses systematically the impact of imbalanced training data on classifier performance. Including all positions, we separately collect two sizes of datasets, 244 and 1937 sample instances. Each instance corresponds to one position and contains twenty-one attributes. The neural networks used in the study were feed-forward neural networks with a single output neuron and one hidden layer consisting of fifty neurons, and the correlative settings are shown in Table 1.

Table 1. All parameters of back-propagation.

Parameters	Value
Epochs	2000
Learning rate	0.3
Momentum	0.85
Initial weight	0.3
Layers	3

5 Simulated Scenario in Warehouse

In this research, we provide an RFID-based warehouse scenario that needs to handle the important rush-orders. Assume that the initial order batches are generated according to random strategy, and consider whether the rush order should be combined with other batches based on MAD. Routes are evolved for a set of customer locations based on sales order information and are frequently modified to include rush orders. When a picker drives a pallet truck in the warehouse, the host system knows the actual position of the picker to calculate the optimal route and arrange adequate picking sequence through our proposed system. No matter which aisle that the picker is crossing through, the location can be known by the above sensing system and deliver the location messages to database. Assume that all these readers can merely cover through the tags on every side of the rack. The ideal deployment is as mentioned in Figure 3, there are some active tags attached on every side of a rack and two ultra high frequency readers under and above the rack respectively.

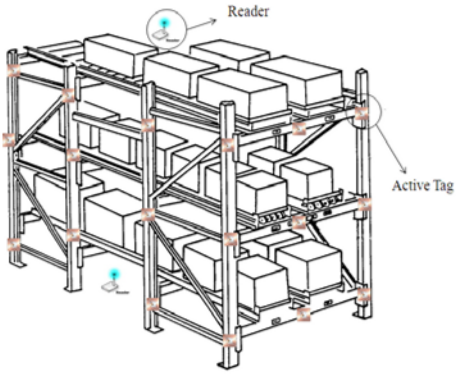


Figure 3. Layout of installing RFID-based sensing system

Figure 4 describes the implementation of re-optimization procedures which aid the dispatcher in rush order processing.

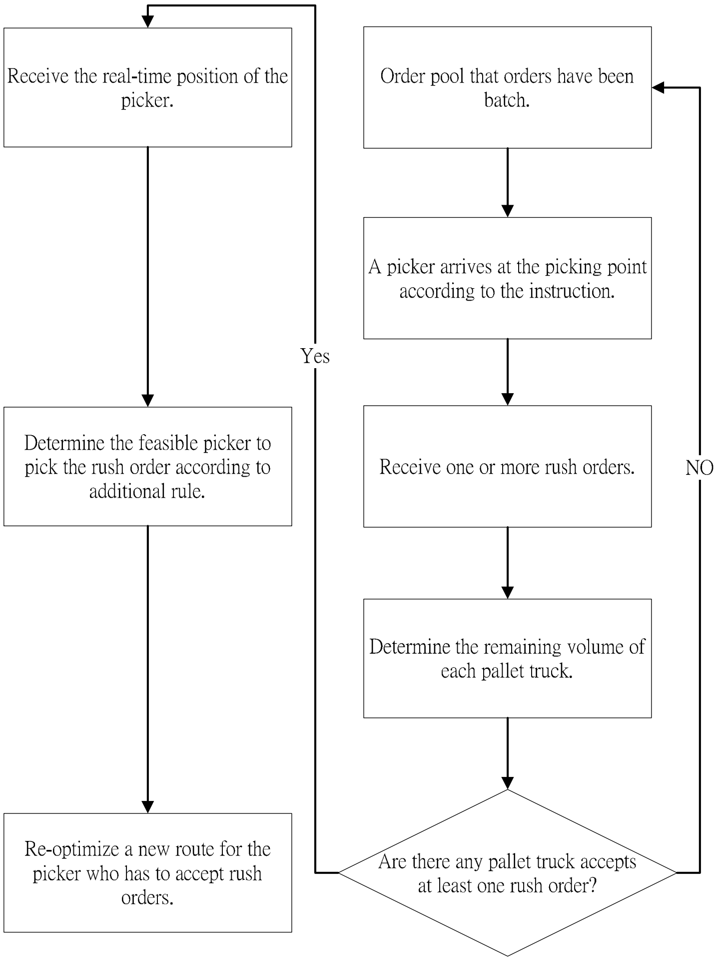


Figure 4. Rush-order processing algorithm.

Supposing that the manager has already routed every customer orders according to seed and additional algorithms and instructed the picker to load orders, and later receives a rush-order (R1 as Table 3) to be delivered along with the routed orders, assume that there are two pickers, A and B, who are picking two orders, O1 and O2, respectively. A is picking SKU3 in O1 and B is picking SKU1 in O2 while the system receives a rush order R1.

In this example, we adopt *Minimal Additional Distance (MAD)* to determine picker A or picker B to pick R1. The first step, it has to calculate the surplus capacity for each pallet truck, the surplus capacity of A is 10-(1+2+3) = 4, the remaining capacity of B is 10-(4+1) = 5, two pallet trucks do not surpass in capacity limit (10 SKU). That is, two pallet trucks can pick up R1 along with the order batch that they have picked, and then we calculate the distance while adding R1 into the present route. Before determining the feasible picker to take this rush order, we can know the actual position via our proposed system, A is at (45, 95) and B is at (175, 45). Separately adding rush order R1 to primal route by branch and bound algorithm, the additional routes of A and B are 250 units and 430 units. The optimal order pick sequences of A is SKU3→SKU2 (R1) →SKU3 (R1) →SKU1 (R1), as shown in Figure 5, the square means SKU3 that is picked by A while the manager receives R1, the triangle means the locations of picking items in R1, and these notations in Figure 6 are identical with Figure 5. The optimal order pick sequences of B is SKU1→SKU2→SKU3 (R1) →SKU2 (R1) →SKU1 (R1). According to the additional distance, we decide to instruct picker A to accept the rush order.

Table 2. Picking orders.

Picker	Order index	SKU Index	Picking location	Block	Coordinates	Volume per SKU	Quantity	Total Volume
A	1	1	13	3	(15,85)	1	1	1
A	1	2	6	2	(35,155)	2	1	2
A	1	3	35	3	(45,95)	1.5	2	3
B	2	1	86	3	(175,45)	2	2	4
B	2	2	179	3	(235,115)	1	1	1

Table 3. Rush order.

Order index	SKU Index	Picking location	Block	Coordinates	Volume per SKU	Quantity	Total Volume
1	1	47	3	(75,55)	1	1	1
1	2	59	3	(75,115)	1	1	1
1	3	61	2	(115,135)	2	1	2

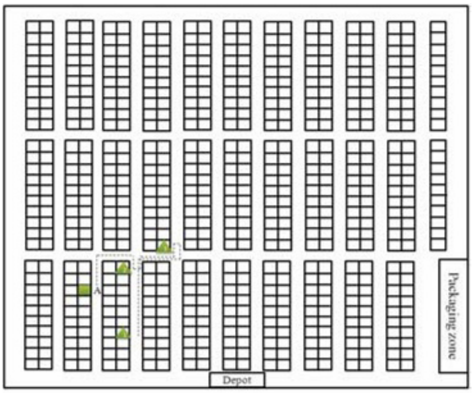


Figure 5. Additional route of picker A.

If we remove the real-time re-optimize system, the rush order has to be considered in another picking route until any picker is idle. For instance, the original optimal picking sequences of A is 0 (Depot) →SKU1→SKU2→SKU3→0 (Depot) and the distance is equal to 620 units. If R1 is accounted to the next route, the additional route is 480 units and the sequence is 0 (Depot) →SKU1→SKU2→SKU3→0 (Depot). The total traveling distance is 620+480=1100 which is longer than the optimal sequences generated by re-optimize system, 800 (0 (Depot) →SKU2→SKU1→SKU3→SKU2 (R1) →SKU3 (R1) →SKU1 (R1) → 0 (Depot)).

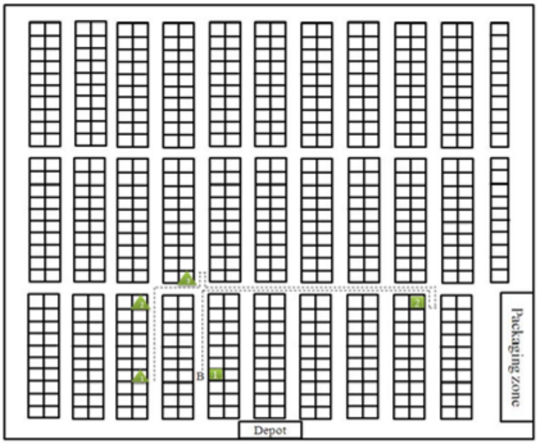


Figure 6. Additional route of picker B.

6 Conclusions

This research investigates the design and development of an RFID-based sensing system and extends the application to order picking activities under rush-order constraints. Back-propagation is adopted to solve insoluble problems with electromagnetic reflections. Despite the difficulty in avoiding the fault which renders RSSI values unstable, we still can revise the error by back-propagation so long as recording RSSI values at each position, for instance, has been done.

However, there is one limiting factor that prevents the application of our proposed system in the context of warehouses. Specifically, the observer needs to collect the features of a given warehouse so as to compare them with the previously observed dataset. Moreover, the shifting of the proposed system to another warehouse, for instance, requires the altering of the perceived features in our pre-build model and the reconstruction of the predefined dataset or the retrieval of an entirely new dataset.

The simulated scenario is an application that we consider extending the use of our system. With the proposed system, we not only obtain the real-time location of a picker to re-optimize the route, but also reduce current operation costs and assists warehouse manager in managing rush order.

7 References

[1] DOUGLAS M. LAMBERT, Supply Chain Management Processes, Partnership, Performance, Supply Chain Management Institute, 2004.

[2] HWANG, H., BAKE, W., and LEE, M. K., Clustering algorithms for order picking in an automated storage and retrieval systems. International Journal of Production Research, vol. 26, no. 2, pp. 189-201, 1988.

[3] KIM, C., KYUNG, H.Y., KIM, J., A strategy for third-party logistics systems A case analysis using the blue ocean strategy. OMEGA vol. 36, pp. 522-534, 2008.

[4] HIGHTOWER, J., BORRIELLO, G., Location systems for ubiquitous computing. Computer, vol. 34, no. 8, pp. 57-66, 2001.

[5] NI, L.M., LIU, Y., LAU, Y.C., PATIL, A.P., LANDMARC Indoor Location Sensing Using Active RFID, Wireless Networks vol. 10, no. 6, pp. 701-710, 2004.

[6] PARAMVIR, B. and VENKATA, N. P., Enhancements to the RADAR User location and tracking system, Microsoft Technical report, 2000.

[7] THIESSE, F., FLEISCH, E., On the value of location information to lot scheduling in complex manufacturing processes. International Journal of Production Economics, vol. 112, no. 2, pp. 532-547, 2008.

[8] YUN, K. and KIM, Y., Robust location tracking using a dual layer particle filter, Pervasive and Mobile Computing, vol. 3, no. 3, pp. 209-232, 2007.

[9] PATIL, A., MUNSON, J., WOOD, D., and COLE, A., Bluebot: asset tracking via robotic tracking crawling, Computer Communication, vol. 31, no. 6, pp. 1067-1077, 2008.

POC of RFID Application in Forest Sample Zone Investigation

Shiang-Shin Lin[a, 1], Teh-Chang Wu[b], Jenn-Sheng Wu[c], Yi-Ping Huang[c], Ming-Hsiung Chang[c], Sheng-Wei Fan[c], Jiun-Jiue Liao[c]

[a]Engineer, Identificaiton and Security Technology Center, Industrial Technology Research Institute, Taiwan, R. O. C.
[b]ISTC, ITRI, Taiwan, R. O. C.
[c]MSL, ITRI, Taiwan, R. O. C.

Abstract. For understanding the ecological behavior of species for afforestation and growing condition of each forest land, it is necessary to widely establish permanent sample zone. However, traditional field investigation process has lots of shortcomings, like identifying the sample tree from the marking label being difficult, the paper document being too heavy to carry, and errors occurrence when keying in the data into computer for further application. Therefore, the purpose of this proof of concept (POC) is to verify that RFID is applicable to the investigation of forest sample zone, capable of enhancing the efficiency and preciseness and reducing error. This POC is divided into three stages: (1) performing basic performance test of RFID equipments; (2) examining the developed prototype system and performing scenario simulation test; (3) performing field test at the sample zone. From the testing results, it is shown that introducing RFID into forest sample zone investigation is really applicable and can indeed improve the investigation process to enhance the preciseness and efficiency as well as reducing human error. It might be concluded that HF frequency band is a better choice for the time being, but near field UHF tag is the future consideration.

Keywords. RFID, POC, HF tag, UHF tag

1 Introduction

Forestry management has been operated over a hundred years in Taiwan. For understanding the ecological behavior of species for afforestation and growing condition of each forest land, it is necessary to widely establish permanent sample zone all over the island, setting up a complete observing system to continuously and effectively collect relative information. Via integrally analyzing this continuous investigation information, the growing mode of particular forest can be

[1] Corresponding Author : Engineer, Identification and Security Technology Center, Industrial Technology Research Institute, Rm 215, Bldg 52, 195, Sec. 4, Chung Hsing Rd, Chutugn, Hsinchu, Taiwan 1040, R. O. C.; Tel: +886-2-23577969 ext. 12; Fax: +886-2-23570854; Email: shunlin@itri.org.tw

well known and an important data bank of Taiwan forestry can be established as the base for carry out the multipurpose and ecological operation of Taiwan forest. Therefore, the Forestry Bureau who is in charge of the forestry practice of Taiwan has already performed planning and establishing permanent sample zone in all business units thereunder and executed tree investigation in each sample zone.

For improving the existing process of tree investigation, the Forestry Bureau intends to study utilizing the characteristics of Radio Frequency Identification (RFID) technology, transmitting identification information via contactless radio wave, and to apply the technology to the forest investigation of permanent sample zone, so as to achieve the following objectives,

- Introducing automated and informative operation process, to enhance the efficiency and preciseness.
- Using automated identification system instead of manpower to search and compare the sample trees, to reduce the error induced by human judgment.
- Combining RFID, handheld information equipment, and data bank/software, to achieve immediate recording and updating in field investigation.
- Being able to synchronously read previous investigation data while reexamining the trees, to provide the investigator real-time reference.
- Providing monitoring and dynamic record capable of being used for management, transportation, and sale of the forest especially for wood production, besides establishing dynamic data bank of tree growth variation.

Since there is no used case of RFID application in forest sample zone investigation internationally, a proof of concept (POC) is performed for reducing the introducing risk, supported by the Initiative Office for Government RFID Application, Ministry of Economic Affairs, Taiwan, in manpower and budget.

2 Literature Survey

Most of the literatures about the RFID application concentrate on supply chain management. Only a few of them are research or report in the application for forestry management. Dennis P. Dykstra stated that RFID labels represent a more advanced technology that holds considerable promise for use in wood chain of custody systems in 2002 [1]. There are several Research papers and reports about forestry application after that [2-5], however, they all apply RFID to log tracking system. Daniel Timpe in 2005 claimed that RFID, in an allied version, could be a good candidate to replace marking stamps, at the same time as possibly enabling the transfer of valuable information in the logistic chain of the forest industry.

Kenetics Group Ltd in their technical white paper on RFID tagging on live trees for forest management [7] summarized the possibilities of RFID forestry management system, comprising:

- Tree identification,
- Endangered tree specie protection,
- Prevention of illegal logging activities
- Supply chain management of felled lumber
- Border delineation and ownership demarcation

They also proposed a way of inserting the RFID tag onto the tree. This technical white paper gives a bright direction to this POC.

3 Analysis of Existing Operational Process

The existing live tree investigation process (shown in Fig. 1) still adopts conventional manual recording manner, having the following shortcomings:

- Using stamped aluminum plate label which is easily blurred by the weather and causing difficulty in identification.
- Requiring a great amount of paper record for each investigation, and requiring to bring data of last investigation for comparison during reexamining.
- Hard to perform comparison between tree to be investigated and the writing data while the tree identification is easy to be confused.
- Requiring a great amount of manpower to key in the measured data after the investigation for performing the follow-up data analysis and management.

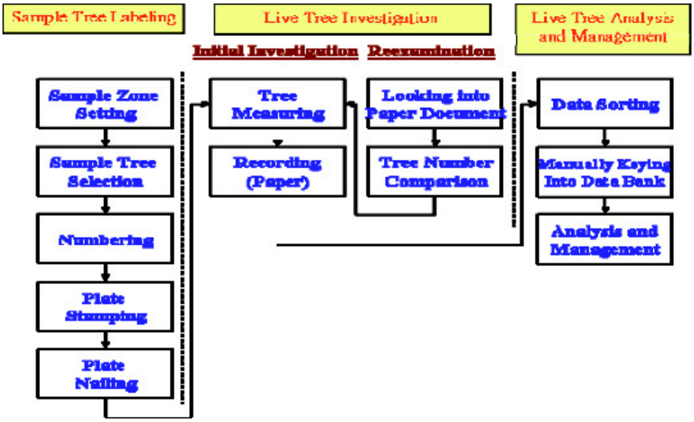


Fig. 1 Existing operational process for live tree investigation of permanent sample zone

4 Application Scenerio Planning

The application scenario is show as Fig. 2, which a passive RFID tag is inserted into a tree to be investigated for the first time; the measured information of the tree is inputted into a handheld RFID reader (a PDA); and relative information about the tree identification is written into the RFID tag by the RFID reader. For reexamination, only RFID reader is required for reading the identification information on the tag of the tree to be investigated, and inputting the updated data into the RFID reader and the tag. After completing the sample zone investigation, the investigated data can be directly transmitted to the data bank for follow-up analysis and management by just connecting the RFID reader with the computer, enabling the objectives of precision and automation in the forest investigation operation to be achieved.

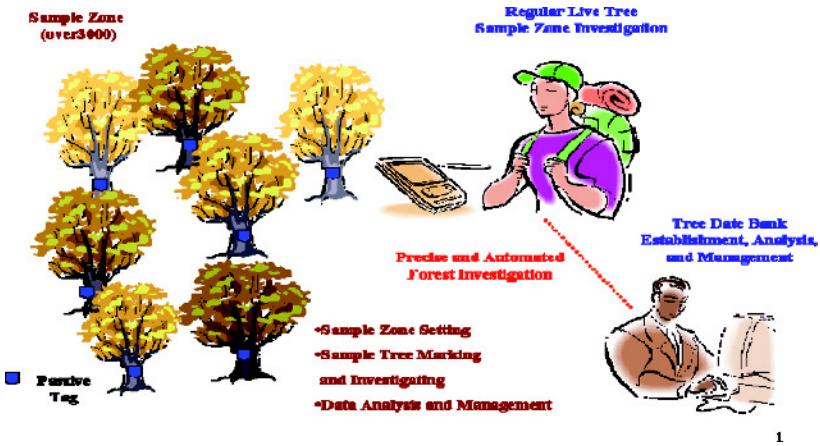


Fig. 2 Application scenario of live tree investigation after introducing RFID

After introducing RFID, the process for live tree investigation of permanent sample zone is changed as shown in Fig. 3.

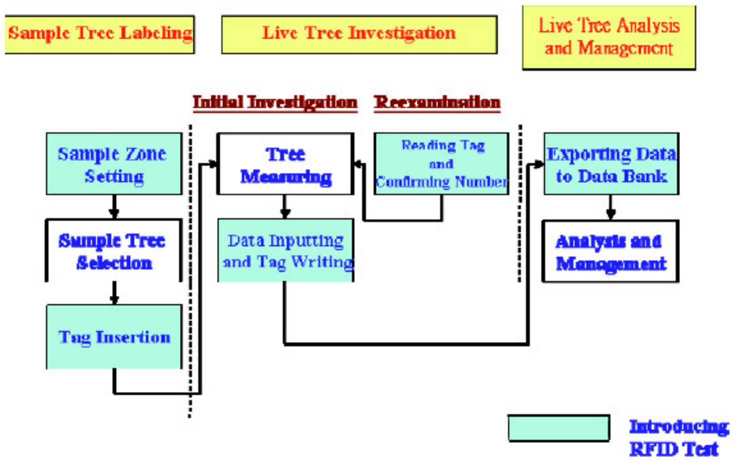


Fig. 3 Process for live tree investigation of permanent sample zone after introducing RFID

The new process after introducing RFID has the following benefits:

- The volume of RFID tag is small and easy to be inserted, without the problem of dropping off or blurred data.
- The tag is packed to be waterproof and anti-freezing
- There is no need to bring great amounts of paper data and stationary

- The problem of comparison between the data and the live tree is avoided, greatly enhancing the trustworthiness and productivity of the investigation results.
- The investigation result is directly inputted into the data bank without the effort of keying in.

5 Requirement Analysis for Software and Hardware

The new system for achieving the aforementioned object comprises RFID tag, handheld RFID reader, and application software for forest investigation. Furthermore, operational process is modified, i.e., the operation of nailing the aluminum plate is changed into inserting tag. Therefore, the requirement of these newly added or modified items need to be explored.

After being promoted in application last few years, there are a lot of RFID products in the market. From the point of view of forest investigation application, the factors needed to be considered for selecting a suitable RFID tag comprises: read distance, size, weather resistance, applicability, cost, weight, and lifetime. Since Taiwan is located in the subtropical zone, it is really damp in the forest in general. It is reported that the damp environment has great effect to the tag using the frequency band of ultrahigh frequency (UHF) or microwave, so that tag of low frequency (LF) or high frequency (HF) should be a better choice for the application in forest investigation. However, tags with different frequency bands will be tested in this POC.

Application site must be considered for selecting RFID reader. The forest investigation characterized in that it is executed in the field, all the equipment are carried by manpower, the weather varys a lot, and it is not easy to charge. Besides, while updating data of various investigated sample trees during sample zone investigation, the information of the sample zone itself needs to be inputted at the same time. According the aforementioned requirement, the specification of the reader considered comprises: light-weight, moisture resistance, impact resistance, GPS function, lasting battery power, spared battery, solar charging manner, replaceable memory card, main-stream operation system, and easiness of software development.

According to the rules of live tree investigation operation manual of the Forestry Bureau, relative investigation in the manual should be included in software design in the future, so that entire process can be accomplished at the same time, so as to achieve the objectives of process simplification and automation. Therefore, the software functions should comprise: sample zone selection, SP setting, PC setting, A tree and B tree setting, sample tree investigation data process (trunk, branch, and tag writing and reading), and data exporting to the computer from handheld RFID reader.

6 Tag Insertion Manner

In the past, an aluminum plates are nailed onto the sample trees by the Forestry Bureau for identification. However, the aluminum plate is wrapped by the tree after the tree is grown up, resulting that the investigator can not find the aluminum plate and can not identify the tree. In the literature, it is reported that the tree grows taller from the top end upward and grows out from the center outward in general. As the tree grows, when the pushing force of the tree bark acting on a disk tag is greater than the friction of the nail to the tree, the tag will be pushed outward and stay on the surface of the tree (shown in Fig. 4). If a tag can stay on the surface of the tree all the time, the read distance of the tag is not an issue anymore. For achieving this effect, the inserting manner needs to consider the following:

- Smooth nail should be adopted to reduce the friction.
- The surface area of the tag should not be too small, otherwise it could result insufficient out-pushing force and shorten the read distance.
- The nail should not be nailed too deep, otherwise it could possibly affect the tree growth and result excessive friction at the same time.
- The outer diameter of the nail should not be too large with the same reason above.
- The nail head should flat and with a suitable contact area to the tag. If the contact area is too large, it will cover the tag and shorten the read distance; and if the contact area is too small, the nail might separate with the tag or the contact portion might be deformed and damaged while the tag is pushed during the tree growth. Rubber washer can be considered to be used on the contact surface to avoid this phenomenon.

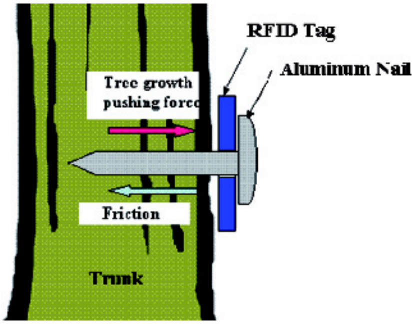


Fig. 4 Schema of force acting on RFID tag inserted on the tree

7 Basic Performance Test

The first stage of the POC is basic performance test. Tags of three different frequency bands are tested, which the first one is HF, considering that the existing packing manner is easy to be applied, the second is UHF, considering that the reading distance is longer, and the third is microwave (2.45 GHz), considering that the antenna can be smaller. The identification code of RFID tags and readers tested are listed in Table 1 and 2, respectively.

Table 1 RFID tags tested

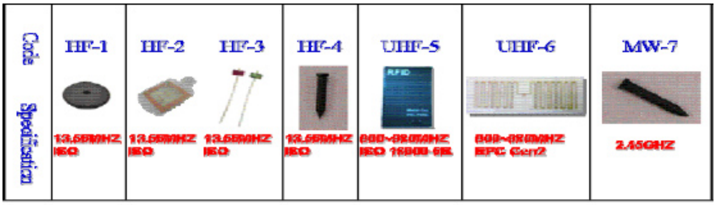


Table 2 RFID readers tested

The test results are listed in Table 3. The original anticipation of the Forestry Bureau is that the normal read distance can be around 50 cm read by handheld reader to facilitate searching. It was shown that two kinds of UHF tags tested can satisfy the aforementioned requirement in normal condition, the read distance of the 3 kinds of HF tags are less than 10 cm wherein the read distance of reader 3 is 4-5 cm and read distance of reader 4 is 6.1-7.5 cm, and the read distance of the microwave tag 11 and reader 5 combination is 9 cm.

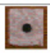
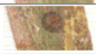
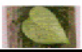
Table 3 Test result of first stage

Combination	Tag Code	Reader Code	Max. Normal Read Distance (cm)
1	HF-1	R3	4.0
2	HF-2	R3	5.0
3	HF-3	R3	4.0
4	HF-4	R3	0.3
5	HF-1	R4	7.5
6	HF-2	R4	6.8
7	HF-3	R4	6.1
8	HF-4	R4	3.0
9	UHF-5	R1	48.0
10	UHF-6	R2	77.0
11	MW-7	R5	9.0

8 Scenario Simulation Test

The second stage is performing scenario simulation test to compare the results of tag and reader combinations at different environments, and tags with different inserting manner. The comparison of tag and reader combination at different simulated environments is shown in table 4. Obviously, the moisture resistance of HF tag is pretty good, which the moisture has nearly no effect on the read distance. However, the read distance of UHF tag under damp environment is greatly reduced. Surprisingly, the microwave tag appears little effected by the moisture.

Table 4 Comparison of tag and reader combination at different simulated environments

Environment	HF-1 R4	UHF-5 R1	UHF-6 R2	MW-7 R5
Normal Condition	7.5 	48.0	77.0	9.0
Wet Paper Rear	7.1 	53.0	42.0	7.5
Wet Paper Front	7.2 	21.5	39.5	7.0
Wood-Dry	7.7 	51.0	46.0	9.0
Wood-Wet	6.7 	44.5	39.0	7.5
Mud Covered	7.2 	27.0	43.5	7.7
Leaf Covered	7.4 	32.0	47.0	8.5

Considering the weather resistance, carrying easiness, applicability, and read distance, HF tag was selected to perform inserting manner test. Two inserting manners are planned as shown in Fig. 5.

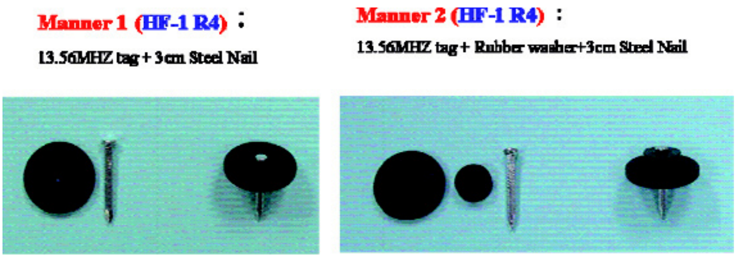


Fig. 5 Two inserting manners planned

The test results of two inserting manners are listed in Table 5 for comparison, which the read distance of inserting manner 1 are 6.7-7.2 cm, and the read distance of inserting manner 2 are 6.2-6.7 cm. Since the tags used in both inserting manner are HF tags, the moisture and mud should not be the factor causing the difference. The reduction in read distance of inserting manner 2 might be caused by the plastic washer covering part of the antenna. However, the reduction is small.

Table 5 Test results of two inserting manners

Environment	Manner 1 (Unit:cm)	Manner 2 (Unit:cm)
Normal Condition (Without Inserting)	6.7 	6.7 
Wood-Dry (Inserting into Wood)	7.2 	6.4 
Wood-Wet (Inserting into Wood)	7.0 	6.4 
Wood-Mud Covered (Inserting into Wood)	7.2 	6.2 
Wood-Leaf Covered (Inserting into Wood)	7.2 	6.4 

9 Verification of Developed Application Software

The major function modules of forest sample zone investigation system are as follows,

- Sample zone setting module: setting fundamental information of the sample zone to be investigated, comprising investigator, SP (starting point), PC (point of center), and GPS coordinate of sample zone, A tree, and B tree.
- Sample tree investigation module: measuring and collecting sample tree data, the process being divided into initial investigation and reexamination, and the content comprising sample tree setup, trunk setup, new branch setup, and read and write RFID Tag.
- Data exporting module: exporting the data of sample zone and sample tree in the PDA to the PC end for follow-up process.

Fig. 6 Appearance picture of reading/writing tree RFID tag

10 Field Test

For effectively verifying the benefit of RFID application in forest sample zone investigation, this team performed several field tests at Wuliao Sample Zone located at Sansia Township, Taipei County, during November, 2007, under the assistant of the Forestry Bureau. The operation process fully refers to the rules of forest permanent sample zone investigation operation manual of the Forestry Bureau to execute the following relative investigation operation in sequence:

- Setting up the investigation instruments and equipments.
- Setting SP, measuring SP to PC, setting the size of sample zone and sample zone line, setting reference trees (A tree and B tree).
- Nailing the aluminum numbering plate, rubber washer, and RFID tag onto the sample tree with aluminum nail.
- Identifying the tree species, measuring the sample tree data of sample zone line distance, sample tree line distance, trunk diameter, tree height, trunk branching height, branch numbering, branch diameter and height, and branch branching height, and inputting into PDA.
- Writing the sample tree data in the PDA into tag.
- Read the sample tree data using RFID reader to confirm the sample tree data being correctly written into the tag.
- Taking a photo of the tree height
- Importing the investigated data into PC from the handheld RFID reader

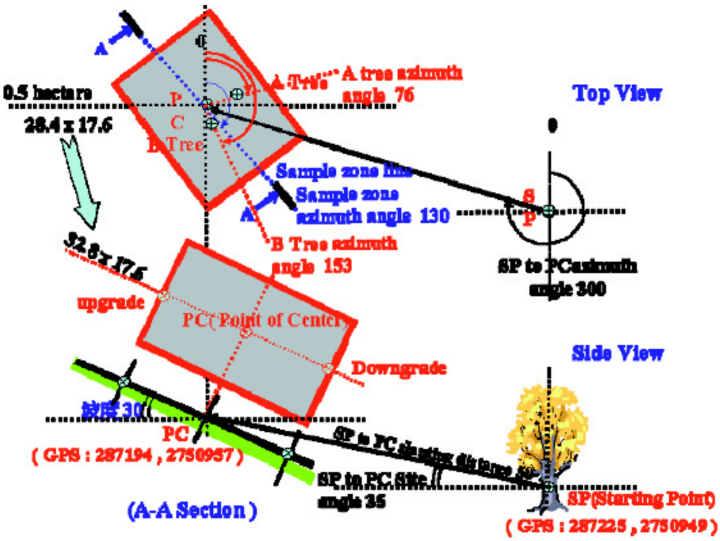


Fig. 7 Schema of sample zone setting of field test

Fig. 8 Sample tree and RFID tag of the field test

For field testing, an aluminum nail of 5 cm is used to nail aluminum numbering plate, rubber washer, and RFID tag in sequence on the trunk, the depth being 2 cm. After measuring the live tree, inputting sample tree data into PDA, and writing into the tag, the read/write are all 100% correct while confirming with RFID reader, even the tag being covered with water or mud.

Using handheld RFID reader in field test, the average read time for a tag is about 2.1 seconds and average write time is about 6.2 seconds, conforming to the anticipation of the user in substance. During testing, heavy raining was encountered. However, Reading or writing the sample tree data from or into the tag with PDA based RFID reader in the rain or after the rain shows that the read/write are all 100% correct.

11 Discussion and Conclusion

From the above testing results, it is shown that introducing RFID into forest sample zone investigation is really applicable and can indeed improve the investigation process to enhance the preciseness and efficiency as well as reducing human error. However, tag selection is important in promoting this application further. It is discussed as follows,

- HF tag: The original anticipation of the Forestry Bureau is that the normal read distance can be around 50 cm read by handheld reader to facilitate searching. Although the read distance of a HF tag is far lower than the anticipated value, the rule for inserting RFID tag and the nature of tree growth make the inserted tag stay the same height and direction which greatly assist the field investigator to identify the location of the tag. Therefore, the read distance is not an issue any more. Besides, the weather resistance is excellent, suitable package is existed in the market, and the power consumption of the HF handheld reader is much lower compared to that of a

UHF handheld reader. Accordingly, HF frequency band is a better choice for the time being.

- UHF tag: Although the read distance of a UHF tag is longer than that of a HF tag either in normal condition or under environmental effect, there are no existing UHF in the market with the package suitable for the tree insertion. However, near field UHF tag could be a possible good choice in the future, since it also has excellent weather resistance, and with the same size of package it has much better read distance than that of HF tag. Besides, since near field UHF tag uses UHF chip which is the mainstream product and the antenna can be printed using conductive ink, the cost will be cheaper. Therefore, near field UHF tag might be the best choice for this application when near field UHF handheld reader is developed and in the market.
- Microwave tag: The read distance is around the same for the nail shaped 2.45 GHz tag and is not suitable for being used on a live tree, since the tag will be covered after the tree grows up. Furthermore, Microwave RFID is not a mainstream product.

Therefore, it might be concluded that HF frequency band is a better choice for the time being, but near field UHF tag is the future consideration.

Since the possible market for this application is huge, further pilot test needs to be conducted to finalize the specification of RFID equipment, and suitable design and packaging manner especially for this application should be developed.

12 References

[1] Dennis P. Dykstra, George Kuru, Rodney Taylor, Ruth Nussbaum, William B. Magrath, Jane Story, Technologies for Wood Tracking: Verifying and Monitoring the Chain of Custody and Legal Compliance in the Timber Industry, Environment and Social Development East Asia and Pacific Region Discussion Paper, December 2002.

[2] Robert Carrigan, Ron Milton, Dan Morrow, Log Tracking System, 2005 COMPUTERWORLD HONORS CASE STUDY

[3] Claire Swedberg, Loggers Use Tags to Track Trucks, Timber, RFID Journal, Nov. 28, 2005

[4] Rhea Wessel, University in Munich Develops RFID-Enabled Log Harvesting, RFID Journal, Dec. 1, 2006

[5] Daniel Timpe, Barcode and RFID Technologies: Alternatives to Log Stamping for Wood Identification in Forestry, Rapportserie FSCN – ISSN 1650-5387 2005:33, FSCN Rapport R-05-61, Sundsvall 2005

[6] Daniel Timpe, RFID In Forestry: Prospects of an RFID-based log tracking system as an alternative to stamping, Rapportserie FSCN – ISSN 1650-5387 2006:39, FSCN Rapport R-06-63, Sundsvall 2006

[7] Technical white paper on RFID tagging on live trees for forest management, Kenetics Group Ltd., 2007

Cost Reduction of Public Transportation Systems with Information Visibility Enabled by RFID Technology

Shuo-Yan Chou[a,1,] Yulia Ekawati[a]

[a] Department of Industrial Management, National Taiwan University of Science and Technology, Taiwan, ROC

Abstract. Radio Frequency Identification (RFID) has been heralded as one of the technologies that will fundamentally transform the industries. After years of hype, the adoption of RFID remains weak. This is due in part to the technology itself but also the economy and the overall strategy of the applications. In this research, the conditions in which RFID adoption is applicable are first synthesized. The merit of RFID is then further exploited for the public transportation applications. By employing RFID along with other communication technologies, public transportation vehicles can be identified and tracked during their services. With the ability to "calibrate" where exactly such vehicles are along their routes, the arrival time for the vehicles to the stops can be estimated more accurately. Passengers can then keep shorter lead time to meet the vehicle schedule; service providers may on the other hand provide the same level of passenger service with less frequent vehicle services.

Keywords: Radio Frequency Identification, Information Visibility, Intelligent Transportation Systems, Cost and Benefit Analysis, Headway

1 Introduction

Radio Frequency Identification (RFID) has been one the most anticipated emerging technologies during the past ten years. However, many companies are still questioning if RFID is really a transformational technology or just an expensive bar code in disguise. In 2003, Wal-Mart notified its top suppliers to start using RFID for managing its supplies (RFID Journal, 2003). In 2004, the US Department of Defense announced a similar RFID mandate for its suppliers (Wyld, 2006). These mandates by important customers have resulted in the booming of RFID technology adoption, where many companies simply follow the mandates and 'slap' the RFID tags on the goods right before they were delivered, without changing business processes nor fully exploiting the true benefit of the implementation. To gain beyond this slap-and-ship process, the capability of RFID in the enhancement of business processes need to be better understood. Without

[1] Corresponding Author: Professor Shuo-Yan Chou, Tel: +886227376327, fax: +886227376344, Email address: sychou@mail.ntust.edu.tw

such justification, companies will remain reluctant in the investment and employment of RFID technologies.

The main focus of this paper is to discuss the justification of RFID application in public transportation systems by analyzing its benefit and cost. Empirical studies showed that the time passengers spend outside the transportation vehicle of choice is more onerous than the time they spend inside the vehicle in motion to the destination (Ben-Akiva *et al,* 1985). The waiting time is therefore equally important to transportation time in the public transportation business process. The expected waiting time can be measured and treated as the service level of public transportation.

Mishalani *et al.* (2006) quantified the relationship between perceived and actual waiting times experienced by passengers waiting for the arrival of a bus at a bus stop. Through the observation of 83 prospective bus passengers, it was found that the passengers perceived waiting time is greater than actual waiting time. Real-time bus location information could potentially reduce the waiting time for the passengers by facilitating better estimation of bus arrival time. Passengers can then plan on their travel more effectively. Furthermore, they may also be able to select a different bus line or an alternative mode of transportation to better manage the travel time. The reduction in waiting time and thus perceived waiting times could potentially be translated to that of operating costs or the increase of passenger satisfaction.

Gentile *et al*. (2005) proposed a general framework for determining the probability of boarding a particular line at a stop (whether to board the arriving bus or to wait for a faster one) when real-time information on the bus location is provided to passengers. It is shown that such information may have a significant impact on the use of the transit line. For example, in the case with real-time information, the expected waiting time increases with the regularity of the service as passengers accept a longer wait to board a line that takes less time to reach the destination. That is why the general principle in evaluating passenger information services is that passengers derive both intrinsic benefit from information and that they may use that information in ways allowing them to make better travel decisions (Hickman, 2006). As previous research concluded on the benefit of online information at bus stops, Menezes *et al.* (2006) proposed the use of RFID technology to solve problems faced by passengers and bus operators in metropolis, and identified challenges in applying the framework, particularly on data management, real time decision making, and scalability.

The performance of public transportation systems can be improved by using RFID technology, as supported by previous research. This research proposes a framework for quantifying the benefit of RFID technology application in public transportation so that the cost of implementation can be justified.

The remainder of this paper is organized as follows. In the following section, an RFID application model for public transportation and the benefits expected from the system are established. The cost model is developed in Section 3. Numerical examples and their results are analyzed in Section 4. Finally, the paper is concluded and directions for future research are provided.

2 RFID Application in Public Transportation

In this section, the application of RFID in the public transportation is discussed. Three-layer information system architecture is established with the front-end data capture, the middleware system and the back-end application layers, as shown in Figure 1. The assumptions for this RFID application model are as follows. Buses operated in the model are all equipped with RFID tags, containing information such as bus ID, type and capacity of the bus, and other information useful for the analysis. RFID readers are installed at specific points along the routes of the buses, including bus terminals, bus stops, traffic lights and lamp posts. When a bus passes by a reader, tag information will be retrieved and delivered to the server and subsequently to enquirers. The reader serves as a front-end data–capturing device. Data in the tag are read by the reader and transferred to the middleware system for filtering, aggregating and routing. The resulting information is then transferred to and used by the back-end applications.

The application layer enables passengers to know the real-time location of the bus they are waiting for, hence giving them the certainty of the waiting time. Two most common ways such information can be provide: one through the telecommunication system and the other through network. The bus location information can be pushed to or pulled from the devices used by the passengers. Advanced planning functions can also be implemented on the passengers' devices to facilitate intelligent transportation.

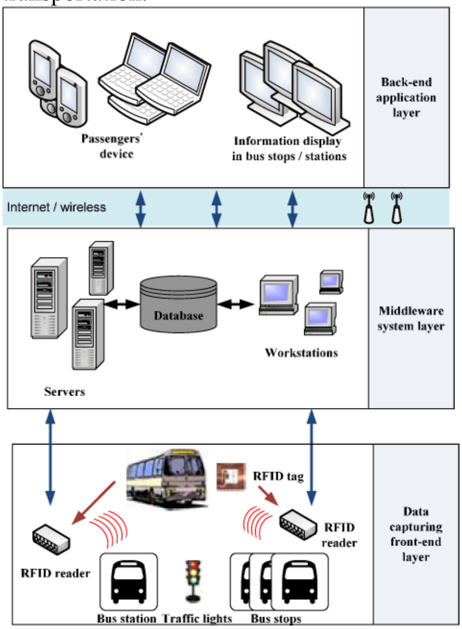


Figure 1. RFID application framework in public transportation system

The proposed information system would provide benefits for both the passengers and the bus operators. For passengers, the system will enable better prediction on bus arrival schedule, and subsequently reducing the waiting time and travel time. In addition, although the perceived waiting time is greater than the actual waiting time, the real-time location information of the buses can instigate the perceived to be equal to the actual one, which in turn increases the level of satisfaction from passengers.

For the bus operators, the objective is to manage the dispatch of buses effectively to meet the passenger service requirement reflected by the length of the headway. With the real-time bus location information, bus operators will be able to improve their operation by varying bus schedules to meet unexpected situations. More importantly, as passengers can learn more accurately the arrival time of the buses, the frequency of the buses will not need to be as high to meet the same level of passenger satisfaction. In other words, with the real-time bus location information, it is possible that a longer headway can still result in the same average waiting time and the same satisfaction level for the passengers. Therefore, the information acquisition and distribution system will either enhance the passenger service level and/or reduce the operation cost for the bus operator. These benefits can be illustrated in Figure 2 below.

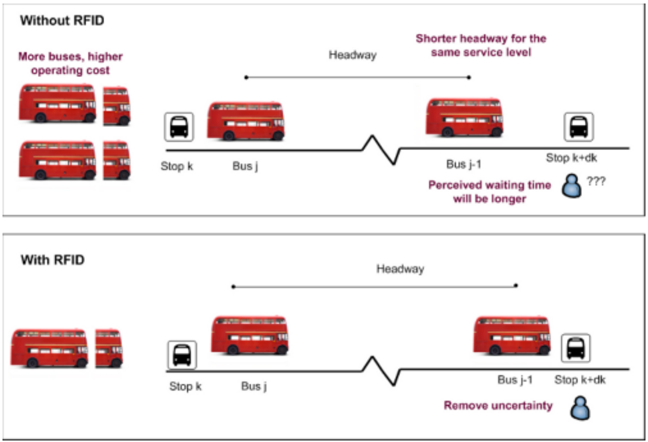


Figure 2. RFID application benefits in public transportation

3 Minimum-Cost Headway Model

In this section, a model which captures the total cost of the system for the justification of the RFID application in public transportation is proposed. By comparing the minimum-cost headway for the system using RFID technology with the one without, the amount in cost reduction can be obtained. Such reduction can then be used as the basis for the justification of RFID implementation. The model is formulated and discussed in the following.

The aim of this model is to find the optimal headways respectively for the two systems. As the bus headway becomes shorter, the service level to passengers increases. However, simultaneously the operating cost for bus operators increases too, because by which it means dispatching a larger number of buses for the same period of time. In contrast, if the headway becomes longer, the number of buss dispatched will be smaller. The passenger waiting time will then be longer and the service level lower. Such relationship can be illustrated in Figure 3 below.

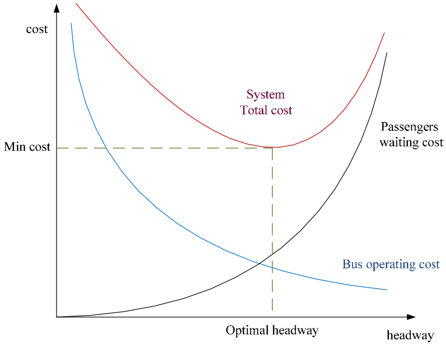


Figure 3. Cost function of the system

The objective function of the model is the minimization of the system's total cost (C_T), which consists of operator cost (C_o) and passenger cost (C_w), that is,

$$C_T = C_o + C_w \quad (1)$$

The operator cost per day is equal to the bus operating cost (p) multiplied by the vehicle service hours per day. The bus operating cost can be estimated from fuel cost, wage rate, maintenance expenses, and the depreciation of the bus. Vehicle service hours per day is the ratio of the average round trip running time *E[t]* to the average headway *E[h]* multiplied by total service hours per day *T*, that is,

$$C_o = p.\frac{2.E[t]}{E[h]}.T \quad (2)$$

The passenger cost per day is equal to the passenger waiting cost function *q(t)* multiplied by the perceived waiting time, total ridership per unit time *R* and total service hours per day *T,* where the perceived waiting time is the perceived waiting time constant μ multiplied by the average waiting time $Ew(t)$, which gives:

$$Cw = q(t).\mu.Ew(t).R.T \quad (3)$$

Understanding the relationship between perceived waiting time and actual waiting time is important. In this model we use perceived waiting time constant to compare both dimensions of waiting time.

Hess *et al.* (2004) did a survey with the bus riders on the duration they usually wait to board the bus. The result of the survey shows that the passengers perceived their average waiting time to be 11.1 minutes. However, the observed average waiting time was 5.8 minutes. Therefore, they proved the hypothesis that the bus riders perceived their wait time to be almost twice what it actually is.

Another research by Mishalani *et al.* (2006) surveyed 83 passengers over a period of approximately one year, and built a simple linear regression model for the perceived waiting time w_p:

$$w_p = \beta_0 + \beta_1 a + \varepsilon \quad (4)$$

where a is the actual waiting time, β_0 is the parameter representing the intercept of the regression line, β_1 denotes the parameter representing the slope of the regression line, and ε is a random variable with a zero mean. The result of their observation again proves that the perceived waiting time is greater than the actual. The result of their observation is shown at the table below:

Table 1. Estimation result of perceived and actual waiting time (Mishalani, 2006)

Variable	*Est. Parameter*	*Standard Error*	*t-statistic*
Intercept	1.33	0.48	2.77
Actual wait time	0.92	0.076	11.96
No. of observations = 83, $\overline{R}^2$= 0.634			

The passenger waiting cost can be estimated by wage rate per unit time. Based on a previous research, the value of bus waiting time was estimated between US$8.5 and US$18. Furthermore, most studies reported that passengers value their waiting time at half of their wage rate per hour (Hess *et al.*,2004). However, the function of the waiting cost with respect to time is not that certain, as occupation, travel time to destination, and many other social aspects of the passenger will affect the value of waiting time. Based on the utility function theory, we can build the waiting cost function as:

$$q(t) = \frac{g}{2}.e^{k.t} \tag{5}$$

where g is the wage rate, t is the waiting time, and k is the growth constant. The value of the growth constant can be derived from the passenger survey, and it indicates the frequency per unit time for the cost increasing by a factor e.

Basic assumptions of the model are as follows:

1. Arrival of passenger at a bus stop is assumed to be random
2. For the system with the information, perceived waiting time is assumed to be the same as the actual one.
3. Every bus will experience the same condition during the trip, the running time is stochastic; and as soon as the bus arrives at the terminal station, it will return to the starting station along the same route.
4. No limit on bus capacity and bus fleet size. For any optimal headway decision, the buses are ready to be operated.

For the passengers arriving randomly at a stop, the probability density function (p.d.f.) of the waiting time $fw(t)$ can be estimated by the p.d.f. of the headway on a specific bus route $fH(t)$; and the relationship is given as below (Larson *et al.,*1981):

$$\mathrm{f}w(t) = \frac{t\int_t^{\infty} fH(u)du}{E(H)} = \lambda\bar{F}H(t) \tag{6}$$

where *FH(t)* is cumulative distribution function of the headway, $\bar{F}H(t)$ = *1-FH(t)*, *E[h]* is the mean headway, and λ is the bus frequency where $\lambda = \frac{1}{E[H]}$.

Furthermore, the mean of the waiting time can be calculated as:

$$Ew(t) = \int_0^{\infty} t.f\ w(t)dt \tag{7}$$

Marguier *et al.* (1984) give the direct result of the mean waiting time as:

$$Ew(t) = \frac{E[H]}{2}\left[1 + \frac{VarH}{E^2(H)}\right] \quad (8)$$

where var$H$ is the variance of the buses headway.
Each trip is a stochastic process with the p.d.f. of the bus service time in one route being r(t). Therefore, the mean running time per trip will be given as:

$$E(t) = \int_0^\infty t.r(t)dt \quad (9)$$

The system total cost can then be obtained by putting together operator costs and passenger costs, and with h as a decision variable, it can be formulated as follows.

$$\min E[Tc\,|\,h] = p.\frac{2.E[t]}{E[h]}.T + q(t).\mu.Ew(t).R.\ T \quad (10)$$

where $h > 0$. Solving the objective function, we can get the system total cost and optimal headway for the system with RFID and the one without. The difference between the two costs can then be derived as the benefit of using RFID in the system.

Furthermore, the verification process should be followed by Cost-Benefit Analysis which is the process of measuring the trade-off between the total cost required for RFID investment and the benefits derived from the project. This is a fundamental and indispensable step, because it is the main analysis behind a "go" or "no-go" decision. The cost for implementing RFID technology consists of direct costs and indirect costs. Smith (2000) suggested taking these six main sources into account: the tag itself, the application of the tags to the products, the purchase and installation of the tag readers, the systems integration, the personnel training and reorganization, and the implementation of application solutions. This analysis can be formulated as a return on investment value, as follow:

$$\text{ROI} = \frac{\text{Benefit of RFID application} - \text{RFID Implementation cost}}{\text{RFID Implementation cost}} \quad (11)$$

When the analysis resulted in the feasible return on investment, the operator can be confidently implement RFID in their system.

4 Numerical Example

A simple case is applied to the minimum-cost headway model to more clearly illustrate the benefit of the proposed RFID system. The passengers arrival is assumed to follow Poisson process, bus operating cost p is 600 NT/hour, wage rate is 600 NT/ hour, the mean ridership R is 150 persons per hour. Perceived waiting time constant μ is 2 for non-RFID system, which means without the information, the passenger perceive their waiting time to be twice of what it actually is. The total service time a day T is 18 hours. The running time is assumed to follow Normal distribution with mean 60 minutes and standard deviation is 5 minutes. The probability density function of the running time will be:

$$r(t) = \frac{1}{\sigma\sqrt{2.\pi}}.e^{-\frac{1}{2}(\frac{t-\mu}{\sigma})^2} = \frac{1}{5\sqrt{2.\pi}}.e^{-\frac{1}{2}(\frac{t-60}{5})^2} \quad (12)$$

The bus arrival is stochastic, and assumed to follow Poisson process, thus Var$H$=$E^2(H)$ and $Ew(t) = E[H]=h$. We also assume that the waiting cost is directly proportional with the waiting time, where $k = 0$, and the equation (10) will be:

$$\min E[Tc \mid h] = p.\frac{2.E[t]}{h}.T + \frac{g}{2}.\mu.h.R.T \quad (13)$$

The assumption of random passenger arrival is valid for short headway. Previous observations of bus passenger arrivals suggested that random passenger arrivals are valid for 12 minutes headway or less (Seddon *et al.,*1974). Therefore, in this case we use 15 minutes headway as the maximum constraints, and we found that the proposed system will have lower system total cost and longer headway.

Table 2. Cost function comparison

Parameter	**Without RFID**	**With RFID**
System total cost	373679.2 NT/day	264275.5 NT/day
Headway	7 minutes	10 minutes

In addition, sensitivity analysis was conducted in order to investigate the system cost model robustness and understand the positive or negative influence of particular model parameters. We used different value for particular parameters; operating cost, waiting cost, passenger ridership and the perceived waiting time coefficient, and the results are as follows:

Table 3. Sensitivity analysis

	Value	**Without RFID**		**With RFID**	
		Cost (NT/day)	Headway (minutes)	Cost (NT/day)	Headway (minutes)
Cw	150	264275.5	10	186839.6	14
	450	458459.1	6	323594.3	8
Cp	300	264275.5	5	186839.6	7
	900	458391.5	8	323594.3	12
λa	75	264275.5	10	186839.6	14
	300	528550.9	5	373679.2	7
μ	1.25	295514.4	9	264275.5	10
	1.5	323594.3	8		
	1.75	350054.2	7		

It is shown that although we change the parameter to the lower or higher value, the system with RFID always have lower system cost and longer headway. Furthermore, when we change the perceived waiting time coefficient, the system with RFID also performs better.

5 Conclusions

In this paper we have proposed an application of RFID in the public transportation. A three-layer system architecture is proposed, consists of data capturing front-end which have the tag attached in every buses and reader located on every terminals bus stops, traffic lights and lamp posts; middleware systems where the data read by the reader is processed; and the application layer where the passengers can utilize the information on real-time bus location.

Potential benefits of the systems are investigated, and it is obvious that knowing the real time bus location will decrease the passengers waiting time, especially the perceived waiting time. The bus operator can utilize the number of buses efficiently, running fewer numbers of buses for the same service level, hence reducing the operating cost. The minimum-cost headway model is generated to verify the benefit. The objective function is to minimize the total system cost, which consists of bus operating cost and passengers waiting cost. The decision variable is to find the optimal service frequency, or known as the headway. Numerical examples are presented to examine the benefit of the proposed framework, and the results imply that the proposed system will benefit and decrease the system total cost.

Finally, our model clearly does not capture many of the complexities involved in the real world. Hence, it is potential to relax some assumptions and make the model close by the real world. It is also potential for future research in the transit route choice analysis which investigates the benefit of the real time information to minimize the total travel time of passengers.

6 References

[1] Ben-Akiva M., SR Lerman. Discrete choice analysis. MIT Press 1985.
[2] Gentile G., Nguyen S., Pallottino S. Route choice on transit networks with online information at stops. Transportation Science 2005;39(3):289-97.
[3] Hess DL, Brown J, Shoup D. Waiting for a bus. Journal of Public Transportation 2004; 7(4) :67-84.
[4] Hickman M. Bus Automatic Vehicle Location (AVL) Systems. In : Assessing the Benefits and Costs of ITS. Springer 2006:59-88.
[5] Larson RC, Odoni AR. Urban Operation Research. Prentice-Hall. Englewood Cliffs 1981.
[6] Margiuer PHJ, Ceder A. Passengers waiting strategies for overlapping bus routes. Transportation Science Journal 1984 ;18(3): 207-30.
[7] McLeod F. Estimating bus passenger waiting times from incomplete bus arrivals data. Journal of the Operational Research Society 2007;58:1518-25.
[8] Menezes B., Laddhad K, Karthik. Challenges in RFID deployment – a case study in public transportation. International Conference on E-Governance 2006.
[9] Mishalani RG, McCord MM. Passenger wait time perceptions at bus stops : empirical results and impact on evaluating Real-Time Bus Arrival Information. Journal of Public Transportation 2006;9 (2):89-106.
[10] Ngai EWT, Moon KKL., Riggins FJ, Yi CY. RFID Research : an academic literature review (1995-2005) and future research directions. International Journal of Production Economics 2007;112:510-20.
[11] Seddon PA., Day MP. Bus passenger waiting time in Greater Manchester. Traffic Engineering Control 1974;15:442-445.
[12] Wyld DC. RFID 101 : The next big thing for management. Management Research News 2006;29(4):154-73.

Competitive Supply Chain Performance

Modeling and Solving the Collaborative Supply Chain Planning Problems

Y. T. Chen[a], Z. H. Che[a], Tzu-An Chiang[b], C. J. Chiang[a] and Zhen-Guo Che[c,1]

[a]Department of Industrial Engineering & Management – National Taipei University of Technology, Taiwan, R.O.C.
[b]Department of Commerce Automation & Management – National Pingtung Institute of Commerce, Pingtung (900), Taiwan, R.O.C.
[c]Institute of Information Management – National Chiao Tung University, Taiwan, R.O.C.

Abstract: Improving overall performance and collaborative value is an important issue to supply chain management. By planning a collaborative supply chain, this paper intends to establish a supplier rating package, which includes production and distribution plans. The aim is to establish a supplier procurement value rating system by combining Analytic Hierarchy Process (AHP) and Rough Sets Theory, and construct a two-objective optimization mathematical model (including procurement value and costs) for planning the supply chain. Finally, a Genetic Algorithm (GA) is applied under capacity constraints to obtain the best solution, thus achieving the planning of collaborative supply chain.

Keywords: Collaborative Supply Chain, Analytic Hierarchy Process, Rough Sets Theory, Genetic Algorithm.

1 Introduction

In today's global market of increased competition and pressure, the relationship between upstream and downstream partners in a supply chain grows more complex, therefore, the management and integration capacity of supply chain partners are crucial components to the success of a company [5]. Through a collaborative supply chain, two or more independent companies could plan and operate a collaborative supply chain, in order to obtain greater operating success [8]. As more investments and resources are available, companies could reduce risks and create more advantages through collaboration, thus, collaborative supply chains play a crucial role in complex manufacturing environments [3]. Moreover, it is important to determine the optimum partners for production, assembly, and completion of final products in a collaborative manufacturing environment [9].

[1] Corresponding Author E-mail: bcc102a@hotmail.com

Dogan et al. [6] considered qualitative and quantitative factors in supply partner evaluation problems, and integrated AHP and the possibilities of multi-objective linear programming, in order select appropriate partners and determine optimum production quantities. Many scholars have extensively applied AHP to solve partner selection problems [2, 7]. Xia and Wu [11] suggested that the Rough Sets Theory can reduce subjective judgments of AHP. In addition, GA is commonly applied to solve single-/multi-objective production and work management problems, given its effective search performance for global solutions [1, 12].

Therefore, this research, based on the concept of a collaborative supply chain, aimed to construct a production and distribution planning and decision-making model that were most appropriate to collaborative supply chain. First, it established a supplier procurement value rating system (Section 2) by combining AHP and the Rough Set Theory. A two-objective optimization mathematical model (Section 3) was constructed under a collaborative supply chain for supplier selection, production, and distribution planning, which considered procurement value and costs. Finally, a GA was employed to solve the optimization mathematical model (Section 4) for obtaining satisfactory decisions within a short period.

2 Procurement Value Rating Model Development

The procurement value rating model was constructed using AHP, after improved by the Rough Sets Theory, in order to reduce the degrees of subjective judgments of a traditional AHP [11]. The detailed steps are as follows, with an example of central assembly plant C_1:

Step 1: Determine the suppliers' rating level and factors, of which the parameters include 3 main criteria: price (pr), quality (qu), and service (se), and 7 sub-criteria: technology level (tl), defect (de), reliability (re), on-time delivery (od), supply capacity (sc), maintenance cycle (mc), and warranty period (wp).

Step 2: Calculate criteria weights.

Sub-step 2.1: Resolve AHP rating errors by attributing significance concepts to the Rough Sets Theory.

Sub-step 2.1.1: Design tables are based on the main criteria, with the fields of price, quality, and service; Table 1 lists definitions represented by 1-3 in the criteria, and Table 2 lists the different groups of criteria.

Sub-step 2.1.2: A rating team completes the decision column in Table 2. Take the main criteria as an example, if "1" is entered into the decision column, it represents that the supplier is selected; if "0" is entered, it denotes that the supplier is not selected. For instance: group 1 indicates a moderate price, and satisfactory levels of quality and service; if the decision column is 1, it indicates that the supplier is selected.

Table 1. The meaning of value 1–3 for different main criteria and sub-criteria

Value	pr	qu	se	tl	de	re	od	sc	mc (Week)	Wp (Month)
1	Low	Good	Good	High	Low	High	Good	Great	Short	Long
2	Middle	Middle	Middle	Middle	Middle	Middle	Middle	Middle	Not too long	Not too long
3	High	Poor	Poor	Low	High	Low	Poor	Small	Long	Short

Table 2. Decision table [11]

	Main criteria				Quality				Service				
combinations	pr	qu	se	Decision	tl	de	re	Decision	od	sc	mc	wp	Decision
1	2	1	1	1	1	1	2	G	1	1	2	2	G
2	3	1	1	0	1	2	1	G	1	2	2	1	G
3	1	2	2	1	1	2	2	M	1	2	3	3	M
4	2	2	2	1	2	1	2	G	2	2	2	2	M
5	3	2	1	0	2	2	1	M	2	1	1	3	M
6	1	2	3	0	2	2	2	M	3	2	1	1	P
7	1	3	1	0	1	1	3	M	3	1	2	2	M
8	1	1	3	1	3	1	2	M	1	3	2	3	M
9	2	2	1	1	3	3	2	P	3	2	2	1	P
10	2	2	3	0	2	3	2	P	3	1	1	1	M
11	3	3	1	0					2	2	1	1	M
12	3	2	2	0					2	3	2	2	P
13	2	3	1	0					2	2	2	3	M
14	2	1	3	0					2	2	3	2	P
15									1	1	3	3	M
16									2	3	1	3	P
17									1	1	1	3	G
18									1	1	3	1	G

G = Good, M = Middle, P = Poor

Sub-step 2.2: According to the decision column in Table 2, the significance of the main criteria could be obtained by the following procedures:
$U|IND$\{pr,qu,se\}=\{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\},\{7\},\{8\},\{9\},\{10\},\{11\},\{12\},\{13\},\{14\}\}; $U|IND$\{Decision\}=\{\{2,5,6,7,10,11,12,13,14\},\{1,3,4,8,9\}\}=\{Y_1,Y_2\}; $U|IND$ \{qu,se\}=\{\{1,2\},\{3,4,12\},\{5,9\},\{6,10\},\{7,11,13\},\{8,14\}\}=\{X_1,X_2,X_3,X_4,X_5,X_6\}; $p(X_1)$=2/14; $p(X_2)$=3/14; $p(X_3)$=2/14; $p(X_4)$=2/14; $p(X_5)$=3/14; $p(X_6)$=2/14; $p(Y_1|X_1)$=1/2; $p(Y_1|X_2)$=1/3; $p(Y_1|X_3)$=1/2; $p(Y_1|X_4)$=1; $p(Y_1|X_5)$=1; $p(Y_1|X_6)$=1/2; $p(Y_2|X_1)$=1/2; $p(Y_2|X_2)$=2/3; $p(Y_2|X_3)$=1/2; $p(Y_2|X_4)$=0; $p(Y_2|X_5)$=0; $p(Y_2|X_6)$=1/2.
The significance of price is SGF(pr,\{qu,se\},\{Decision\})=H(\{Decision\}|\{qu,se\})-H(\{Decision\}|\{pr,qu,se\})=-2/14[(1/2)log(1/2)]×3-3/14[(1/3)log(1/3)+(2/3)log(2/3)]= 0.1882 and the significance of other main criteria and sub-criteria can be obtained by the same procedures.
Sub-step 2.3: Conduct pairwise comparison for relative significance according to a hierarchical structure. In AHP, the weight of every option is originated from eigenvectors placed in a pairwise comparison matrix. Take the main criteria for example, form matrix J by the significance of criteria, then turn the matrix into maximum eigenvectors, and obtain the weights of price, quality, and service: 0.4321, 0.2346, and 0.3333, respectively, and the maximum eigenvector (λ_{max}) is 3. The weights of the sub-criteria are calculated in the same way.

$$J=\begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ & & \cdots & \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{bmatrix}=\begin{bmatrix} 1 & 1.8415 & 1.2961 \\ 0.5430 & 1 & 0.7039 \\ 0.7715 & 1.4207 & 1 \end{bmatrix}$$

Sub-step 2.4: Check whether the pairwise comparison matrix constructed by the Rough Sets Theory is consistent through the consistency index $CI=(\lambda_{max}-n)/(n-1)$, where n is the rank of the judgment matrix; if CI=0, it represents complete consistency.
Step 3: Calculate the final assessment of suppliers for central assembly plant C_1, as listed in Table 3.

Sub-step 3.1: Multiply the weights of the main criteria and sub-criteria to obtain the global weights.
Sub-step 3.2: Consider the quantitative information of suppliers, and then convert by certain proportions to obtain the suppliers' rating.
Sub-step 3.3: Combine the global weights and suppliers' rating, then multiply the total to obtain the final rating of the suppliers.

Table 3. Final rating of central assembly plant C_1

	Selection factors (Global weights)																
	pr (0.43)		tl (0.06)		de (0.12)		re (0.05)		od (0.13)		sc (0.10)		mc (0.07)		wp (0.03)		Final rating
Supplier	Raw data	Rating	Raw data	Rating	Raw data	Rating	Raw data	Rating	Raw data	Rating	Raw data	Rating	Raw data	Rating	Raw data	Rating	
M_{11}	0.7	0.44	3	0.29	0.05	0.27	0.80	0.33	0.80	0.31	120	0.48	4	0.23	1	0.25	0.370
M_{12}	1.2	0.26	3	0.29	0.05	0.27	0.80	0.33	0.85	0.33	50	0.20	3	0.31	1	0.25	0.271
M_{13}	1	0.31	2	0.43	0.03	0.45	0.85	0.35	0.90	0.35	80	0.32	2	0.46	2	0.50	0.358

3 Assumptions and Mathematical Foundation

The assumptions are: 1) the preceding sequences of central assembly plants are known; 2) no shortages of inventory occur; 3) the suppliers face the constraint of maximum capacity; 4) the raw material/processing cost of the suppliers does not vary with the delivery quantity, but is calculated by units of raw material/processing costs; 5) the transportation costs of suppliers does not vary with the delivery quantity, and is calculated by unit transportation costs.

The notations for the mathematical model are listed below:

i, l Central assembly plant index, i=1,2,…, I; l=1,2,…, L
I, L Total number of central assembly plant
j Supplier index, j=1,2,…, J_i
J_i Total number of suppliers in the central assembly plant i
QS_{ij} Quantity to be ordered by supplier j in the central assembly plant i
YT_{il} $\begin{cases} 1 & \text{Central assembly plants } i \text{ and } l \text{ have delivery relationship} \\ 0 & \text{otherwise} \end{cases}$
QT_{il} Order quantity transported from central assembly plant i to l
W_{ij} Weight of supplier j in the central assembly plant i
PP_{ij} Raw material/processing cost of supplier j in the central assembly plant i
TP_{il} Transportation price from central assembly plant i to l
C_{ij} Max. supply capacity of supplier j in the central assembly plant i
D_i Total demand of central assembly plant i

The optimal mathematical model for collaborative supply chain planning is as:

$$Max \quad Z = Z_1 - Z_2 \tag{1}$$

$$Z_1 = \sum_{i=1}^{I} \sum_{j=1}^{J_i} \left(W_{ij} \cdot QS_{ij}\right) \tag{2}$$

$$Z_2 = \sum_{i=1}^{I} \sum_{j=1}^{J_i} \left(PP_{ij} \cdot QS_{ij}\right) + \sum_{i=1}^{I} \sum_{j=1}^{J_i} \left(TP_{il} \cdot QT_{il}\right) \tag{3}$$

$$\text{st}: \; QS_{ij} \le C_{ij} \;,\; i = 1,2,3 \ldots I \;,\; j = 1,2,3 \ldots J_i \tag{4}$$

$$\sum_{j=1}^{J_i} QS_{ij} = D_i \quad , \quad i = 1,2,3...I \tag{5}$$

Objective function Eq. (1) is the total procurement value (Z_1) minus total cost (Z_2), thus, maximizing the target function (Z). Eq. (2) is the total procurement value of suppliers. Eq. (3) is the total raw material/processing costs, as well as transportation costs. Eq. (4) meets the capacity constraints of suppliers in the central assembly plant. Eq. (5) ensures that the demand of every central assembly plant is equal to the total supply of the suppliers.

4 GA Solving Model for Mathematical Model

The GA solving model proposed in this research was used to efficiently formulate supplier selection, production, and distribution planning in a collaborative supply chain. The GA computational procedures are described below:

Step 1: Chromosome coding: binary coding is used for chromosomes selected by the supplier; a central assembly plant may cooperate with several suppliers, as shown in Figure 1.

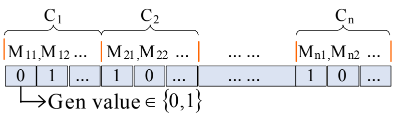


Figure 1. Chromosome sturcture

Step 2: Generate the initial population: the initial supplier combination generates feasible populations under capacity constraints (Eq. (4)), and supply and demand constraints (Eq. (5)).

Step 3: Calculate the fitness function: substitute individual values into the optimization mathematical model for calculation. As two objects have different units, the objects are standardized according to (Eq. (6)). Of which, $f_{i,\min}$ and $f_{i,\max}$ are min and max values of the i-th target in existing generations.

$$f_i' = (f_i - f_{i,\min})/(f_{i,\max} - f_{i,\min}) \quad , \quad i = 1,2,...,n \tag{6}$$

Step 4: Reproduction: the common Roulette Wheel Selection method of GA is adopted for reproduction; the percentage of individual fitness function values, as calculated from the previous step, to the total fitness function is taken as selection probability. The individuals of larger fitness functions are easily selected.

Step 5: Crossover: single-point crossover [12] is employed to randomly extract two chromosomes from the population; a tangent point is randomly generated on the chromosome, and then the gene codes, after the tangent point of two chromosomes are exchanged.

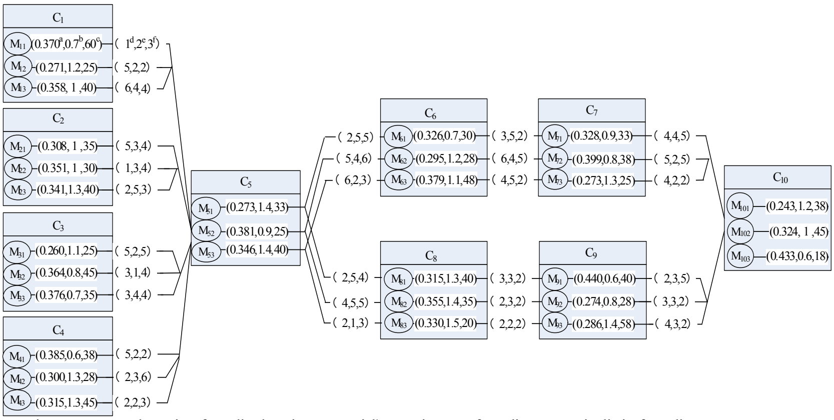
Step 6: Mutation: single-point mutation of the common GA is adopted. First, randomly extract a chromosome, then randomly generate a mutation position, and randomly change the gene codes of the position within a reasonable range (under capacity constraints (Eq. (4)), and supply and demand constraints (Eq. (5)).

Step 7: Generate new offspring by gene evolution modes, as shown in Steps 4-6; if the optimal fitness value of the offspring is superior to that of the current population, replace it within the population as the new value for the next-generation evolution; otherwise, retain the original population to the next generation.

Step 8: Check whether the stop conditions are met: the stop conditions indicate the executed generations; the predetermined generation number is set and entered prior to operation, and it is required to stop when the evolution number reaches the generation.

5 Illustrative Example

The case of a supply chain network [4] is shown in Figure 2, where C_1-C_{10} are different central assembly plants in the supply chain network, of which every central assembly plant may cooperate with 3 suppliers. The unit procurement value, costs, and capacity of every supplier are shown in Figure 2. A single product is considered in this research, while a collaborative supply chain plan is implemented under 60 products required by the central assembly plant.



a: unit procurement value rating of supplier; b: unit raw material/processing cost of supplier; c: capacity limit of supplier

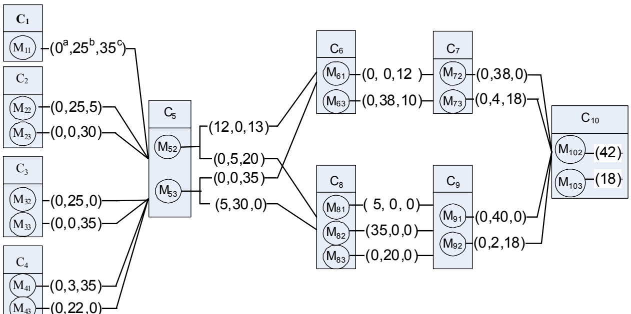
d, e, f: unit transportation costs from upstream supplier to the 1st, 2nd, and 3rd suppliers of downstream central assembly plant

Figure 2. Collaborative supply chain network

The procurement value of every supplier is calculated by the rating system in Section 2, and the relevant data are entered into the optimization mathematical model. Next, according to the parameter settings from the GA of [10], this research proposed 16 solution models, which parameters are listed in Table 4. Every model is subjected to 15 times of calculation to ensure that all suppliers' capacity constraints and demands are met. The calculated results are listed in Table 4. The results indicate that model 15 has better calculated result, with the mean gross procurement value as 235, and mean gross cost as 2345. The optimum production and distribution plan of a collaborative supply chain is shown in Figure 3. As seen, the suppliers selected by central assembly plant C_6 are M_{61} and M_{63}, of which M_{61} has to process 12 units from supplier M_{52} of upstream central assembly plant C_5, and M_{63} has to process 13 units from supplier M_{52}, and 35 units from M_{53} of upstream central assembly plant C_5. In addition, M_{61} has to transport 12 units to supplier M_{73} of downstream central assembly plant C_7, and M_{63} has to transport 38 of 48 units to downstream supplier M_{72}, and the remaining 10 units to downstream supplier M_{73}.

Table 4. Parameter combinations and execution results

Model	Parameter combination				Result	
	Generation	Population	Crossover rate	Mutation rate	Mean total procurement value	Mean total cost
1	200	10	0.3	0.03	230	2481
2	200	10	0.3	0.05	231	2478
3	200	10	0.6	0.03	230	2500
4	200	10	0.6	0.05	230	2486
5	200	50	0.3	0.03	234	2386
6	200	50	0.3	0.05	235	2358
7	200	50	0.6	0.03	234	2384
8	200	50	0.6	0.05	235	2366
9	500	10	0.3	0.03	229	2485
10	500	10	0.3	0.05	229	2449
11	500	10	0.6	0.03	229	2492
12	500	10	0.6	0.05	230	2453
13	500	50	0.3	0.03	233	2397
14	500	50	0.3	0.05	235	2408
15	500	50	0.6	0.03	235	2345
16	500	50	0.6	0.05	235	2375



a, b, c : quantities transported from upstream supplier to the 1st, 2nd, and 3 rd suppliers of downstream central assembly plant

Figure 3. Production and distribution plan for collaborative supply chain

6 Conclusions

This research proposed a collaborative supply chain planning model, which allows decision-makers to select suitable suppliers, and determine the production and distribution quantities of all suppliers. According to the model, the supplier's procurement value rating system is established by combining Rough Sets Theory and AHP. An optimal mathematical model is constructed by considering the supplier's procurement value, cost, capacity constraints, and demand. Finally, GA is used to solve the mathematical model and the experimental results indicated that GA can equal to the production and distribution planning problems of a collaborative supply chain.

In the current form of the proposed GA approach, however, it may not be effective in dealing with more complex problems. For example, it cannot find the quality result in the problem when more events such as inventories of products are taken into consideration. For further research we thought about extending this developed approach to more complex problems such as collaborative supply chain planning problems involving different weights of considered factors, inventories, etc.

7 References

[1] ALTIPARMAK, F.; GEN, M.; LIN, L.; PAKSOY, T. A genetic algorithm approach for multi-objective optimization of supply chain networks. Computers and Industrial Engineering, Vol. 51, No. 1, pp. 196 –215, 2006.

[2] CHAN, F.T.S. Interactive selection model for supplier selection process: An analytical hierarchy process approach. International Journal of Production Research, Vol. 41, pp. 3549–3579, 2003.

[3] CHEN, M.-C.; YANG, T.; LI, H.-C. Evaluating the supply chain performance of IT-based inter-enterprise collaboration. Information and Management, Vol. 44, No. 6, pp. 524-534, 2007.

[4] GAMBERINI, R.; GRASSI, A.; RIMINI, B. A new multi-objective heuristic algorithm for solving the stochastic assembly line re-balancing problem. International Journal of Production Economics, Vol. 102, No. 2, pp. 226-243, 2006.

[5] LAMBERT, D.M.; COOPER, M.C. Issues in Supply Chain Management. Industrial Marketing Management, Vol. 29, pp. 65–83, 2000.

[6] O ZGEN, D.; O NU T, S.; GU LSU N, B.; TUZKAYA, U.R.; TUZKAYA, G. A two-phase possibilistic linear programming methodology for multi-objective supplier evaluation and order allocation problems. Information Sciences, Vol. 178, pp. 485–500, 2008.

[7] SEVKLI, M.; KOH, S.C.L.; ZAIM, S.; DEMIRBAG, M.; TATOGLU, E. An application of data envelopment analytic hierarchy process for supplier selection: A case study of BEKO in Turkey. International Journal of Production Research, Vol. 45, No. 9, pp. 1973–2003, 2007.

[8] SIMATUPANG M.; SRIDHARAN R. The Collaborative Supply Chain. The International Journal of Logistics Management, Vol. 13, No. 1, pp. 15-30, 2002.

[9] TSENG, Y.-J.; JHANG, J.-F.; HUANG, F.-Y. Multi-plant assembly planning models for a collaborative manufacturing environment. International Journal of Production Research, Vol. 45, No. 15, pp. 3333-3349, 2007.

[10] WANG, H.S.; CHE, Z.H. A multi-phase model for product part change problems. International Journal of Production Research, Vol. 46, No. 10, pp. 2797-2825, 2008.

[11] XIA, W.; WU, Z. Supplier Selection with Multiple Criteria in Volume Discount Environments. Omega, Vol. 35, pp. 494-504, 2007.

[12] XU, J.; LIU, Q.; WANG, R. A class of multi-objective supply chain networks optimal model under random fuzzy environment and its application to the industry of Chinese liquor. Information Sciences, Vol. 178, No. 8, pp. 2022-2043, 2008.

A Bi-Objective Model for Concurrent Planning of Supplier Selection and Assembly Sequence Planning

Y. Y. Lin[a], Z. H. Che[a], Tzu-An Chiang[b], Zhen-Guo Che[c] and C. J. Chiang[a,1]

[a]Department of Industrial Engineering & Management – National Taipei University of Technology, Taiwan, R.O.C.
[b]Department of Commerce Automation & Management – National Pingtung Institute of Commerce, Pingtung (900), Taiwan, R.O.C.
[c]Institute of Information Management – National Chiao Tung University, Taiwan, R.O.C.

Abstract: This study discussed parts supplier selection for central assembly plant. In addition to the delivery state of suppliers, the production conditions of the central assembly plant should be considered. A bi-objective optimal mathematical model was thus established for supplier selection in this study, which incorporates assembly sequence planning (ASP). At the same time, the multi-objective particle swarm optimization (MPSO) was used to solve the model for a case of stapler assembly, in order to validate the solving effects of MPSO. The results showed that MPSO not only has faster convergence, but can also obtain better Pareto-optimal solutions.

Keywords: Assembly sequence planning, supplier selection, particle swarm optimization, multi-objective optimization.

1 Introduction

To enhance their competitiveness, businesses must enhance their capacity and productivity by incorporating parts suppliers into the production process. Both parts supply status and productivity conditions should be considered during production planning for timely responses to the market demands. In this regard, ASP is one of the key issues in production planning. Boothroyd *et al.* [2] and Corana *et al.* [5] suggested that if assembly requirements are considered and ASP is effectively implemented during the planning phase, it is possible to reduce planning errors, real assembly time, and production costs. Thus, this research incorporated ASP into the supplier selection model in order to achieve better concurrent planning. Tseng *et al.* [7] indicated that a connector-based assembly-planning model could efficiently reduce the complexity of search. Thus, the

---

[1] Corresponding Author E-mail: wind112233@yahoo.com.tw

preceding relationship of parts assembly was discussed based on connector concept.

Since ASP is a combinatorial optimization problem, many scholars have applied genetic algorithms (GA) to solve ASP [3]. Particle swarm optimization (PSO) has been mainly applied to continuous optimization problems, rather than ASP, due to continuity characteristics of PSO. Mohemmed *et al.* [6] found an optimal path with PSO, and its solving effect was significantly better than a GA-based search algorithm. Thus, this paper aims to solve the problem with PSO in a bi-objective situation.

The three main purposes of this paper are as follows: (1) to propose a multi-objective optimal mathematical model for ASP and supplier selection concurrent planning, which objectives include minimized entire timeframe and minimized gross cost for completing orders; (2) to incorporate a non-domination sorting approach into PSO, and develop the MPSO to solve bi-objective optimal model; (3) to discuss the applicability of MPSO on a case study of stapler assembly, and compare the solving effects of MOGA to prove the superior solving capability of MPSO. The comparison criteria include operation time, convergence algebra, mean value of minimum fitness function, and average number and average ratio of Pareto-optimal solutions [1].

The remainder of this paper is organized as follows: Section 2 describes the assumptions and bi-objective optimal mathematical model; Section 3 introduces the computing processes of a MPSO; Section 4 presents the case study, with comparative and analytical results of the MOGA; Section 5 gives conclusions and suggestions for further researches.

2 Problem formulation

This research makes the following assumptions: (1) each part is procured from a single supplier, and no shortage occurs during the procurement process; (2) the parts are not assembled for production until complete delivery to the central assembly plant; (3) the assembly sequence is applied to ASP based on connector concept, parts information, and the preceding relations of connector are known; (4) the assembly time and costs are known, while the connector time and cost are sequence-dependent; (5) the combination, assembly direction, assembly tools are not considered in the assembly work.

The notations available for the mathematical model:

i , j	Parts index, i =1,2,3,…, I , j =1,2,3,…, I
I	Total number of parts
s , t	Supplier index of part i and part j , s =1,2,3,…, S_i , t =1,2,3,…, S_j
S_i , S_j	Total number of suppliers of part i and part j
m , n	connector index, m =1,2,3,…, M , n =1,2,3,…, M
M	Total number of connectors
p	Assembly sequence index, p =1,2,3,…, P
P	Total number of assembly sequences, $M = P$

D	Order demand
A_{pmn}	$\begin{cases}1 & \text{From connector } m \text{ to } n \text{ under the assembly sequence } p\\ 0 & \text{otherwise}\end{cases}$
B_{mij}	$\begin{cases}1 & \text{The assemby of part } i \text{ and part } j \text{ under the connector of } n\\ 0 & \text{otherwise}\end{cases}$
C_{is}	$\begin{cases}1 & \text{Supplier } s \text{ of part } i\\ 0 & \text{otherwise}\end{cases}$
F_{pm}	$\begin{cases}1 & \text{Connector } m \text{ under assembly sequence } p\\ 0 & \text{otherwise}\end{cases}$
G_m	Preceding connector set of connector m, m =1,2,3,…, M
MT_{is}	Manufacturing time of supplier s of part i
PC_{is}	Procurement cost of supplier s of part i
DT_{is}	Delivery time from supplier s of part i to the central assembly plant
DC_{is}	Delivery cost from supplier s of part i to the central assembly plant
AT_{isjt}	Assembly time of supplier s of part i and supplier t of part j
AC_{isjt}	Assembly cost of supplier s of part i and supplier t of part j
OT_{mn}	Operation time from connector m to n
OC_{mn}	Operation cost from connector m to n

Bi-objective optimal mathematical model:

Objective functions:

Objective 1 - Minimized entire timeframe for completing an order. The entire timeframe for completing the order refers to the time required for ordering and delivery of supplies to the central assembly plant until the assembly is completed, it includes: lead time, assembly time, and operation time, of which the lead time contains the manufacturing and delivery time of suppliers. The lead-time is set as the longest because the assembly cannot commence until all parts have arrived.

$$Min \quad f_1 = \max\{(MT_{is}D + DT_{is})C_{is} \mid i = 1,2,3,...,I; s = 1,2,3,...,S_i\}$$
$$+\sum_{m=1}^{M}\sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{s=1}^{Si}\sum_{t=1}^{Sj} AT_{isjt}B_{mij}C_{is}C_{jt}D + \sum_{p=1}^{P}\sum_{m=1}^{M}\sum_{n=1}^{M} OT_{mn}A_{pmn}D \tag{1}$$

Objective 2 - Minimized gross costs for completing an order. The gross cost for completing the order refers to the costs required for order supply and delivery to the central assembly plant until the assembly is completed, it includes: lead cost, assembly cost, and operation cost, of which the lead cost includes procurement and delivery costs of parts.

$$Min \quad f_2 = \sum_{i=1}^{I}\sum_{s=1}^{Si} (PC_{is} + DC_{is})C_{is}D + \sum_{m=1}^{M}\sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{s=1}^{Si}\sum_{t=1}^{Sj} AC_{misjt}B_{mij}C_{is}C_{jt}D$$
$$+\sum_{p=1}^{P}\sum_{m=1}^{M}\sum_{n=1}^{M} OC_{mn}A_{pmn}D \tag{2}$$

Constraints:

Ensure that assembly is conducted only after the completion of all preceding connector set procedures.

$$\sum_{p=1}^{P} F_{pm}p - \sum_{p=1}^{P} F_{pn}p \geq 0 \qquad \forall m,n \in G_m \tag{3}$$

0-1 integral constraint of decision variable A_{pmn}, B_{mij}, C_{is}, F_{pm}.

$$A_{pmn} = \{0,1\} \qquad \forall p,m,n \tag{4}$$

$$B_{mij} = \{0,1\} \qquad \forall m,i,j \tag{5}$$

$$C_{is} = \{0,1\} \qquad \forall i,s \tag{6}$$

$$F_{pm} = \{0,1\} \qquad \forall p,m \tag{7}$$

3 The proposed MPSO approach

In this paper, a MPSO is proposed to present and update a feasible assembly sequence, which uses real-value solution-particle mapping and a real-value velocity model, in which the suppliers' data could be updated through mutation. In the subsequent process, the particles of minimal fitness values are selected as the global optimal value, leading to an updated assembly sequence. The steps of MPSO are described as follows:

3.1 Particle coding

Each particle represents a group of feasible solutions, while the particle structure is represented by chromosomes; it includes: assembly sequence floating-point vector, assembly sequence velocity, assembly sequence, combination of suppliers, and optimized objectives. The assembly sequence is restrained by the precedence relation (Eq. (3)), and the assembly sequence floating-point vector indirectly determines the initial assembly sequence.

3.2 Generation of initial group

The initial floating-point vector of the assembly sequence, assembly sequence velocity, and suppliers' data are randomly generated. The initial random vector of the assembly sequence is set between [-100,100]; and the initial random vector of assembly sequence velocity is set between [-10,10].

3.3 Calculation of fitness function

Substitute the assembly sequence and combination of suppliers into the population of the appraisal model, calculate the objective functions for particles by Eqs. (1) and (2), and then the fitness function of particles by Eq. (8). When the objective functions are minimized, the minimal fitness function is optimum.

$$Fitness = w_1 f_1 + w_2 f_2 \qquad w_1 = w_2 = 1 \tag{8}$$

3.3 Updating *pBest* and *gBest*

Set initial particles as initial *pBest*, compare *pBest* of subsequent generations with the fitness value of the particles of the preceding generation, select the first 50% as *pBest*; then, select *gBest* by taking the particles of the minimal fitness function obtained from the preceding step to guide the evolution of the next generation assembly sequence.

3.4 Updating particle positions

Update the assembly sequence velocity by Eq. (9), then update the assembly sequence floating-point vector by Eq. (10) [4], thus generating the new assembly sequence. The population of suppliers is updated by multi-point mutation. The gene number and gene position of particles to be mutated are randomly set, while the gene codes of the gene position are randomly changed within the range of total parts suppliers.

$$v_{id}^{j+1} = k(v_{id}^{j} + \phi_1 \times rand() \times (p_{id} - x_{id}^{j}) + \phi_2 \times rand() \times (g_{id} - x_{id}^{j}))$$

In which, $k = \dfrac{2}{\left|2 - \phi - \sqrt{\phi^2 - 4\phi}\right|}$, $\phi = \phi_1 + \phi_2$, $\phi > 4$ (9)

$$x_{id}^{j+1} = x_{id}^{j} + v_{id}^{j+1} \tag{10}$$

3.5 Stop conditions

The stop conditions refer to the generation number executed.

3.6 Optimal strategies

When the remaining population is subjected to the non-dominated sorting approach, the non-dominated solutions are the optimal strategies.

4 Case Study and Results Analysis

The case study concerns the staplers produced by Company A. The spare parts of staplers are shown in Figure 1. There are 18 parts, and each is available from 4 suppliers. Company A has received an order of 1000 staplers, thus, there are 18 parts and 9 assembly modes. Figure 2 shows the preceding relations of staplers.

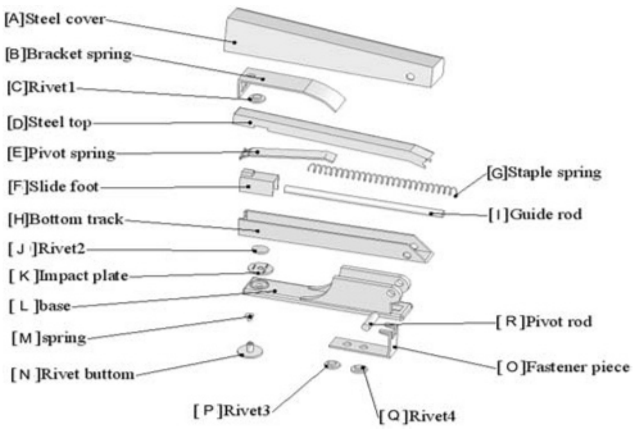


Figure 1. Decomposition diagram of a stapler [7]

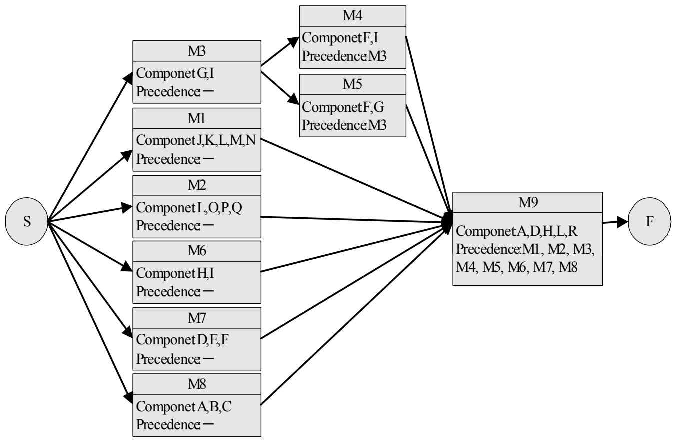


Figure 2. Connector-based preceding relation

The MPSO proposed in this research was used for solving problems. First, the parameters were experimentally planned to obtain an optimum combination of parameters. The experimental results indicated that, the optimum parameter combination of MPSO was population number Po , generation number G , cognitive parameter ϕ_1 , social parameter ϕ_2 , maximum velocity $V\max$)=(30, 5000, 2.8, 1.3, $\pm$ 3000). In this research, the operation time, convergence function, mean value of minimum fitness function, average number of Pareto-optimal solutions, and average ratio of Pareto-optimal solutions were taken as indicators of the computational performance. The experiments were repeated 30 times to achieve optimized results for comparison, with the significance level of $\alpha = 0.05$.

The executed results of MPSO and MOGA are listed in Table 1. As seen, though the operation time of MPSO is slower than MOGA, MPSO can be quickly converged with little evolutionary algebra, as shown in Figure 3. Additionally, MPSO has better performance than MOGA in terms of the average number of Pareto-optimal solutions and average ratio of Pareto-optimal solutions, indicating that the Pareto-optimal solution obtained by MPSO is superior to that obtained by MOGA. Finally, this research selected 3 solutions from the Pareto-optimal solutions obtained from the MPSO. The optimal strategies are listed in Table 2. The optimal strategy includes: optimal assembly sequence, combination of optimal suppliers, and minimized order time and costs.

Table 1. The results of experiment

	MPSO	**MOGA**[A]
Operation time	37.66667	33.53333
Convergence algebra	2125	4897
Mean value of minimum fitness function	408083.3	408858.3
Average number of Pareto-optimal solutions	30[B]	28.6
Average ratio of Pareto-optimal solutions	86.67%[B]	66.67%
Performance ranking	1. MPSO; 2. MOGA	

A. Crossover method ORDER1, the assembly sequence adopts a single gene insertion mutation method, and the supplier adopts multi-point mutation method. (*Po*, *G*, crossover rate *CR*, mutation rate *MR*)=(30, 5000, 0.5 , 0.25)

B. Indicates p-value<α.

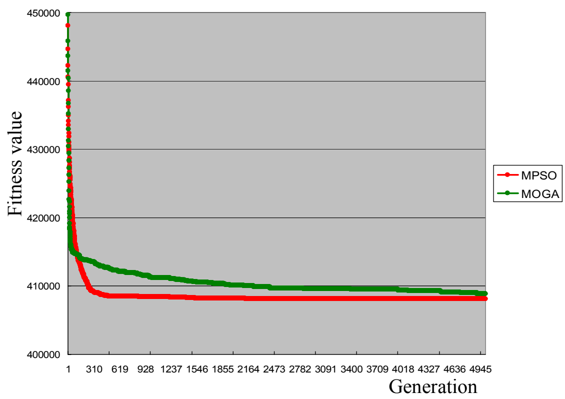


Figure 3. Comparison of converged generations of MPSO and MOGA

5 Conclusions and Suggestions

This research developed a bi-objective optimization decision-making model, and applied it to ASP and supplier selection planning. Given that the market demands are met, this planning could determine the combinations of optimal suppliers, and provide the optimal assembly sequence, in order toe minimize the time and costs for completing the orders. In addition, MPSO was used for solving the optimal mathematical decision model. The results indicated that MPSO not only could allow for rapid convergence, but also obtain better Pareto-optimal solutions when dealing with the assembly sequence and supplier evaluation planning. Based on the research findings, future studies can discuss the planning of suppliers or assembly involving production loss.

Table 2. Combinations of optimal strategies

Solutions	Objectives (f_1, f_2)	Assembly sequence	Parts supplier
1	(133000, 274250)	$M_3 \rightarrow M_2 \rightarrow M_1 \rightarrow M_5 \rightarrow M_7 \rightarrow M_8 \rightarrow M_6 \rightarrow M_4 \rightarrow M_9$	$A_2,B_4,C_1,D_3,E_4,F_4,G_4,H_4,I_2, J_4,K_1,L_3,M_3,N_1,O_2,P_1,Q_2,R_3$
2	(133000, 274250)	$M_2 \rightarrow M_8 \rightarrow M_3 \rightarrow M_6 \rightarrow M_5 \rightarrow M_7 \rightarrow M_1 \rightarrow M_4 \rightarrow M_9$	$A_2,B_4,C_1,D_3,E_4,F_4,G_4,H_4,I_2, J_4,K_1,L_3,M_3,N_1,O_2,P_1,Q_2,R_3$
3	(137500, 271000)	$M_2 \rightarrow M_1 \rightarrow M_3 \rightarrow M_5 \rightarrow M_6 \rightarrow M_7 \rightarrow M_4 \rightarrow M_8 \rightarrow M_9$	$A_2,B_4,C_1,D_3,E_4,F_4,G_4,H_4,I_2, J_3,K_2,L_3,M_1,N_2,O_2,P_1,Q_2,R_3$

6 References

[1] ALTIPARMAK F, GEN M, LIN L, PAKSOY T. A genetic algorithm approach for multi-objective optimization of supply chain networks. Computers & Industrial Engineering 2006;51:197–216.

[2] BOOTHROYD G, DEWHURST P, KNIGHT W. Product design for manufacturing and assembly. Marcel Dekker, New York, 1994.

[3] CHEN RS, LU KY, YU SC. A hybrid genetic algorithm approach on multi-objective of assembly planning problem. Engineering Applications of Artificial Intelligence 2002;15:447–457.

[4] CLERC M. The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization. In: Proc CEC 99. IEEE Service Center, Washington, 1999; 1951–1957.

[5] CORANA A, MARCHESI M, MARTINI C, RIDELLA S. Minimizing multimodal functions of continuous variables with the simulated annealing algorithm. ACM Transactions on Mathematical Software 1987;13:262–280.

[6] MOHEMMED AW, SAHOO NC, GEOK TK. Solving shortest path problem using particle swarm optimization. Journal of Applied Soft Computing 2008;8:1643–1653.

[7] TSENG HE, LI JD, CHANG YH. Connector-based approach to assembly planning using genetic algorithm. International Journal of Production Research 2004;42:2243–2261.

Automobile Manufacturing Logistic Service Management and Decision Support Using Classification and Clustering Methodologies

Charles V. Trappey[a], Amy J.C. Trappey[b,c,1], Ashley Y.L. Huang[b] and Gilbert Y.P. Lin[b]

[a] Department of Management Science, National Chiao Tung University, Hsinchu, Taiwan
[b] Department of Industrial Engineering and Engineering Management, National Tsing Hwa University, Hsinchu, Taiwan
[c]Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei, Taiwan

Abstract. Given the growing complexity of consumer preferences and the underlying market advantages of addressing these preferences, manufacturers and logistic service providers constantly monitor supply chain efficiency and quality requirements. Third-party logistic services are offered as a means to attract customers and enhance competitiveness as long as these services are effectively integrated into the order fullfilment processes. This research uses customer preference attributes to define distinctive dilivery and distribution of orders. The clustering and classification methods provide decision support capabilities to logistics providers so that they can adapt processes to satisfy specific customer preferences. A K-means clustering algorithm clusters customers' orders using demand attributes. Second, a decision tree classification approach analyzes each cluster segment using the history of consumer order preferences. Thus, the cluster results are the input data for the classification of logistics operations. The logistics service provider's delivery services are tailored to satisfy each customer's order requirements and preferences.

Keywords. Automotive manufacturing industry, third party logistics provider, K-means clustering, decision tree classification

1 Introduction

Automobile and automobile parts manufacturers recognize the strategic importance of flexible and efficient manufacturing as well as fully integrated delivery systems

[1] Corresponding Author : College of Management, National Taipei University of Technology, Taipei (10608), Taiwan; Tel: +886-2-2771-2171 ext. 4541; Fax: +886-2-2776-3996; Email: trappey@ntut.edu.tw and trappey@ie.nthu.edu.tw.

for the global market. Current research over the last decade has demonstrated many ways to integrate and synchronize materials and information flows, enhance supply chain efficiency, and reduce logistic management costs for supply chain intermediaries. An important development is the outsourcing of logistics services. Research [1, 2, 3] shows that third-party logistics (3PL) providers enhance product and materials availability and facility the consolidation of orders. Using 3PL services, the supply chain intermediaries increase information sharing, build more robust working relations, and decrease the uncertainty of the production life cycle [4]. The scope of logistics management has been extended from intra-organization to inter-organizational coordination and combines customized services with the building of long-term customer loyalty [5, 6]. In order to achieve satisfactory outsourcing goals, the participating supply chain members must re-evaluate and often re-engineer their logistic process models and business strategies.

Modern business research promotes customer relationship management (CRM) as a means to preserve the value of customers, target profitable segments, and cultivate high-quality relationships to ensure customer loyalty [7, 8]. Logistic service providers apply CRM as a means to analyze orders and customize delivery preferences. By dividing target customers (i.e., manufacturers, wholesalers, retailers, and customers) into sub-markets, suitable logistic services can be selected that best meet customer requirements and preferences. One of the most useful tools for managing market diversity is the application of customer segmentation techniques [9, 10, 11]. Customer segmentation is a core function of CRM and frequently uses geographic data, demographic data, and behavioral variables to group customers [12]. Even with the applications of market segmentation, not all customer classifications satisfy profitability requirements. Thus, companies are categorizing customers in terms of their profitability and target customers that contribute the most towards profit objectives. Many different methods and techniques are utilized to precisely identify different levels of customer profitability and service suitability.

Data mining is a technique which enables the extraction of information from large amounts of historical data [13]. Through data mining, companies identify valuable customers, predict behavior, and support customized service strategies. Several researchers use data mining techniques to extract meaningful patterns and build predictive customer relationship models [14, 15, 16]. Clustering, as a data mining approach [17, 18, 19], groups members that are fairly homogeneous within clusters (minimum variance) but significantly heterogeneous between clusters (maximum variance). Since analysts do not need to predefine customer group membership, businesses can concentrate marketing efforts using readily available attributes to create the clusters

This research provides a methodology for 3PLs companies to customize services for the Taiwan automobile parts and components manufacturing industry. In the first phase of the study, the 3PL firms are surveyed to define the current or as-is logistic models. In the second phase, the 3PL customers are classified into distinct clusters so that specific logistic services can be defined for members of a given cluster. Following the customers' preferences and the recorded product distribution attributes for delivery orders, the goal of the cluster classification method is to provide decision support, to improve logistics effectiveness, and to lower distribution costs. There are two levels of analysis for the methodology. First, K-means clustering segments customers' orders according to the demand attributes. Second, a decision tree classifies each sub-segment based on the customer's order preferences. That is, the first level clustering results are the data input for the classification of logistic operations and modeling. Using these results, the 3PLs provide customized deliver services for each customer group based on their order requirements and preferences.

This research paper is organized as follows. Section 2 discusses the current logistic management processes in the automotive supply chain from the viewpoint of third-party logistic service providers). In Section 3, the two-phase clustering and classification methods are developed and discussed. Finally, the conclusion and future research are described in Section 4.

2 Logistic Services in the Automobile Manufacturing Industry

This section describes the supply chain participants and the products provided by the automobile parts and component manufacturers. The logistic services offered by 3PLs and the corresponding logistic models are also discussed in this section.

2.1 The Logistic Processes of he Automobile Manufacturing Industry

The automobile industry channel intermediaries consist of suppliers, third party logistic service providers (3PLs), automobile manufacturers, automobile parts and component manufacturers, dealers, independent distributor sites, and customers as shown in Figure 1. The supply chain enables manufacturers to receive knock-down parts from suppliers, manufacture the finished goods (vehicles or after-market parts), and deliver the goods to the end customers. The product types provided by the automobile industry include completed vehicles, automotive parts and components. The parts manufacturers are classified as Original Equipment Manufacturers (OEM) and the After Market (AM) parts suppliers. The OEM suppliers sell parts and components directly to the automobile manufacturers that assemble the final vehicles. The AM parts suppliers provide parts and components for maintenance service providers and the retail marketplace. The logistics activities conducted among the supply chain members include automotive parts delivery, parts assembly, vehicle distribution, warehouse management, and other value-adding processes. The auto-makers and parts suppliers maintain their own warehouses to consolidate goods before shipment to intermediary and final destinations. The 3PLs are contracted to by the manufacturers to help manage all aspects of the intermediary logistic tasks.

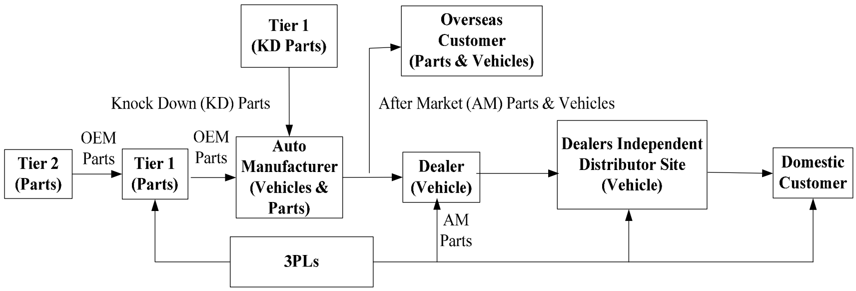


Figure 1. The business hierarchy and participants for the automobile industry [20]

The logistic processes for sellers of domestic vehicles are classified into four process types. Type one is the case where a 3PL receives delivery notice from a manufacturer and manages deliveries from the inventory of cars kept in the 3PL's warehouses. Type two is the case where a 3PLs receives new cars from a manufacturers' assembly line and delivers directly to a dealer's service location according to the manufacturer's instructions. For case three, the 3PL delivers new cars to an assigned location according to the dealers' instructions. Finally, case four occurs when the 3PLs receive the manufacturers' delivery notice to manage all logistic services for imported vehicles including transportation and storage at the manufacturers' warehouse. For new vehicle export, the 3PL receives the delivery requirements from the auto-maker, transports the vehicle to the seaport.

The logistic processes for domestic OEM automobile parts and components are classified into four types. These processes include cases where OEM parts and/or components are delivered to the assembly line just-in-time (JIT), deliveries are made to the central manufacturer's warehouse in advance to cover unexpected demand, parts are delivered to the 3PL warehouse and then the 3PL manages the inventory and coordinates delivery to the manufacturers. Finally, there is a milk run delivery case where 3PLs collect parts from OEM parts suppliers on a regular basis and deliver the parts to the auto-maker's warehouse.

2.2 The Logistic Operation Models of Third-party Logistic Service Providers

The automobile manufacturing industry customers, as participants in the supply chain processes, have distinct and variable order requirements. Therefore, 3PLs that provide customized services enhance the logistic efficiency of their customers and gain a competitive advantage.

The logistics service functions offered by 3PLs are divided into two categories, including basic logistic services and the high value-added services (Table 1). The basic service category includes loading and unloading goods (1), packaging goods (2), warehousing (3), and delivering on request (4). The high value added services includes circulation processing (5) and information reporting (6). According to each customer's needs, the 3PLs offer specific service combinations for each operation model.

Table 1. 3PL logistic operation services

Operation Model \ Service Category	Service Combination
Auto-maker logistics model	(1), (2), (3), (4), (5), (6)
OEM parts supplier logistics model	(1), (2), (3), (4), (6)
AM parts supplier logistics model	(1), (4), (6)
Dealer logistics model	(1), (3), (4), (5), (6)

3 Decision Support Methdology

This section describes the decision support methdology which combines K-means clustering and decision tree classification (Figure 2). The automobile manufacturer's customer order preferences are the input parameters which classify different customer groups' order requirements. Ten variables are selected from the customer requirements and product attributes to create to distinct cluster (Table 2). The data used as input, the K-means clustering algorithm, the decision tree classification algorithm, and the customized stratigic logistics decision support are discussed in this section.

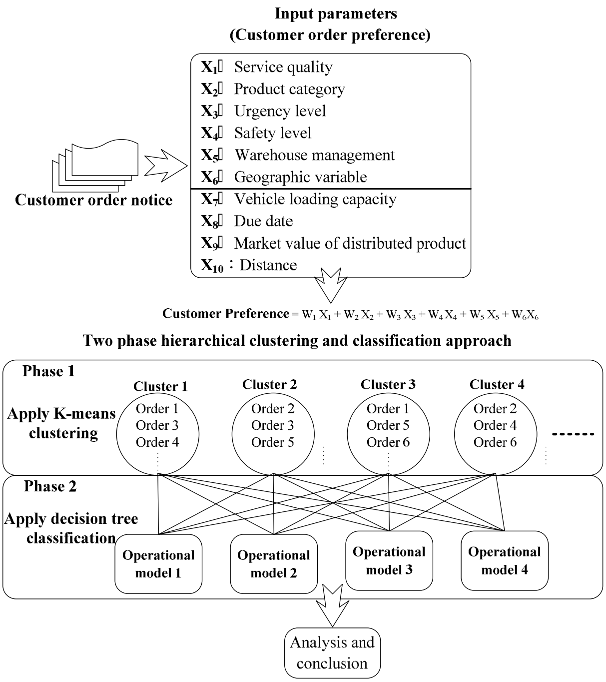


Figure 2. The two phase hierarchical analysis approach

Table 2. Customer preference indices

Variable / Category	Parameter	Index description	Attribute
Customer Requirements	X_1	Service quality	Qualitative value
	X_2	Product category	Qualitative value
	X_3	Urgency level	Qualitative value
	X_4	Safety level	Qualitative value
	X_5	Warehouse management	Qualitative value
	X_6	Geographic variable	Qualitative value
Attributes of Distributed Product	X_7	Vehicle loading capacity	Quantitative value
	X_8	Due date	Quantitative value
	X_9	Market value of distributed product	Quantitative value
	X_{10}	Delivery distance	Quantitative value

3.1 Data Collection

Before applying the decision support methodology, the customer's preferences and requirements are collected and organized. The 3PLs use order management systems, ERP systems, telephone calls and faxes to receive customer's orders as shown in Figure 3. The customer orders describe the attributes of the product to be delivered, the delivery quantity, the due date, the degree of urgency, the pick-up location, and the delivery destination. The customer's transactions and behaviors are recorded and consolidated in a data warehouse. The data are then used to analyze and segments the customers into groups. The results enable the 3PLs to support strategic service decision making and offer greater customization.

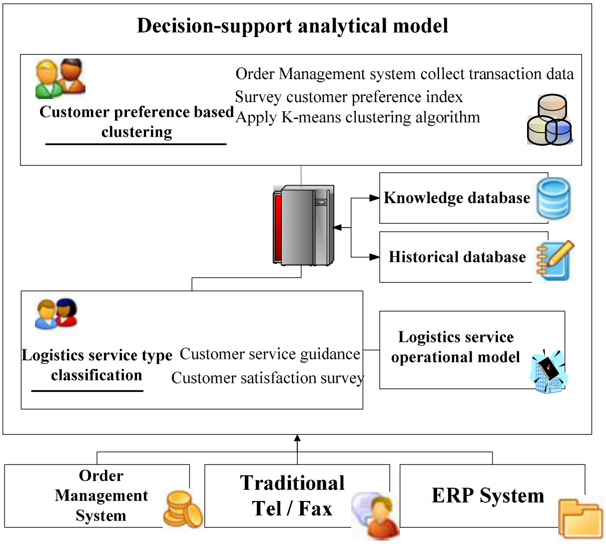


Figure 3. The two phase analytical model for logistic service decision support

3.2 Clustering to Segment and Target 3PL Customers

The K-means algorithm is implemented using STATISTICA 7 to creates sub- 3PL customer segments. The customers are clustered based on their order preferences and the attributes of the products managed and delivered. Figure 4 illustrates the clustering process flow.

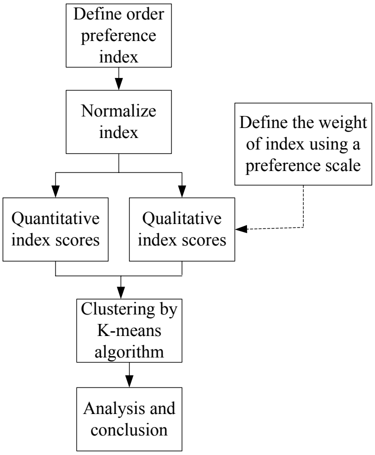


Figure 4. Clustering process flow

- Step 1: The K random seed data points are selected. The left side of Figure 5 shows that K equals the number of clusters and the number of seed data points.
- Step 2: Three seeds are used to define the initial cluster boundaries.
- Step 3: The centroids for each cluster are used to determine the initial cluster boundaries.
- Step 4: The process of assigning entities to clusters is repeated until the cluster boundaries cease to change.

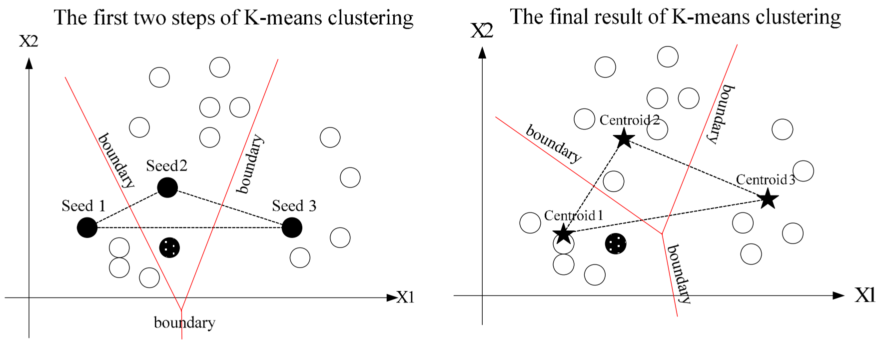


Figure 5. The first two steps for K-means clustering methodology

The customer preference (CP) value is calculated using the ten attributes of the customer order requirement. Customer preference scores are computed using the following equation where X_1 through X_6 are qualitative indices and X_7 through X_{10} are quantitative indices. W_1 through W_6 are the weights of the parameters.

$$CP = W_1X_1 + W_2X_2 + W_3X_3 + W_4X_4 + W_5X_5 + W_6X_6 \quad (1)$$

The weights of the qualitative parameters are adjusted using neural network training. A Likert-like questionnaire is used to record the customers' order preferences. The five interval scale for the qualitative indices ranges from most important to most unimportant. After training a weight matrix assigns customer preferences to the range 8 through 10 if the customer chooses "most important," 6 through 8 if the customer chooses "important" and downward to 0 through 2 if the customer chooses "most unimportant." The flow path for assigning weights is shown in Figure 6.

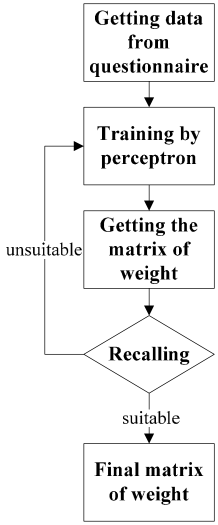


Figure 6. The weight training procedure for qualitative parameters

Explain the process from training weights:

- Step 1. Produce weight matrix between 0~1 randomly.
- Step 2. Input the normalized data.
- Step 3. Compute the net_j

$$net_j = \sum_i w_{ji} x_i - \theta_j \tag{2}$$

θ_j is the limited value and is selected by the customer. If the customer chooses "most important" in the questionnaire, his θ_j equals 9. If customer chooses "important," his θ_j equals 7. If customer chooses "medium," his θ_j equals 5. If the customer chooses "unimportant," his θ_j equals 3. If customer chooses "most unimportant," his θ_j equals 1.

- Step 4. Compute the ΔW_{ij} (η is the learning rate)

$$\Delta W_{ij} = \eta \text{x } net_j \text{ x } (\theta_j - X_j) \tag{3}$$

- Step 5. Update the weight W_{ij}

$$W_{ij} = W_{ij} + \Delta W_{ij} \tag{4}$$

- Step 6. Repeat step 2 ~ step 5 in a specific number.

3.3 Decision Tree Classification

A decision tree classification approach further classifies the customer segments with homogeneous order preferences into the appropriate logistic operation model as shown in Figure 7.

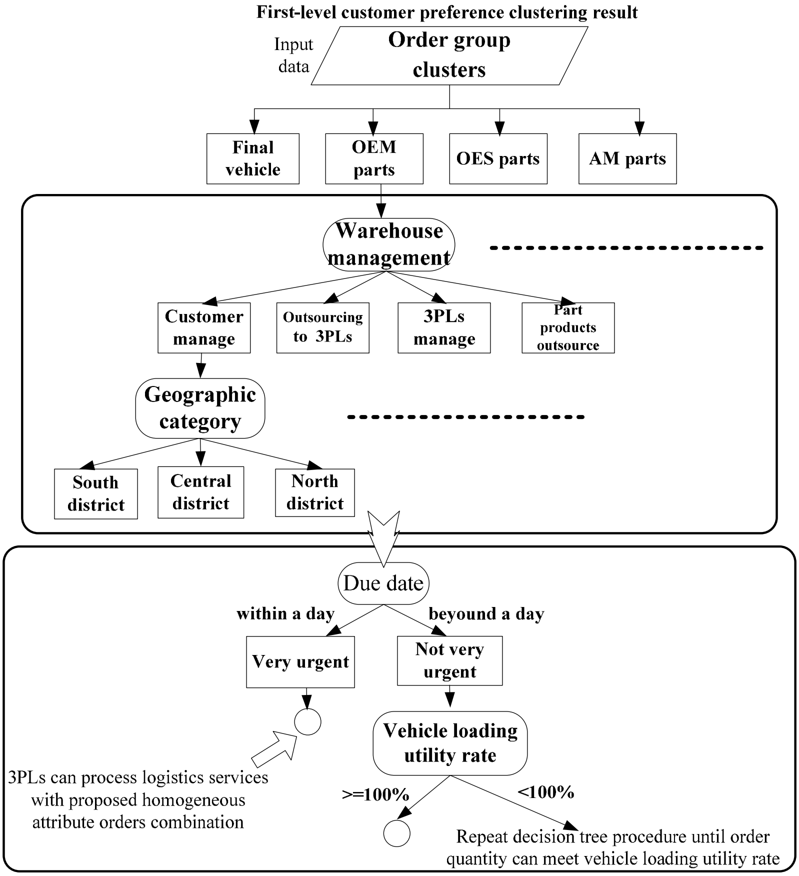


Figure 7. Decision tree classification used to select operation model for 3PL customers

The distinctive segments are derived according to customer order preferences as described in Section 3.2 and then the clustering results are placed into suitable 3PL service models. The decision tree algorithm is implemented at this level. The classification approach uses the specified criteria as a partition node to generate correlative rules. For example, if two orders' warehouse management type is the same (e.g., managed by the 3PL) and their delivery geographic location is also held in common (e.g., north region), then, the study classifies the two orders into the same service policy model. Similarly, the study implements correlation rules using pre-classified orders which satisfy the two conditions in the same manner. For example, if the due date is not very urgent (beyond a certain day) and the vehicle loading utility rate does not exceed 100 %, then, we suggest that 3PL processes these orders at the same time. Thus, 3PLs can utilize the methodology to develop support decisions, customize services, and improve the management of their logistic services.

4 Conclusions

This paper presents a logistics decision support method to help third party logistics service providers customizing services for customers in the automobile manufacturing supply chain. Site visits were used to collect customer data and train a neural network of preferences. Then, a two phase (clustering and classification) methodology was developed to target customer segments, analyze their preferences and provide best case service models. This method simultaneously satisfies customers' requirements and coordinates deliveries. As a result, the 3PLs can better support decisions and standardize service that builds customer loyalty.

5 References

[1] Sohail, M.S. and Sohal, A.S. The use of third party logistics services: a Malaysian perspective. Technovation, vol. 23, no. 5, pp. 401-408, 2003.
[2] Carbone, V. and Stone, M.A. Growth and relational strategies used by the European logistics service providers: Rationale and outcomes. Transportation Research Part E, vol. 41, no. 6, pp. 495-510, 2005.
[3] Kim, C., Yang, K.H. and Kim, J. A strategy for third-party logistics systems: A case analysis using the blue ocean strategy. Omega, vol. 36, no. 4, pp. 522-534, 2008.
[4] Trappey, C.V., Trappey, A.J.C., Lin, G.Y.P., Liu, C.S. and Lee, W.T. Business and logistics hub integration to facilitate global supply chain linkage. Journal of Engineering Manufacture (Proceedings of the Institution of Mechanical Engineers Part B), vol. 221, no. 7, pp. 1221-1233, 2007.
[5] Sihn, W. and Schmitz, K. Extended multi-customer supplier parks in the automotive industry. CIRP Annals-Manufacturing Technology, vol. 56, no. 1, pp. 479-482, 2007.
[6] Gol, H., Third-party logistics provider selection: insights from a Turkish automotive company. Supply Chain Management: An International Journal; vol. 12, no. 6, pp.379-384, 2007.
[7] Chen, I.J. and Popovich, K. Understanding customer relationship management (CRM) People, process and technology. Business Process Management Journal, vol. 9, no. 5, pp. 672-688, 2003.
[8] Peppard, J. Customer relationship management (CRM) in financial services. European Management Journal, vol. 18, no. 3, pp. 312-327, 2000.
[9] Freytag, P.V. and Clarke, A.H. Business to business market segmentation. Industrial Marketing Management, vol. 30, pp. 473-486, 2001.
[10] Kim, S.Y., Jung, T.S., Suh, E.H. and Hwang, H.S. Customer segmentation and strategy development based on customer lifetime value: A case study. Expert Systems with Applications, vol. 31, pp. 101-107, 2006.
[11] Chu, C.C.H. Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer. Expert Systems with Applications, vol. 34, no. 4, pp. 2754-2762, 2008.
[12] Lee, J.H. and Park, S.C. Intelligent profitable customers' segmentation system based on business intelligence tools. Expert Systems with Applications, vol. 29, no. 1, pp. 145-152, 2005.
[13] Rygielski, C., Wang, J.C. and Yen, D.C. Data mining techniques for customer relationship management. Technology in Society, vol. 24, no. 4, pp. 483-502, 2002.

[14] Magnini, V.P., Honeycutt, E.D. and Hodge, S.K. Data mining for hotel firms: use and limitations. Cornell Hotel and Restaurant Administration Quarterly, vol. 44, no. 2, pp. 94-104, 2003.
[15] Crespo, F. and Weber, R. A methodology for dynamic data mining based on fuzzy clustering, Fuzzy Sets and Systems, vol. 150, no. 2, pp. 267-284, 2005.
[16] Lau, K.N., Lee, K.H. and Ho, Y. Text mining for the hotel industry. Cornell Hotel and Restaurant Administration Quarterly, vol. 46, no. 3, pp. 344-362, 2005.
[17] Crespo, F. and Weber, R. A methodology for dynamic data mining based on fuzzy clustering. Fuzzy Sets and Systems, vol. 2, no. 1, pp. 267-284, 2005.
[18] Kuo, R.J., Wang, H.S., Hu, T.L. and Chou, S.H. Application of ant K-means on clustering analysis. Computers and Mathematics with Applications, vol. 50, no. 10-12, pp. 1709-1724, 2005.
[19] Hsu, F.C., Trappey, A.J.C., Trappey, C.V., Hou, J.L. and Liu, S.J. Technology and knowledge document cluster analysis for enterprise R&D strategic planning. International Journal of Technology Management, vol. 36, no. 4, pp. 336-353, 2006.
[20] Trappey, C.V., Trappey, A.J.C., Liu, C.S., Lee, W.T. and Hung, Y.L. The design and evaluation of a supply chain logistic hub for automobile and parts distribution. Materials Science Forum, vol. 594, pp 119-131, 2008.

Lead Time Reduction by Extended MPS System in the Supply Chain

JrJung Lyu[a,1] and Hwan-Yann Su[b]

[a] Department of Industrial and Information Management,
National Cheng Kung University, Tainan, Taiwan, R.O.C.
[b] Department of Asia-Pacific Industrial and Business Management,
National University of Kaohsiung, Kaohsiung, Taiwan, R.O.C.

Abstract. The problems which frequently challenge the management of a supply chain including uncertainty of customers' demands, high inventory levels and costs, inaccurate due date estimation, and slow response to customer inquiries are addressed. Lead time reduction is a critical issue which enables manufacturers to solve problems. The proposed extended master production scheduling (MPS) system, developed using Internet technology, can be deployed in a supply chain environment. A case study is used to illustrate the implementation process of the system and accompanying benefits.

Keywords. Lead Time, supply chain management, master production scheduling

1 Introduction

Uncertain demands from the customer such as urgent orders, sudden increase of ordered units and requests to deliver before the due date are common challenges for the management of a supply chain. In order to respond quickly to these uncertainties, manufacturers in supply chains usually adopt a strategy of increasing their inventory level. However, high inventory levels may result in other problems such as high inventory costs, the increased risk of inventory depreciation and the unnecessary occupation of space. In addition, slow response to customer inquiries about due date, inaccuracy in promised due date and late delivery schedule create unhappy customers. Lead time reduction plays a key role in dealing with the problems mentioned above. As pointed out by Wedel and Lumsden [10], short lead times may improve customer service and reduce inventory level. Moreover, it is also suggested that lead time reduction is itself a very significant productivity improvement driver and a prime metric for performance [8]. In many situations, lead time is controllable and can be shortened by adding extra cost to reduce

[1] Corresponding Author Address: Department of Industrial and Information Management, National Cheng Kung University, Tainan 701, Taiwan, R.O.C.; Tel: (886) 62080576; Fax: (886) 62501734; E-mail: jlyu@mail.ncku.edu.tw

investment in safety stocks held to fulfill the unanticipated demand during lead time [3].

In this research, the authors argue that extended master production scheduling (MPS) provides good opportunity for lead time reduction in the supply chain. Traditionally, MPS is used for the arrangement of orders, manufacturing and production control of individual products in many companies. MPS is the key to balancing sales and production. This research develops an effective extended MPS that is necessary for the achievement of customer satisfaction and efficient utilization of production capacity in a supply chain.

2 Literature review

The application of information technology to enable rapid and effective intra-organizational and inter-organizational communication and coordination, to remove intra-organizational and inter-organizational barriers and to achieve customer satisfaction and lead time reduction, has been getting more and more emphasis [2, 6]. Rho and Yu [9] argued that lead time reduction might eliminate the unnecessary work-in-process (WIP) materials in production lines and unnecessary activities. They stressed that lead time reduction was a new world-class strategy which might increase productivity and competitivity, improve the speed of service provision and enhance profits and sales. Emphasizing the reduction of information transmission lead time from one end of the supply chain to the other, Mason-Jones and Towill [7] argued that the essence of an agile supply chain was its ability to respond quickly and efficiently to a fast-changing market. The key feature of an agile supply chain is how each of its members reduces the lead time from receiving demand information from the customer to transmission of the information to the suppliers. They pointed out that the shortening of information transmission lead time might be achieved by the deployment of information technology. Agrawal et al. [1] indicated that accurate lead time estimation was very important for production in accordance with schedule and reduction of inventory levels. Overall, as suggested by Hameri and Paatela [4], improving productivity through the reduction of process lead times is a well-known fact in manufacturing studies. Hnaien et al. [5] dealt with the problem of planned lead time optimization under stochastic lead times. In this paper, the focus is on exploring the opportunities for the reduction of order processing lead time, supply management lead time and manufacturing lead time.

A MPS system is developed to assist the decisions regarding what, when and how many end items will be manufactured. Taking MPS output as the input for materials requirements planning (MRP) and capacity requirements planning (CRP), these three systems are used for separate decision making. In the literature, suggestions for improving MPS by estimating, correcting and reducing lead time include:

- adjust MPS until MRP is consistent with the availability of production capacity;
- manage lead time, manually intervene when necessary and increase short-term capacity;

- monitor the deviation of lead time and correct it;
- utilize MRP-based systems to aid the estimation of lead time and the release of production orders to the shop floor.

This paper argues that it is necessary to extend MPS to be integrated with MRP and CRP in a supply chain, from the perspective of lead time and inventory level reduction and efficient utilization of production capacity.

3 The e-MPS system

The proposed extended master production scheduling system (e-MPS) aims to deal with the problems of uncertainty of customer demand, slow response to customer inquiries into due date, inaccuracies in due date estimation, inadequate schedules for production and inefficient utilization of production capacity through the deployment of information technology and new ways of thinking about business processes. It can be integrated with the Enterprise Resources Planning (ERP) system or the Manufacturing Execution System (MES) system of a business.

The proposed framework of the e-MPS system integrates with the existing internal information system as shown in Figure 1. It is assumed internal information systems such as an order management system, production management system and production control system have already been implemented in the business. If these information systems have not been implemented, it is assumed manual operation may offer similar information as required.

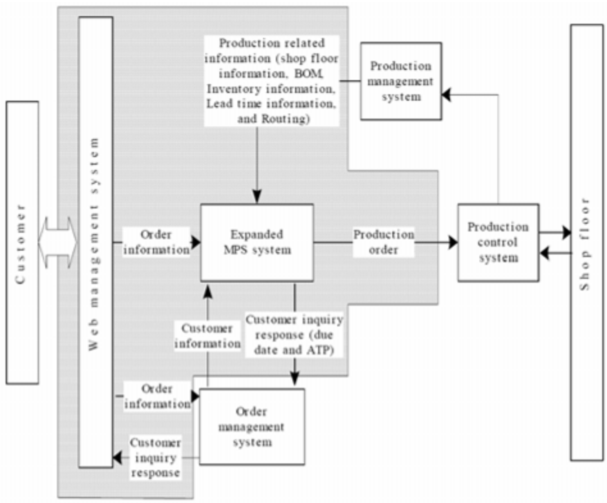


Figure 1. Proposed framework of the e-MPS system

The e-MPS system includes two major modules – a web management system and an expanded MPS system - and their connection and interaction with external customers and three internal information systems. The modules of the e-MPS system and how the related internal information systems interact with them are detailed below:

- Web management system: This system is the interface between the business and its customers. It provides the functions of account management, dialogue management and customer inquiry information management.
- Expanded MPS system: This system provides the basic functions of MPS, MRP, CRP, as well as the estimation of due date and available-to-promise (ATP).
- Order management system: This system provides the functions of order management and customer information management.
- Production control system: This system provides the functions of shop floor management and the management of shop floor related information such as inventory status, production capacity availability, load and the status of schedules.
- Production management system: This system provides management information regarding production planning such as bills of material (BOM), routing and inventory information and management information regarding the shop floor provided by the production control system.

4 A practical application for e-MPS system

4.1 Background

The case study was a steel company that produced mainly stainless steel and structural steel. In the stainless steel product group, the company had a yearly capacity of 20,000 tons of steel plate and it also manufactured steel ingots and hot-rolled straight bars. In the structural steel product group, it produced stainless steel cord, seamless pipes and bars with superior purity, stability, and precision tolerance. The case had earned many product quality and processing certifications including ISO 9001:2000 accreditation and American Association for Laboratory Accreditation (A2LA) laboratory and machinery certification.

The case company has a tradition of investment in information technology. The original information system implementation is shown in Figure 2. SAP, an ERP system, had been installed. The modules implemented included the SD module for order management, PP module for production planning, MM module for material inventory management, FI and CO module for finance and accounting, and PM for plant maintenance. There was also a MES developed internally. The PRP module in this MES had the function of issuing production orders. The other modules include the shop floor control module for control of shop floor, inventory management model for the function of managing inventory, and shipping module for transportation management. The LAB, MTS and MIC modules were developed to meet the quality regulations. CIM (Steel rolling) and CIM (Steel refining) modules had been developed for specific production management. There was also a legacy scheduling system that was written in Microsoft Excel.

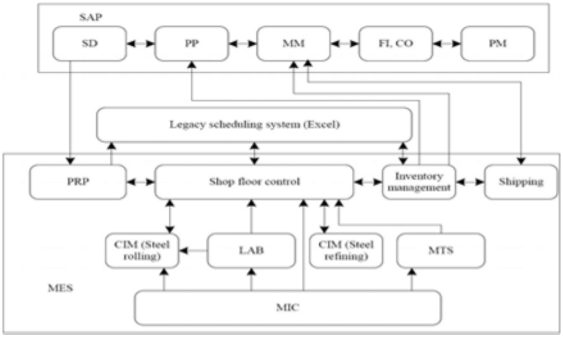


Figure 2. The original internal information systems of the case

Despite the available systems, the managers, however, faced some serious problems including inaccurate production scheduling resulting in frequent postponement of manufacturing and inaccurate due date estimation. These problems made the case company difficult to quickly respond to customer inquiries and to prepare needed raw materials. Inventory costs were also high compared to its competitors.

4.2 The implementation plan

In order to meet growing customized needs regarding various kinds of metals to be included for refinery and the sizes to be rolled, the case had developed a three-phase implementation plan for the introduction of the e-MPS system as shown in Figure 3. Details of each phase are described below.

Phase one was the planning phase. The first step was to form a project team and to determine an appropriate MPS software. Existent MPS software such as Mimi or i2 were suggested to quickly enable the case to determine ATP accurately and to reduce order processing lead time. The index for lead time and other related indicators required determination and utilization to monitor the new system and its effects. The order response time and on time shipping rate, which were related to customer service level, supply management lead time and manufacturing lead time, were selected in this case.

Phase two was the system design phase. The scope of the system had to be defined. Thus, the production management process had to be clearly re-defined and production related information also had to be collected. A conceptual model regarding how the production management process should work was well discussed in this phase. Furthermore, the input and output of the data flow of the system had to be clearly defined in this phase. The related data such as customer information or production capacity had to be collected and confirmed. When the relevant information was collected and the conceptual model regarding how the production management process would work in detail was defined, a complete structure of the system was designed.

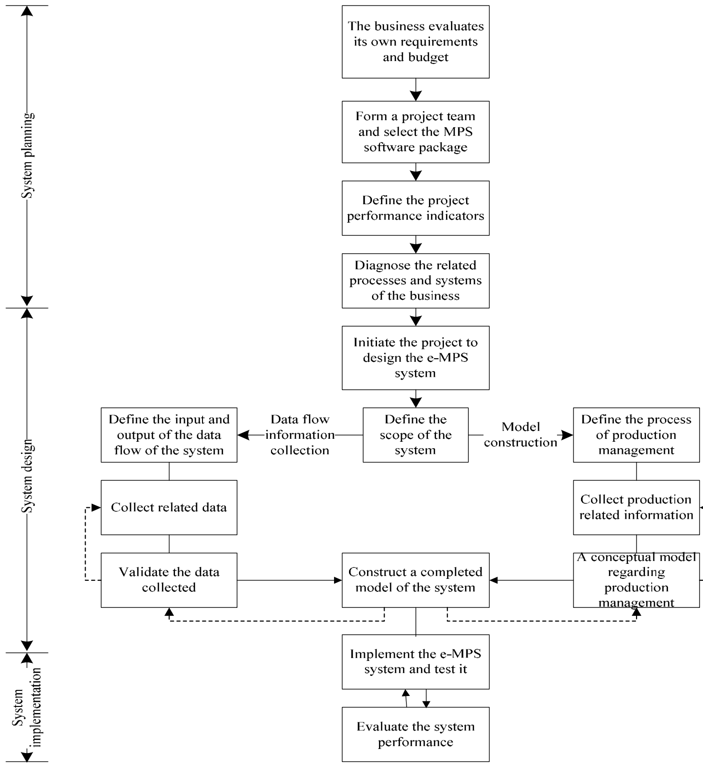


Figure 3. Implementation process of the e-MPS system in the case

Phase three was the system implementation phase. As shown in Figure 4, consider the SD module as an order management system, the PP, MM, PRP, Inventory management, CIM modules as a completed production management system, the shop floor control as a production control system. In this case, it was suggested that with in mind the functions of MPS, MRP and CRP must be considered as a whole, the e-MPS system be implemented by an existent MPS software mentioned above which can integrate with the original information systems. Thus, the e-MPS system, which includes a web management system and the expanded MPS system, connected and interacted with the rest of the systems. The integrated system was validated and verified by the project team carefully according to the specified functions and the set indexes.

4.3 Benefits of the System

Compared to the original process, the modified order management and production planning processes are different in the way that when a customer issues an order through the web management system, the order information is directly inputted to the order management system and simultaneously passed to the e-MPS system. Due date and ATP are calculated by the software as a whole and interact with the order management system. The expected lead time reduction in the order and production planning processes is achieved because order management and production planning are done at the same time. Since the manual operation between the customer and the sales department is replaced by the Internet interface of the web management system and that the manual operation between the sales department and the production department is also replaced by the automatic e-MPS system, the unnecessary queuing between the customer, the sales department and the production department can be effectively eliminated. The manufacturing lead time is reduced because production schedules become accurate in the e-MPS system and the materials, components, or WIP needed for production can be prepared based upon the production capacity where the supply management lead time can be further reduced.

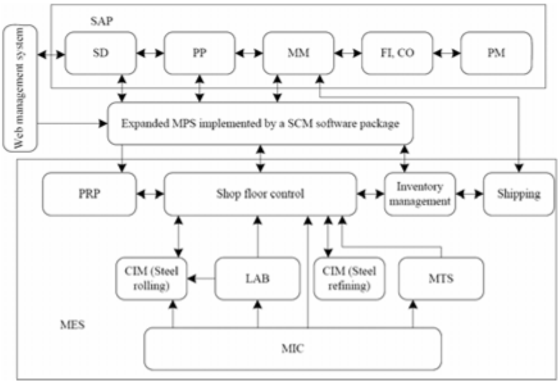


Figure 4. The e-MPS that integrates with the original internal information systems in the case

After a two years project, it was estimated that the manufacturing lead time was reduced from an average 11 days to around 4 days due to the smoothing of the bottleneck processes. The WIP materials between the rolling and the acid processing processes were also reduced from an average 2000 to 2500 tons of steel to about 500 tons for the same reason. The on-time shipping rate was increased from the current 65% to about 90% and the response time to customers regarding due date inquiries was shortened to about half an hour from the original 2 days on average. It was expected that the performance of the system would be even better if the system was implemented under an agreement involving supply chain collaboration with customers regarding a pre-planned production schedule where the production system could prepare and reserve the required materials and production capacity according to the collaborated production schedule and modify it according to the actual demands. It was decided by the case company to leave this initiative as a project for the next phase.

5 Concluding remarks

To respond quickly to the uncertainties of customer demand, manufacturers in supply chains should adopt a better strategy and use a modern Internet-based operations management system. This research proposes an extended master production scheduling system that can deal with the uncertainty problem through lead time reduction. The strength of the proposed e-MPS system lies in its consideration of MPS, MRP and CRP as a whole as well as its development of the supply chain idea to connect the business with the customers over the Internet while integrating with the original internal information systems. As a result, the proposed framework offers a new way of thinking about business processes so as to allow production planning and order management to proceed together in order to reduce lead time further. The illustrated case demonstrated the implementation process of the developed e-MPS system, concerning how it may be planned, designed and implemented. While the benefits of the implementation of an e-MPS system are case dependent and further research is needed to assess and enhance the robustness of the proposed framework by evaluating its implementation in different industries, it is safe to state that it is a very useful means for many enterprises to dramatically improve their performance.

6 References

[1] Agrawal A, Minis I, Nagi R. Cycle time reduction by improved MRP-based production planning. Int J Prod Res 2000:38:4823-4841.
[2] Calosso T, Cantamessa M, Vu D, Villa A. Production planning and order acceptance in business to business electronic commerce. Int J Prod Econ 2003;85:233–249.
[3] Chang CT. 2005, A linearization approach for inventory models with variable lead time. Int J Prod Econ 2005;96:263–272.
[4] Hameri A, Paatela A. Supply network dynamics as a source of new business. Int J Prod Econ 2005;98:41-55.
[5] Hnaien F, Dolgui A, Ould Louly M. Planned lead time optimization in material requirement planning environment for multilevel production systems. J Syst Sci Syst Eng 2008;17:132-155.
[6] Lim D, Palvia PC. 2001, EDI in strategic supply chain: impact on customer service. Int J Inform Manage 2001;21:193-211.
[7] Mason-Jones R, Towill DR. Total cycle time compression and the agile supply chain. Int J Prod Econ 1999;62:61-73.
[8] Naylor B, Naim M, Berry D. Leagility: integrating the lean and agile manufacturing paradigms in the total supply chain. Int J Prod Econ 1999;62:107-118.
[9] Rho B. Yu Y. A comparative study on the structural relationships of manufacturing practices, lead time and productivity in Japan and Korea. J Oper Manag 1998;16:257-270.
[10] Wedel J, Lumsden K. The influence of lead-time reductions on decisions and rules in the production planning process. Int J Prod Econ 1995;41:399-404.

A Multi-Products EPQ Model with Discrete Delivery Order: a Lagrangean Solution Approach

Gede Agus Widyadana[a,b,1] , Hui Ming Wee[a]

[a]Department of Industrial Engineering, Chung Yuan Christian University,Chung Li 320, Taiwan.
[b]Department of Industrial Engineering, Petra Christian University, Surabaya, Indonesia.

Abstract: In this paper, a multi products EPQ vendor buyer integrated model in Just in Time (JIT) philosophy with budget constraint is developed. Buyer asks the vendor to deliver products in small batch with discrete delivery order. The problem is modeled as a Mixed Integer Non Linear Program (MINLP) with constraint. Langrangen method and branch and bound procedure are developed to solve the problem. One numerical example is shown to illustrate the model and solution procedures and the result shows that the Lagrangean solution approach can derive global optimum solution.

Keywords: EPQ, JIT, Lagrangean approach, Branch and bound.

1 Introduction

Inventory cost is one of significant cost in many companies. The inventory cost mainly consists of ordering cost, holding cost and delivery cost. Deciding how many order quantity and when to order is still a challenge to managers in both manufacturing and service industries. Some industries implement Just in Time (JIT) system to reduce their inventory cost. They ask vendor to send item when it is needed, so the item inventory can be kept as low as possible.

The most challenging aspects of inventory decision are multitude of items involved and sets of constraints. Many researchers developed single product item inventory model and assumed no constraints. In reality, there are some constraints in inventory problem such as warehouse capacity, transportation capacity and budget capacity.

In this paper a model of integrated single vendor single buyer with multi products is developed. Some research showed that vendor buyer integrated model has better performance than non integrated model in supply chain. Vendor and the

1 Corresponding Author : Department of Industrial and System Engineering, Chung Yuan Christian University, Chung Li 32023, Taiwan; Tel. 886-3-265-4458, Fax: 886-3-265-4499, Email: gedeaw@yahoo.co.id

buyer can share their costs to achieve global optimum supply chain. The model is operated in just in time system where buyer can set how many items should be ordered and how many items should be delivered in one shipment. The buyer pays the transportation cost. The model is an extension of the model by Pasandieh and Niaki [1]. Buyer budget is considered to be a constraint and the model is solved using Lagrangean methods. This research gives contribution by adding realistic constraint to buyer vendor production inventory problem and solves the problem efficiently.

Vendor - buyer integrated production inventory model has been developed in some research in recent years. Ouyang et al. [2] introduced single vendor-single buyer integrated production inventory model. They assumed lead time demand is stochastic. Sajadieh et al. [3] extended Ouyang et al. [2] by releasing an assumption that supplies lead time is constant. They assumed that supply lead time between vendor and buyer is stochastic and shortage is allowed. Ben-Daya and Hariga [4] developed a similar single vendor-single buyer production model like Ouyang et al. [2]. In their model, they assumed that the lead time is varying linearly with lot size. Hsiao [5] modified Ben-Daya and Hariga‘s [4] model using different assumptions. They assumed that there are two different reorder points and service levels.

The vendor-buyer production inventory models above assume the buyer requesting the vendor to send the products in small batches. This assumption adopts concept of just in time system. Study of just in time system has received much attention in recent years. Yang and Wee [6] developed multi lot size production and inventory model of deteriorating item under JIT philosophy. Wang and Sarker [7] developed a mixed-integer nonlinear programming (MNILP) model with kanban controlling a multi stage supply chain system. Rau and Ouyang [8] developed integrated single vendor single buyer production-inventory policy. They assumed that the planning horizon is finite and demand has linear trend.

In our literature search, most researchers did not consider constraints that might occur in vendor-buyer production system. In reality, transportation equipment capacity, warehouse capacity and buyer budget are limited. Hoque and Goyal [9] developed a single vendor-single buyer integrated production system with limited transportation equipment. Multi constraint model is introduced by Haksever and Moussourakis [10]. They presented a mixed-integer programming model for optimizing multi product inventory system with multiple constraints.

This paper is divided to five sections. In the first section, background of the problem and some related literature reviews are introduced. The problem and formulation are explained in section two. In section three, Lagrangean solution approach and branch and bound procedure are developed to solve the problem and then one example is shown in section four to illustrate the model development and the solution procedures. Conclusion is presented in the subsequent section.

2 Problem Definition and Formulation

The mathematical model developed in this paper is based on assumptions:

a. Single vendor and single buyer are considered.

b. Set-up and transportation times are insignificant and can be ignored.
c. Demand rate is constant.
d. Shortages are not allowed.
e. Time horizon is infinite.
f. All costs are known and constant.
g. Buyer pays transportation cost

Notations:

l	number of products
T_i	cycle time of product i
Q_i	order quantity of product i
m_i	shipment quantity of product i
K_i	number of shipments placed during a period T of product i
w_i	number of shipments placed during the production time of product i
P_i	vendor production rate of product i, units/year
D_i	the buyer's demand rate of product i, units/year
A_i	buyer ordering cost of product i
Av_i	vendor setup production cost of product i
b_i	transportation cost of product i
h_i	buyer holding cost per unit of product i
hv_i	vendor holding cost per unit of product i
c_i	product unit cost of product i
IP_i	average vendor inventory of product i
Bg	buyer budget
$TBUC_i$	total buyer unit cost of product i
$TVUC_i$	total vendor unit cost of product i
TUC	total vendor-buyer unit cost

Buyer Inventory level of product i can be represented in Fig. 1. When inventory level reach zero,instantaneous replenishment of m units is done.

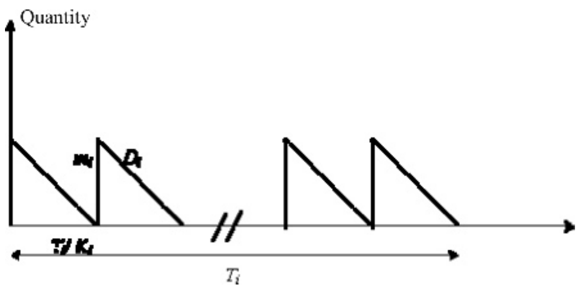


Figure 1. Buyer inventory level

The buyer‘s total inventory cost consists of setup cost, transportation cost, and holding cost. The total inventory cost per unit of product i can be shown as:

$$TBUC_i = \frac{A_i D_i}{Q_i} + \frac{b_i K_i D_i}{Q_i} + \frac{h_i m_i}{2} \tag{1}$$

Substitute Q_i with $m_i K_i$ to (1), one has:

$$TBUC_i = \frac{A_i D_i}{K_i m_i} + \frac{b_i D_i}{m_i} + \frac{h_i m_i}{2} \tag{2}$$

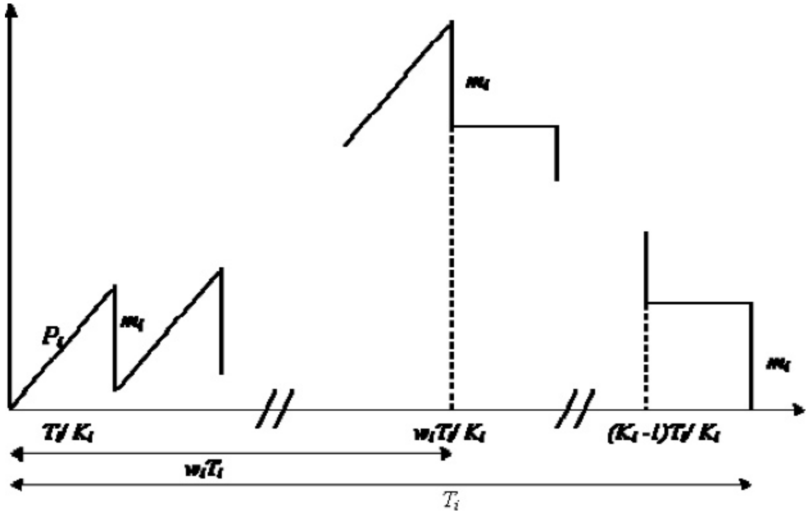


Figure 2. Vendor inventory model

The vendor inventory level can be shown in Fig 2. The vendor produces product *i* until w_iT_i/K_i period and deliver *m* unit of product *i* every T_i/K_i period in K_i delivery times.The vendor's average inventory can be represented as:

$$IP_i = \frac{m_i(K_i - w_i + 1)}{2} \tag{3}$$

The vendor total cost consists of setup cost and inventory holding cost. The vendor total cost is:

$$TVUC_i = \frac{Av_i D_i}{Q_i} + \frac{hv_i m_i (K_i - w_i + 1)}{2} \tag{4}$$

Substitute Q_i with $m_i K_i$ to (4), one has:

$$TVUC_i = Av_i \frac{D_i}{m_i K_i} + \frac{hv_i m_i (K_i - w_i + 1)}{2} \tag{5}$$

Total inventory cost is equal to total buyer cost and total vendor cost, one has:

$$TUC = \sum_{i=1}^{l} \frac{A_i D_i}{K_i m_i} + \frac{b_i D_i}{m_i} + \frac{h_i m_i}{2} + Av_i \frac{D_i}{m_i K_i} + \frac{hv_i m_i (K_i - w_i + 1)}{2} \tag{6}$$

When the buyer set up the quantity order, she has to spend cQ dollar to pay for the order. The buyer has a budget of Bg, one has:

$$\sum_{i=1}^{v} c_i m_i K_i \leq Bg \tag{7}$$

The problem can be formulated by minimizing the total unit cost of vendor and buyer with budget constraint. The vendor-buyer model can be represented as:

$$\text{Min } z = \sum_{i=1}^{l} \frac{D_i}{m_i K_i}\left(A_i + Av_i\right) + \frac{b_i D_i}{m_i} + \frac{m_i K_i}{2}\left(\frac{h_i + hv_i}{K_i} + hv_i\left(1 - \frac{D_i}{P_i}\right)\right)$$

$$\text{s.t. } \sum_{i=1}^{v} c_i m_i K_i \leq Bg$$

3 Lagrangean Multiplier Method and Computation Procedure

Lagrangean problem is developed by multiplying Lagrange multiplier λ to (2) and adding the result to the objective function. The Lagrangean expression, $LM(\lambda)$ is:

$$\text{Min } \sum_{i=1}^{l} \frac{D_i}{m_i K_i}\left(A_i + Av_i\right) + \frac{b_i D_i}{m_i} + \frac{m_i K_i}{2}\left(\frac{h_i + hv_i}{K_i} + hv_i\left(1 - \frac{D_i}{P_i}\right)\right) + \lambda\left(Bg - \sum_{i=1}^{v} c_i m_i K_i\right) \tag{8}$$

The problem can be defined as Mixed Integer Non Linear Programming (MINLP) with constraint because the objective function is non linear and K_i are integer. A solution procedure is introduced using the Lagrangean method as follows:

1. Solve the vendor-buyer model by relaxing budget constraint
2. If the optimal solution do not violate budget constraint and all K_i are integer then the optimal solution is achieved. If the optimal solutions violate budget constraint then go to 3.
3. Solve (8) and derive λ^*.
4. Use λ^* from step 3 to solve m_i^* and K_i^*. If K_i^* is not an integer then round it down to $K_i^{(lb)}$ and round it up to $K_i^{(ub)}$.
5. Set K_i^* as constants, solve (8) and check the optimality for the Lagrangean with $\lambda = \lambda^*$.
6. If $TUC(K_i^{(lb)}) < TUC(K_i^{(ub)})$ then set $K_i^* = K_i^{(lb)}$.
   If $TUC(K_i^{(lb)}) > TUC(K_i^{(ub)})$ then set $K_i^* = K_i^{(ub)}$.
7. For all i values of K_i^*, do branch and bound analysis.

After the Lagrangean method , proceed using branch and bound procedure for better result. The branch and bound steps to solve the problem is:

1. Start by letting $K_i = K_i^*$ for all i.
2. Branch all K_i to K_i -1 and K_i +1 then solve (8) to get λ^*, m_i^* and the optimal objective function for first stage is derived.

3. If all the objective function for each combination in stage $j+1$ has greater value than stage j, then bound the branch and set the objective function in stage j as the optimal objective function for that branch. Otherwise, continue to branch the solution to the next stage.
4. Return to step 3 until all combination have been calculated or bounded.
5. Choose the best objective function for each branch to derive the optimal objective value.

4 Numerical Example

In this section, an example is introduced to illustrate the model and the solution procedures in section 2 and 3. A buyer wants to buy four products from the vendor. Buyer order Q units and ask the vendor to deliver in m units every batch. Data of the example are shown in Table 1.The buyer has a budget of $ 20000.

Table 1. Data for the example

Product	A_i	D_i	h_i	c_i	Av_i	hv_i	P_i	b_i
1	47	1361	5	17	68	3	2444	14
2	49	1039	9	13	71	5	2355	15
3	58	1434	8	16	75	6	2392	13
4	49	1113	6	14	55	5	2440	18

The problem is solved by relaxing the budget constraint. The optimal total inventory cost is equal to $ 5193.35. Substitute the decision variables, the buyer need $ 25394.23 budget. The optimal solution need 27% higher budget than the budget costraint, so Lagrangean method should be used.

Substitute the optimal variables to (8) one has λ = -0.0602. The new optimal value of m_i and K_i are obtained by substituting λ to (8). The optimal values of m_i and K_i are shown in Table 2.

Table 2. The optimal variables solution of Lagrange method

Product	m_i	K_i
1	54.464	5.58
2	70.940	4.29
3	47.244	5.57
4	52.968	5.41

Set K_1 equal to 5 and the other K_i values as in Table 2 and solve (8) for λ^*, one has:

$$TUC = \frac{46937}{m_1} + 9.615m_1 + \frac{51098.4}{m_2} + 7.407m_2 + \frac{43294.6}{m_3} + 15.009m_3 + \frac{41059.1}{m_4} + 10.907m_4 - 871.820 \quad (9)$$

Global optimum solution for a constant value of K_i can be achieved since (9) is convex.

Applying Lagrangean method in Section 3, one can derive the optimal integer number of shipment for all products as K_1 = 6, K_2 = 5, K_3 = 6, and K_4 = 6. When the optimal integer numbers of shipments for all products have been found, then branch and bound procedure is applied. The branch and bound procedure for the problem are shown in Figure 3.

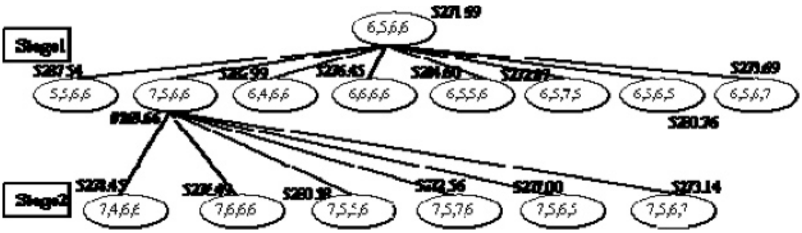


Figure 3. Branch and bound for the example

The optimal total inventory cost is equal to \$ 5269.66 and it is found when m_1 = 52.585, m_2 = 70.624, m_3 = 48.426, m_4 =53.607, K_1 = 7, K_2 =5, K_3 = 6, and K_4 = 6. It is obvious that the optimal total inventory cost for feasible solution is higher than the problem without constraints. The difference is (5269.66 – 5193.35) * 100% / 5193.35 = 1.47%.

The optimal total inventory cost are obtained when the number of shipment is greater than one. It means that the vendor-buyer inventory system can be more efficient when JIT philosophy is applied. The budget constraint model is more practical, eventhough it is more complicated. With the buyer's limited budget, the total inventory cost is higher than without limited budget. In this example, the budget needed by the buyer to derive the optimal inventory cost is 27% higher than the one without budget constraint, and the optimal inventory cost with the buyer budget constraint is 1.4% higher than the problem without constraint.

5 Conclusion

A mixed integer non linear programming with constraint has been developed to solve multi products EPQ model with JIT philosophy and discrete delivery. A Lagrange method and branch and bound procedure are used to derive the optimal condition. The result shows that the global optimum solution can be derived. Introducing of budget model make this model more practical. The implementation of branch and bound procedure in this model results in long computation time which will increase rapidly when problem size increase.

Since the computation time can be very high for large problem, other methods can be explored for future research.

6 References

[1] PASANDIEH, S.H.R.; NIAKI S.T.A. A genetic algorithm approach to optimize a multi products EPQ model with discrete delivery orders and constrained space. Applied Mathematics and Computation, vol. 195, pp. 506-514, 2008.

[2] OUYANG L.Y.; WU, K.S.; HO C.H. Integrated vendor-buyer cooperative models with stochastic demand in controllable lead time. Int. J. Production Economics, vol. 92, pp. 255-266, 2004.

[3] SAJADIEH M.S.; JOKAR M.R.A; MODARRES M.X Developing a coordinated vendor buyer model in two stage supply chains with stochastic lead times. Computers & Operations Research, doi:10.1016/j.cor.2008.10.001, 2008.

[4] BEN-DAYA M.; HARIGA M. Integrated single vendor single buyer model with stochastic demand and variable lead time. Int. J. Production Economics, vol. 92, pp. 75-80, 2004.

[5] HSIAO Y.C. A note on integrated single vendor single buyer model with stochastic demand and variable lead time. Int. J. Production Economics, vol. 114, pp. 294-297, 2008.

[6] YANG P.C.; WEE H.M. An integrated multi-lot-size production inventory model for deteriorating item. Computers & Operations Research, vol. 30, pp. 671 – 682, 2000.

[7] WANG S.; SARKER B.R. Optimal models for multi-stage supply chain system controlled by kanban under just-in-time philosophy. European Journal of Operational Research, vol. 172, pp. 179-200, 2006.

[8] RAU H.; OUYANG B.C. An optimal batch size for integrated production-inventory policy in supply chain. European Journal of Operational Research, vol. 185, pp. 619-634, 2008.

[9] HOQUE M.A.; GOYAL S.K. An optimal policy for a single-vendor single buyer integrated production inventory system with capacity constraint of the transport equipment. Int. J. Production Economics vol. 65, pp. 305-315, 2000.

[10] HAKSEVER C.; MOUSSOURAKIS. A model for optimizing multi-product inventory systems with multiple constraints. Int. J. Production Economics, vol. 97, pp. 18-30, 2005.

A Study on Impact Factors of Retailing Implementing CPFR – A Fuzzy AHP Analysis

Hsin-Pin Fu[a] , Sheng-Wei Lin[b,c,1], and Chi-Ren Chen[d]

[a] Department of Marketing and Distribution Management, National Kaohsiung First University of Science and Technology, Taiwan.
[b] Department of Tourism management, Shih Chien University Kaohsiung Campus, Taiwan.
[c] Graduate institute of Management, National Kaohsiung First University of Science and Technology, Taiwan.
[d] Logistic manager, POYA Co.,Ltd.

Abstract: Numerous researches have focused on exploring the factors of success or failure in the implementation of collaborative, planning, forecasting, and replenishment (CPFR) in retailing since CPFR was proposed in 1998. Despite of considerable research primarily used statistical analysis methods, such as the analysis of variance (ANOVA) and multivariate analysis of variance (MANOVA) approach, can discover significant impact factors. However, it still cannot understand the importance of each impact factor. This result can not provide sufficient information as a useful reference for retailing to allocate optimal resource while implementing CPFR. Therefore, the factors of adopting CPFR or B2B information system were collected and further construct a three level of hierarchical structure in this paper. Sequentially, the pair-wise questionnaire was designed and distributed to experts who are familiar to CPFR implementation of retailing. Finally, fuzzy analytic hierarchy process (Fuzzy AHP) is applied to find out the weights (importance) of each impact factors of CPFR implementation. The paths of critical factors were analyzed and suggestions for retailing in practice were provided. This result indicate that to use fuzzy AHP method can provide more precise information as a valuable reference for retailing to allocate optimal resource with more efficiently while implementing CPFR.

Keywords: CPFR, Key impact factors, Fuzzy AHP

1 Introduction

To assist retailing to increase the accuracy of forecasts and to establish a capability to quickly response to market, VICS (Voluntary Inter-industry Commerce Standards association, 1998) proposed the Collaborative Planning, Forecasting and

[1] Corresponding Author: Doctoral candidate, Graduate Institute of Management, National Kaohsiung First University of Science and Technology, Taiwan, R.O.C; Tel: +886-7-6011000 ext.3929; Fax: +886-7-6011093; E-mail: u9528914@ccms.nkfust.edu.tw

Replenishment (CPFR) model which is a scheme integrating enterprises' internal and external information systems to assist retailing in establishing a collaborative forecast and replenishment scheme between retailers and suppliers [1].

Although CPFR can help companies increase sales and operating performance [2], there have been many failed implementations. Despite of previous research focused on the factors of success or failure in the implementation of CPFR in retailing, these researches primarily used statistical methods, such as the analysis of variance (ANOVA) and multivariate analysis of variance (MANOVA), that can offer retailing to determine significant impact factors of CPFR implementation. These results, however, can not provide sufficient information and further realize the weights (importance) of each factor for retailing to allocate optimal resource with efficiently when they would like to implement CPFR. Therefore, this research collected the factors of adopting a CPFR or B2B information system to construct a hierarchical structure. Second, a pair-wise questionnaire was designed and distributed to experts who are familiar with CPFR implementation in the field of retailing. After questionnaires were returned, fuzzy analytic hierarchy process (Fuzzy AHP) was applied to calculate the weights of each factor to understand the importance and priority of impact factors which influenced CPFR implementation. The results of this study would provide more accurate information as a valuable reference for retailing to allocate optimal resource while implementing CPFR.

2 Literature Review

Since VICS [1] first proposed the CPFR model in 1998, many enterprises have implemented it, and researchers have paid more attention to CPFR-related issues. Aviv [3] considered CPFR can work out majority of the problems that are encountered in adaptation of vendor-managed inventory (VMI) since all members need to jointly develop demand forecasts, production and purchasing plans, and inventory replenishments in supply chain. Williams [4] described how Procter and Gamble (P&G) took advantage of CPFR in a supply chain to create value for the corporation, trade partners, and consumers. Foote and Malini [5] found that the incorporation of Data Warehouse with Wal-Mart in CPFR enabled more accurate forecasting of operational processes.

Barratt and Oliveira [6] reviewed the literature on the subject and a broad range of case studies; they identified the potential difficulties that can arise in the implementation of CPFR and proposed five possible solutions to the identified difficulties. Albright [7] demonstrated that CPFR implementation can reduce inventory levels, increase sales, and improve trade partnerships. Sagar [8] reviewed Whirlpool Corporation implementation of CPFR in 2000, and found that it significantly enhanced sales forecasting between the company and its suppliers. Steermann [9] investigated the case of Sears, a major American retailer, and its supplier, Michelin, who collaborated in applying the CPFR model in 2001—producing a reduction in their inventory level of 25%. Esper and Williams [10] worked on a case study and discovered that, by implementing CPFR, Collaborative Transportation Management (CTM) could bring about better outcomes and profits.

Smaros [11] adopted an exploratory case study to examine CPFR operation in the European grocery.

Previous literatures demonstrate that related research of CPFR has emphasized in cost down, increasing sales and forecasting ability in practice. It is noted that there is a lack of research on the importance of factor impacted CPFR implementation.

3 Methodology

There are many factors that affect the adoption of CPFR by firms. If enterprises can appreciate the degree of importance of any given factor, this will allocate the optimal enterprises resources to adopt CPFR model with lower costs and greater efficiency. The purpose of this paper is to find the degree of importance of each factor and the key impact factors. Therefore, this is a multi-criteria analysis (MCA) problem. The analytic hierarchy process (AHP) is a multi-criteria technique, or more generally, a philosophy proposed by Saaty [12], which is widely applied to solve MCA problem. AHP is suitable for solving the following 12 decision-making problems: (i) setting priorities; (ii) generating a set of alternatives; (iii) choosing a best alternative/policy; (iv) determining requirements; (v) allocating resources; (vi) predicting outcomes/risk assessment; (vii) measuring performance; (viii) designing systems; (ix) ensuring the stability of a system; (x) optimization; (xi) planning; and (xii) resolving conflict [12].

According to the objective of this study, the priority of impact criteria is the issues of setting priorities and resources allocation, which are expected to realize more precisely for decision-maker to allocate limited enterprise resources while lead in CPFR model. However, AHP does not involve the fuzziness of human thoughts or the uncertainties of the environment so that could not objectively reveal the actual situation. Thus, Laarhoved and Pedrycz [13] included fuzziness in AHP, and proposed the Fuzzy AHP to solve the disadvantage. Fuzzy AHP involved the uncertainty and multi-criteria aspects of the problems and the experts' opinions. When there were many decision-making criteria and substitute plans, Fuzzy AHP could avoid the subjectivity of pair wise comparison and inaccurate results. That is this study utilized Fuzzy AHP as the research methodology in this paper. Based on literature review [14,15], this study generalized the calculation procedures of Fuzzy AHP as follows:

(1) Construction of hierarchical table

According to the topic, a factor hierarchy is constructed. The second hierarchical factor was the precise description of the previous one.

(2) Questionnaire design

When designing the questionnaire, this study conducts a pair wise comparison on the factors in the same hierarchy based on a factor hierarchical table, and probes into the importance of each factor, to develop the questionnaire of AHP format.

(3) Construction of Fuzzy Numbers

Past studies tended to have discrete answers in questionnaires; this study was based on interval answers and constructed Fuzzy Numbers of all respondents' answer intervals. The most common Fuzzy Numbers were Trapezoidal Fuzzy Numbers and Triangular Fuzzy Numbers.

(4) Construction of the Fuzzy Positive Reciprocal Matrix

Based on the Fuzzy Numbers from the questionnaire results, a Fuzzy Positive Reciprocal Matrix is constructed. When there were n sub-criteria studied in one criterion, this study could establish $n \times n$ Fuzzy Positive Reciprocal Matrix A.

(5) Consistency test

Before calculating the weights, this study conducted Consistency Index (CI) and Consistency Ratio (CR) tests on the fuzzy weights of a Positive Reciprocal Matrix as following functions:

$$CI = (\lambda_{\max} - n)/(n-1), \quad CR = (CI/RI_n) \times 100\%$$

The consistency test was accomplished after confirming Fuzzy Numbers, through geometric means. Saaty [12] suggested that $CI \leq 0.1$ is an acceptable scope, and $\lambda_{\max}$ is the Maximized eigenvector of a pair wise comparison matrix, n is an attribute of the matrix, and RI_n is a random index [16] and the values are shown in Table 1.

(6) Construction of start matrix (X matrix)

After consistency tests, this study constructed a start matrix by the Fuzzy Positive Reciprocal Matrix previously established.

(7) Defuzzification (α-cut)

Defuzzification was based on α-cut proposed by Buckley and Csutora [15] to determine fuzzy weights.

(8) Normalization

This study normalized fuzzy weight intervals, and after calculating the weights of all factors, connected the local weights to acquire the global weights of the factors in each hierarchy.

Table 1. Randomized index of RI_n

n	3	4	5	6	7	8	9	10	11	12	13	14	15	16
RI_n	0.525	0.882	1.115	1.252	1.341	1.404	1.452	1.484	1.513	1.535	1.555	1.570	1.583	1.595

4 Data Collection

The purpose of this study is to analyze the difference and importance of factors which affect enterprises to implement CPFR model. Data collection, therefore, was based on the steps of the Fuzzy AHP. After collecting literature and the experts' opinions, this study allocated the factors of enterprises implementing CPFR into a three level of hierarchy structure, that is goal level, criteria level and sub-criteria level, which shown in Table 2.

Subsequently, we transformed the hierarchy structure into a pairwise comparison questionnaire of fuzzy analytic hierarchy process format. Furthermore, this study selected 10 experts who had much experienced in implementing CPFR model. The respondents were requested to fill in the area of 1-9 point scale with desired answers for a total of thirty-nine questions. As the processes mentioned above, there were 10 valid questionnaires were returned. Logical examinations were performed to avoid related errors from the returned questionnaires while we conducted the data analysis.

Finally, this study conduct data analysis based on Matlab, which to test the consistency and calculate the weights of each fuzzy internal. The results of the consistency ratio, all of them not exceed 0.1, were acceptable and then it defuzzified the fuzzy weights into explicit weight by α-cut (α=0.5, λ=0.5), which proposed by Buckley and Csutora [15], to obtain the relative weight for each factor as shown in Table 3.

5 Data Analysis

5.1 Goal Level

The results of goal level in Table 3 show that the priority of adopting CPFR are organization indicator (58.96%), technology indicator (24.29%), and environment indicator (16.76%), respectively. This finding indicated organization is the most critical indicator for vendor to consider adopting CPFR.

Table 2. The hierarchy structure of adopting CPFR for retailer

Goal	Criteria	Sub-criteria	Source
A: Technology indicator	A1: System efficacy	A11: System function integrity	[17,18]
		A12: Electronic data interchange	[17]
		A13: Information technology service	[17]
	A2: Information technique	A21: Reliability	[4,19]
		A22: Maturity	[19]
		A23: Stability	[4]
	A3: Implementation risk	A31: System compatibility	[20,21]
		A32: System security	[22]
		A33: System complexity	[17]
B: Organization indicator	B1: Collaboration effect	B11: Mutual objective	[4]
		B12: Information sharing	[23]
		B13: Perceived benefit	[20]
	B2: Organization leadership	B21: Top management support	[24]
		B22: Change management	[25]
		B23: Education training	[26,27]
	B3: Organization management	B31: Amalgamation capability	[28]
		B32: Innovation capability	[29]
		B33: Communication capability	[25]
C: Environment indicator	C1: Micro-environment condition	C11: Industrial competitiveness	[30]
		C12: Information readiness	[31]
		C13: Industrial know-how and experience	[17]
	C2: Macro-environment condition	C21: Political policy	[32]
		C22: Competition pressure	[31]
		C23: National foundational construction	[33,34]
	C3: Partnership	C31: Organization scale	[31]

condition	C32: Resource support	[35]
	C33: Trust and communication	[17,33]

5.2 Criteria Level

In the aspect of organization indicator, organization management (53.17%) is the most critical element than organization leadership (29.95%) and collaboration effect (16.88%). This implies that the compatibility of technology and culture, organization innovative capability, and cross-department operation in coordination are critical criteria which impact on retailer implementation CPFR model. Second, respondents consider the importance of implementation CPFR model from the technology indicator view are implementation risk (58.25%), system efficacy (28.78%), and information technique (12.98%). The results indicate that implementation risk, such as system security and system complexity, are likely to influence the intention to implement CPFR model.

Finally, from the aspect of environment indicator, the priorities of critical criteria are partnership condition (54.81%), micro-environment condition (27.2%), and macro-environment (17.99%). This results show that partnership condition is crucial criteria such as attributes of trust and communication. In whole, the top five important criteria are organization management (31.34%), organizational leadership (17.65%), implementation risk (14.14%), collaboration effect (9.95%), and partnership condition (9.18%) in the criteria level.

Table 3. Synthesis the final results (α=0.5，λ=0.5)

Goal (E)	Criteria (F)	Streamed (G)	Sub-criteria (H)	Globe weights (I)
A: 0.2429【2】	A1:0.2878【2】	0.0699【6】	A11 0.1822【3】	0.0127【22】
			A12 0.5180【1】	0.0362【10】
			A13 0.2998【2】	0.0209【16】
	A2:0.1298【3】	0.0315【8】	A21 0.1603【3】	0.0050【27】
			A22 0.2762【2】	0.0087【24】
			A23 0.5635【1】	0.0177【17】
	A3:0.5825【1】	0.1414【3】	A31 0.1688【3】	0.0238【14】
			A32 0.2869【2】	0.0405【9】
			A33 0.5443【1】	0.0770【4】
B:0.5896【1】	B1:0.1688【3】	0.0995【4】	B11 0.5620【1】	0.0559【5】
			B12 0.3065【2】	0.0305【11】
			B13 0.1315【3】	0.0130【21】
	B2:0.2995【2】	0.1765【2】	B21 0.2879【2】	0.0508【7】
			B22 0.5489【1】	0.0969【2】
			B23 0.1633【3】	0.0288【12】
	B3:0.5317【1】	0.3134【1】	B31 0.1659【3】	0.0520【6】
			B32 0.2723【2】	0.0853【3】
			B33 0.5619【1】	0.1761【1】
C:0.1676【3】	C1:0.2720【2】	0.0455【7】	C11 0.3145【2】	0.0143【20】
			C12 0.1762【3】	0.0080【25】

		C13	0.5093【1】	0.0232【15】
C2:0.1799【3】	0.0301【9】	C21	0.1730【3】	0.0052【26】
		C22	0.2897【2】	0.0087【23】
		C23	0.5372【1】	0.0161【18】
C3:0.5481【1】	0.0918【5】	C31	0.1643【3】	0.0150【19】
		C32	0.3041【2】	0.0279【13】
		C33	0.5313【1】	0.0488【8】

Note:
1. E: Initial weights of goal level; F: Initial weights of criteria level; G: Streamed weights of criteria level and goal level (G=E×F); H: Initial weights of sub-criteria level; I: Globe weights (I=H×G)
2. Number in 【】 represents ranking

5.3 Sub-criteria Level

First, respondents considered the priority of performance criteria were electronic data interchange information (51.80%), information technology service (29.98%), and system function integrity (18.22%) in the system efficacy criteria. The results show that retailer have to pay much attention on a standard form of computer to computer communication, such as electronic data interchange (EDI), since it is critical to implement CPFR model success. Form the information technique criteria, the ranking of sub-criteria were stability (56.35%), maturity (27.62%), and reliability (16.03%). That is, respondents focus on the stability of CPFR system since it is critical factor for retailer to operate and communicate in order. In the implementation risk aspect, the crucial sub-criteria is system complexity (54.43%) rather than system security (28.69%) and system compatibility (16.88%). The results show that respondents consider an acceptable range of system complexity with operational training that would influence retailer intention to implement CPFR model.

Viewed in collaboration effect criteria, the ranking of sub-criteria is mutual objective (56.20%), information sharing (30.65%), and perceived benefit (13.15%), respectively. This implies that the mutual objective in collaboration effect, such as reduction out-of-stock, order accuracy and timeless, reduction of purchasing period, and quick response and order tracing, would facilitate CPFR model success. From the organization leadership criteria, the most important sub-criteria is change management (54.89%), others are top management support (28.79%) and educational training (16.33%). The results indicate that change management is a capability conception that implies organization should adjust business operation condition as soon as possible to keep the competitive advantage. Among the criteria level of organization management, communication capability (56.19%) is most crucial factor than innovative capability (27.23%) and amalgamation capability (16.59%). This reveals that cross-department operation in coordination, such as information department, marketing department, financial department, and procurement department, would facilitate the CPFR system success.

In the micro-environment condition criteria, industrial know-how and experience (50.93%) is key sub-criteria, others are increase industrial competitiveness (31.45%) and information readiness (17.62%). The results indicate that, based on the industrial experienced implementation of CPFR model, retailer would decrease the related lead-in fumble cost and free of implementation fail risk. Form the macro-environment condition criteria, prioritized importance of sub-criteria is national foundation construction (53.72%), competitive pressure (28.97%), and political policy (17.3%). The results show that the foundations of information technology infrastructure such as Internet and network cable are critical factor for CPFR model success. In the partnership condition criteria, respondents consider the importance ranking of sub-criteria is trust and communication (53.15%), resource support (30.41%), organization scale (16.43%). This implies that trust and communication would play a crucial role in the partnership condition.

6 Discussion

We present top ten critical factors which the value is greater than 0.035 in weights as shown in Table 4. The sum of weights accumulates closely 72% which account for these factors play a powerful impact. Further, among the hierarchy structure, there are six sub-criteria which are highly relevant to organization factor. Among them, two sub-criteria are closely connection with implementation risk, as well as two sub-criteria are mainly associated with organization leadership. In addition, the criteria of organization management take possession of key reality influence such as cross-department communication capability, organization innovative capability, and amalgamation capability. This is an implication that the supplier should pay attention to these critical factors within limited resource, as well as lead in CPFR model through the need of customer- oriented.

As mentioned above, we propose the path patterns of implementation CPFR based on top seven sub-criteria whose value are greater than 0.05 in weights as shown in Figure 1. There are main two kinds of paths to implement CPFR for retailer. For technology indicator aspect, the key impact factors in criteria level and sub-criteria level are implementation risk and system complexity, respectively. In the aspect of organization indicator, the criteria level should highlight on organization management, organization leadership, and collaboration effect. The crucial factor of sub-criteria level is mutual objective in collaboration effect criteria level. The crucial factors of sub-criteria level are top management support and change management in organization leadership criteria level. The crucial factors of sub-criteria level are amalgamation capability, innovation capability, and communication capability in organization management criteria level. It is worth noted that decision-maker should pay attention on critical paths which would facilitate an efficient use of limited enterprise resources and capital investment while leading in CPFR model.

Table 4. The sub-criteria of ranking for retailer implementation CPFR

Sub-criteria	Globe weights	Ranking
Communication capability	0.176160	1

Sub-criteria	Globe weights	Ranking
Change management	0.096928	2
Innovation capability	0.085363	3
System complexity	0.077013	4
Mutual objective	0.055933	5
Amalgamation capability	0.052008	6
Top management support	0.050839	7
Trust and communication	0.048806	8
System security	0.040593	9
Electronic data interchange	0.036212	10
Sum of weights	0.719855	

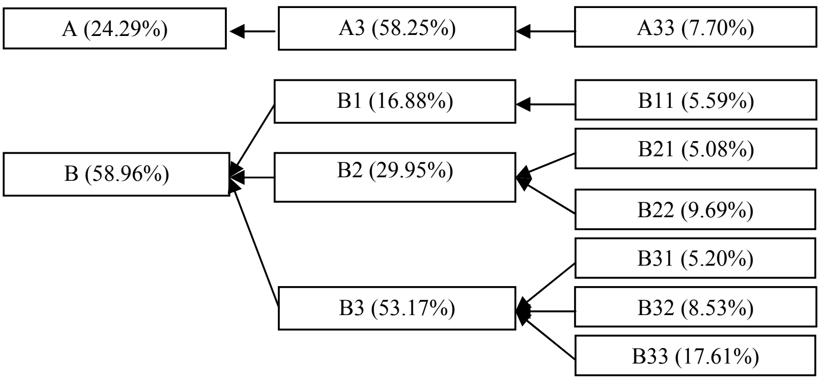


Figure 1. Path analysis of critical factor

7 Conclusion

This study is to explore the factors, which retailer consider implementation CPFR, as well as the importance of each factor through data analysis. The results of this study show that retailer consider the key impacted factors for adopting CPFR model are such as cross-department support, change management, innovation capability, system complexity, mutual objective, amalgamation capability, top management support, partner trust and communication, and integration of exchange data. The strategy for retailer, therefore, is to focus on these criteria, in particular leading importance indictors, within limited resource. For instance, in organization perspective, retailer should consider an assignment project which conducts relative work such as the integration of cross-department capability, organizational innovation capability, and technical amalgamation capability. Moreover, in technology and environment perspective, simplify of system function, the humanity of operation interface, and educational training, trust, communication mechanism and cooperation partner selection should be considered within CPFR success.

8 References

[1] VICS, Voluntary Interindustry Commerce Standards Association. Available at: <http://www.vics.org/>. Accessed on Mar. 30th 2008.

[2] Sherman, RJ. Collaborative planning, forecasting and replenishment (CPFR): Realizing the promise of efficient consumer response through collaborative technology. Journal of Marketing theory and Practice 1998; 6:6-9.

[3] Aviv Y. Gaining benefits from joint forecasting and replenishment process: The case of auto-correlated demand. Manufacturing and Service Operations Management 2002;4:55-74.

[4] Williams SH. Collaborative planning, forecasting, and replenishment. Hospital Materiel Management Quarterly 1999;21:44-58.

[5] Foote PS, Malini K. Forecasting using data warehousing model: Wal-Mart's experience. The Journal of Business Forecasting Methods & Systems 2001;20: 13-17.

[6] Barratt M, Oliveira A. Exploring the experiences of collaborative planning initiatives. International Journal of Physical Distribution & Logistics Management 2001; 31:266-289.

[7] Albright B. CPFR's secret benefit. Frontline Solutions 2002; 3:30-35.

[8] Sagar N. CPFR at Whirlpool Corporation: two heads and an exception engine. The Journal of Business Forecasting Methods & Systems 2003;22:3-8.

[9] Steermann, H. A practical look at CPFR: The Sears-Michelin experience. Supply Chain Management Review 2003;7:46-53.

[10] Esper TL, Williams LR. The value of collaborative transportation management (CTM): Its relationship to CPFR and information technology. Transportation Journal 2003;42:55-65.

[11] Smaros J. Forecasting collaboration in the European grocery sector: Observations from a case study. Journal of Operations Management 2007;253:702-716.

[12] Saaty TL. The analytic hierarchy process. New York, McGraw- Hill,1980.

[13] Laarhoven PJM, Pedrycz W. A fuzzy extension of Saaty's priority theory. Fuzzy Sets and Systems 1983;11:199-227.

[14] Buckley JJ. Fuzzy hierarchical analysis. Fuzzy Sets and Systems 1985;17:233-247.

[15] Csutora R, Buckley JJ. Fuzzy hierarchical analysis: The Lambda-Max method. Fuzzy Set and Systems 2001;120:181-195.

[16] Aguarón J, Moreno-Jiménez JM. The geometric consistency index: Approximated thresholds. European Journal of Operational Research 2003;147:137-145.

[17] Fliedner G. CPFR: An emerging supply chain tool. Industrial Management Data systems 2003;103:14-21.

[18] Rosemary S, Craig S. A framework for the selection of electronic marketplace: A content analysis approach. Internet Research 2002;12:221.

[19] Angeles R, Corritore, CL, Basu, SC, Nath, R. Success factors for domestic and international electronic data interchange implementation. International Journal of Information Management 2001;21:329-347.

[20] Beatty RC, Shim JP, Jones, MC. Factors influencing corporate web site adoption: A time-based assessment. Information and Management 2001;38:337-354.

[21] Kaefer F, Bendoly E. Measuring the impact of organizational constraints on the success of business-to-business e-commerce efforts: a transactional focus. Information and Management 2004;41:529-541.

[22] Attaran M. Nurturing the supply chain. Industrial Management 2004; 46:16-20.

[23] Kwon IWG, Suh T. 2004. Factor affecting the level of trust and commitment in supply chain relationship. Journal of Supply Chain Management 2004;40:4-14.

[24] Byrd, TA, Davidson NW. Examining possible antecedents of IT impact on the supply chain and its effect on firm performance. Information and Management 2003;41: 243-255.

[25] Hoque F. E-enterprise: Business models, architecture and components. Cambridge University Press, 2001.

[26] Bingi P, Sharma MK, Godla JK. Critical issues affecting an ERP implementation. Information Systems Management 1999;16:7-14.

[27] Sakaguchi T, Dibrell CC. Measurement of the intensity of global information technology usage: Quantitating the value of a firm's information technology. Industrial Management and Data Systems 1998;98:380-394.

[28] Sharifi, S. Organizational learning and resistance to change in Estonian Companies. Human Resource Development International 2002,5:313-331.

[29] Scupola, A., 2003, "The adoption of Internet commerce by SMEs in the South of Italy: An environmental, technological and organizational perspective. Journal of Global Information Technology Management 2003;6:52-71.

[30] Sohal AS, Moss S, Ng, L. Comparing IT success in manufacturing and service industries. International Journal of Operations and Production Management 2001;21:30-45.

[31] Zhu K, Kraemer K, Xu S. Electronic business adoption by European firms: A cross-country assessment of the facilitators and inhibitors. European Journal of Information Systems 2003;12:251-268.

[32] Tan M, Teo TSH. Factors influencing the adoption of Internet banking. Journal of the AIS 2000;1(5):1-42.

[33] Kmpstra RP, Ashayeri J, Gattorna JL. Realities of supply chain collaboration. The International Journal of Logistics Management 2006;17:312-330.

[34] Eid R, Trueman M, Ahmed AM. 2002. A cross-industry review of B2B critical success factors. Internet Research 2002, 12: 110-123.

[35] Deeter-Schmelz DR, Bizzari A, Graham R, Howdyshell C. Business-to-Business online purchasing: Suppliers' impact on buyers' adoption and usage intent. The Journal of Supply Chain Management 2001;37:4-10.

Autonomous Capacity Planning by Negotiation Against Demand Uncertainty

Shih-Min Wang [a,] and Kung-Jeng Wang [b,1]

[a] Department of Industrial Management, National Taiwan University of Science and Technology (Taiwan Tech), Taiwan

[b] Department of Industrial Management, National Taiwan University of Science and Technology (Taiwan Tech), Taiwan

Abstract. The investment on facilities for high-tech manufacturing requires a large amount of budget. One of the challenging issues for a firm is how to allocate finite budget to its factories and to fully utilize limited manufacturing resources. In this study, we develop a model to negotiate budget allocation among factories. During the negotiation, factories are modeled as intelligent agents and exchange a serie of messages with each others as bargaining over capital. A negotiation mechanism is designed in the study to mimic the attitude of a factory. Intrinsic and budget utility are used as performance indices to observe the behaviors of negotiation trajectories. The influence of the demand uncertainty are discussed under different settings of the negotiation-tactic combinations. Experimental outcomes have empirically justified that, the demand uncertainty significantly harms the social-welfare under the information asymmetric environment.

Keywords. Capacity planning, genetic algorithm, negotiation.

1 Introduction

As the competition of high-tech industries increases, finite budget allocation of a firm has become an important issue to improve business performances, especially in a multiple-factories environment. Resources utilization is largely damaged by asymmetric information (i.e., production and order information) in local factories which myopic attitude would force them to maximize the overall profit of the firm. In such situation, an automatic negotiation process for rapid response to demands and effective share resources among factories is critical to the success of a firm. A negotiation model is designed in the study to negotiate the budget allocation among factories. Motivated by the potential profits, an agent representing a factory exchanges a series of messages with other agents so that an appropriate local

[1] Corresponding Author : Department of Industrial Management, National Taiwan University of Science and Technology (Taiwan Tech); #43, Sec.4, Keelung Rd., Taipei, 106, Taiwan, R.O.C.; Tel: +886-2-2737 6769; Fax: +886-2-2737 6344; Email: kjwang@mail.ntust.edu.tw; http://tom.im.ntust.edu.tw/

resources portfolio under market demand finally can be set. After receiving a budget allocation plan (i.e., an "offer") from its counterpart agent, an agent evaluates the influence of the plan based on the local capacity model and responses an offer. A Negotiation Decision Function (NDF) mechanism is employed herein to mimic the negotiation attitudes of a factory. The negotiation is ended either a deal is obtained or the time limit is reached. Two performance indices, intrinsic and budget utility, are used to observe the performances of the negotiation for deterministic and stochastic demand cases. The influence of demand uncertainty is also examined under different settings of the negotiation-tactic combinations. This research focuses on the following issues:

-How to develop a mutually acceptable budget allocation plan for factories under an information asymmetry environment?

-What is the influence of the different negotiation tactics on budget allocation?

-What is the influence of demand uncertainty on the negotiation outcomes?

The rest of this article is organized as follows. Section 2 reviews related research into negotiation framework, capacity planning and allocation, and solution methodology. In Section 3, the problem of the inter-factory budget allocation is justified and a negotiation model is designed. In Section 4, two performance indices, intrinsic utilities and budget utilities, are used to examine the negotiation trajectories. Comprehensive experiments are conducted to examine the influences of different negotiation attitudes under both deterministic and stochastic demand cases. Finally, Section 5 is the conclusions and the future research directions.

2 Related Work

Single-shot negotiation models (such as contract net) first developed by David and Smith [5] are commonly employed, but no guarantee is given by this type of models to reach the global optimum owing to the myopic nature of the class [17]. Negotiation is processed by iterations among agents and focuses on the exchange of partial plans to attain a global goal in a computer science domain [6]. Information is shared extensively among agents and the intentions of opponents are uncover through such communication. Several scholars [11, 12] address the issue of negotiation time and establish an artificial diplomat system to address negotiation issues. While negotiation is promising in resolving conflicts of multi-agent systems in the field of distributed artificial intelligence, this study make a contribution to devote in the multiple-plant capacity planning domain.

Multiple-agents negotiation by negotiation-decision-function (NDF) first proposed by Faratin et al. [7], which are derived from bilateral negotiation functions [14], has showed a promising applications in intelligent, collaborative production control systems [13, 16]. Conventional negotiation models based on disclosure of information among agents (such as game theory) is limited to several real applications, but NDF-based negotiation is characterized by its autonomous (private) behavior, consideration of timing, and issues, and thus can be applied to numerous real world application domains, such as industry production planning and control. In addition, it provides a solid basis to build an incentive mechanism in which agents use certain negotiation parameters to achieve socially desirable outcomes. Wang and Chou [18] investigate the properties of the NDF mechanism in an agent-based system. Other applications of the NDF negotiation also can be founded [15]. This study will adopt this type of negotiation mechanism as negotiation tactic to reflect the attitude of each factory during the negotiation process as well as to reduce the confliction of budget allocation by a headquarters.

In terms of capacity planning level, several studies in literature directly deal with short-term capacity adjusting through autonomous coordination and negotiation among factories. Jiang [8] proposes an architecture and methodology of capacity trading to solve capacity shortage incurred in wafer foundries. A bidding scheme using auction among the foundry factories is developed to solve the capacity shortage of bottleneck tools. Chang [3] develops a simple internet-based auction scheme to sell foundry capacity. The scheme acts a capacity manager of foundry factory that can automatically negotiate with customers. In our study, we focus on the long-term finite budget allocation negotiation. After an autonomous negotiation process, each factory thus can develop capacity plan according to the local information under the compromised budget allocation plan. Several existing literatures dealing in multiple-factories are characterized by assuming a centralized decision maker, using a simplified negotiation, scheme and/or having not considered the influence of the demand uncertainty. Agrawal et al. [1] develop an optimal model for managing capacity in a multi-vendor environment. In the study, given demand forecasts, vendor's capacity and constraints, the solution with maximal retailer's expected gross profit can be obtained. Chiung et al. [4] develop a genetic algorithm to solve the integrated scheduling problem in supply chain composed of a production facilities network with the character of multiple-products manufacturing. Chan and Chung [2] present a hybrid genetic algorithm for a production problem in a multiple-factories supply chain. Jolayemi and Olorunniwo [9] address the issue of multiple-plants with extensible capacities under deterministic demand. The objective of their model is to maximize the total profits considering production mix in a finite planning horizon. Karabuk and Wu [10] design a capacity allocation game coordination mechanism to implement capacity allocation in semiconductor manufacturing environment. In their study, the headquarters allocates capacity of resources to each product line on the base of its privately owned demand information. However, having not well discussed demand uncertainty and the attitudes of the negotiation limit those models in practice and trapped in local optima.

In summary, effective use of limited resources for a headquarters is highly important to the competition of high-tech industries, but academic studies have solved these problems either by centralized model or by simplified negotiation. In terms of demand uncertainty, most of the related studies only deal with them using deterministic demand; but it cannot precisely reflect the situations of the industry in real practice. Furthermore, negotiation statics are significant factors that affect the result of negotiation on resource planning and allocation among factories have been neglected or simplified. Thus, a negotiation model is required to describe precisely the problem facing the industries in view of the above factors. To support decision-making locally of each agent in our study, a negotiation model with NDF mechanism is designed to negotiate the budget allocation. A local model with stochastic planning characteristics proposed by Wang et al. [19] is modified to support the decision of each agent during the negotiation.

3 Problem Formulation

This section proposes a budget negotiation model, NDF negotiation tactic mechanism, and individual capacity model against asymmetric information environment. In such environment, all the factories are profit-centered business units and receive their local demands while the finite budget is provided by the same headquarter. Motivated by the profits, each factory is represented by an agent and prone to fully utilize allocated budget for purchasing resources as to maximize its own profits. Given that such a myopic decision will harm global gain of the enterprise, a negotiation model is designed in Section 3.1 to reduce the confliction among factories as to improve the entire enterprise-level profit through coordinated budget allocation and collaborative capacity planning. To fulfill a successful negotiation, in Section 3.2, the attitude of an agent is represented accurately by NDF mechanism. Once the budget allocation plan is obtained, each factory thus can evaluate or develop its resulting resource investment portfolio and capacity plan by the local capacity planning model presented in Section 3.3.

3.1 A Negotiation Model for Resource Planning Among Factories

Since profit-centered agents who have local demand information intend to maximize local profits and the resources are limited by finite budget provided by the headquarters, budget allocation becomes the major negotiable issue among the factories. Motivated by potential benefits, every agent involves in the negotiation. In the model, we allow each agent/factory to propose budget usage plans as offers to its opponents. The negotiation attitude of a factory is represented by an NDF mechanism. Agents who receive an offer from other agents then calculate its own potential benefit using the local capacity planning model and determine whether to accept a deal. The agent then generates a counter offer to its negotiating opponents by using NDF mechanism. The potential profit of such an offer is compared to the profit of the previous received offer. If the newly received offer can produce a higher potential profit than the potential profit that the offer is preparing to send back to the offer provider, the received offer is accepted as a compromised plan. The negotiation procedure with NDF mechanism and local capacity planning model among the agents is presented as follows:

Step 1. In the beginning, the negotiation starter, named Agent *a*, provides an offer produced by applying NDF mechanism. The NDF mechanism reflects the attitude of this agent as a tactic of the negotiation.

Step 2. The offer receiver, named Agent *b*, first checks the negotiation time. If the negotiation time is expired, the agent sends a negotiation failure message to the offer provider and goes to Step 5. Otherwise, the agent uses the local capacity planning model to evaluate its potential profits that are gained from local demand and the budget allocation plan (i.e., offer) provided from its counter parties.

Step 3. Agent *b* compares the potential profit obtained from Step 2 with the offer value V^b, which is the potential profit of the budget allocation plan generated by the local NDF mechanism. If the received offer has a higher potential profit than the local generated one, Agent *b* sends a deal message back to Agent a and goes to Step 5. Otherwise, the Agent *b* sends its local offer value to Agent *a*.

Step 4. After receiving the counter-offer from Agent *b*, Agent *a* conduct the procedure Steps 2 and 3 of Agent *a*.

Step 5. Negotiation is over and output the results.

3.2 Negotiation Decision Function (NDF) Mechanism

In the study, eight tactics are applied during the negotiation for reflecting the attitude of an agent. Those tactics are Boulware tactic, Linear tactic, Conceder tactic, Dynamic deadline, Resource Estimation, Relative tit for tat, Random tit for tat ,and Average tit for tat. It also can be referred to Faratin et al. [7] for more details.

3.3 Capacity Planning Model of Individual Factories

An individual factory resource planning and capacity allocation model modified with a budget constraint is used in the study to maximize the profit of a local factory. In our experiments, two cases (deterministic and stochastic demand) are considered separately to investigate the influences of the demand uncertainty. In the stochastic demand case, a set of random variables of realized demand of product $\boldsymbol{p}$ in period $\boldsymbol{t}$**,** $d^{\eta}_{t,p}$ are drawn to represent the realized demand scenario η. In the deterministic demand case, the constant values of demand of each product in each period are drawn. This model maximized the total profits and considered several constraints — multiple resources capacity, production capability, and simultaneous resources configuration— of an individual factory. The details of the proposed mathematical model and data cases can be referred to Wang et al. [19].

4 Performance Evaluation of the Negotiation based Capacity Planning Model

4.1 Performance Indices of the Negotiation

Two performance indices, budget utility of the agent and agent intrinsic utility, are applied for observing the trajectory of negotiation. All the results of the proposed model are compared to the "perfect-information" situation.

(1) The Budget Utility of the Agent (U^a)

The budget range of agents *a* and b can be identified as [min^a *budget*, max^a *budget*] and [min^b*budget*, max^b*budget*], separately. If the budget ranges of both agents are overlapping, the deal would be obtained. The budget utility U^a is a normalized value of budget obtained by agent *a* from the negotiation. The budget utility of againt *a* uses tactic *i* and opponent replies tactic *j* can be presented by *Ua* ($game^a[i,j]$).

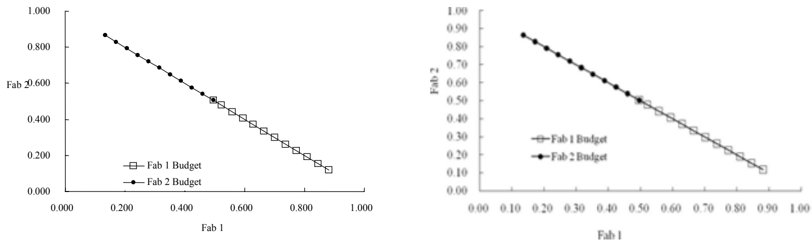
(2) Intrinsic Agent Utility (Gain) and the Perfect Information Game
The potential profits of againt *a* uses tactic *i* and opponent replies tactic *j* can be normalized to an intrinsic agent utility, Gain [i,j]. Note that, the intrinsic agent utility of an agent depends on the marginal profit as using budget; hence, the intrinsic agent utility of an agent is not a linear function of budget use. In addition, the optimal solution H of deterministic demand obtained from a perfect-information game is set as a benchmark. The difference between the H and the compromised solution represents the influence of the asymmetric information, and tactic combinations while the social welfare of a negotiation can be evaluated by Dist[i,j]=Gain(game[i,j])-Gain(H).

4.2 Parameters of the NDF Mechanism

Since finite budget is the negotiable issue among all the factories to maximize their local profit, each agent applies preferred tactics during the negotiation to reflect its attitudes. The parametrical information of the NDF mechanism is shown in Table 1.

4.3 Negotiation Trajectory

Assume that headquarters with 40 million of budget and there are two agents, representing factories 1 and 2 in an information asymmetric environment to share the finite budget. In the case, both agents use linear tactic as negotiation tactic to negotiate the 40 millions with the opponent. As shown in the Figure 1, the negotiation trajectory follows a linear tendency because the budget provided by the headquarters is a constant. In both deterministic and stochastic demand cases, as Figure 1 shown, each agent finally make a compromised budget allocation plan around 0.5 of budget utility. It implies that, in both cases, each factory thus can develop its capacity planning plan under 20 million budget.



(a) Negotiation with deterministic demand model

(b) Negotiation with stochastic demand model

Figure 1. Negotiation trajectory with respect to budget utility (U^a) with deterministic and stochastic demand

Table 1. Tactics and corresponding parameters settings.

Categories	Tactics	Abbreviation	Parameters	Note
Time Dependent Tactics	Boulware	TB	β=0.5	The offer value is depends on the parameter $\boldsymbol{\beta}$
	Linear	TL	β=1	
	Conceder	TC	β=10	
Resource Dependent Tactics	Dynamic Deadline	RDD	μ=480	The offer value is depends on the parameter $\boldsymbol{\mu}$
	Resource Estimation	RRE		
Behavior Dependent Tactics	Relative Tit for Tat	BRE	δ=1	The offer value is depends on the related relationship of the opponent's offers
	Random Tit for Tat	BRA	δ=1, $G \in \{-1,1\}$ million	The offer value is depends on the absolute relationship of the opponent's offers and a random factor ***G***
	Average Tit for Tat	BAV	γ=2	The offer value is depends on the average related relationship of the opponent's offers

Figure 2 shows the trajectory of intrinsic agent utility of two agents with deterministic (shown as solid line) and stochastic demand (shown as dash line) cases during the negotiation. In the case, both agent use time depend linear tactics. Note that the trajectory is nonlinear due to the variant marginal intrinsic profit of the budget for both agents. In the Figure, points **H** and **H'** represent the optimal solutions of deterministic and stochastic demand case under a perfect-information assumption respectively. Points **A** and **A'** represent the compromised solution of the two cases. As can be seen, the compromised budget allocation plans (**A** and **A')** are located at 70% and 50% of the **H**. The gap between point **A** and **H** can be used to explain the influence of the asymmetric information. Also, the 20% social welfare gap between point **A** and **A'** can be explained that the influence of the demand uncertainty. In the deterministic (stochastic) demand case, points **B** (**B')** and **D** (**D')** are the maximal intrinsic agent utility of Factory 2 and Factory 1 when fully utilized budget. The distance between point **A** **(A')** imaged in *y*-axis and point **B** (B') is the profit loss of Factory 2 due to the negotiation, also it can be explained the distance between point **A** (A') imaged in *x*-axis and point **D** (**D'**).

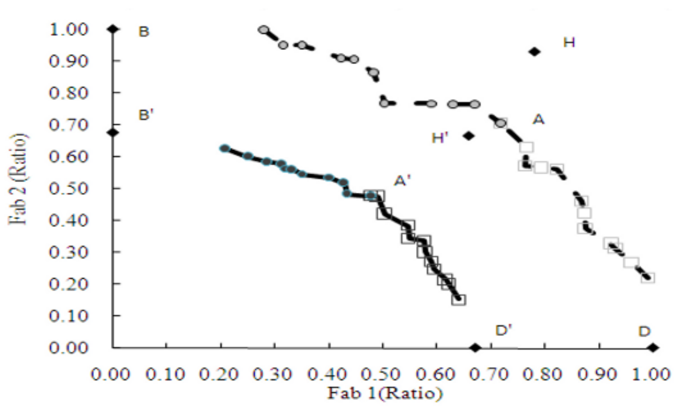


Figure 2. Negotiation trajectory with respect to intrinsic agent utility.

4.4 Experiments on Negotiation Tactics Performances

In the study, a set of experiments are conducted to justify the proposed budget negotiation model. Eight tactics are considered. The negotiation time limit is set at 50 units of time.

In Figure 3, the compromised budge allocation plans under different setting of tactic combinations of two-agent experiment (i.e., Fab 1 and Fab 2) can be generally divided into four groups. Groups A and C represent the results that Factory 1 and Factory 2 takes advantage after negotiation respectively. Group B reflects the fair results between the factories. The worst cases are the plans that located at group D, which represents no compromised solution can be obtained before the negotiation time limit. Moreover, it can be observed significantly that the social welfare of developed capacity plan under deterministic demand (solid compromised points) outperforms the plans developed under the stochastic demand (hollow compromised points). The details of the negotiation results of the two cases (deterministic and stochastic demand) can be obtained in Tables 2 and 3.

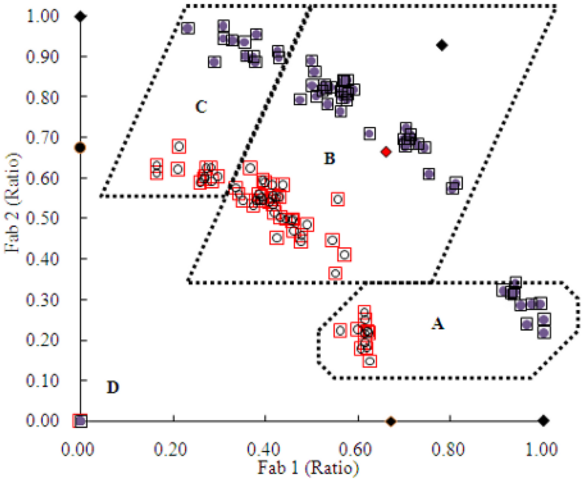


Figure 3. Performance grouping of tactic combinations with respect to intrinsic agent utility

Table 2. Negotiation Results Under Deterministic Demand Using Single Tactic.

Acceptance party (Negotiation#) [U^a, U^b] [*$gain^1$*, *$gain^2$*]		Fab 2 Agent							
		TB	TL	TC	RDD	RRE	BRE	BRA	BAV
Fab 1 Agent	TB	Fab2(34) [.51, .49] [.73, .68]	Fab1(29) [.60, .40] [.81, .59]	Fab 1(9) [.85, .15] [.96, .24]	Fab2(20) [.77, .23] [.94, .34]	Fab2(46) [.17, .83] [.31, .95]	Fab1(41) [.35, .65] [.50, .83]	Fab2(34) [.51, .49] [.71, .70]	Fab (40) [.35, .65] [.53, .83]
	TL	Fab1(29) [.38, .62] [.59, .82]	Fab1(23) [.50, .50] [.71, .71]	Fab 2(6) [.81, .19] [.95, .29]	Fab2(18) [.59, .41] [.80, .57]	Fab2(40) [.20, .80] [.33, .94]	Fab1(33) [.35, .65] [.53, .82]	Fab1(23) [.52, .49] [.74, .68]	Fab2(30) [.38, .62] [.57, .84]
	TC	Fab2 (6) [.19, .82] [.35, .94]	Fab 1(5) [.17, .83] [.31, .98]	Fab 2 (2) [.29, .71] [.43, .90]	Fab2(4) [.22, .78] [.37, .90]	Fab2(10) [.14, .85] [.28, .89]	Fab1(7) [.21, .79] [.38, .89]	Fab1(5) [.21, .79] [.36, .90]	Fab 1(7) [.22, .78] [.38, .96]
	RDD	Fab1(21) [.24, .76] [.43, .91]	Fab (17) [.39, .61] [.58, .79]	Fab 1(7) [.83, .17] [.99, .29]	Fab2(16) [.48, .52] [.70, .72]	Fab1(21) [.10, .90] [.23, .97]	Fab1(19) [.34, .66] [.52, .82]	Fab2(16) [.48, .52] [.70, .68]	Fab2(18) [.36, .64] [.54, .82]
	RRE	Fab1(45) [.80, .20] [.94, .32]	Fab1(45) [.89, .11] [.99,.22]	Fab1(7) [.83, .17] [.97, .29]	Fab1(21) [.85, .15] [.99,.25]	No(50) [.00,.00] [.00,.00]	No(50) [.00,.00] [.00,.00]	No(50)$^{\#}$ [.00,.00] [.00,.00]	No(50)$^{\#}$ [.00,.00] [.00, .00]
	BRE	Fab1(25) [.31, .69] [.50, .89]	Fab1(15) [.35, .65] [.53, .78]	Fab 2(6) [.80,.20] [.94,.31]	Fab2(12) [.32, .68] [.47, .79]	No(50) [.00,.00] [.00, .00]	Fab2(42) [.38, .62] [.58, .82]	Fab 2(8) [.38, .62] [.56, .82]	Fab1(13) [.34, .66] [.51, .86]
	BRA	Fab2(34) [.47, .53] [.70, .70]	Fab2(22) [.54, .46] [.75, .61]	Fab 1(5) [.80, .20] [.93, .32]	Fab1(15) [.47, .53] [.62, .71]	No(50) [.00,.00] [.00, .00]	No(50) [.00,.00] [.00,.00]	Fab1(31) [.50, .50] [.71, .69]	No(50) [.00,.00] [.00,.00]
	BAV	Fab2(22) [.37, .64] [.58, .84]	Fab2(16) [.37, .63] [.57, .84]	Fab 1(5) [.80, .20] [.91, .32]	Fab1(13) [.37, .63] [.57, .81]	No(50) [.00,.00] [.00, .00]	Fab2(46) [.34, .66] [.51, .80]	Fab2(10) [.42, .58] [.56, .77]	Fab1(23) [.37, .63] [.56, .80]

Table 3. Negotiation Results Under Stochastic Demand Using Single Tactic**.**

Acceptance party (Negotiation#) $[U^a, U^b]$ $[gain^1, gain^2]$		Fab 2 Agent							
		TB	TL	TC	RDD	RRE	BRE	BRA	BAV
Fab 1 Agent	TB	Fab2(34) [.51,.49] [.48,.44]	Fab1(29) [.60,.40] [.57,.41]	Fab1(9) [.85,.15] [.62,.19]	Fab2(20) [.77,.23] [.61,.27]	Fab1(47) [.12,.88] [.21,.68]	Fab2(40) [.35,.65] [.37,.55]	Fab1(35) [.50,.50] [.56,.55]	Fab2(34) [.35,.65] [.39,.56]
	TL	Fab2(28) [.41,.59] [.43,.50]	Fab1(23) [.50,.50] [.48,.46]	Fab2(8) [.77,.23] [.62,.25]	Fab1(17) [.58,.42] [.53,.40]	Fab1(45) [.10,.90] [.17,.61]	Fab1(33) [.35,.65] [.41,.58]	Fab2(24) [.49,.51] [.49,.48]	Fab2(28) [.41,.59] [.41,.54]
	TC	Fab1(9) [.12,.88] [.21,.62]	Fab 2(4) [.22,.78] [.28,.63]	Fab2(2) [.29,.71] [.34,.56]	Fab 2(4) [.22,.78] [.27,.60]	Fab 2(4) [.22,.78] [.30,.60]	Fab 2(6) [.19,.82] [.27,.60]	Fab 1(5) [.22,.78] [.26,.59]	Fab 1(7) [.22,.78] [.27,.63]
	RDD	Fab1(21) [.24,.76] [.29,.59]	Fab1(17) [.39,.61] [.42,.56]	Fab1(5) [.80,.20] [.62,.22]	Fab2(16) [.48,.52] [.45,.50]	Fab1(21) [.10,.90] [.17,.63]	Fab1(19) [.34,.66] [.34,.58]	Fab1(17) [.54,.46] [.46,.47]	Fab1(19) [.37,.63] [.38,.53]
	RRE	Fab2(48) [.86,.15] [.61,.18]	Fab1(41) [.82,.18] [.63,.22]	Fab1(7) [.83,.17] [.62,.18]	Fab2(22) [.90,.10] [.63,.15]	No(50) [.00,.00] [.00,.00]	No(50) [.00,.00] [.00,.00]	No(50) [.00,.00] [.00,.00]	No(50) [.00,.00] [.00,.00]
	BRE	Fab1(25) [.31,.69] [.35,.55]	Fab1(15) [.35,.65] [.40,.59]	Fab1(5) [.80,.20] [.62,.22]	Fab1(11) [.29,.71] [.37,.63]	No(50) [.00,.00] [.00,.00]	Fab2(42) [.38,.62] [.43,.55]	Fab 1(8) [.38,.62] [.44,.58]	Fab1(13) [.34,.66] [.40,.60]
	BRA	Fab1(33) [.47,.53] [.42,.45]	Fab1(21) [.46,.54] [.42,.52]	Fab1(5) [.80,.20] [.60,.23]	Fab2(16) [.48,.52] [.46,.50]	No(50) [.00,.00] [.00,.00]	No(50) [.00,.00] [.00,.00]	Fab2(24) [.53,.47] [.55,.45]	No(50) [.00,.00] [.00,.00]
	BAV	Fab2(22) [.38,.62] [.39,.55]	Fab 2(6) [.37,.63] [.40,.54]	Fab 1(5) [.80,.20] [.56,.22]	Fab1(13) [.37,.63] [.42,.55]	No(50) [.00,.00] [.00,.00]	Fab2(42) [.38,.63] [.42,.54]	Fab1(10) [.48,.52] [.46,.49]	Fab1(24) [.35,.65] [.39,.56]

According to the observation of the Tables 2 and 3, we conclude from the experiments as follows.

(1) Demand uncertainty harms the factory intrinsic utility about 20% to 30%.

(2) In Table 2 and 3, it can be observed that the majority of deals are made between the groups **B** and **C**, implying that the negotiation starter (i.e., Fab 1 herein) would loss little advantages to its opponent (i.e. Fab 2).

(3) Most fair negotiation results can be obtained when both agents apply the same tactic during the negotiation. As Tables 2 and 3 shown, in the case of both agents apply the same tactics (i.e., TB, and TL), fair deals can be obtained.

(4) The agent who uses a conceder tactic usually leads to a bad deal to itself. For instance, Fab 1 agent applies time dependent conceder tactic (the **TC** row of the Tables 2 and 3) and the final budget allocation plans located between the groups **B** and **C**. Furthermore, Fab 2 agent applying conceder tactic (the **TC** column of the Table 2 and 3) and the final budget allocation plan located between the groups **A** and **B**.

(5) The agent who applies the resources dependent tactics are risky since it can get either a very good or bad deal, depending largely on the attitudes of its opponent. For instance, The deal of Fab 1 agent applying a resource-estimation tactic (the “RRE” row of Tables 2 and 3) is either located at group A or reaches a failure negotiation. Furthermore, the deal of Fab 2 agent applying resource-estimation tactic (the “RRE” column of Tables 2 and 3) is either located at group C or reaches a failure negotiation.

5 Conclusions

The study makes a contribution to successfully build a mutually budget allocation negotiation model for a firm, so that the violations among the factories can be resolved. In the study, a negotiation model with eight alternative tactics is designed to mimic the negotiation attitude of each factory. In the experiments, the negotiation trajectories are traced and observed through the budget utility and intrinsic agent utilities. The results of single negotiation tactic combinations demonstrate that the negotiation attitudes will influence the compromised result significantly. Also, the different settings of demand variation in local capacity model have revealed that demand uncertainty would harm the overall firm welfare significantly. According to the compromised budget plan, each factory with private information regarding company objectives, demand and capacity developed a long-term capacity plan so that a set of local demand thus can be fulfilled.

In our future work, the proposed framework can be easily extended to a multiple-factories negotiation case. That is, during the negotiation, the potential profit of each agent can be calculated according to the finite budget minus latest offered values provided by all the other negotiation attendants. Note that any attendant can accept the plan during the negotiation. At the time, the budget value of this agent is booked and the remaining budget can be further negotiated by other agents. The negotiation is continuous until either the negotiation time is expired or all the attendants accept the plan.

6 Citations and References

[1] Agrawal N, Smith SA, Tsay AA. Multi-Vendor Sourcing in a Retail Supply Chain. Production of Operations Management 2002;11:157–182.

[2] Chan, FTS, Chung SH. A Hybrid Genetic Algorithm for Production and Distribution. The International Journal of Management Science 2005;33:345–355.

[3] Chang TS. An Auction System for Resource Capacity. Taiwan Patent 2001;29;No.503361

[4] Chiung M, Jongsoo K, Sun H. Integrated Process Planning and Scheduling with Minimizing Total Tardiness in Multi-Plants Supply Chain. Computers and Industrial Engineering 2002;43:331–349.

[5] David R, Smith RG. Negotiation as a Metaphor for Distributed Problem Solving. Artificial Intelligence 1983:20:63-109.

[6] Durfee EH, Lesser VR. Negotiating Task Decomposition and Allocation Using Partial Global Planning. Distributed Artificial Intelligence 1989;2:229–243.

[7] Faratin P, Sierra C, Jennings RN. Negotiation Decision Functions for Autonomous Agents. Robotics and Autonomous Systems 1998;24:159–182.

[8] Jiang AX. Capacity Trading Among Semiconductor Manufacturing Factories. Master thesis, National Taiwan University, 2000.

[9] Jolayemi JK, Olorunniwo FO. A Deterministic Model for Planning Production Quantities in a Multi-Plant, Multi-Warehouse Environment with Extensible Capacities. International Journal of Production Economics 2004;87:99–113.

[10] Karabuk S, Wu SD. Incentive Schemes for Semiconductor Capacity Allocation : A Game Theoretic Analysis. Production and Operation Management 2005;14:175–188.

[11] Kraus S, Lehmann D. Designing and Building an Negotiating Autonomous Agent. Computational Intelligence 1995;11:132–171.

[12] Kraus S, Wilkenfeld G, Zlotkin G. Multiagent Negotiation Under Time Constraints. Artificial Intelligence 1995;75:297–345.

[13] Parunak HVD. Agents in Overalls: Experiences and Issues in the Development and Deployment of Industrial Agent-Based Systems. International Journal of Cooperative Information Systems 2000;9:209–227.

[14] Raiffa H. The Art and Science of Negotiation, Cambridge, Mass.: Belknap Press of Harvard University Press, 1982.

[15] Rau H, Tsai MH, Chen CW, Shiang WJ. Learning-based Automated Negotiation Between Shipper and Forwarder. Computer & Industrial Engineering 2006;51:464–481.

[16] Shen W, Norrie DH. Agent-Based Systems for Intelligent Manufacturing: a State-of-The-Art Survey. International Journal Knowledge and Information Systems 1999;1:129–156.

[17] Veeramani D, Wang KJ. Performance Analysis of Auction-based Distributed Shop-floor Control Schemes from the Perspective of the Communication System. International Journal of Flexible Manufacturing System 1997;9:121–143.

[18] Wang KJ, Chou CH. Evaluating NDF-based Negotiation Mechanism Within An Agent-based Environment. Robotics and Autonomous Systems 2004;43:1–27.

[19] Wang KJ, Wang SM, Chen JC. A Resource Portfolio Planning Model Using Sampling-Based Stochastic Programming and Genetic Algorithm. European Journal of Operational Research 2008;184:327–340.

A Negotiation-Based Approach to Supply Chain Planning and Scheduling Problems in a Fully Distributed Environment

K. Robert Lai[a] and Bo-Ruei Kao[a,1]

[a]Department of Computer Science & Engineering, Yuan Ze University.

Abstract. This paper presents an agent-based fuzzy constraint-directed negotiation mechanism for planning and scheduling in supply chain. The supply chain scheduling problem is modeled as a set of fuzzy constraint satisfaction problems (FCSPs), interlinked together by inter-agent constraints. For converging each distinct firm's interests, the conflicts among the set of FCSPs are resolved through negotiation by iteratively exchanging offers/counters with limited sharing of their perspectives and preferences. During the negotiation, proposing offers/counter-offers takes not only firm's self-interest and preferences but also opponents' perspectives into consideration. By sharing perspectives between agents to gradually uncovering the intent of opponents, consensus can be obtained and the quality of consensus can be guaranteed at satisfactory level. Experimental results suggest that the proposed approach obtains a superior solution for supply chain scheduling than other negotiation models in a fully distributed processing.

Keywords. Supply Chain Planning and Scheduling, Fuzzy Constraint-directed Negotiation, Fuzzy Constraint Satisfaction Problem, Multi-Agent System

1 Introduction

Facing with global economic downturn and intense competition, firms have to rapidly respond to the variations in market situation. Business entities (suppliers, manufacturers, distributors, retailers and customers) are urged to integrate their operations into multi-layer supply chain for product/service provision [1]. However, supply chain usually handles multiple projects concurrently with shared components, facilities, and capacities governed by distinct firms. Collaborating with supply chain partners in planning and scheduling in a cost effective manner is the key to survive in today's fierce market competition.

Since the supply chain environment in nature is distributed, autonomous, and heterogeneous, agent-based approaches which are characterized by decentralization of computation and information processing are particularly attractive for supply chain modeling and problem solving [2, 3]. To ensure the global performance with

[1] Corresponding Author Email: s909402@mail.yzu.edu.tw

limited interaction among agents, a mediator or a third party agent is used in some systems for coordination. Negotiation processes can be facilitated by a mediator which helps the parties understand their needs, suggests possible agreements and/or supports them in the implementation of the agreement. But, agents may have to share sensitive strategic information that should not be revealed to competitors or even to a third party agent, to the mediator for coordination.

So, instead of coordinating with a third party agent, coordination/negotiation in a fully distributed environment for privacy concern has been addressed in some agent based approaches. Those papers address the dynamic nature of supply chain and emphasize the flexibility and responsiveness. Two popular and fully distributed negotiation models, contract net protocol (CNP) and market-based protocol, have been proposed for these purposes [4]. The CNP can rapidly produce a feasible solution to overcome the frequently changes in supply chain. With simplified negotiation protocol, however, agents with very limited interaction and information sharing can only make their decisions independently and optimize their local objectives in a myopic way. Accordingly, more sophisticated negotiation mechanism is needed and thus bringing us to a market-based approach. Market-based approaches employ a bargaining or auction process characterized by iterative bidding among agents. During the bidding process, agents will adjust bids according to the direction of surplus and deficit of demand. The bids which imply the degree of competition can be used to resolve the conflicts caused by contention and improve the system performance. But the bid indicates only the demand that is in high or low contention, not where the contention can be resolved. Thus, the process might oscillate and not achieve the convergence. It also affects the quality of solution.

Accordingly, facilitating the convergence and guaranteeing the system performance in supply chain are the critical challenges and they are highly influenced by the level of information sharing [5, 6, 7, 8] as well. To cope with these problems, this paper proposes an agent-based fuzzy constraint-directed negotiation mechanism (AFCN) in a fully distributed environment.

In AFCN, a supply chain problem is modelled as a set of fuzzy constraint satisfaction problems (FCSPs), interlinked together by inter-agent constraints. Each FCSP represents the firm's perspectives and issues in supply chain. Constraints which are characterized by non directional, declarative, and intuitive properties are closely related to the real-world problem descriptions. Besides, the subjective, imprecision and qualitative knowledge (e.g., human cognition, preferences, or even opponent's perspectives) which are frequently used for business decision making in supply chain can easily be coped by fuzzy constraints with the levels of consistency. Fuzzy constraints also can be used to rank the solutions by specifying the possibilities prescribing to what extent the solutions are suitable. Additionally, it even provides a measure of similarity between solutions and opponents' counteroffers and a basis for the selection of multi-objective decision from a set of feasible alternatives. Thus, AFCN can be a practical and effective methodology for the supply chain problem modelling.

To coordinate supply chain parties, the conflicts among set of FCSPs are resolved through AFCN negotiation protocol by iteratively exchanging offers/

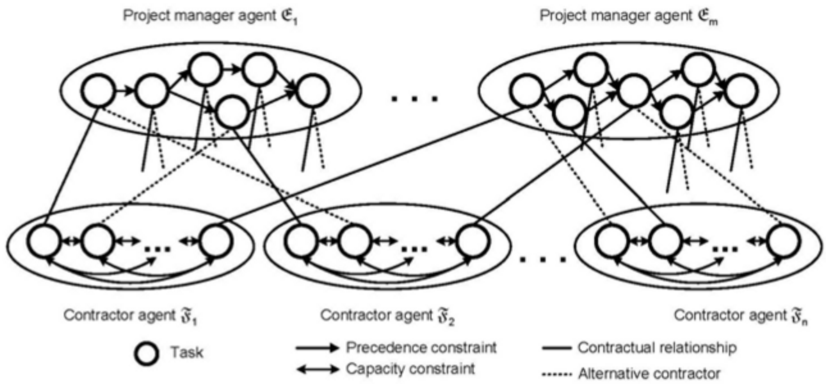


Fig. 1. Supply chain project scheduling with m project agents and n contractor agents.

counter-offers with limited sharing of their perspectives. For converging each distinct firm's interests in supply chain, proposing an offer/counteroffer takes not only firm's self-interest and preferences but also opponents' perspectives into consideration. In a proposed offer/counteroffer, it indicates not only the region of acceptable solutions and preference degrees but the possibility of conflicts in the region. For each FCSP, incremental propagation eliminates the redundant decision values and infeasible combination of solutions. According to the ranking of solutions by fuzzy constraints, the set of feasible solutions can be further restricted in a preferred region with an acceptable/ controlled threshold. It supports agent to quick response the changes in the environment and promises the proposed offers/counteroffers to be focused within the interest/attention area.

By sharing limited information among agents to incrementally uncovering the intention of opponents, consensus can be obtained and the quality of solution can be guaranteed at a satisfaction level as well. Moreover, the framework of AFCN also provides flexibility to incorporate negotiation strategies, such as self-interested, cooperative, and win-win, for various global performance measures.

The remainder of this paper is organized as follows. Section 2 introduces the theoretical basis for modelling supply chain planning and scheduling problem as a distributed fuzzy constraint satisfaction problem (DFCSP). Section 3 then presents an agent-based fuzzy constraint-directed negotiation mechanism in detail. Then, Section 4 provides the experimental results to show the effectiveness of the proposed approach followed by conclusions in Section 5.

2 Modeling Supply Chain Planning and Scheduling as DFCSP

During the process of production/service provision from acquiring material to deliver to the end customers, it passes through the stages of raw material supply, intermediate supply, manufacturing, distribution, and retail. This paper focuses on the stage of manufacturing in supply chain. The supply chain project scheduling problem, as shown in Figure 1, is to schedule multiple projects in a network of manufacturers and suppliers (or contractors). Each project involves a set of tasks or operations with complex precedence relationships in which the task can be performed by a set of alternative suppliers. Suppliers differ from each other in terms of resource capacity, processing times, and costs of performing tasks. The process of project planning and scheduling specifies how the suppliers are selected from alternatives and tasks are scheduled. The problem is multi–objective in nature. For example, manufacturers wish to optimize the scheduling performance as well as to minimize the operation cost. On the other hand, the suppliers are more inclined to maximize the utilization and profit.

The problem can be represented as a triple $(\mathbf{E}, \mathbf{F}, \mathbf{I})$, where $\mathbf{E}$ is a set of m project manager agents (PAs) for the manufacturers, $\mathbf{F}$ is a set of n contractor agents (CAs) for the suppliers, and $\mathbf{I}$ is a set of inter-agent constraints between two class of agents. Each PA consists of a chain of tasks which are specified further by precedence constraints, including the processing time on the set of alterative suppliers for each task, due date, arrival (release) date, and tardiness cost.

Each CA consists of capacity constraints and processing cost. Thus, the problem can be modeled as a distributed constraint satisfaction problem (DFCSP) in that coming up to a mutually acceptable solution between two classes of agents is the same as uncovering a consistent solution satisfying all the constraints in a distributed fuzzy constraint network specifying the fuzzy relationships inside each agent and between agents. Adapted from [9], a distributed fuzzy constraint network (DFCN) can be defined as below.

Definition 1. A distributed fuzzy constraint network $(\mathbf{U}, \mathbf{X}, \mathbf{C})$ for a supply chain project scheduling problem $(\mathbf{E}, \mathbf{F}, \mathbf{I})$ can be defined as a set of $m+n$ fuzzy constraint networks $\{\mathbf{N}_1, ..., \mathbf{N}_{m+n}\}$, $\mathbf{N}_l$ representing an agent $l \in \{\mathbf{E}, \mathbf{F}\}$, where

- $\mathbf{U}_l$ is a universe of discourse for an FCN $\mathbf{N}_l$;
- $\mathbf{X}_l$ is a tuple of non-recurring objects of agent l;
- $\mathbf{C}_l$ is a set of fuzzy constraints which involves a set of internal fuzzy constraints existing among objects in $\mathbf{X}_l$, and a set of external fuzzy constraints $\mathbf{I}_l$ between agent l and opponent agents;
- $\mathbf{N}_l$ is connected to other FCNs by a set of external fuzzy constraints $\mathbf{I}_l$;
- $\mathbf{U}$ is a universe of discourse;
- $\mathbf{X} = (\bigcup_{l=1}^{m+n} \mathbf{X}_l)$ is a tuple of all non-recurring objects;
- $\mathbf{C} = (\bigcup_{l=1}^{m+n} \mathbf{C}_l)$ is a set of all fuzzy constraints.

In Definition 1, a set of $\mathbf{X}_l$ of agent l is corresponded to its beliefs, its knowledge of the environment (tasks, resources, etc.), and any other attitudes (desires, intentions, opponents' response, etc.). A set of fuzzy constraints $\mathbf{C}_i$ for project agent $\mathbf{E}_i$ is corresponding to a set of restrictions (e.g. precedence constraints), objectives (e.g. flowtime, operating cost) and inter-agent constraints between project agent $\mathbf{E}_i$ and related contractor agents; $\mathbf{C}_k$ for contractor agent $\mathbf{F}_k$ is corresponding to a set of restrictions (e.g. capacity constraints), objectives (e.g. utilization) and inter-agent constraints between contractor agent $\mathbf{F}_k$ and related project agents.

By Definition 1, a consistent solution of a DFCSP for scheduling is an instantiation of all time allocation of the tasks such that all the constraints of the agents are satisfied at a degree of membership function. No agent knows about other agents' feasible solutions and possible agreements a priori. An agent negotiation mechanism is to explore potential agreements and then to move the negotiation toward a globally beneficial solution.

3 Agent-based fuzzy constraint-directed negotiation mechanism

Fuzzy constraint-directed approach has been demonstrated in [10, 11, 12, 9, 13] as an effective framework for agent negotiation. During the negotiation, each agent plans the solution by solving its own FCSP, and exchanges the proposals among the agents for resolving the inconsistency of time allocation of activities. As a proposal cannot be accepted by its opponents, a counter-proposal will be made according to the negotiation strategies which consider alternative solutions at the same constraint satisfaction level or offer a solution with a lower constraint satisfaction level. Exchanging proposals and counter-proposals will continue until termination conditions (e.g., achieving a consensus or a failure) are met.

In the proposed negotiation mechanism, the process of proposing a offer/counteroffer in an agent is regarded as the inference process for solving its own FCSP. The process includes the following steps: opponent responsive state evaluation, internal state update, behavioral state determination, set of feasible solutions generation, prospective solution selection and offer/counter-offer generation. The process of proposing an offer $\mathbf{A}_{i,j}^{*}$ with the FCN $\mathbf{N}_{i,j} = (\mathbf{U}_{i,j}, \mathbf{X}_{i,j}, \mathbf{C}_{i,j})$ for the task j in project manager agent $\mathbf{E}_i$ in a negotiation round is described as follows.

While receiving the counter-offer $\mathbf{B}_{i,j}$ over negotiation issues $\mathbf{I}_{i,j} \in \mathbf{X}_{i,j}$ from opponent agents, the opponent responsive state $\sigma_{i,j}$, which indicates the difference between current self-interest and opponents' perspectives, is obtained by

$$\sigma_{i,j} = 1 - ((\hat{\mathbf{D}}_{i,j} - \mathbf{D}_{i,j}) / \hat{\mathbf{D}}_{i,j}), \tag{1}$$

where $\mathbf{D}_{i,j}$ is the distance between the offer $\mathbf{A}_{i,j}'$ generated in last round and the latest counter-offer $\mathbf{B}_{i,j}$ and $\hat{\mathbf{D}}_{i,j}$ is the basis of distance for normalizing which is obtained at the first negotiation round.

The distance $\mathbf{D}$ between an offer $\mathbf{A}$ and a counter-offer $\mathbf{B}$ over set of issues $\mathbf{I}$ is obtained from

$$\mathbf{D} = \frac{1}{N_I}\sqrt{\sum_{k=1}^{N_I} \mathsf{L}(\mathbf{A}^k,\mathbf{B}^k)^2}, \tag{2}$$

where L is a distance measure for two fuzzy set, $\mathbf{A}^k$ and $\mathbf{B}^k$ is the fuzzy set for the offer $\mathbf{A}$ and the counter-offer $\mathbf{B}$ over issue $\mathbf{I}^k \in \mathbf{I}$, respectively.

As none of the counter-offers can be accepted by the project manager agent E_i, agent E_i has to decide the behavioral state to proposal alternative solutions in the current preferred region or offer a solution with a lower constraint satisfaction level. The internal states of agent and the opponent responsive state are involving to determine the behavioral state. In this manner, the internal states include the satisfaction level $\rho_{i,j}$ for task j and degree of tightness in solution space $\delta_{i,j}$. The aggregated satisfaction value $\rho_{i,j}$ for task j indicates the satisfaction level for the current prospective solution $\mathbf{S}'_{i,j}$ and is obtained by

$$\rho_{i,j} = \Psi(\mathbf{S}'_{i,j}) = \frac{1}{N_G}\sum_{k=1}^{N_G} \mu_{\mathbf{C}_k}(\mathbf{S}^k_{i,j}), \tag{3}$$

where $\mathbf{S}'_{i,j}$ is the prospective solution of task j obtained in last negotiation round, $\mathbf{\Psi}(\mathbf{S}'_{i,j})$ denotes the aggregated satisfaction value for the prospective solution, $\mu_{\mathbf{C}_k}(\mathbf{S}^k_{i,j})$ is the satisfaction degree of the fuzzy constraint $\mathbf{C}_k$ over a set of E_i's goals $\mathbf{G}$.

The tightness of solution space $\delta_{i,j}$ indicates the remaining feasible solution space between the aggregated satisfaction value $\rho_{i,j}$ and the acceptable threshold $\varepsilon'_{i,j}$ and is obtained by

$$\delta_{i,j} = 1 - (\rho'_{i,j} - \varepsilon'_{i,j}), \tag{4}$$

where $\rho'_{i,j}$ is the aggregated satisfaction value, and $\varepsilon'_{i,j}$ is the threshold of aggregated satisfaction determined in last negotiation round.

According to the internal state and the opponent responsive state, the level cut τ for the desire of concession $\mathbf{V}$ obtained by

$$\tau = (\mu_\rho(\rho'_{i,j}) \wedge \mu_\delta(\delta'_{i,j}) \wedge \mu_\sigma(\sigma_{i,j}))^{\mathbf{W}_c}, \tag{5}$$

where $\mu_\rho(\rho'_{i,j})$, $\mu_\delta(\delta'_{i,j})$ and $\mu_\sigma(\sigma_{i,j})$ denote the desire of concession according to the degree of tightness, aggregated satisfaction and degree of difference, respectively; $\mathbf{W}_c$ denotes the weight associated with the desire of concession.

Then, the behavioral state regarded as the threshold of aggregated satisfaction $\varepsilon_{i,j}$ is obtained from

$$\varepsilon_{i,j} = \varepsilon'_{i,j} - \Delta\varepsilon = \varepsilon'_{i,j} - \mathsf{D}(\mathbf{V}_\tau), \tag{6}$$

where $\varepsilon'_{i,j}$ is the threshold of aggregated satisfaction obtained from last negotiation round, and $\Delta\varepsilon$ is the concession value transformed from the fuzzy set of the desire of concession $\mathbf{V}$ with τ level cut. D is the defuzzification method.

To evaluate and generate the offer and counter-offer, an agent has to plan its prospective solution first. Each agent can only plan the prospective solution from its individual area of interests limited by the threshold of aggregated satisfaction $\varepsilon_{i,j}$. That is, a set of feasible solution $\mathsf{P}_{i,j}$ for task j in agent E_i can be defined as

$$\mathsf{P}_{i,j} = \{\mathbf{S}_{i,j} \big| (\mathbf{S}_{i,j} \in_{\alpha} \Pi_{i,j}) \wedge (\Psi(\mathbf{S}_{i,j}) > \varepsilon_{i,j})\}, \tag{7}$$

where $_{\alpha}\Pi_{i,j}$ is the solution set with α level cut, $\mathbf{\Psi}(\cdot)$ is the aggregated satisfaction value, and $\varepsilon_{i,j}$ is the threshold of aggregation satisfaction degree of objectives.

Given the counter-offer $\mathbf{B}_{i,j}$ and the feasible solution set $\mathsf{P}_{i,j}$, the prospective solution $\mathbf{S}^*_{i,j}$ is generated by solution selection method

$$\mathbf{S}^*_{i,j} = \arg_{\mathbf{S}}(\max_{\mathbf{S}\in\mathsf{P}_{i,j}} \mathsf{H}\,(\mathbf{S},\mathbf{B}_{i,j})), \tag{8}$$

where H is a utility function to evaluate the appropriateness between the feasible schedule $\mathbf{S}$ and the counter-offer set $\mathbf{B}_{i,j}$ in which the utility function H can be defined by

$$\mathsf{H}\,(\mathbf{S},\mathbf{B}) = \frac{1}{N_I}\sqrt{\sum_{k=1}^{N_I}(\mathsf{P}\,(\mathbf{S}^k)^{w_p} \wedge \mathsf{S}\,(\mathbf{S}^k,\mathbf{B}^k)^{w_q})^2}, \tag{9}$$

where P is a satisfaction function over the issue $\mathbf{I}_k$, and S is a similarity function (a distance measure) between the solution $\mathbf{S}$ and the counter-offer $\mathbf{B}$; w_p, and w_q denote the weight associated with the satisfaction and the similarity of the solution of task, respectively.

For the task j, given the feasible solution set $\mathsf{P}_{i,j}$ and the prospective solution $\mathbf{S}^*_{i,j}$ which involves a set of values $\{\mathbf{S}^1_{i,j},\mathbf{S}^2_{i,j},\ldots,\mathbf{S}^{N_I}_{i,j}\}$ for NI issues, the offer $\mathbf{A}^*_{i,j}$ for task j is a tuple of fuzzy set $\{\mathbf{A}^1_{i,j},\mathbf{A}^2_{i,j},\ldots,\mathbf{A}^{N_I}_{i,j}\}$ in which each fuzzy set $\mathbf{A}^k_{i,j}$ for issue $\mathbf{I}^k_{i,j} \in \mathbf{I}_{i,j}$ is the marginal particularized possibility distribution

$$\mathbf{A}^k_{i,j} = Proj_{\mathbf{X}^k_{i,j}}(\mathsf{P}_{i,j} \cap \overline{\Pi}_{\mathbf{X}^1_{i,j}} \cap \ldots \cap \overline{\Pi}_{\mathbf{X}^{k-1}_{i,j}} \cap \overline{\Pi}_{\mathbf{X}^{k+1}_{i,j}} \cap \ldots \cap \overline{\Pi}_{\mathbf{X}^{N_I}_{i,j}}), \tag{10}$$

where $\overline{\Pi}_{\mathbf{X}^k_{i,j}}$ is the cylindrical extension of $\Pi_{\mathbf{X}^k_{i,j}}$ in the space $(\mathbf{X}^1_{i,j},\mathbf{X}^2_{i,j},\ldots,\mathbf{X}^{N_I}_{i,j})$, $\Pi_{\mathbf{X}^k_{i,j}} = \mathbf{S}^k_{i,j}$, and $\mathbf{X}^k_{i,j}$ is the object of issue $\mathbf{I}^k_{i,j}$ for task j.

Meanwhile, when project manager agent E_i receives a counter-offer $\mathbf{B}_{i,j} = \{\mathbf{B}^1_{i,j},\mathbf{B}^2_{i,j},\ldots,\mathbf{B}^{N_I}_{i,j}\}$ for the task j, the opponent agents' preferred solution $\hat{\mathsf{P}}_{i,j}$ can be obtained by

$$\hat{\mathbf{P}}_{i,j} = (\overline{\mathbf{B}}_{i,j}^{1} \cap \overline{\mathbf{B}}_{i,j}^{2} \cap \ldots \cap \overline{\mathbf{B}}_{i,j}^{N_l}), \quad (11)$$

where $\overline{\mathbf{B}}_{i,j}^{k}$ is the cylindrical extension of $\mathbf{B}_{i,j}^{k}$ in the space $\mathbf{X}_l$. Each element in $\hat{\mathbf{P}}_{i,j}$ represents a solution preferred by opponent agents and the membership degree of the element represents the acceptability of solution for all opponent agents.

The project manager agent $\mathbf{E}_i$ will accept the counter-offer $\mathbf{B}_{i,j}$ proposed by its opponent as an agreement if

$$(\mathbf{P}_{i,j} \cap \hat{\mathbf{P}}_{i,j}) \neq \{\}, \quad (12)$$

The negotiation process for task allocation will be terminated when the project manager agent reach an agreement with one of contractor agents or the project manager agent or all of their contractor agents withdraws the negotiation. A project manager agent will terminate the negotiation when each of tasks in project is assigned to a contractor agent or the project manager agent withdraws the negotiation while failing to reach agreement. And the supply chain planning/scheduling process will be terminated when all of project manager agents are termination.

4 Experiments

To demonstrate the utility of the proposed model for project planning and scheduling problem in supply chain, the experiments are meant to compare AFCN with centralized heuristic (CTR) [14], conventional CNP (CNP) [15], modified CNP (MCNP) [4], market-based auction CNP (MA-CNP) (i.e., extended from CNP) and market-based auction MCNP (MA-MCNP) (extend from MCNP) (i.e., extended from MCNP) in terms of the number of project managers over makespan, total operating cost, and computational time.

The experiment was implemented on a PentiumMPC with 0.8 GHz running windows XP and 256 MB RAM. In the experiment, each project has a linear sequence of tasks, and each task is then specified by its precedence constraints, required resources, processing time and tardiness cost. Each supplier is specified by its unique capacity and processing cost. The total operating cost is the sum of the total processing cost and tardiness cost of each project. For simplicity, the number of tasks per project is identical at five, alternative suppliers for a task are identical at three, and 50 instances are generated for each number of project managers. Figures 2 to 4 show the performance comparisons over makespan, total operating cost, and computational time, respectively.

The conventional CNP supports one PA to assign a task at a time and another PA can start only after the PA finishes selecting CAs for all its tasks. In this form of negotiation model, a large number of slacks will be generated since that the conflicts among PAs cannot be coordinated. Each PA's performance highly depends on the order of contracting.

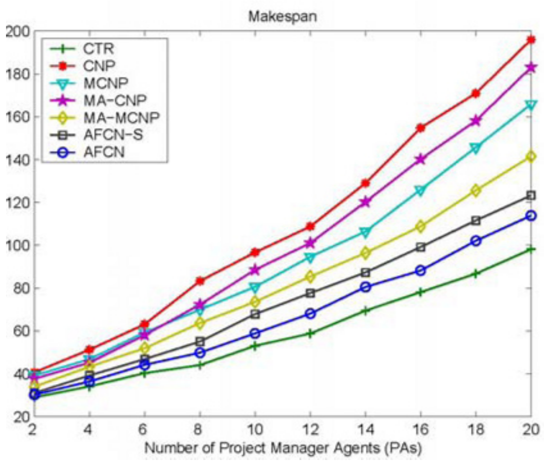


Fig. 2. Makespan of various negotiation mechanisms over a number of project managers.

Instead of one PA assigning a task, MCNP supports multiple PAs to contract tasks to multiple CAs simultaneously. In the manner, PAs propose the whole time window of task and price to CAs. CAs resolve the conflicts among tasks according to the urgency of task be completed and the price of the task. It reduces the possibility of violating the due date of the PAs and decreases the tardiness when a project is delayed.

Since the simplicity of negotiation protocol, the single-shot negotiation models (CNP and MCNP) have the best time efficiency in Figure 4. With very limited information sharing and myopic decision-making, however, the global performances can easily get trapped at local optima and are highly unstable. It also can be observed in Figures 2 and 3, these models worse results than other approaches over makespan and total operating cost.

For reflecting the service preference in PA and the task contention in CA, MACNP and MA-MCNP incorporate the market-based auction mechanism [16] which is characterized by iterative bidding among agents. Agents will adjust the prices of bid according to the direction of surplus and deficit of demand. Task allocation processing in CA can be adjusted iteratively according to time boundary and the updating price of tasks. In this way, MA-MCNP with iterative bidding have better makespan and total operating cost than the single-shot negotiation models. However, due to the characteristic of protocol in MA-CNP, MA-CNP even worse than MCNP over makespan and total operational cost when the number of project managers is larger than 8 in Figure 2 and 3, respectively.

However, the price of task proposed by CA indicates the contention degree of the desired time slots of task, but not indicates where the contention can be resolved. PA might blindly increase price but the schedule of task can not be improved (overpricing). Besides, sharing full of task's time boundary in PA is not practical in reality.

In AFCN, PA generates the offer not only considers how much suppliers can provide this service, but how pressing the task be finished. Instead of the bid

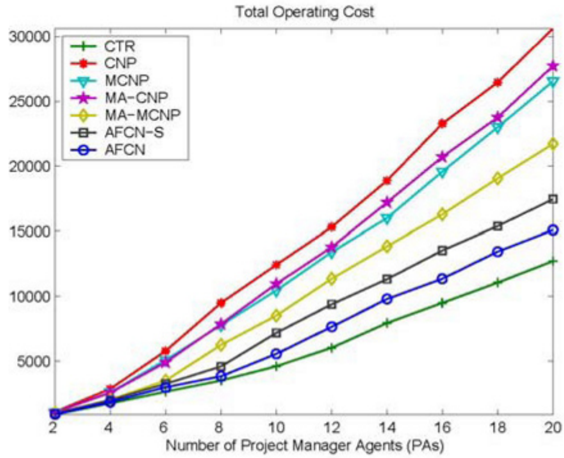


Fig. 3. Operating cost of various negotiation mechanisms over a number of project managers.

information in the market-based mechanism, the negotiation information in the AFCN varies not only depending on the particular combination of task demands/resource contracts, but the objectives and preferences of agent. Both of PA and CA proposes offers/counter-offers which involve the current feasible time slots of task with preference and price. The time slots with preference in offer/counter-offer indicate not the current time boundary of task, but the time window for conflict-free. According to time slots with preference from PA, CA can arrange the task allocation efficiency and can response the conflict in counter-offer directly. Meanwhile, CA generates the counter-offers considering not only the tasks of PAs are satisfied but the profit obtained in the transaction. Based on the counter-offers from CA, PA can avoid requiring the high contention area in CA or can abort lower quality service of suppliers. Thus, PA can bid the desired time slots to preferred suppliers with controllable budget.

In Figure 2 and 3, it can be observed that the AFCN has superior performance than other approaches in makespan and total operating cost. Meanwhile, AFCN performs better in time efficiency than that of market based mechanism. Additionally, as the problem size is increasing, the performance improved by the AFCN grows more significantly.

5 Conclusions

We have presented an agent-based fuzzy constraint-directed negotiation protocol for project planning and scheduling in supply chain. Constraint modeling gives a more direct fusion to the real-world problem descriptions and the impreciseness of knowledge in supply chain can easily be represented by fuzzy constraints with the levels of consistency as well. Experimental results suggest that the AFCN indeed can provide a practical and efficient framework for supply chain planning and scheduling.

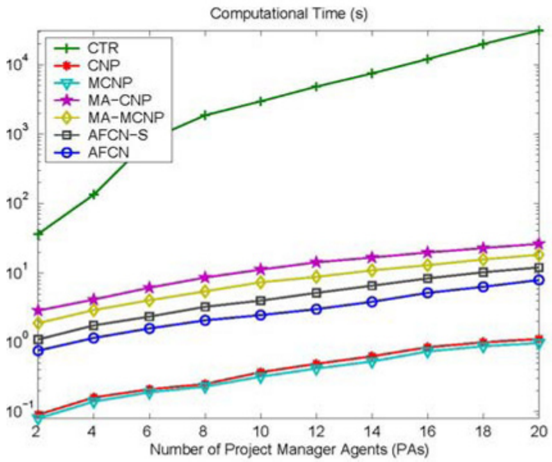


Fig. 4. Computational time of various negotiation mechanisms over a number of project managers.

6 References

[1] B. M. Beamon. Supply chain design and analysis: Models and methods. International Journal of Production Economics, 55(3):281–294, 1998.

[2] F. R. Lin, H. C. Kuo, and S. M. Lin. The enhancement of solving the distributed constraint satisfaction problem for cooperative supply chains using multi-agent systems. Decision Support Systems, 45(4):795–810, 2008.

[3] T. Calosso, M. Cantamessa, and M. Gualano. Negotiation support for make-to-order operations in business-to-business electronic commerce. Robotics and Computer-Integrated Manufacturing, 20(5):405–416, 2004. Cited By (since 1996): 7.

[4] J. S. K. Lau, G. Q. Huang, K. L. Mak, and L. Liang. Agent-based modeling of supply chains for distributed scheduling. IEEE Transactions on Systems, Man, and Cybernetics, Part A, 36(5):847–861, 2006.

[5] Y. Aviv. The effect of collaborative forecasting on supply chain performance. Management Science, 47(10):1326–1343, 2001.

[6] P. J. Byrne and C. Heavey. The impact of information sharing and forecasting in capacitated industrial supply chains: A case study. International Journal of Production Economics, 103(1):420–437, 2006.

[7] G. P. Cachon and M. Fisher. Supply chain inventory management and the value of shared information. Management Science, 46(8):1032–1048, 2000.

[8] L. Yan and T. G.Woo. Information sharing in a supply chain with dynamic consumer demand pattern. In Proceedings of the Hawaii International Conference on System Sciences, volume 37, pages 1143–1152, 2004.

[9] K. R. Lai. Fuzzy Constraint Processing. PhD thesis, NCSU, Raleigh, N. C., 1992.

[10] K. Y. Cai. Robustness of fuzzy reasoning and ±-equalities of fuzzy sets. IEEE Transactions on Fuzzy Systems, 9(5):738–750, 2001.

[11] D. Dubois, H. Fargier, and H. Prade. Propagation and satisfaction of flexible constraints. In Fuzzy Sets, Neural Networks, and Soft Computing, pages 166–187. Van Nostrand Reinhold Company, New York, 1994.
[12] D. Dubois, H. Fargier, and P. Fortemps. Fuzzy scheduling: Modelling flexible constraints vs. coping with incomplete knowledge. European Journal of Operational Research, 147(2):231–252, June 2003.
[13] J. Moura-Pires and H. Prade. Specifying fuzzy constraints interactions without using aggregation operators. In Proceedings of the Ninth IEEE International Conference on Fuzzy Systems, volume 1, pages 228–233, 2000.
[14] C. G. Li, J. B. Yu, and Liao. X. F. Fuzzy tabu search for solving the assignment problem. In Communications, Circuits and Systems and West Sino Expositions, IEEE 2002 International Conference on, volume 2, pages 1151–1155, June-1 July 2002.
[15] R. G. Smith. The contract net protocol: High-level communication and control in a distributed problem solver. IEEE Transactions on Computers, C-29(12):1104–1113, 1980.
[16] M.P. Wellman, W.E. Walsh, P.R. Wurman, and J.K. MacKie-Mason. Auction protocols for decentralised scheduling. Games and Economic Behavior, 14(5):429–439, 2001.

Environmental Transparency of Food Supply Chains – Current Status and Challenges

Nel Wognum[1], Harry Bremmers[1]

Wageningen University – WUR, The Netherlands

Abstract: Food chains need to become more sustainable to regain and retain consumer trust after recent food incidents and scandals. One of the key components of sustainability is environmental care. To what extent do supply chains invest in environmental care and to what extent are consumers willing to pay for products from a sustainable production chain? In other words, can environmental care contribute to supply chain value? Transparency of environmental care in the entire food chain is necessary to contribute to increased consumer awareness. Currently, though, transparency is mainly limited to traceability for food safety for human health. In this paper significance and current status of environmental care systems in supply chains and chain actor companies are described. Challenges to be addressed to improve environmental care are identified.

Keywords: Food supply chain, sustainability, value creation.

1 Introduction

In the recent past several food scandals have occurred, which have increased consumer awareness with respect to food safety, health, and environmental issues. These food scandals showed that the food industry has not cared very much about health and safety aspects of the food chain leading to reduced consumer confidence in the food products they buy. To regain and retain consumer trust, the food system needs to become more sustainable [1].

Managers in the food industry and agribusiness become increasingly aware of the fact that consumers differentiate with respect to the quality of their products. Consumers not only value food safety and physical characteristics, but also more and more tune their buying behaviour to their judgment of environmental and hygiene dangers of the food production system. In this way consumers provide the (different) standards for food quality. Environmental care and sustainability are increasingly regarded as distinct product characteristics.

---

[1] Corresponding Author : Wageningen University, Social Sciences Department, Management Studies Group, Hollandseweg 1, 6706 KN Wageningen, The Netherlands, Phone: +31317485577, E-mail: {nel.wognum;harry.bremmer}@wur.nl.

Increased transparency is needed to reassure consumers, assure quality and provide an incentive to buy. Transparency can be defined as the extent to which all stakeholders of a food system have a shared understanding of, and access to, product-related information requested by them without loss, noise, delay, and distortion [2]. Transparency is not possible without shared understanding of the information [3], for which standards and agreed formats are necessary.

Transparency has a horizontal and vertical dimension [1]. The vertical dimension refers to legislation and quality demands for traceability along the supply chain. The horizontal dimension refers to content demands of information per chain stage from quality standards and legislation. Sustainability information encompasses human health and safety, animal health and safety, and environmental impact [1]. Until now, mainly traceability for food safety for human health has been addressed, as required by European and national legislation and standards. Although animal health and environmental issues are considered important by all food chain actors [1] many food chains do not yet sufficiently address these issues, in particular large pork chains (see also [4]).

In this paper we focus on the environmental component of transparency. While food safety has a direct impact on human health, environmental impact of the food production process is not directly measurable in the final product. Environmental impact can be direct, for example in cases of water pollution, but is often indirect and long-term. In particular impact on climate by deterioration of rain forest areas by soy or palm oil production is not directly visible for food buyers. Moreover, investments in environmental care systems are large without direct visible contribution to the value of the company and food supply chain. In section 2, we will define transparency in more detail and present the current status of transparency in the food chain. In section 3 we explore environmental transparency into more depth. We present results of a study into current actor and supply chain level environmental management in food chains. Section 4 presents challenges that need attention to improve environmental transparency of food chains.

2 Transparency in the food chain

Recent food crises have increased consumer awareness of the impact of intensive bio-industrial food production, which occupies the largest part of food production in Europe and beyond. Destruction of animals associated with the BSE crisis, for example, has induced ethical concerns [5]. The BSE crisis and other ones, like dioxin, classical swine fever, and Avian Influenza, have fuelled consumer concerns about quality and safety in food production systems (see e.g., [6]). Consumers have become more critical and want more information about the food they buy, like origin of the food, safety levels, production means, hygiene, use of genetically modified feed, application of pesticides, and other environmental issues, like food miles [7].

Food systems have become increasingly complex. First of all, consumers not only want their food to be safe, produced without too much damage to the environment, and still be affordable, but also desire a wide choice of fresh and processed food. Secondly, food systems have become increasingly international, involving more, small and large, companies with long transportation times to transport feedstuffs, life cattle, or intermediate and ready food products. Measures to safeguard consumers against any food hazards need, hence, be formulated internationally.

In January 2002, the EU government formulated the General Food Law (GFL - 178/2002/EC), applicable from January 1, 2005, which imposes demands for the vertical dimension of transparency, in particular traceability, to all actors in a food production chain. Article 18 of the GFL requires that food and feed chain actors should be able to identify incoming and outgoing material flows. To this end systems and procedures are needed to store and retrieve information. In addition, food and feed need to be adequately labelled to facilitate its traceability. Traceability needs to be ensured at all stages of the food production chain, including third countries.

Traceability is currently mainly focused on food safety for human health. An international benchmark study [8] showed that performance levels are quite different between chains as well as within chains. For example, costs of traceability are unclear as is added value. Moreover, consumer do not seem to be willing to pay more for higher traceability. Full traceability can only be achieved by means of suitable infrastructures which are costly to install. Legislation and quality management systems for traceability do not focus on the complete supply chain yet. As indicated before, they are mainly related to incoming and outgoing flows, not on how companies need to manage traceability within the company. There is a lack of standardisation which hampers smooth exchange of information throughout the chain. In addition, food supply chain actors have different characteristics making standardisation very difficult.

Full traceability has only been achieved in food chains with large vertical integration and coordination. Information systems and internet technologies play a key role in these chains [9]. Such chains are mostly specialty chains with high added value or small food chains. Chain actors in these chains are willing to invest in traceability to create competitive advantage. In large supply chains like commodity pork chains as the fresh pork meat chain in the Netherlands with mainly market relationships [10] and a large number of chain actors full traceability does not yet exist.

Administrative costs of traceability are very high. Moreover, large investments are needed. Since the EU does not yet require electronically readable devices, a lot of paper work is still needed or extensive manual data entry. Electronic devices are used, but not at a large scale yet. Transponders, for example, are still fairly expensive and mostly used for sows only. Efforts are going on to develop low-cost and robust electronic devices that can pass the slaughterhouse stage. Some specialty food chains have already solved traceability across the slaughterhouse stage, like the National Livestock Identification System (NLIS) in Australia [11]. For the large intensive pork supply chains this has not yet been achieved.

With respect to animal health and safety and the environment, also only small food chains address these issues currently. For example, organic chains use environmental and animal welfare information to market their food products to consumers. Again, in large food chains this does not yet happen, although most chain actors favour attention to animal health and environmental impact [1].

However, incentives to install traceability for animal health and environment are still lacking and division of responsibilities between chain actors, government and other stakeholders are not yet clear. Transparency could provide competitive advantage when consumers are willing to pay more if they know quality and provenance of food products [8].

In the next section, we will explore in more depth the issue of environmental impact in the current food chain.

3 Environmental impact of the food chain

To limit environmental impact of food chains, not only individual supply chain actors need to foster environmental care, but also concerted action is needed on the supply chain level. Companies need to change their organisation and social system to improve environmental care within and across their borders. There is value in increased environmental attention, but barriers towards change in this direction also exist.

3.1 Barriers and opportunities

Environmental impact can be regarded as a product quality dimension comparable to technical product characteristics, safety, price and social circumstances under which products are made. Not all quality dimensions are clearly visible to actors in the supply chain. Dimensions like product composition, hygiene and safety can be measured with more ease than environmental or social impact of products and production processes. Although technical (ISO)-systems, with which environmental impacts and technical product quality can be monitored and controlled, are complementary, the actual impact on profitability can be adverse. Creating transparency in a supply chain is costly, while the effects on turnover are questionable. This will obstruct the implementation of systems that make the environmental impact of products and processes along a supply chain visible. Other barriers exist in addition that obstruct environmental monitoring and control in food supply chains.

First of all “impact” is a negative product characteristic, while the environmental benefits of pro-active companies is at best neutral. It is questionable whether customers are willing to pay for products from an improved sustainable production chains. The question in this respect is not whether “pollution prevention pays” but whether it pays more than competing business projects. Investments to improve environmental care need to be balanced against other investments. Evaluating investments for tacit intangible assets is difficult. The reason for this is the connectedness of care systems like quality, social responsiveness and environment with the existing business structure. Creating transparency could increase the value of a company as a whole of which tacit assets are a part.

Second, if environmental impact is caused along the supply chain and measures are taken at one stage, it could easily be that advantages of environmental pro-activeness are harvested in another. Burning of packaging and waste to generate energy could be beneficial for the energy plant, while other stages involved in the supply chain carry the transport costs. Vertical redistribution of costs and benefits is necessary to engage supply chain actors in costly environmental care measures. Transparency of costs and benefits that makes such redistribution possible is lacking to a large extent.

Third, social and technical rigidities can obstruct creating transparency to improve environmental pro-activeness. Especially companies in food supply chains are embedded in social and legal ties which have a negative impact on innovativeness. In terms of Miles et al. [12], food companies are – in general - more defender than prospector. Innovations in food are mainly process-oriented, picking “low hanging fruit” of technical adjustments to comply with changing regulations or covenants. The companies strive for cost-effectiveness in an extremely competitive market with pressure on margins. Under such circumstances improved image of environmental pro-activeness does not pay. What pays is taking internal measures to reduce costs.

3.2 Current situation

In 2002, data were collected among in total 2620 Dutch agri-food companies with five employees or more on their attitude towards internal and external (horizontal and vertical) environmental care measures. After elimination of incomplete and invalid response, the analyzed sample consisted of 492 companies (19.2%). The results showed remarkable effects of the company chain and network environment on environmental management system development. We made a distinction between internally (I-EMS) and externally oriented measures (E-EMS). Relevant correlations between these dependent variables and stakeholder influences are given in table 1.

Table 1. Environmental measures and stakeholder influences (Source: [13])

	Mean	SD	GO	CO	SUP	CLI	ENV	HAB	MED	INT
Government (GO)	3.62	1.07	- .							
Competitors (CO)	1.68	.805	.28	-						
Suppliers (SUP)	2.00	.981	.26	.53	-					
Clients (CLI)	2.47	1.21	.34	.57	.59	-				
Environmental organizations (ENV)	2.02	1.09	.28	.33	.30	.26	-			
Inhabitants (HAB)	2.72	1.22	.34	.33	.30	.34	.31	-		
Intermediaries (MED)*	1.59	1.31	.24	.27	.25	.27	.15	.18	-	
I-EMS development (INT)*	1.05	1.37	.22	.30	.17	.19	.19	.27	.30	-
E-EMS development (EXT)*	.90	1.04	.08-/	.30	.32	.32	.21	.14	.20	.26

*Correlations significant at p<0.01 level (two-tailed), except -/ (at p=.08). * at a 0-5 point scale.*

Table 1 shows, that externally oriented measures (such as a tracking & tracing system for environmental impact, data collection in the supply chain, etc.) were only meagrely developed (0.90 on a 0-5 point scale, SD = 1.04), while internally oriented measures were deployed to a larger extent (1.05, SD = 1.37). Government appears to have the biggest impact on the environmental policy of companies, next to inhabitants and clients. To create transparency, therefore, of environmental measures in a supply chain, governmental influence is needed. At the same time negative effects on administrative burdens need to be limited.

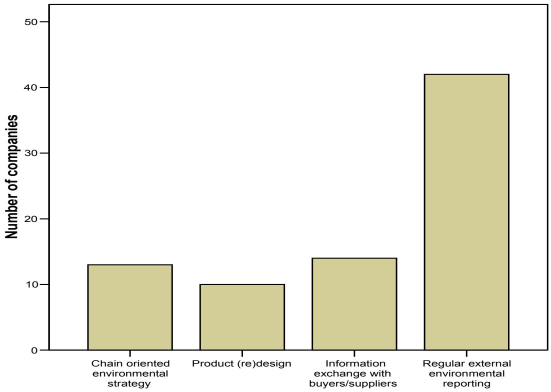


Figure 1. Number of companies (N=99) that implement the various different elements of product-oriented environmental management (Source: [14]).

Follow-up research by Haverkamp and Bremmers [14] in 2005 showed that especially large companies will use environmental product dimensions as a marketing device for improving their image. The research confirmed that radical, product-oriented innovations occurred only scarcely within the Dutch agri-food sector. Also, information exchange with buyers and suppliers (i.e., creating transparency) was only performed to a limited level. Figure 1 addresses product-oriented environmental management issues.

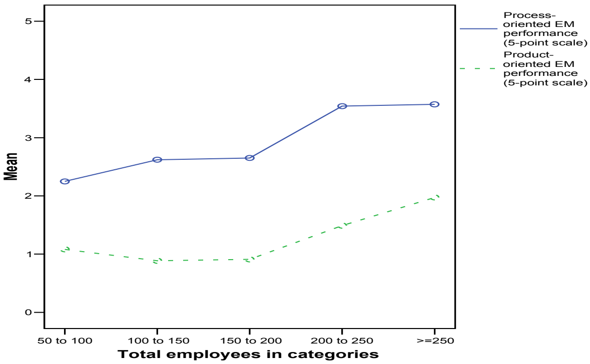


Figure 2. The relationship between size and environmental concerns (Source: [14])

The mean score on process-oriented environmental measures was 2.78 (SD = 1.57) and on product-oriented measures on average 1.24 (SD = 1.08) (with 5 being the maximum score). Interestingly, process- and product-oriented measures appear to be heavily dependent on the size of the companies (see figure 2).

4 Challenges to be addressed

Creating transparency along the supply chain is necessary, considering the above. Despite its intangibility, transparency can be beneficial, because:

- It can increase the visibility and comparability of companies in an industrial sector, like food and beverage;
- It can reduce risks by communicating the impact of environmental measures on costs and benefits;
- Information exchange can improve flexibility and cooperativeness, since shirking and opportunistic behavior within a supply chain is suppressed;
- Technical rigidities can be overcome by improved information exchange on technological innovation opportunities, and sharing of costs and benefits.

Results above show that transparency and environmental pro-activeness are related to organizational size. The question remains which measures can be taken to increase size and, hence, favour attention to improved environmental behaviour. Suggestion are:

- Copy existing supply chain oriented safety systems, like food safety tracking and tracing, for environmental tracking and tracing. Attributes to that are necessary to track causes of environmental impact in a supply chain are similar to those necessary for tracing food safety issues.
- By concluding upon public-private agreements to cooperate in technological innovation while sharing costs and benefits. Achievement levels can be formulated for chain actors. An example is the multi-year agreement made in the Dutch public-private agreement on energy reduction [15].
- By forming vertical environmental cooperatives, which provides a governance structure for addressing environmental goals in a cooperative way.
- By sharing and distributing knowledge and expertise via branch organizations that mediate between public demands and private concerns.

Increasing size, especially of animal farms, though, may be in conflict with citizens' adverse inclination to accept large farms in their environment, like is visible in social pressure groups.

The role of the government is important as shown above and in various other research results (see e.g., [1;16]). This observation may be in contrast with the current emphasis of EU and local governments to stress self-regulation capabilities of chain actors.

Implementation of environmental care systems may dramatically increase the already large administrative pressure on companies. Ways to limit administrative burdens might be certification and branding [8;9]. Branding, in addition, is a good way to make food product quality and provenance transparent for consumers and to achieve competitive advantage.

6 References

[1] KALFAGIANNI, Agni. Transparency in the food chain. Policies and politics, PhD Thesis, University of Twente, Enschede, The Netherlands, 2006.

[2] HOFSTEDE, Gert Jan. Hide or confide? The dilemma of transparency, in: Hide or confide? The dilemma of transparency, HOFSTEDE, Gert Jan; SCHEPERS, Hans; SPAANS-DIJKSTRA, Lia; TRIENEKENS, Jacques; BEULENS, Adrie (Eds.), Reed Business Information, Den Haag, The Netherlands, 2004.

[3] HOFSTEDE, Gert Jan. Trust and transparency in netchains: a contradiction? In: Supply chain management: issues in the new era of collaboration and competition, Wang, William Y.C.; Heng, Michael S.H.; CHAU, Patrick Y.K. (Eds.), Idea Group Publishing, Hershey, USA, 2007.

[4] TRIENEKENS, Jacques; PETERSEN Brigitte; WOGNUM Nel; BRINKMANN Detert (Eds.). European pork chains-Diversity and quality challenges in consumer-oriented production and distribution, Wageningen Academic Publishers, Wageningen, The Netherlands, 2009.

[5] KLEEF, van E.; FREWER, L.J.; CHRYSSOCHOIDIS, G.M.; HOUGHTON, J.R.; KORZEN-BOHR, S.; KRYSTALLIS, T.; LASSEN, J.; PFENNING, U.; ROWE, G. Perceptions of food risk management among key stakeholders. Results from a cross-European study, Appetite, Vol. 47, No. 1, pp. 46-63, 2006.

[6] PLAGGENHOEF, van W. Integration and Self Regulation of Quality Management in Dutch Agri-Food Supply Chains, Doctoral Thesis, Wageningen University, The Netherlands, 2007.

[7] TRIENEKENS, Jacques; WOGNUM Nel; NIJHOFF-SAVVAKI Rannia; WEVER Mark. Paper presented at the IAMA 2008 Symposium, Monterey, USA, 2008.

[8] VORST, van der Jack G.A.J.; BEURDEN van Jan; FOLKERTS Henk. Tracking and tracing of food products – An international benchmark study in food supply chains, Report Vrt Mwn 2003010 - Management Summary, Rijnconsult, Velp, The Netherlands, 2003.

[9] BUHR, Brian L. Traceability and information technology in the meat supply chain: implications for firm organization and market structure, Journal of food distribution research, Vol. 34, No. 3, pp. 13-26, 2003.

[10] WEVER, Mark; WOGNUM Nel; TRIENEKENS Jacques; OMTA Onno. Alignment between chain quality management and chain governance in EU pork chains: a transaction cost economics perspective, in: Proceeding of Congreso Mundial del Jamón, Spain, 2009.

[11] http://www.dpi.qld.gov.au/cps/rde/dpi/hs.xsl/4790_6000_ENA_HTML.htm

[12] MILES, R. E.; Snow, C.C.; MEYER E.D.; COLEMAN, H.J. Jr. Organizational Strategy, Structure, and Process, The Academy of Management Review, Vol. 3, No. 3, pp. 546-562, 1978.

[13] BREMMERS H.J.; KEMP, R.; HAVERKAMP, D.J.; OMTA, S.W.F. Do stakeholder groups influence environmental management system development in the Dutch agri-food sector? Business Strategy and the Environment, Vol. 16, No. 3, pp. 214-231, 2007.

[14] HAVERKAMP, Derk-Jan; BREMMERS, Harry. Environmental management in the Dutch agri-food industry - Preliminary results of a survey conducted in 2005, Wageningen University-Business Administration Group, WUR, 2005.

[15] MJA-2,. http://www.senternovem.nl/mmfiles/ezorgalg_energiek_tcm24-123317.pdf, Multi-year agreement – SenterNovem, 2003 (in Dutch)

[16] PORTER, M.E.; LINDE, van der C. Toward a new conception of the environment-competitiveness relationship, The journal of economic perspectives, Vol. 9, No. 4, pp. 97-118, 1995.

Multi-Product Min-Cost Recycling Network Flow Problem

Chiao-Lin Deng[1], Chun-Mao Shao

Department of Industrial Management
National Taiwan Unerversity of Science and Technology.

Abstract: Recently, the growth of environmental concerns has given increasing attention to the concept of "recycle". An efficient product recycling process is required. For this reason, this research establishes a minimum cost recycling network configuration model. This general model is based on the logistic transportation model, while being applied in a reverse fashion. The recycling system network is constructed with four process layers, including a: collection site, pre-processing site, processing site, and termination site (for reuse, resell, or disposal). Each process layer (or site) has a unique and distinct function. The EU Waste Electrical and Electronic Equipment Directive (WEEE) standard is used to determine the recycling rate for the product. The proposed model attempts to minimize the total cost, which consists of transportation cost, operating cost, and final disposal cost, as well as the sale revenue of reclaimed materials.

Keywords: Recycling, Reverse Logistics, WEEE, Minimum Cost Model.

1 Introduction

In recent years, several issues such as global warming, resource shortage, energy conservation, and resource recovery have become a central concern. An increasing number of countries have made concrete actions (ex. enacting legislations, advocating energy savings, and creating recycling systems) in order to reduce the impact of these environmental issues such as acid rain, air pollution, global warming, etc.

One of the main sources of environmental pollution in this industrial era stems from the extensive use of electronic components and equipment. Rapid technical innovation and market changes have shortened product life cycles. Due to the fact that consumers are using and replacing older products at a higher rate, this is resulting in increased volumes of waste that may be a factor leading to potentially severe environmental hazards. In order to contain this trend, the European Union (EU) enacted two specific directives: the Waste Electrical and Electronic Equipment Directive (WEEE Directive) [1] and the Restriction of Hazardous Substances Directive (RoHS Directive) [2].

[1] Corresponding Author Email : ciao1119@gmail.com

The aim of the WEEE Directive is to reduce the amount of electrical waste going to landfills, and to encourage the recycling and reuse of such hazardous waste. The WEEE Directive distinguishes ten different categories of equipment, and establishes a minimum reuse and recycling rate as well as a minimum recovery rate by which an *average weight per appliance* requirement should be reached for different equipment categories [1]. In addition, the RoHS Directive restricts the use of six hazardous materials in the manufacture of various types of electronic and electrical equipment. These two directives successfully established the recycling requirements for product design and end-of-life processing, as a key issue and focal concern.

Typical supply chains that researchers considered are the forward supply chain where the strategy is to distribute the products from manufacturing plants or factories to customers. The capacitated plant location problem (CPLP) is one of the most well-researched problems in this area [4-8]. On the other hand, the reverse supply chain flows backward from the customers upward all the way to manufacturers and was established to meet the needs of reducing, substituting, reusing, and recycling of disposed products. The reverse logistics system is concerned with activities associated with the handling and management of recycled materials, components, products, equipment, information or even recovery of entire technical systems. From a logistics perspective, recovery activities increase additional flow of goods from the consumers back to producers. The management of this opposite flow forwarding the supply chain is the concern of the recently evolved field of “reverse logistics” [9]. The operations of reverse logistics are receiving increased attention as rampant solid waste pollution, frequent energy shortages, and serious materials scarcity are recognized as realities of modern age.

Fleischmann *et al.* emphasized that reverse logistics is not necessarily a symmetric representation of forward logistics [3]. Hence, the objective of our research is to establish a minimum cost recycling network flow model so that the recycling can be performed cost effectively. This model is based on the logistic transportation model but in a reverse fashion, which is a variation of multi-commodity min-cost flow (MMCF) problem. With a given logistic network configuration, the model finds the optimal units that should be transferred between facilities until reaching the destination in order to minimize the total cost of the network, while considering each facility’s capacity constraint. Notice that the processing efficiency at each site may be different, resulting different amount of regenerated materials.

In general, there are many differences between forward logistics supply chains and reverse logistics supply chains; recycling of products here cannot be handled in the same manner as in a regular channel. Therefore, an ideal reverse logistics system network design for recycling products is required.

2 Literature Review

In recent years, terms like recycling, reuse, resource reduction, environmental manufacturing responsibility, and green products are increasingly becoming familiar to all of us. Since the mid-nineties, especially in Europe, this was accompanied by legal enforcement of product and material recovery or proper disposal. In the USA, landfill tolls are more expensive and restrictions on cross-state transport of waste are rising substantially. More recently, practical examples from the remanufacturing of mobile phones have validated the profitability of recovery activities and its value creation as oppose to being merely a cost-consuming environmental conscious activity [10]. In addition, competition, marketing, and strategic arguments have pushed companies into providing more generous policies as incentives to customers in order to encourage exchange and recycle of used products. Recovery of used products has become a field of growing importance. The characterization and impact of product recovery networks were investigated [11-14]. Others have proposed quantitative models for different industries that take into account those changes in the logistics environment. For example, Louwers *et al.* considered the design of a recycling network for carpet waste [15]; Barros *et al.* reported on a case study addressing the design of a logistics network for recycling sand [16]; and Krikke *et al.* presented a model for recycling durable consumer products [17].

Sarkis *et al*. [18] argued that three characteristics can be used to differentiate a reverse logistics system from a traditional supply chain system. First, most logistics systems are not equipped to handle recycled product movement in a reverse channel. Second, reverse distribution costs may be higher than moving the original product from the manufacturing site to the consumer. Third, returned goods often cannot be transported, stored, or handled in the same manner as in a regular channel. Kokkinaki *et al*. provided three reasons to design and to develop reverse logistics networks diverging considerably from forward logistic [19]. First, the supply chain composition and structure is changed due to new parties involved and new roles being resumed by existing parties. For example, the customer is now also the "supplier" of disposed of goods. Second, planning and control becomes more complicated because the high supply uncertainty both in quality and quantity and the legislation constraints. Third, producer's responsibilities extend over the whole life cycle of their products.

3 Problem Definition

The main activities in a reverse logistics network are collection, inspection or separation, reuse, remanufacturing, recycling, refurbishing, re-distribution and disposal [19]. But for the recycling industry, it more focuses on recycling and reuse processes, normally those recycling company didn't handle the product remanufacturing and refurbishing processes. Therefore, we focus on the product recycling process, which the main recycling process activities include collection, separation, extraction, transportation and disposal.

This paper addresses the problem of efficiently transporting multiple types of recycled products from multiple sources and in the process satisfying the demand on the recycled items at a number of destinations. The recycling system problem can be formulated as follows:

Given: (a) Locations of pre-processing sites and extraction sites
(b) Capacities of these facilities
(c) Process efficiency of these extraction facilities.
(d) Cost structures (transporting, processing, operating, and disposal costs),
(e) Location of the reuse/resell sites
(f) Availability and demand of reusable material for each reuse/resell site

Find: Which facility should be used (opened) and how the resource waste should be delivered so that the total cost of the system is minimized.

In the next section, we address a recycled goods processing sequence first, than on the basis of the recycling sequence develop a recycling system network Model.

3.1 Recycling Sequence

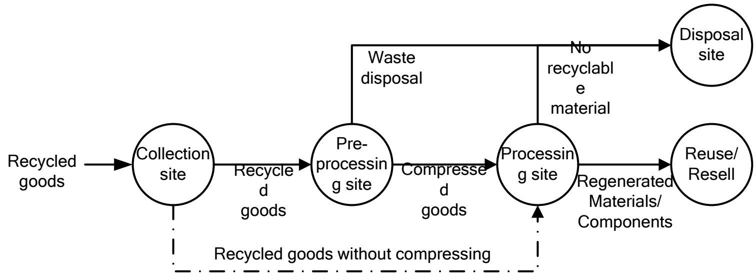


Figure 1. Recycled Goods Processing Sequence

Figure 1 shows the recycled goods processing sequence. There are four layers of processing sequences from left to right. Each layer has its own function which can be described as follows:

- **First Layer (Collection Site):** This is usually a garbage truck, waste yard, or local recycling group. All resource waste is received from customers at collection sites and passed to pre-processing site or directly sent to processing site.
- **Second Layer (Pre-processing Site):** This site receives the collected waste from collection sites. Its main function is to reduce product volume in order to increase truckload utility by manual disassembly of recycled products first, compression of recycled material, and disposal of the waste. As the volume of recycled products becomes smaller, the transportation utility increases. The pre-processing site can also collect large quantities of waste, to ensure that transportation vehicles are fully loaded.
- **Third Layer (Processing Site):** A processing site receives the sorted waste from collection sites and pre-processing sites. Some components could be reuse directly after the products have been disassembled; rest of the waste will be transform into regenerated materials. If the recycled product is defined according to the ten categories of the WEEE Directive, it is required that the recycling rate must be higher than the regulation. All processing sites recycling rates will comply with WEEE regulation.

- **Fourth Layer (Termination Site):** This layer is for the reuse or resell of the regenerated material. For a non-recyclable material, it will be disposed properly.

We summarize the sequences as following: an end-of-life product will be sent to an initial collection site which is the closest point to the customer. After a period of time, all products from the collection site will be transported to a pre-processing site for volume reduction. Notice that the recycled products from the collection site could be sent to processing site directly if the total cost is lower than passing through the pre-processing site. Next, the compressed products will be transported to a processing site to extract raw materials from those recycled products. Finally, the extracted raw materials will be sold to manufacturers or material suppliers for reuse or reselling. The non-recyclable waste will be properly disposed.

3.2 Recycling System Network Model

Based on the recycled goods processing sequence mentioned in the previous section, we construct a recycling system network for solving recycling management problems.

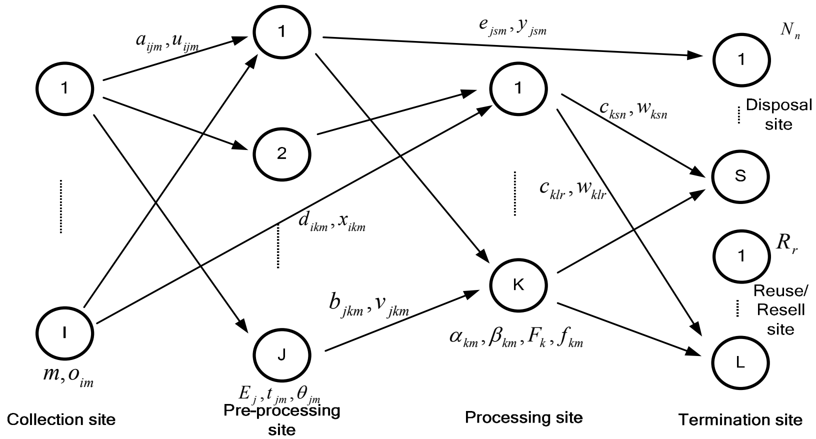


Figure 2. Recycling System Network

There are four layers of nodes in this recycling system network: collection site, pre-processing site, processing site, and termination site. The arcs represent the flow of resource waste, sorted waste and regenerated material, and hence there are arcs between each adjacent level. In our problem, *m* kinds of products are to be transferred from a set of source nodes I to a set of destination nodes L (or S). The *m* represents the index of product type. o_{im} is unit of product *m* be collected at collection site *i*. a_{ijm} is the cost of transporting product *m* from collection site *i* to pre-processing site *j*. b_{jkm} is the unit cost of transporting product *m* from pre-processing site to processing site *k*, c_{klr} is the unit cost of transporting regenerated recyclable material or component *r* from processing site *k* to reuse/resell site *l*, c_{ksn}

is the unit cost of transporting regenerated no recyclable material n from processing site k to disposal site s, d_{ikm} is the unit cost of transporting product m from collection site to processing site directly, e_{jsm} is the unit cost of transporting waste from pre-processing site to disposal site s.

u_{ijm} is the unit of product m at collection site i transported to pre-processing site j. v_{jkm} is the unit of product m at pre-processing site j transported to processing site k. w_{ksn} is the unit of no recyclable material n, w_{klr} is the unit of recyclable material or component r. x_{ikm} is the unit of product m at collection site transported to processing site. y_{jsm} represents the unit of waste at pre-processing site transported to disposal site.

E_j is the fixed cost of pre-processing site j. t_{jm} is the unit operation cost of product m at pre-processing site j. θ_{jm} is the disposal rate of product m at pre-processing site j. F_k is the fixed cost of processing site k. f_{km} is the unit operation cost of product m at processing site k. At processing site, α_{km} is the material extracted rate, β_{km} is the product component reuse rate. If the recycled product which defines in the ten categories of WEEE Directive, $\alpha_{km}+\beta_{km}$ (reuse and recycling rate) must be higher than the regulation. N_n is the unit cost of disposing no recyclable material n. R_r is the revenue of recyclable material r per unit.

4 Model Formulation

In this section, we formulate the problem of minimum cost for the recycling system. Considering the transportation cost, capacity and operating cost of the collection site and recycling facility, recycling rate, the lowest recycling material demand quantity of manufacturing, and the profit of reuse those recycling material. This formulation is based on a single objective function of minimizing the cost of the recycling system.

To formulate a mathematical model for the recycling problem, we make the following assumptions:

1. All waste must enter the recycling system through collection sites.
2. The transporting vehicles among each node are the same.
3. The location of collection sites, pre-processing sites, processing sites and final sites are fixed.

The objective function minimizes total reverse logistics cost.

Total cost = transport cost + operation cost + disposal cost
- revenue of sell reclaimed materials and components

$$
\begin{aligned}
\text{Min } Z = & \sum_i \sum_j \sum_m a_{ijm} u_{ijm} + \sum_j \sum_k \sum_m b_{jkm} v_{jkm} + \sum_k \sum_l \sum_r c_{klr} w_{klr} + \sum_k \sum_s \sum_n c_{ksn} w_{ksn} + \\
& \sum_i \sum_k \sum_m d_{ikm} x_{ikm} + \sum_j \sum_s \sum_m e_{jsm} y_{jsm} + \sum_j \sum_m P_j (E_j + u_{ijm} t_{jm}) + \\
& \sum_k \sum_m Q_k (F_k + v_{jkm} f_{km}) + \sum_s \sum_n w_{ksn} N_n - \sum_l \sum_r w_{klr} R_r
\end{aligned}
$$

Subject to:

(1) Flow balance equations:

$$\sum_j u_{ijm} + \sum_k x_{ikm} = o_{im} \text{ , for all } i, m$$

$$\sum_i u_{ijm}\theta_{jm} = \sum_s y_{jsm}, \text{ for all } j,\ m$$

$$\sum_i u_{ijm} = \sum_k v_{jkm} + \sum_s y_{jsm}, \text{ for all } j,\ m$$

$$\sum_i \sum_m x_{ikm}(1-\alpha_{km}) + \sum_j \sum_m v_{jkm}(1-\alpha_{km}) = \sum_s \sum_n w_{ksn}, \text{ for all } k$$

$$\sum_i \sum_m x_{ikm}(\alpha_{km}+\beta_{km}) + \sum_j \sum_m v_{jkm}(\alpha_{km}+\beta_{km}) = \sum_l \sum_r w_{klr}, \text{ for all } k$$

(2) Demand completion constraint:

$$\sum_k w_{klr} \geq D_{lr}, \text{ for all } l,\ r$$

(3) Capacity constraints:

$$\sum_i u_{ijm} \leq \overline{G}_{jm} P_j, \text{ for all } j, m$$

$$\sum_j v_{jkm} + \sum_i x_{ikm} \leq \overline{H}_{km} Q_k, \text{ for all } k, m$$

$$\sum_i u_{ijm} \geq \underline{G}_{jm} P_j, \text{ for all } j, m$$

$$\sum_j v_{jkm} + \sum_i x_{ikm} \geq \underline{H}_{km} Q_k, \text{ for all } k, m$$

$$\underline{P} \leq \sum_j P_j \leq \overline{P}$$

$$\underline{Q} \leq \sum_k Q_k \leq \overline{Q}$$

(4) Binary restrictions:

$$P_j \in \{0,1\}$$

$$Q_k \in \{0,1\}$$

(5) Non-negative decision variables:

$$u_{ijm}, v_{jkm}, w_{klm}, x_{ikm}, y_{jsm} \geq 0 \text{ for all } i, j, k, l, s, m$$

Where D_{lr} represents the lowest demand of regenerated material r for manufacturer l; $\underline{P}$ and $\overline{P}$ represents the minimum and maximum numbers of pre-processing sites to open and operate; $\underline{Q}$ and $\overline{Q}$ represents the minimum and maximum numbers of processing sites to open and operate; $\underline{G}_{jm}$ and $\overline{G}_{jm}$ is the minimum and maximum storage and processing capacity of handling product m at pre-processing site j; $\underline{H}_{km}$ and $\overline{H}_{km}$ is the minimum and maximum storage and processing capacity of handling product m at processing site k.

In this research, we use Matlab to build and solve this multi-commodity network flow problem. At first, we design a small recycling problem in which the pre-processing site and processing site open and operate, and we also design the value of all parameters such as transportation cost, operating cost, the capacity of pre-processing site and processing site in the mathematical model. Then, Matlab software is used to solve the problem. Problem solution will find a feasible solution which is the minimum cost for the model. This provides helpful information for reverse logistic management to manage the logistic system. Moreover, we found that compression is an important factor in recycling reverse supply chain; products with a small volume could bring down the unit transportation cost which refers to the transportation path schedule.

5 Conclusion

In this research, we have established a recycling network configuration model; and determined cost-optimal flows of goods through the network under the given capacity constraints. Also, the results of the study indicate that waste compression at pre-processing sites is an important function for the whole recycling system. This may provide an initial idea of product design for recycling.

Last but not least, supply uncertainty both in quantity and quality appears to be a major distinction between reverse logistic network and the traditional forward logistic network. In this research, we didn't consider the uncertainty of the recycled product quality and quantity. This may be a reason for a more complex network structure, which more study in this field is needed.

6 References

[1] Directive 2002/96/EC of the European parliament and of the council on waste electrical and electronic equipment (2003). Official Journal of the European Union, 1 (27), pp 24-37. European Parliament/Council of the European Union.

[2] Directive 2002/95/EC of the European Parliament and of the Council on the restriction of the use of certain hazardous substances in electrical and electronic equipment (2003). Official Journal of the European Union, 1 (27), pp 19-23. European Parliament/Council of the European Union.

[3] Fleischmann, M., Bloemhof-Ruwaard, J.M., Dekker, R., van der Laan, E., van Nunen, J.A.E.E., van Wassenhove, L.N., (1997). Quantitative models for reverse logistics: A review. European Journal of Operational Research 103, 1-17.

[4] Davis, P.S., Ray, T.L., (1969). A branch-and-bound algorithm for the capacitated facilities location problem. Naval Research Logistics 16, 331-344.

[5] Geoffrion, A.M., Graves, G.W., (1974). Multi-commodity distribution system design by benders decomposition. Management Science 20, 822–844.

[6] Guignard,M., Spielberg, K., (1979). A direct dual method for the mixed plant location problem. Mathematical Programming 17, 198-228.

[7] Barcelo, J., Casanovas, J., (1984). A heuristic Lagrangian algorithm for the capacitated plant location problem. European Journal of Operational Research 15, 212-226.

[8] Tragantalerngsak, S., Holt, J., Ronnqvist, M., (1997). Lagrangian heuristics for the two-echelon, single-source, capacitated facility location problem. European Journal of Operational Research 102, 611-625.

[9] Stock JR. Development and implementation of reverse logistics programs. Oak Brook, IL: Council of Logistics Management, 1998.

[10] Guide Jr. V.D.R. and L.N. Van Wassenhove, (2001), Business Aspects of Closed Loop Supply Chains, Carnegie Mellon University Press, forthcoming.

[11] Bloemhof-Ruwaard, J.M., Salomon, M., van Wassenhove, L.N., (1994). On the coordination of product and by-product flow in two-level distribution networks:

Model formulations and solution procedures. European Journal of Operational Research 79, 325-339.
[12] Fleischmann, M., Krikke, H.R., Dekker, R., Flapper, S.D.P., (2000a). A characterisation of logistics networks for product recovery. Omega 28, 653-666.
[13] Fleischmann, M., Beullens, P., Bloemhof-Ruwaard, J.M., van Wassenhove, L.N., (2000b). The impact of product recovery on logistics network design. INSEAD working paper 2000/ 33/TM/CIMSO.
[14] Krikke, H.R., van Harten, A., Schuur, P.C., (1999). Business case OCE: Reverse logistics network re-design for copiers. OR Spektrum 21, 381-409.
[15] Louwers D, Kip BJ, Peters E, Souren F, Flapper SDP. Afacility location allocation model for re-using carpet materials. Computers and Industrial Engineering 1999;36(4),1-15.
[16] Barros AI, Dekker R, Scholten V. A two-level net work for recycling sand: acase study. European Journal of Operation Research 1998; 110:199-214.
[17] Krikke HR, Van HartenA, Schuur PC. Business case oce:reverse logistic network re-design for copiers. OR Spektrum 1999;21(3):381-490.
[18] Sarkis, J., Darnall, N., Nehman, G., Priest, J., (1995). The role of supply chain management within the industrial ecosystem. In: Proceedings of the 1995 IEEE International Symposium on Electronics and the Environment, Orlando, FL, 229-234.
[19] Kokkinaki A.I., Dekker R., Koster M.B.M. de, Pappis C., Verbeke W., (2002). E-business models for reverse logistics: Contribution and Challenges. Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC’02), 470- 476.

Service Solutions

Applying RFM Model and K-Means Method in Customer Value Analysis of an Outfitter

Hsin-Hung Wu[a,1], En-Chi Chang[b] and Chiao-Fang Lo[c]

[a] Department of Business Administration – NCUE, Taiwan.
[b] Manchester Business School – UK.
[c] Graduate Institute of Marketing and Logistics Management – NCUE, Taiwan.

Abstract. This case study applies RFM model and K-means method in the value analysis of the customer database of an outfitter in Taipei, Taiwan. By considering gender, birth date, shopping frequency, and the total spending, six clusters have been found among 675 member customers from the company's database. In addition to the clustering analysis, different promotion strategies for the members of different clusters are provided. The analyses show that Clusters 5 and 6 are the two most important groups that the company has to devote resources into. Moreover, the company might ration resources for the customers in Clusters 1 and 2 because they do not contribute enough values to the outfitter.

Keywords. RFM model, K-means method, customer value analysis, data mining.

1 Introduction

According to the Commerce Industrial Services Portal of the Ministry of Economic Affairs, R.O.C., there were 52 outfitters, most of which were located in Northern and Central Taiwan areas at the end of 2006. The large number of outfitters brought competition, which led to the decrease of profits. Thus, the outfitter must identify and retain customers of high value and profit potentials in order to survive competition and sustain profits. Achieving the aforementioned goals will require the outfitter to customize marketing strategies and fulfill the needs of different customers and also to allocate resources effectively and efficiently, based on a well-managed customer database.

As the transactions become larger in size, managing customer database is not an easy task. The better approach might be to divide all customers into appropriate number of clusters (groups) based on some similarities by data mining techniques, particularly the clustering techniques. The values of different groups can then be estimated and evaluated to provide useful decisional information for management to utilize resources rationally. Besides, recency, frequency, and monetary (RFM)

[1] Corresponding Author : No. 2 Shida Road, Changhua City, Changhua, Taiwan, R.O.C.; Tel: +886-4-723-2105-7412; E-mail: hhwu@cc.ncue.edu.tw

model is a simple but effective tool that can be applied for market segmentation by examining when (recency), how often (frequency), and the money spent (monetary) in a particular item or service [11].

A real case of an outfitter in Taipei, Taiwan will be analyzed based on RFM model along with the K-means method, one of the most popular clustering techniques, to manage customer data and help draft promotion strategies. The transaction data consist of 675 customers who shopped at the outfitter's store from April 2004 to March 2006. The profile for each customer includes the membership number, gender, birth date, shopping frequency, and the total spending at the store.

This paper is organized as follows. Section 2 reviews RFM model and K-means method briefly. The case study of the outfitter based on RFM model and K-means method is analyzed in Section 3 amd marketing implications are provided. Finally, conclusions are drawn in Section 4.

2 Review of RFM Model and K-Means Method

RFM model and K-means method will be reviewed briefly in this section.

2.1 RFM Model

Recency, frequency, and monetary is a simply but effective tool that can be applied for market segmentation [3, 5]. Hughes [3] described that RFM analysis can segment the customer base and optimize the purchase response rates of the marketing efforts. RFM analysis improves the market segmentation by examining when (recency), how often (frequency), and the money spent (monetary) in a particular item or service [11]. Yang [11] has summarized that the more recent, more frequent, and high money spent customers would be much more likely to respond to the future promotions. Therefore, RFM analysis has been widely used by direct marketers.

The philosophy of RFM analysis is that customers are to be arrayed in terms of recent purchase, frequent purchases, and average purchase amount [3, 6, 13]. For recency, the technique is to sort the purchase date data in descending order, and then assign numerical values of 1 to 5 to the data set. The original data set with purchase date is transformed into a value of 1 to 5 depending upon the most recent purchase date. Therefore, the value of 5 is assign to the top 20% of the data set in terms of the most recent purchase date. The value of 4 is given to the next 20% of the data set and so on. For frequency, the technique is to sort the number of transactions in a certain period of time, such as the number of transactions per month, in descending order. The top 20% of the data set is given by the value of 5. The value of 4 is assigned to the next 20% of the data set and so on. The original data set with the number of transactions is transformed into a value of 1 to 5. Finally, for monetary, the technique is to sort the amount of the money spent such as average by month, year, or transaction in descending order. Assign the value of 5 to the top 20% of data set. Assign the value of 4 to the next 20% of the data and so on. The original data set with the amount of the money spent is transformed into a value 1 to 5.

In this study, the classifications of the data set in recency, frequency, and monetary are no longer in terms of 1 to 5. The recency is first reorganized in ascending order, and then the oldest and newest data are assigned values to be 0 and 458, respectively. The frequency is the total number of purchases between April 2004 and March 2006. The maximum and minimum values of frequency in this database are 47 and 1, respectively. For monetary, the average amount spending per visit is used.

2.2 K-Means Method

K-means method is a non-hierarchical method and a very popular approach for classification because of its simplicity of implementation and fast execution and has been widely used in market segmentation, pattern recognition, information retrieval, and so forth [2, 4, 10]. Yoon and Hwang [12] have pointed out that the typical K-means method uses Euclidean distance, and the formula of K-means method is as follows. The distance between two points X_r and X_s is given by the square root of the sum of the squared distance over each coordinate, where X_r = (x_{r1}, x_{r2}, x_{r3}, …, x_{ri}, …, x_{rn}) and X_s = (x_{s1}, x_{s2}, x_{s3}, …, x_{si}, …, x_{sn}), and each c_i in Equation (1) represents the weight. If the weights are normalized, then $\sum_{i=1}^{n} c_i = 1$ [1].

$$d(X_r, X_s) = \left[\sum_{i=1}^{n} c_i (x_{ri} - x_{si})^2 \right]^{1/2} . \tag{1}$$

K-means method consists of the following two major steps [8,9]. First, the assignment step where the instances are placed in the closest class. Second, the re-estimation step where the class centroids are recalculated from the instances assigned to the class.

One of the major problems of K-Means method is to select the best value of K [1]. Kuo *et al.* [4] have pointed out that non-hierarchical methods, such as K-means method, can have higher accuracy if the starting point and the number of clusters are provided. Punj and Steward [7] suggested a two-stage method by deploying Ward's minimum variance method to determine the number of clusters for K-means method. In contrast, Kuo *et al.* [4] have proposed a modified two-stage method by applying self-organizing feature maps (SOFM) to determine the number of clusters for K-means method. The reason is that self-organizing feature maps can converge very fast since it is a kind of learning algorithm that can continually update or reassign the observations to the closest cluster. Therefore, self-organizing feature maps is applied to determine the number of clusters in this study.

3 A Case Study

This section provides analyses based on RMF model model and K-means method, along with marketing implications based on the framework of Reinartz and Kumar [8]. The data set consists of 675 member customers who shopped between April 2004 and March 2006. The profile for each customer includes a membership number, gender, birth date, all transactions and the total spending at the store. Table 1 summarizes the total spending information for different genders, where the number of male customers is twice more than the number of female customers. Besides, the average amount spending per male customer is much larger than that for each female customer.

Table 1. Total Spending informtion ofor different genders

Gender	Size	Average Amount Spending	Standarfd Deviation	Minimum Total Spending	Maximum Total Spending
Male	462	10,230.20	14,501.50	80	126,934
Female	213	8,069.80	10,664.72	135	89,196
Sum	675	9,548.49	9,548.49	80	126,934

The birth date information of 675 customers can be classified into seven age groups, i.e., Group 1 aged 25 and below, Group 2 aged between 26 and 30, Group 3 aged 31 and 35, Group 4 aged between 36 and 40, Group 5 aged betweem 41 and 45, Group 6 aged betweem 46 and 50, and Group 7 aged 51 and above. Table 2 shows the distribution of these 675 customers in terms of age groups. The majority of customers fall in the age range of 26-40. It is worth to note that almost 60% of the customers are at the age of 35 or below.

Table 2. The distribution of seven age groups

Age Group	Number	Percentage
25 and below	72	10.67
26-30	164	24.30
31-35	163	24.15
36-40	121	17.93
41-45	69	10.22
46-50	48	7.11
51 and above	38	5.62
Sum	675	100.00

The information of recency, frequency, and monetary is stated below. The last purchase dates (recency) were arranged in an ascending order, where the farthest purchase date is set to 0, while the nearest purchase date is 458. The frequency is counted by the times of purchase from April 2004 to March 2006, where the maximum and minimum values of frequency are 47 and 1, respectively. Finally, monetary is defined as the average amount of spending per visit since each purchase varies significantly from 80 to 126,934 dollars. Moreover, the frequencies vary significantly from 1 to 47. Therefore, it would be more applicable to use the average amount of spending per visit instead of the total money spent in this time period by observing the patterns of 675 customers. The specific information regarding recency, frequency, and monetary is provided in Table 3.

Table 3. The distribution of seven age groups

	Size	Maximum	Minimum	Average	Standard Deviation
Recency	675	458	0	308.25	121.60
Frequency	675	47	1	2.48	3.55
Monetary	675	89,196	80	6,110.35	9,777.19

To determine the number of clusters for the K-means method, Kuo *et al.* [4] have proposed a modified two-stage method by deploying SOFM to partition the data set. When the number of clusters has been decided by SOFM, then assign this value to the K-means method for data clustering [9]. In this study, SOFM has partitioned the data into six clusters. Therefore, Table 4 summarizes the information of six clusters generated by K-means method. Cluster 3 has the smallest sample size, second largest average recency value, largest average frequency value, but smallest average amount of money. The average frequency value in Cluster 3 is 10 times more than the frequency of the second largest frequency group , i.e., Cluster 4. In contrast to Cluster 3, Cluster 6 has the smallest average frequency value among all clusters, but the largest average amount of money and the largest average recency value.. More importantly, the average amount of money in Cluster 6 is over 10 times more than the average. In addition, Cluster 5 has the second largest average amount of money, and the average value is 3 times larger than the average of six clusters. The specific information for each cluster regarding recency, frequency, and money is summarized in Tables 5, 6, and 7, respectively.

Table 4. The information of six clusters

Cluster	Size	Average Recency	Average Frequency	Average Amount of Money
1	108	90.85	1.32	2,292.36
2	194	249.96	1.89	1,978.60
3	4	431.75	35.75	232.75
4	275	397.19	3.28	4,437.83
5	88	411.56	1.27	20,631.92
6	6	432.33	1.17	76,024.00
Sum	675	308.25	2.48	6,110.39

In Table 5, the recency distributions in Clusters 3-6 are quite similar, all with the value of more than 300. In contrast, the recency in Cluster 1 concentrates on the range of 200 and less. This indicates that customers in Cluster 3, 4, 5 and 6, are recent customers. Particularly, Cluster hosts a large number of customers who just shopped recently. However, it is not clear whether these customers are new customers or old customers coming back recently.

Table 5. The distribution of recency for each cluster

Recency	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
100 and below	65	---	---	---	---	---
101 – 200	43	30	---	---	---	---
201 – 300	---	138	---	---	---	---
301 – 400	---	26	1	152	33	1
401 and above	---	---	3	123	55	5
Sum	108	194	4	275	88	6

The structure of frequency distribution for all clusters in Table 6 are similar, skewing towards 20 times or less, except for Cluster 3, which has two customers with the frequency value above 20. Thus Cluster 3 can be defined as customers with exteremely high loyalty. They 'fall in love' with the store and visit the store quite often.

Table 6. The distribution of frequency for each cluster

Frequency	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
10 and below	108	192	1	265	88	6
11 – 20	---	2	1	10	---	---
21 – 30	---	---	1	---	---	---
31 – 40	---	---	---	---	---	---
41 and above	---	---	1	---	---	---
Sum	108	194	4	275	88	6

Finally, the distributions of average amount of money per visit in Clusters 5 and 6 are different from those in Clusters 1-4. All customers in Cluster 1-4 spent 15,000 dollars or less during each visit but customers in Cluster 5 and 6 are high spenders in general. All six members in Cluster 6 show the extreme and spent more than 40,000 dollars in average during each visit. The average amount of money spent in Cluster 5 ranges from 10,000 dollars to over 40,000 dollars but most fall in the range between 10,001 dollars to 25,000 dollars.

Table 7. The distribution of average amount of money per visit for each cluster

Average Amount	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
5,000 and below	92	176	4	170	---	---
5,001 – 10,000	16	16	---	74	---	---
10,001 – 15,000	---	2	---	31	23	---
15,001 – 20,000	---	---	---	---	27	---
20,001 – 25,000	---	---	---	---	18	---
25,001 – 30,000	---	---	---	---	9	---
30,001 – 35,000	---	---	---	---	5	---
35,001 – 40,000	---	---	---	---	3	---
40,001 and above	---	---	---	---	3	6
Sum	108	194	4	275	88	6

The sysnthesis of the above tables shows that Cluster 3 has the highest average frequency, the second-highest average recency, but the lowest average amount of money. This group of customers brings the outfitter limited value due to their low share of wallet. The results in Table 5, Table 6 and Table 7 indicate that customers in Cluster 3 may be new customers who are curious about store offers. The outfitter may design promotions for up- and cross-selling to increase the average amount of money spent by these customers. For Cluster 4, the marketing strategies are similar to those designed for Cluster 3. The major objectives of marketing stratgeis for Cluster 4 will be bot to increase the shopping frequency and to persuade customers to spend more money. Thus the outfitter may use coupons or bundling to encourage customers to visit more often and spend more money.

Customers in Cluster 5 are relatively new customers but they spend more money than Cluster 3 and Cluster 5, The similar situation can be found in Clusrer 6. They are parallel to 'butterflies' termed by Reinartz and Kumar [8]. Thus they have higher profit potentials if the outfitter can turn these customers into long-term and loyal customers. For example, the outfitter can keep the customers informed by regularly mailing the catalogs and observe whether they can possibly become loyal customers. If they avoid building long term relationship, the outfitter should consider to limit its investment in this cluster. However, customers in Cluster 6 have the highest average amount of money and thus the outfitter may consider to devote more resources into this cluster. For instance, use promotion events such as outdoor recreational activities which introduce products sold in the store at the same time to allure these customers to build strong relationships with the outfitter. Besides, the outfitter may grant VIP status to customers to increase their shopping frequencies and to distinguish them from other customers.

Based on Table 4, the outfitter should not devote too many resources to the customers in Clusters 1 and Cluster 2 since their average recency scores are the lowest among six clusters. Moreover, their average frequencies and their average amount of money spent are far below the average. These customers have very low recency and frequency scores. They may already leave the outfitter. The outfitter may contact these customers occasionally, i.e., sending out annual sale message, to check whether these customers are likely to come back. However, The contribution of customers in Cluster 2 seem unclear since they have higher recency scores. Therefore, the outfitter may try to increase the number of contacts with these customers by sending mails about major promotional activities.

4 Conclusions

A case study is presented by applying RFM model and K-means method to analyze customer values of an outfitter. Six clusters have been generated. Based on the analysis, customers in Cluster 5 and Cluster 6 have higher potentials to become loyal customers in the long run but the outfitter should observe these two clusters of customers carefully. The objective is to establish strong relationship and eventually consolidate customer loyalty for high profitable long-term customers. For customers of Clusters 3 and 4, the outfitter's objectives will be to increase customers' shopping frequency and the amount they spend in the store; promotion mechanisms such as coupons, cross-selling and up-selling should be considered. Finally, the company might only use limited resources on customers in Clusters 1 and 2 or even consider to give up customers in Cluster 1, since they seem already out of touch. For customers in Cluster 2, the outfitter may consider to post promotion information and check whether these customers will come back.

5 References

[1] Buttrey SE, Karo C. Using k-nearest-neighbor classification in the leaves of a tree. Computational Statistics and Data Analysis 2002;40:27-37.

[2] Cheung Y-M. K-means: a new generalized k-means clustering algorithm. Pattern Recognition Letters 2003;24:2883-93.

[3] Hughes AM. Boosting response with RFM. Marketing Tools 1996;5:4-7.

[4] Kuo R-J, Ho LM, Hu CM. Integration of self-organizing feature map and K-means algorithm for market segmentation. Computers & Operations Research 2002; 29(11):1475-93.

[5] Marakas GM. Decision Support Systems in the 21st Century, Second Edition. Prentice Hall, Upper Saddle River, NJ, 2003.

[6] Newell F. Loyalty.com: Customer Relationship Management in the New Era of Internet Marketing. McGraw-Hill Professional, New York, 2000.

[7] Punj G, Steward DW. Cluster analysis in marketing research: review and suggestions for applications. Journal of Marketing Research 1983;20:134-48.

[8] Renartz W, Kumar V. The mismanagement of customer loyalty. Harvard Business Review 2002;7:86-94.

[9] Wu K-H, Wu H-H, Kou K-Y. Applying clustering techniques in traffic accidents analysis. Journal of Quality 2006;13(3):305-12.

[10] Wu H-H, Shieh J-I, Liao AYH, Lin S-Y. An application of the generalised K-means algorithm in decision-making processes. International Journal of Operational Research 2008;3(1/2):19-35.

[11] Yang AX. How to develop new approaches to RFM segmentation. Journal of Targeting, Measurement and Analysis for Marketing 2004;13(1):50-60.

[12] Yoon K, Hwang CL. Multiple Attribute Decision Making: An Introduction. Sage Publication, California, 1995.

[13] Zikmund WG, McLeod R, Gilbert FW. Customer Relationship Management: Integrating Marketing Strategy and Information Technology. John Wiley and Sons, 2003.

An Investigation of Community Response to Urban Traffic Noise

Ghorbanali Mohammadi[1]

Department of Industrial Engineering, college of Engineering, University of Kerman, Iran.

Abstract. Road transportation plays a major role in the economic and social development of a nation. With the increase in vehicle population, noise will increase. This paper describes the reaction to environmental noise of the city of Kerman. A total of 250 questionnaires were processed. The finding of this study revealed that the main isolated noise sources were traffic (50%), street noise (34%) and construction (16%). This research also revealed that the main outcomes of exposure to noise were: irritability (40.8%), insomnia (24%), difficulty to concentrate (16%) and conservation disruption (16%). In this study A-weighted continuous equivalent sound level values L_{Aeq} , $L_{\max}$, $L_{\min}$ and statistical levels L_{50}, L_{90} , L_{99} were manually measured at each site separately.

Keywords. Noise pollution; traffic noise; noise measurements; urban city; noise annoyance

1 Introduction

Noise in cities is considered by the World Health Organization to be the third most hazardous type of pollution, right after air and water pollution [11].

This study presents the results obtained from a social surveyed and noise measurement carried out in the city of Kerman, Iran. Kerman is one of the 30 provinces of Iran. Kerman city had an estimated population of 533,799 in 2005. The city of Kerman (population: 400,000) embraces about 80% of the urban population, being the most developed and largest city of the province. In Kerman, the increases in the population and in the number of circulating vehicles have led to an increase in the urban noise levels. The need for studies regarding the urban noise pollution and its consequences for the community has motivated various have various impacts on mental and physical exposure standards for free field established by Iranian government [4] health and disturbance of daily activities. It may affect sleep, conservation, lead to perception of annoyance, causes hearing

[1] Corresponding Author : Tel. &fax: +98-341-2112861. E-mail address:gmohammadi@mail.uk.ac.ir

loss, cardiovascular problems as well as affects task performance [7].Table 1 shows the environmental noise

The objective of this survey is to investigate noise pollution and its impact on the community. This research represents the first study of community noise levels in city of Kerman located south eastern in Iran

Table 1. Environmental noise exposure standards for free fields in Iran (IDE).

Type of areas	Daytime(dBA) 07:00-22:00	Night-time (dBA) 22:00-07:00
Residentia	l50	30
Residential-Commercial	60	50
Commercial	65	55
Residential-Industrial	70	60
Industrial	75	65

2 Methodology

2.1 Sound level measurements

This study was focused on one of the busy and the most crowded street in downtown area along with its 6 connecting streets, , which have heavy traffic during the day. Total of 20 measuring points were selected along the roads, pavements and in the shopping areas of the city center to adequately represent the different acoustically residential-commercial situations. In this study the A-weighted continuous equivalent sound level values L_{Aeq} , $L_{\max}$, $L_{\min}$ and statistical levels L_{50}, L_{90} , L_{99} were manually measured at each site separately. The measurements were taken on various days of the week from 07:00 to 22:00. Noise levels were measured 1.5 m above ground using a Class I 2231 Bruel and Kjaer integrating sound level meter with a 4230 B and K calibrator. During the study there was no wind and no rain.

2.2 Questionnaire description

In order to knowing the opinion of the citizens from the area we surveyed about how the noise levels have affected their daily life, a specific questionnaire has been designed. The survey questionnaire contained four different parts. The first part identified sex, age. The second part contained information about accommodation. The third part has questions where information about noise levels and its effects on people's habit is obtained. In the last part, the main noise type and its variation with time is evaluated.

The respondents were randomly selected on a one-person per family basis. A total of 250 questionnaires were processed.

2.3 Definitions

In this section, some terms used in this survey will be explained

2.3.1 Decibel (dBA)

The ear sensitivity to sound is dependent on the frequency of the sound being detected. Frequencies at the extreme of the hearing frequency range are not detected as well as frequencies in the middle of the range. To account for this, when measuring sound, a weighting curve is used to place more emphasis on frequencies to which humans are more sensible. The "A" weighting curve is generally used for the purpose of measuring sound levels. The sound weighted by the "A" curve approaches the perception of the human ear and its value is given by dBA.

2.3.2 Sound pressure level

It is given by the formula below, where the reference sound pressure is the minimum sound pressure that can be perceived by the human ear. Its value is equitant to Pa (0.00002Pa).

$$SPL = 10\log\left(\frac{P}{\mathrm{Pr}e}\right)^2$$

Where P - sound pressure; Pre - reference sound pressure. The unit used is decibel (dBA).

2.4 Noise- human repose

The human ear detects sound pressure changes in the air and transmits a signal, which is related to the sound pressure changes, to the brain where it is perceived as sound. The signal which is "perceived" by the person is not directly proportional to the sound pressure stimulus which first entered the ear. As well as level and frequency content of any noise an important consideration is the duration of exposure. Duration must therefore be included in any environmental limits or guidelines for assessing human response to noise.

3. Results and Discussions

3.1 Measurement results

Table 2 displays the mean values for L_{Aeq} , $L_{\max}$, $L_{\min}$ and statistical values L_{50}, L_{90} and L_{99} for shariatte street and 6 connecting streets at all measurement stations. The overall mean value of L_{Aeq} was 66.55dBA. The maximum and minimum values were 73.87 and 55.1 dBA respectively. An overview of the Table 2 reveals that L_{Aeq} have a maximum value of 66.55 dBA, indicating that they are over to the governmental legislations for residential-commercial area (see table 1).

Table 2. The mean values for L_{Aeq}, L_{max}, L_{min} in Downtown area of Kerman.

Station No.	Measurement stations	L_{Aeq}	L_{max}	L_{min}	L_{50}	L_{90}	L_{99}
1	8	68.42	79.2	63	69.3	65.7	64.9
2	2	65.64	73.4	64	68.3	60.2	62.3
3	2	67.82	72.3	65	66.3	67.2	68.3
4	2	66.76	73.8	63	64.3	65.4	67.3
5	2	65.92	72.4	63.2	63.9	64.3	65.8
6	2	65.64	73.2	61.3	62.3	65.1	66.3
7	2	65.62	72.87	60.1	64.5	65	65.9
Total	20	66.55	73.87	55.1	65.5	64.5	65.8

3.2 Questionnaire results

Among the participants 40% were men and 60% were female, and they were between 18 and over 50 years old age. Of those polled, 32% among age group 18-30 years old, 40% between 30 and 50 and 28% were over 50.

The frequency distributions of the subjects concerning the noise types in their home are present in Figure 1. The major parts of the respondents (50%) have answered that traffic noise, about 34% have answered that street noise and 16% have answered that construction noise. In another question, the subjects have classified the noise in his/her home as "very high" (70%), "high" (18%), and "normal" (12%).

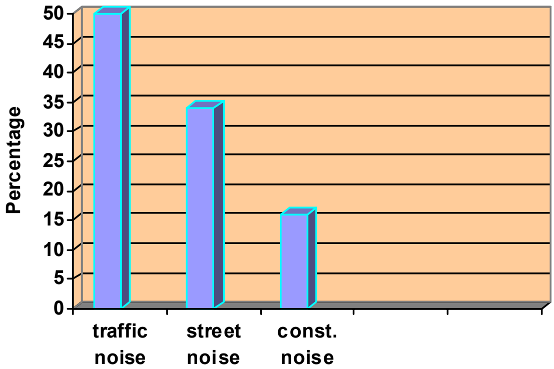


Figure 1. Frequency distribution concerning types of noise

Figure 2 displays frequency distributions of subjects concerning the noise levels. In more than half out of the respondents have affirmed that they had been living at the same home for more than 5 years and 40% had been living at the same home for 1 to 5 years and 10% had been living at the same location for less than one year.

The respondents have been asked "what time does noise bother them more". For this question, 52% out of them answered that noise bother them in the morning and 58% answered that noise bother them in the evening.

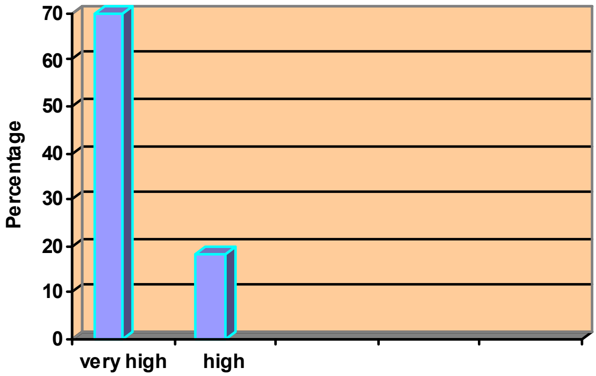


Figure 2. frequency distributions of question concerning levels of street noise

Asked about what noise sources have caused annoyance, the majority of the respondents (50%) have pointed the traffic, followed by the street (34%) and (16%) construction. No doubt the traffic is a continuous noise source, as well as street and construction in many cases.

The majority of the respondents have answered that they sometimes felt annoyed by noise in his/her home and have pointed out that at least one of these non-contiguous sources as the cause of annoyance. The frequency distributions of the subject's answered that traffic noise produces physical and psychic annoyance to them are displays in figure 3. Nearly 86% of the respondents have answered that traffic noise produce physical and 86.8% have answered that noise produce psychic annoyance to them.

Among the participants 40% were men and 60% were female, and they were between 18 and over 50 years old age. Of those polled, 32% among age group 18-30 years old, 40% between 30 and 50 and 28% were over 50.

The respondents have been asked "what time does noise bother them more". For this question, 52% out of them answered that noise bother them in the morning and 58% answered that noise bother them in the evening.

Asked about what noise sources have caused annoyance, the majority of the respondents (50%) have pointed the traffic, followed by the street (34%) and

(16%) construction. No doubt the traffic is a continuous noise source, as well as street and construction in many cases.

The majority of the respondents have answered that they sometimes felt annoyed by noise in his/her home and have pointed out that at least one of these non-contiguous sources as the cause of annoyance. The frequency distributions of the subject's answered that traffic noise produces physical and psychic annoyance to them are displays in figure 3. Nearly 86% of the respondents have answered that traffic noise produce physical and 86.8% have answered that noise produce psychic annoyance to them.

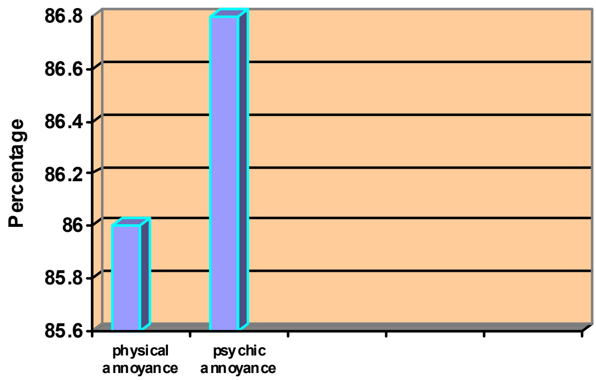


Figure 3. Frequency distributions of physical and psychic annoyance produce by high noise levels

As expected, the majority of the respondents (80%) have answered that traffic noise has negative impacts on their health and safety. By considering this new focus, it is possible to construct a new point of view in this analysis. For example, the respondents who felt irritability by the noise (40.8%), and the respondents who have insomnia (24%), the respondents have difficulty to concentrate (16%) and the respondents have conservation disruption are (16%).

Maschke [6] and Babisch et al. [3] pointed out some effects of the noise over urban inhabitants: irritability, insomnia, difficulty to concentrate. It has been observed that everybody within the group "Annoyed by Urban Noise" had declared that they felt at least one of the effects related above, predominantly irritability, insomnia and difficulty to concentrate, as displays in Figure 4.

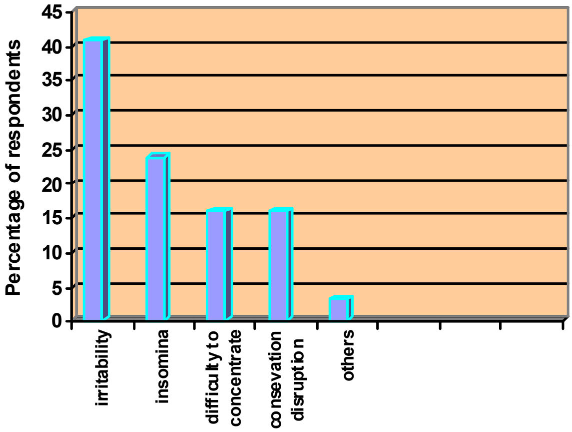


Figure 4. Respondents answers to question "what does the noise cause on you?"

4 Conclusions

Traffic noise is a major environment source of pollution in the whole planet, both in developed and in undeveloped nations. Pollution in cities is a growing problem due to the fact that the urban environment is becoming increasingly crowded and busy.

In this research, noise levels were measured, for the first time, in the city center along with its 6 connecting streets in Kerman.

Simultaneously with measurements of noise, the opinion of 250 residential who levied in area concerning their level of annoyance, traffic noises, as well as awareness concerning the health impact of noise, were surveyed.

The measured noise levels and the traffic noise index as well as the noise pollution level all indicted that the noise levels in community were generally high. L_{Aeq} , for example, ranged between a maximum of 73.87 (dBA) and a minimum of 55.1 (dBA).

Nearly 52% of subjects were beliefs noise bother them in morning and 23.3% beliefs noise bother them in evening. The more unpleasant noise was traffic, street and construction.

The respondents have pointed out the following effects of the urban noise: irritability (40.2%), insomnia (24%), difficulty to concentrate (16%) and conservation disruptions (16%).

5. References

[1] Abdel-Raziq, I R, Zeds, Q and Seh, M. Noise measurements in the community of Nablus in Palestine. Acoustical 2000:86, 578–580.
[2] Marana, M and Garcia, A. A social survey on the effects on environmental noise on the residents of Pamplona, Spain. Applied Acoustics 1998:53,245- 253.
[3] Babisch W., Ising H., Gallacher JEJ, Sweetnan PM, Elwood PC. 1999. Traffic noise and cardiovascular risk: the Caerphilly and Speedwell studies third Phase-10 year to follow up. Arch Environ Health; 54(5): 210-216
[4] [Iran Department of Environment, (IDE); 1998. Environmental Standards and regulations, IDE. Press Tehran.
[5] Khilman, T., 2004. Noise pollution in cities, Curtiba and Totenberg as examples. In: proceedings of the seminar-Environmental Aspects of Urbanization-Seminar in Honor of Dr. Motafa Kamal Tolba, Gothenburg, Sweden, in CD.
[6] Maschke C. Preventive medical limits for chronic traffic noise exposure. Acustica 2001: 85(5):448.
[7] Piccolo, A. Plutino, D. Cannistraro, G. Evaluation and analysis of the environmental noise of Messin, Italy. Apl. Acoust. 2005:66(40):447-465.
[8] Sunscald, T, Sulkies, P, Tabucanon, S M, Aoi, I, Shirai, K and Tanaka, H Road traffic noise prediction model in Thailand. Applied Acoustics1999: 58,123–130.
[9] Zannin, P H T, Diniz, F, Calixto, A and Barbosa, W. Environmental noise pollution in residential areas of the city of Curitiba. Acustica 2001:87, 1–4.
[10] Zannin, P H T, Calixto, A, Diniz and F, Ferreira, J A. A survey of urban noise annoyance in a large Brazilian city: the importance of a subjective analysis in conjunction with an objective analysis. Environmental Impact Assessment review, 2003:23:245-255.
[11] World Health Organization, 2005. United Nations Road Safety Collaboration: A Handbook of Partner Profiles (Geneva: World Health Organization). alter the size settings as shown in Figure 1 (the defaults are too large).

A Market Segmentation System for Consumer Electronics Industry Using Particle Swarm Optimization and Honey Bee Mating Optimization

Chui-Yu Chiu[1], I-Ting Kuo and Po-Chia Chen

Department of Industrial Engineering and Management, NTUT, Taiwan.

Abstract. The use of information technologies in various business areas is emerging in recent years. With the development of information technology, how to find useful information existed in vast data has become an important issue. The most broadly discussed technique is data mining, which has been successfully applied to many fields and analytic tools. Clustering analysis which tries to segment data into homogeneous clusters is one of the most useful technologies in data mining methods. Market segmentation is among the important issue of most companies. Market segmentation relies on the data clustering in a huge data set. In this study, we propose a clustering system which integrated particle swarm optimization and honey bee mating optimization methods. Simulations for a benchmark test functions show that our proposed method possesses better ability to find the global optimum than other well-known clustering algorithms. The results show that system through PSHBMO can effectively find the global optimum solution, and extend the application of market segmentation to solve the RFM model.

Keywords. Market segmentation, Particle swarm optimization, Honey bee mating optimization.

1 Introduction

Recently, the progress of information technology has transformed the way of marketing and information management in companies. With the large number of data from customer behavior, it has become possible to realize the consumer insight by the variety of data mining techniques and CRM tools. In general, market segmentation is the most important issue for companies to recognize different group of customers, who have some similar characteristics explaining different customer value.

Market segmentation divides a market into distinct subsets of buyers, each has some similar attributes. It's most important variable is purchasing behavior. In order to describe customer's purchasing behavior, the RFM analytic model

[1] Corresponding Author : Department of Industrial Engineering and Management, National Taipei University of Technology, 1, Section 3, Chung-Hsiao East Road, Taipei 106, Taiwan, ROC; Tel.: +886 2 2771 2171x2365; fax: +886 2 2731 7168. E-mail address: cychiu@ntut.edu.tw (C. Y. Chiu).

proposed by Hughes [1], is usually applied. It is a model that tells important customers from large data by three attributes, i.e., interval of customer purchasing, frequency and amount of money. The definitions of RFM model are summarized as follows: (1) R-Recency of the last purchase; (2) F-Frequency of the purchases; (3) M-Monetary value of the purchases.

Data clustering is a useful technique for the discovery of some knowledge from a dataset. Clustering is the process of grouping a set of abstract or physical objects into classes of similar objects. The purpose of clustering analysis is to find the difference among each groups and the similarity in the same group. Aldenderfer and Blashfield [2] concluded five basic steps that characterized all clustering studies. Each of these steps is essential to clustering as follows. (1)Selection of a sample to be clustered; (2) Definition of a set of variables on which to measure the entities in the sample: (3) Consumption of the similarities among the entities; (4) Use of clustering method to create groups of similar entities, and (5) Validation of the resulting cluster solution.Clustering techniques are relatively popular for market segmentation due to its short computation time and easy accommodation. For examples, Shin et al. [3] used three clustering algorithms, including K-means, FCM and SOM, to find properly graded stock market trading brokerage commission rates based on the 3-month long total trades of two different transaction modes and concluded that FCM is the most robust approach. Huang et al. [4] used support vector clustering for market segmentation by the drink company. The results can be seen that the proposed method can outperform to the k-means and the SOFM methods. Liao et al. [5] proposed the apriori algorithm and clustering analysis as methodologies for data mining. The results was illustrated what functionalities best fit the consumers' needs and wants for life insurance products by extracting specific knowledge patterns and rules from consumers and their demand chain.

Particle swarm optimization (PSO) is an evolutionary computation technique developed by Kenney and Eberhart [6]. The method has been developed through a simulation of simplified social models. Each particle also has its own coordinates and velocity to change the direction of flying in predefined search domain with *D*-dimensional. All particles fly through the search domain by following the direction of current optimal particle. He et al. improved the standard PSO with passive congregation (PSOPC), which can improve the convergence rate and accuracy of the SPSO efficiently [7]. Liu [8] proposed CPSO (chaotic particle swarm optimization) algorithm. It applies PSO to perform global exploration and chaotic local search to perform local search on the solutions produced in the global exploration process. Over the years, many successful applications of PSO, to image registration [9], have been reported. PSO algorithm is a powerful optimization technique for solving multimodal continuous optimization problems [10].

Afshar [11] brought up a swarming relation in society model assuming a polygynous colony called honey bee mating optimization (HBMO). It studies the behavior of social insects and uses their models to solve the optimum problems. Each bee performs sequences of actions according to genetic, environmental, and social regulation. Result of each action itself became a portion of the environment and greatly influences the subsequent actions of both single bee and many drones. The marriage process represents one type of action that was difficult to study because the queens mate during their mating-flight far from the nest. Fathian and Amiri [12] applied honeybee mating optimization in clustering problems. It compared HBMO means with other heuristics algorithm in clustering, such as GA, SA, TS, and ACO, by implementing them on several well-known datasets. It was found that the HBMO algorithm works than the best one.

2. Methodology

In most customer relationship management problems, one may notice that no tool for data mining is perfect because there are many uncertain variables. In this study, we propose an integrating particle swarm optimization with honey bee mating optimization (PSHBMO) market segmentation system based on the structure of decision support system to solve the clustering problem.

The proposed method was divided into four phases: (1) Data pre-processing (2) Using particle swarm optimization to search initial solution (3) Using honey bee mating optimization to search the best solution (4) Comparing clustering performance evaluation, as follows:

2.1 Data pre-processing

Before feeding data into the clustering algorithms, database variables should be normalized to eliminate scale effects. Normalization entails relatively minor additional computations during application of a solution to new data, which must also be normalized. For some attributes whose preferences are monotonically increasing, such as frequency and population et al., a simple positively linear normal function is shown in Eq. (3.1). However, exhibits a monotonically decreasing preference. An inverse function is shown in Eq. (3.2) is then applied.

$$NormalizedValue = (log_{Xi} - log_{Xmin}) / (log_{Xmax} - log_{Xmin}) \quad (3.1)$$

$$NormalizedValue = 1-(log_{Xi} - log_{Xmin}) / (log_{Xmax} - log_{Xmin}) \quad (3.2)$$

2.2 Using particle swarm optimization to search initial solution

This step utilizes particle swarm optimization (PSO) to decide the initial vector of cluster center. A single particle represents the *k* cluster centroid vectors. That is, each particle X_{id} has its vectors V_{id} that be constructed as follows:

$$X_{id} = (z_{i1}, \ldots\ldots, z_{ik}) \quad (3.3)$$

where z_{ij} refers to the *k*-th cluster centroid vector of the *i*-th particle

(1) Initialize each particle to contain k randomly selected cluster centroids.

(2) Calculate the Euclidean distance between all of data to cluster centroids, and assign for the minimum distance.

(3) Calculate the fitness using equation follow:

$$J_e = \frac{\sum_{j=1}^{k} \left[\sum_{\forall x \in n_{ij}} d(x, z_{ij}) / |n_{ij}| \right]}{k} \quad (3.4)$$

(4) Update the global best, local best positions and cluster centroids (X_{id} and V_{id}) as follow, and recalculate back to step (2):

$$V_{id}^{new} = W \times V_{id}^{old} + c_1 \times rand_1 \times (P_{id} - X_{id}) + c_2 \times rand_2 (P_{gd} - X_{id}) \quad (3.5)$$

$$X_{id}^{new} = X_{id}^{old} + V_{id}^{new} \quad (3.6)$$

(5) Proceed until meet epochs equal to a parameter. Keep the global best solution and locate best solution.

2.3 Using honey bee mating optimization to search the best solution.

Having got the initial solution of PSO, this step was searching the optimization solution. This algorithm is constructed with the following five main stages:

(1) Initialize mutation rate, cross rate and each parameter. Let the global best solution be queen's chromosome and location best solution be drone's chromosome.

(2) Use simulate annealing and roulette wheel selection to select the set of drones from the list for the creation of broods. After each transition in space, the queen's speed and energy decays according to the following equations:

$$S(t+1)=\alpha(t)\times S(t) \quad (3.7)$$

where α is a factor[0,1] and is the amount of speed reduction after each transition.

(3) Create new set of broods by crossover the drone's genotypes with the queens.

(4) Calculate the Euclidean distance between all of data to cluster centroids, and assign for the minimum distance.

(5) Calculate the fitness using equation as formulate 3.4. Determining that is the new best solution better than the previous one.

(6) Randomly mutate chromosome and replace weaker queens by fitter broods.

(7) Proceed until meet epochs equal to a parameter.

2.4 Comparing clustering performance evaluation.

This algorithm tries to minimize the error function – Mean Square Error (MSE). The main purpose is to compare the quality of the respective cluster, where quality is measured according to the following three criteria:

(1) The mean square error as defined in Eq.(3.8).

$$\mathrm{MSE}=\sum_{i=1}^{n}\sum_{j=1,x_i\in c_j}^{k}\left|x_i - m_j\right|^2 \quad (3.8)$$

where x_i (i=1,2,…n) is a data set X with n objects,k is the number of clusters,m_j is the centroid of cluster C_j (j=1,2,…k).

(2) The intra-cluster distances: the distance between data vectors within a center, where the objective is to minimize the intra-cluster distances. It is defined in Eq. (3.9).

$$\text{intra-clusterdistance} = \sum_{i=1}^{n}\sum_{j=i+1}^{n} d(x_i, x_j) \quad (3.9)$$

where x_i (i=1,2,…n) is a data set X with n objects, x_j (j=1,2,…n) is a data set X with n objects.

(3) The inter-cluster distances: the distance between the centroids of the clusters, where the objective is to maximize the distance between clusters. It is defined in Eq. (3.10).

$$\text{inter-clusterdistance} = \sum_{i=1}^{k}\sum_{j=i+1}^{k} d(m_i, m_j) \quad (3.10)$$

where m_i is the centroid of cluster C_i (i=1,2,…k), m_j is the centroid of cluster C_j (j=1,2,…k),k is the number of clusters.

This research considered inter-cluster and intra-cluster distances at the same time to make sure that the latter ensures compact clusters with little deviation from the cluster centroids, while the former ensures larger separation between the different clusters. In order to get the value which maximize the distance between clusters and minimize the intra-cluster distances, this study according to the notion of Intra−cluster Distance and Inter−cluster Distance.

3 Evaluation of proposed model on data sets

We presented a set of experiments to show the performance of the PSHBMO algorithm. The experiment was conducted on a Pentium 3.40 GHz, 512 MB RAM computer and coded with Borland C++ 6.0 Builder software.

3.1 Experiment-IRIS

A well-know database, the Iris Plant, is utilized to test the performance of our proposed method. Iris plant is a database with 4 numeric attributes, 3 classes and 150 instances. We compared the PSHBMO algorithm with PSO+K-means and SOM+K-means by Chiu [13]. Finally, we provided the clustering results and used three criteria to evaluate the quality of the results as in following table.

Table1 summarizes the results obtained from the clustering algorithm for the problem above. The values reported are averaged 30 simulations, for which standard deviations indicates the range of values where the algorithms converge. If the algorithm could cluster data with a lower MSE value, the similarity within a segment increases. For the problem, the PSHBMO algorithm had a smallest average MSE. When considering Intra −cluster Distance and Inter −cluster Distance, the PSHBMO algorithm also had a smallest value comparing with PSO +k-means and SOM +k-means.

Table 1. Comparison of three methods of proposed system

Algorithm	MSE	Inter-cluster Distance	Intra-cluster Distance	$\frac{\text{Intra - cluster Distance}}{\text{Inter - cluster Distance}}$
PSO+k-means	0.218	0.281	0.803	0.350068
SOM+k-means	0.224	0.290	0.792	0.366445
PSHBMO	0.194	0.264	0.804	0.329403

3.2 Case study-RFM model

For the case study, a real-world database is collected from Podak Co., an authorized agent of Panasonic that provides passive and active electronic components for consumer electronics, telecommunications, computers etc. The period of the business transaction data is from 2003/1/1 to 2006/6/15.

In this study, we employ a two-stage clustering suggested by Punj and Stewart [14]. In the first stage, PSHBMO is used to cluster the normalized data set into different groups. The result is shown in Table2. It is found that the intra-cluster distance /inter-cluster distance is the lowest at Group=9 and the distance is relatively decreasing flatly when the number of clusters is more than nine. Therefore, it is implies that the best number of clusters should be nine.

Table 2. Clustering Result Form Group=3 to Group=10

Index	Group=3	Group=4	Group=5	Group=6	Group=7	Group=8	Group=9	Group=10
Intra-cluster distance	0.305664	0.305267	0.274965	0.265611	0.248516	0.223385	0.218951	0.239033
Inter-cluster distance	0.585759	0.612778	0.580098	0.598582	0.578599	0.571645	0.575373	0.591121
$\frac{\text{Intra} - \text{cluster}}{\text{Inter} - \text{cluster}}$	0.521826	0.49817	0.473997	0.443734	0.429513	0.390776	0.380537	0.404371

In the second stage, we compared the results of the PSHBMO, SOM+*k*-means and PSO+*k*-means algorithms showed the best number of groups is six. The result is presented in Table 3.

Table 3. Comparison of three methods

Algorithm	MSE	Inter-cluster Distance	Intra-cluster Distance	$\frac{\text{Intra - cluster Distance}}{\text{Inter - cluster Distance}}$
PSO+k-means	0.185	0.238	0.580	0.410162
SOM+k-means	0.190	0.246	0.574	0.429780
PSHBMO	0.168	0.242	0.603	0.401326

As a result, PSHBMO posseses lowest MSE and $\frac{\text{Intra-cluster Distance}}{\text{Inter-cluster Distance}}$. PSHBMO is better than PSO+ *k*-menas and SOM +*k*-means. The clustering results of PSHBMO is further utilized for make marketing strategies; Table 4 shows the clustering results for customers.

Table 4. Clustering Results of PSHBMO

Cluster	Customer Counts	Recency (Avg.)	Frequency (Avg.)	Monetary (Avg.)	RFM Status
1	2	253.5000	1773.0000	89,582,024.00	R↑F↑↑M↑↑
2	10	16.0000	440.3000	4,420,767.00	R↓↓F↑M↑
3	9	152.8889	194.8889	10,984,008.00	R↓F↑M↑
4	12	12.6667	8.5000	183,222.40	R↓↓F↓↓M↓
5	16	118.0000	23.5000	93,436.13	R↓F↓M↓↓
6	29	331.1034	2.7931	15,980.83	R↑F↓↓M↓↓
7	16	355.0625	10.8125	314,885.10	R↑F↓M↓
8	12	1.4167	402.2500	11,089,204.00	R↓↓F↑M↑
9	14	20.7143	74.7143	491,144.30	R↓↓F↓M↓
Total Avg.		163.94167	135.9000	3,928,079.2833	

In table 4, the sign ↑ denotes that the value was greater than an average, and the sign ↑↑ denotes that the value was the much greater than an average. the sign ↓ denotes that the value was smaller than an average, and the sign ↓↓ denotes that the value was the much smaller than an average. Cluster 2, Cluster 3 and Cluster 8 which has R↓(↓)F↑M↑ can be considered as loyal ones who frequently deal with and make a large purchase. Cluster 4, Cluster 5 and Cluster 9 who has R↓F↓M↓ was probably a new customer who recently dealt with. Cluster 1 who has R↑F↑↑M↑↑ is promising one who might be promoted to the loyal customer. Cluster 6 and Cluster 7 who have R↑F↓M↓ is likely to be vulnerable customers who have not dealt with for a long time.

Among the nine clusters, Cluster 3 is selected as a target customer segment with the first priority, followed by Cluster 6. It is because that the effect to these target segments might become potentially greater than the effect to others from the RFM point of view.

4 Conclusion

Data mining is the process of posting various queries and extracting useful information, patterns, and trends often previously unknown from large quantities of data possibly stored in databases. Database technology has been used with great success in traditional business data processing. Through this study, we built that an effective and accurate market segmentation system based on intelligent clustering methods integrated particle swarm optimization and honey bee mating optimization methods. We used three clustering algorithms to solve the Iris dataset. For the problem, the PSHBMO algorithm had a smallest value of MSE and Intra–cluster Distance/Inter–cluster Distance which compared with PSO+k-means and SOM+k-means. In the business database, we decided RFM variables are to be used for analyzing. First, we ran PSHBMO for this normalized data to find which group had the lowest of the intra-cluster distance /inter-cluster distance at nine groups. Then, the result found that Cluster 3 was a target customer segment. With the above clustering information, the proposed method can help marketers to develop proper tactics for their customers.

In order to segment customers precisely, many enterprises have installed customer clustering system. However, most of customer clustering systems in operation are lack of an accuracy performance evaluation analysis based on logical inference or just provide a rough clustering method. An improper clustering result usually leads to a conclusion that can't meet firms and customers' requests and expectation. Consequently, this study proposes a system using the technique of data mining to serve as decision support system by finding useful information hided out.

Furthermore, the procedure proposed in this study can be also applied to other industries. According to the clustering analysis, enterprises of different industry can design various marketing strategies and advertisement models to aim at different group of customers with common features to achieve the needs of customers, enhance the strength with customers for company, satisfy their demands, increase customer loyalty and finally obtain better profits.

5 References

[1] Hughes AM. Strategic Database Marketing: The masterplan for starting and managing a profitable customer-based marketing program. Chicago: Probus Publishing. 1994

[2] Aldenderfer MS, Blashfield RK. Cluster Analysis: Quatitative applications in the social sciences. Beverly Hills: Sage Publication, 1984.

[3] Shin HW, Sohn SY. Segmentation of stock trading customers according to potential value. Expert Systems with Applications. Vol. 27, pp.27-33, 2004.

[4] Huang, J. H., Tzeng GH, Ong CS. Marketing segmentation using support vector clustering. Expert Systems with Applications. Vol.32, pp.313-317,2007.

[5] Liao SH, Chen YN, Tseng YY, Mining demand chain knowledge of life insurance market for new product development. Expert Systems with Applications. Vol. 36, pp.9422–9437,2009.

[6] Kennedy J, Eberhart RC. Particle swarm optimization. International Joint Conference,Neural Networks, Vol. 4, pp. 1942-1948, 1995.

[7] He S, Wu QH, Wen JY, Saunders JR, Paton RC. A particle swarm optimizer with passive congregation. Biosystem,78:135-47,2004.

[8] Liu B, Wang L, Jin YH, Tang F, Huang DX. Improved particle swarm optimization combined with chaos. Chaos, Solitons &Fractals,25:1261-71,2005.

[9] Yin PY. Particle swarm optimization for point pattern matching. J. Visual Commun. Image Representation 17, 143-162,2006.

[10] Paterlini S, Krink T. Differential evolution and particle swarm optimisation in partitional clustering. Computational Statistics & Data Analysis 50 (5): 1220–1247, 2006.
[11] Afshar A, Bozorg O. Honey-bee mating optimization (HBMO) algorithm for optimal reservoir operation. Journal of the Franklin Institute, Vol.344, pp.452-462 , 2007.
[12] Fathian M, Amiri B. Application of honey-bee mating optimization algorithm on clustering, Applied Mathematics and Computation. Vol.199, pp.1502-1513, 2007.
[13] Chiu CY, Chen YF, Kuo IT, Ku HC. An Intelligent Market Segmentation System Using K-Means and Particle Swarm Optimization. Expert Systems with Applications, 2008.
[14] Punj G, Stewart DW. Cluster Analysis in Marketing Research: Review and Suggestions for Application. Journal of Marketing Research, Vol. 20, pp. 134-148, 1983.

Why the Big Three Decline Despite Their Lean Management - A Study Based on the Theory of Constraints

Simon Wu[1] and H. M. Wee

Industrial and Systems Engineering Department, Chung Yuan Christian University, Chungli, Taiwan

Abstract: For the first time since early 1930s, General Motors cannot call itself the world's largest automaker. Its sales fell behind Toyota in 2008. The famous Lean Production (LP), a brainchild of Toyota Production System (TPS) has played a part in Toyota's success story. The Big 3 (the Three Big Motor in U.S.) and other enterprises have imitated all the Toyota's production system in a tentative to become more competitive. It took more than ten years for the Big 3 to learn TPS or LP related production strategy, but without the same power of Toyota Co. Why did the Big 3 fail to achieve their expected growth despite their adoption of lean management? The aim of this paper is to answer two significant questions. The first question is "what is the comprehensive action for the Big 3 to grow up their market share?" In order to address this question and get reasonable answer, this study tries to use the Theory of Constraints (TOC) to discover the root causes and countermeasures for the Big 3 to break through their paradigms. The second question is "TOC can really deal with this kind of complicated problem effectively?" From the case study, we successfully answered the two preceding questions, and present some findings about TOC problem solving strategy.

Keywords: the Big 3, Toyota, Lean Production, Theory of Constraints (TOC), Problem Solving

1 Introduction

In 1903, Henry Ford organized Ford Motor Company in Michigan with 11 partners. Then till 1917, the production volume was increased to 700,000 units [2]. Even though General Motor (GM) was established later than Ford, but it overtook Ford in 1931 and dominated global auto market over 70 years [6]. That time, GM, Ford

[1] Corresponding author: wusimon@ms4.hinet.net

and Chrysler were the three biggest automakers (the Big 3) worldwide through their automated mass production system.

In 1950s, Eiji Toyoda arranged a 12-week study tour in U.S. plants for his management team. They found several inherent flaws from the mass production system, such as high inventory due to large batches, discrete process steps with uneven flow, high equipment cost, overproduction, and hidden defects. During their study tour, Toyota management had seen an opportunity to learn from Ford and GM. They learnt the original quality, standardizing processes, continuous material flow and waste elimination, as well as the pull system (Kanban) [4, 13], which was inspired by American supermarkets [14].

Within the next half decade, the Big 3 continued the mass production while Toyota grew to become number six automaker in 1970 from out of the top ten worldwide; and squeezed Chrysler out to become number three in 1994. In 2003, Toyota replaced Ford to become the second-largest auto manufacturer in the world because of its continuous improvement of the Toyota Production System [1]. Because of the Lean Production system [5, 12, 13, 20], Toyota has replaced GM to be the largest automaker in 2008 [6].

As can be seen nowadays, the U.S. automakers, Ford, GM, Chrysler, have been facing their most critical financial crisis since late 2008 [10].

In order to improve the market share, it took more than ten years for the Big 3 to learn TPS or LP related production strategy, but without the same power of Toyota Co. The interesting question is that why the Big 3 still fails to achieve their expected growth despite their long term benchmarking with Toyota? Is there any comprehensive action for Big 3 to overcome their market decline at this severe downtime? In order to address this question, this paper applies the Theory of Constraints (TOC).

2 Problem solving model and Theory of Constraints

Kepner & Tregoe (1985) defined "Problem Solving" as the activity associated with changing the AS IF state to the TO BE state. Wang (2002) defined "Model" as a scheme used for describing system operation. "Problem Solving Model (PSM)" is useful not only for solving problems through comprehensible processes, but also is applicable to monitor, evaluate and manage them by problem solvers.

The TOC is a problem solving model developed by Eliyahu M. Goldratt [11]. Since mid-1970s, Goldratt has used scientific methods to create concepts in management, which have been proven to be of great value to industry. Goldratt also presented his own expression of the scientific method and the structured TOC thinking processes using common sense. The problem solving process takes the form of Theory of Constraints thinking processes and a family of five Theory of Constraints logic tree diagrams [8, 17].

In order to induce effective output from TOC thinking processes, Goldratt used three basic questions namely (1) What to Change? (2) What to change to? And (3) How to cause the change? They are linked with five logic trees: Current Reality Tree

(CRT), Evaporating Cloud Tree (EC), Future Reality Tree (FRT), Prerequisites Tree (PT) and Transition Tree (TT) [8, 17].

3 A case study using TOC

In this case study, we employ the basic analysis tools of four TOC tree diagrams except Evaporating Cloud Tree (EC). They are fitted into the following 4-step process adapted from a generalized problem solving process: (1) Problem finding, (2) Idea finding, (3) Obstacle finding, and (4) Solution finding.

Step (1) Problem finding

In order to effectively find the root causes and solutions, the first priority is to correctly identify and clearly define the confronted problem under Big 3 Motors. For the purpose, this paper summarized the major issues of customer complaints on Big 3's products and the main reasons of market decline into four portions: Product, Cost, Delivery and Business Strategy in order to effectively identify the undesirable effects (UDE). So as to the Current Reality Tree (CRT) is developed based on the following list of UDE, see Table 1:

Table 1. the undesirable effects (UDE) of Big 3 Motors

A. Product	1. High oil consumption 2. High rework 3. Complicated product design 4. Neglect local requirements
B. Cost	5. Cost oriented due to Finance domination 6. High inventory (caused by global Part-by-Part system) 7. Create non-value added works, due to lean thinking not in place 8. No sufficient training program because of cost issue
C. Delivery	9. Batch build design, in order to reduce cost but longer lead time 10. Frequently change schedule results in bullwhip effect
D. Business Strategy	11. Decision making failures based on low cost purchasing 12. Merging smaller companies to promote economic scale

The completed current reality tree is shown in Figure 1. From this diagram, the root causes are clearly identified as:

(1) Without Lean culture

(2) Focus on short term & low cost oriented philosophy

According to the above root causes, we define the task as how to promote auto market share – from the perspective of Big 3 Motors. There is a core problem which must be dealt with: how to transform cost oriented short term strategy to long term philosophy based on lean thinking culture?

Step (2) Idea finding

Having defined the objectives of the problem solving project, the next immediate task is to develop a potential solution through collecting a set of ideas for further evaluation. For this purpose, we exploited the TOC future reality tree (FRT) to build the road map. The gap between the objective and the reality is bridged with a series of intermediate tasks. The completed future reality tree diagram is shown in Figure 2. Basically, the future reality tree is also a logic tree diagram depicting the cause and effect relationship. The interpretation of FRT diagram follows the same manner as CRT.

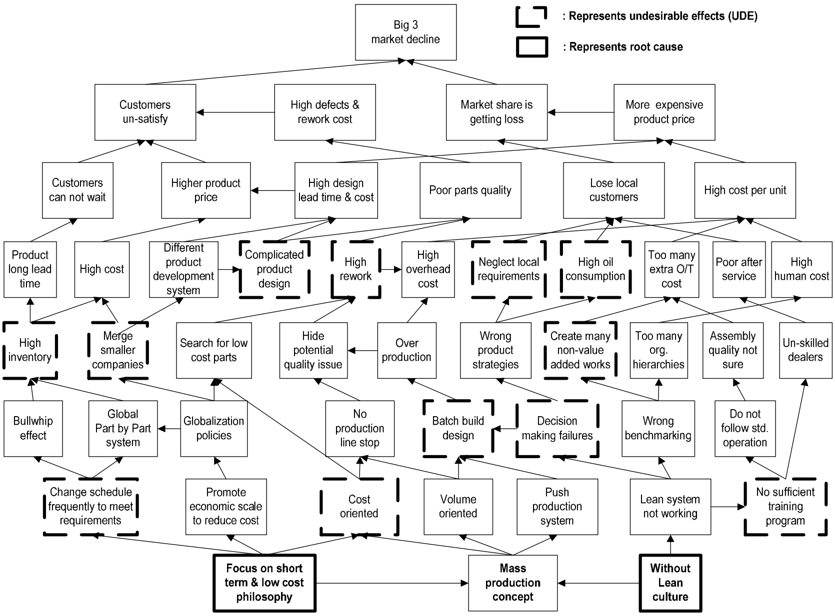


Figure 1. Current reality tree for "Big 3 market decline"

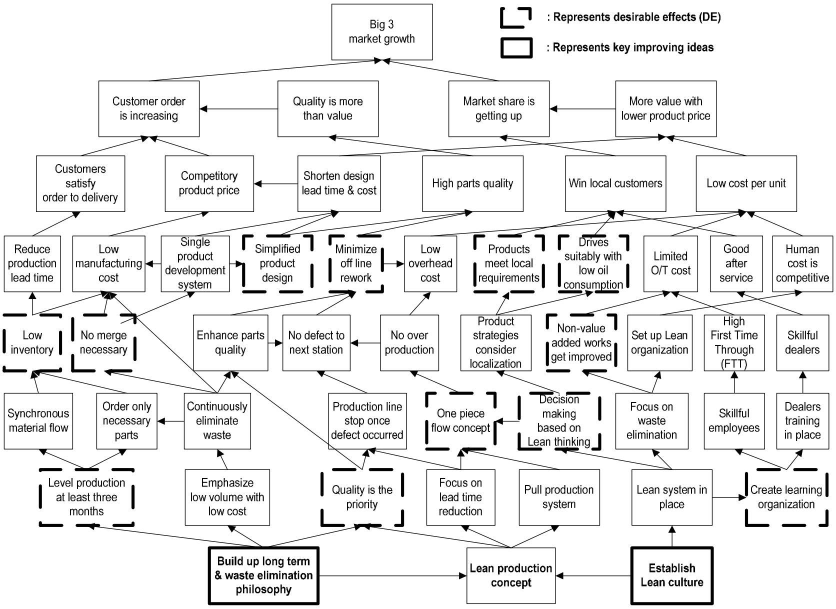


Figure 2. Future reality tree for “Big 3 market growth”

Examining both current reality tree and future reality tree, it is clear that the 12 undesirable effects can be effectively eliminated by the desirable effects through the scheme as shown in Figure 2. After deploying the FRT, this paper finds out two key ideas for the Big 3 Motors to improve their market share, which are: (1) Build up long term business philosophy with continuous waste elimination, (2) Establish globalization strategy with lean thinking culture at each hierarchy.

Step (3) Obstacle finding

The set of solution ideas generated through the future reality tree is of preliminary nature. There might be a number of potential problems hindering its realization. In order to discover the hidden obstacles, the TOC prerequisite tree is exploited to map out a prerequisite plan for examination. Based on the future reality tree diagram of Figure 2, the prerequisite tree diagram for this case study is completed as shown in Figure 3. A total of 12 obstacles are identified and shown as hexagons to obstruct the achievement of each intermediate objective (IO).

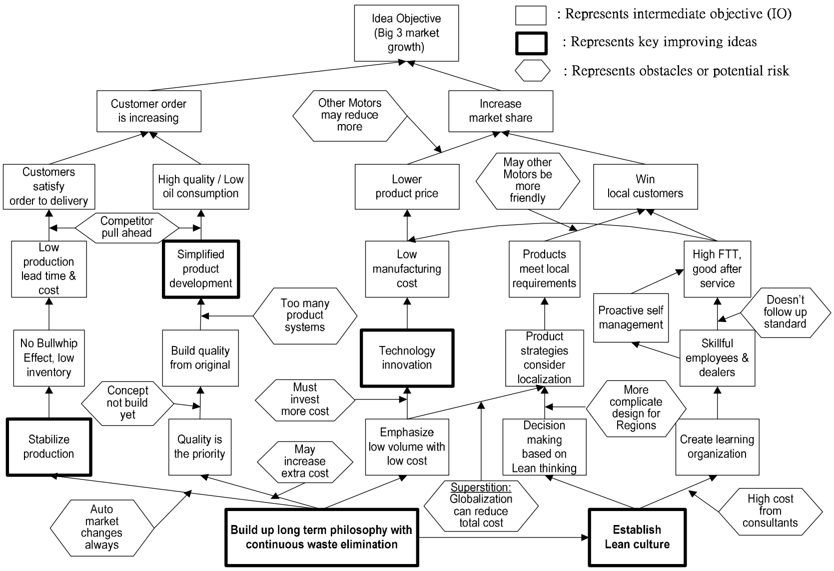


Figure 3. Prerequisite tree for "Big 3 market growth"

Step (4) Solution finding

Having identified the intermediate objectives and obstacles, we are ready to formulate the solution scheme. In order to achieve this purpose, two separate issues need to be dealt with: (1) develop action plans for achieving intermediate or ideal objective, and (2) take necessary actions to remove those obstacles.

The Transition Tree diagram is developed based on the prerequisite tree diagram. Figure 4 depicts the transition tree with respect to the five key issues (the darkened box of Figure 3) for Big 3 market growth. Enveloped in the darkened boxes with dotted line are the proposed solution schemes.

The solution scheme so developed proposes the following action plans:

A. Build up long term philosophy with continuous waste elimination [1]

(1) Set up a long term business plan with budget control
(2) Eliminate wastes & non-valued works as an organization culture [5]
(3) Develop quality rather than volume by total cost concept [12]

B. Establish Lean culture

(1) Set up learning organization by itself
(2) Follow up standard process
(3) Develop proactive attitude of self management
(4) Build continuous improvement culture [5]
(5) Simple, visual & stable design with Lean concept
(6) Decision making based on Lean thinking [20]

C. Simplify product development system
(1) Integrated different product development systems
(2) Product development by knowledge management [3]
(3)Design simple & stable products with low cost but meet regional requirements
D. Stabilize production
(1) Apply level production
(2) Use pull system to reduce inventory and prevent parts shortage [4, 13]
(3) Parts supply based on JIT concept and electronic Kanban system [13, 20]

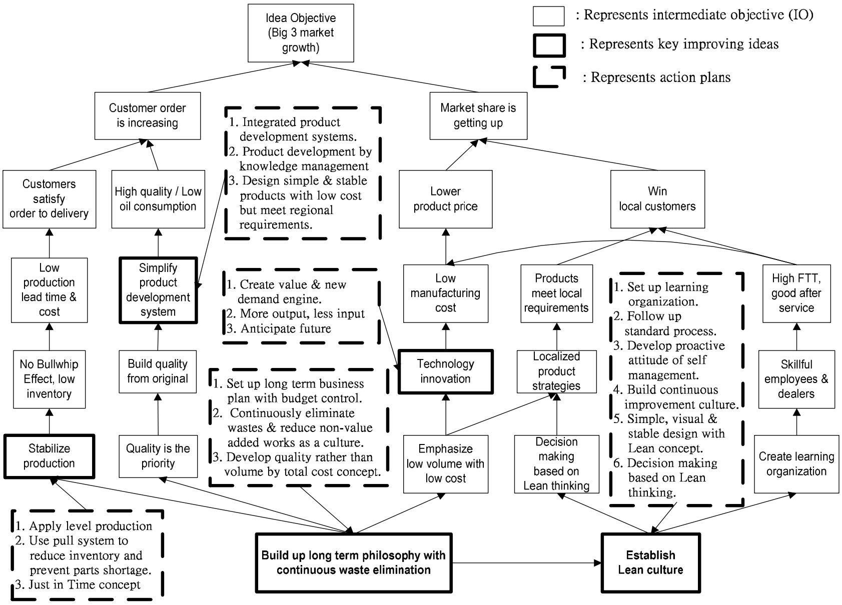


Figure 4. Transition tree for "Big 3 market growth"

E. *Technology Innovation* [9]
(1) Create value & new demand engine
(2) Emphasize on more output and less input
(3) Anticipate future requirements

In summary, there are all together 18 activities highlighted in the Transition Tree for increasing Big 3's market shares, but they may need to be broken down in detail according to the current severe situation of auto market.

4 Findings and discussions

In this article, we have successful demonstrated how a sophisticated case such as global competition in the motor market can be solved with only four logic trees.

This is contrary to the conventional practice where the five logic trees are almost always in use. A similar case was reported by Taylor III et al. [22].

Furthermore, it is interesting to note that the four TOC logic trees fit perfectly well with each of the four problem solving steps. Through this logic thinking process, it not only supports production to do problem solving; but also provides industrial insight to deal with versatile managerial implications.

As illustrated earlier, the problem of “Why Big 3 failed to achieve their expected growth despite their lean management?” is solved by starting with undesirable effects (UDE) as the fundamental stepping stone. It is further developed by exploiting cause-effect-cause relationship through “if…then…” thinking process. It is obvious that a different set of UDE may lead to a different core problem and thus a totally different solution scheme. Therefore, correct finding of the critical UDE determines the success in the identification of the root cause or core problem for further development. That’s the reason why this paper summarized the major issues of customer complaints on Big 3’s products and the main reasons of market decline into four portions: Product, Cost, Delivery and Business Strategy at beginning in order to effectively and correctly identify the undesirable effects.

The TOC approach has advantages at least two: it is easy to follow owing to its clear thinking process, and the solution scheme has a structured presentation using logic tree diagrams.

5 Conclusion

This study provides managerial insights for the Big 3 Motor Companies to break through their paradigms to counteract market decline; and identify how a complex problem beyond production field can be analyzed and dealt with effectively. The proposed use of four TOC trees instead of five TOC logic trees demonstrates that this model is more flexible. This originality is contrary to the conventional practice where the five logic trees are almost always in use. We have made an attempt to answer the question raised in the outset: “Is there any comprehensive action for Big 3 Motors to counteract their market decline?” TOC has been shown to deal with this kind of complicated problem effectively. Future research may consider the use of the proposed TOC model for different industries.

6 References

[1] ANONYMOUS. Leaner and fitter: Growth through waste elimination, Strategic Direction, Volume: 24 Issue: 5 Page: 18 – 21, 2008
[2] BANHAM, R. Ford Motor Company 100 Years, Tehabi Books, ISBN 1-887656-96-0, 2002
[3] BECKER, M. C.; ZIRPOLI, F. Organizing new product development: Knowledge hollowing-out and knowledge integration – the FIAT Auto case, International Journal

of Operations & Production Management, Vol. 23 Issue: 9 Page: 1033 – 1061, 2003
[4] BLACKHURST, J. V.; SCHEIBE, K. P.; Johnson, D. J. Supplier risk assessment and monitoring for the automotive industry, International Journal of Physical Distribution & Logistics Management, Volume: 38 Issue: 2 Page: 143 – 165, 2008
[5] BROWN, A.; EATOCK, J.; DIXON, D.; MEENAN, B. J.; ANDERSON, J. Quality and continuous improvement in medical device manufacturing, The TQM Journal, Volume: 20 Issue: 6 Page: 541 – 555, 2008
[6] BUNKLEY, N. Toyota Ahead of G.M. in 2008 Sales, The New York Times, January 21, 2009
[7] CAI, J.; HWANG, S. Generalized and generative thinking in US and Chinese students mathematical problem solving and problem posing, Journal of Mathematical Behavior 21: pp. 401-421, 2002
[8] DETTMER, H. W. Goldratt's Theory of Constraints: A Systems Approach to Continuous Improvement, ASQC Quality Press Milwaukee, Wisconsin, 1997
[9] DILK, C.; GLEICH, R.; WALD, A.; MOTWANI, J. State and development of innovation networks: Evidence from the European vehicle sector, Management Decision, Volume: 46 Issue: 5 Page: 691 – 701, 2008
[10] GHOSN, C. BusinessWeek, Autos April 29, 2008, 12:01 AM EST, 2008
[11] GOLDRATT, E. M. Theory of Constraints, The North River Press Publishing Corporation, 1990
[12] LEE, J.; PECCEI, R. Lean production and quality commitment: A comparative study of two Korean auto firms, Personnel Review, Volume: 37 Issue: 1 Page: 5 – 25, 2008
[13] LEE-MORTIMER, A. A continuing lean journey: an electronic manufacturer's adopting of Kanban, Assembly Automation, Volume: 28 Issue: 2 Page: 103 – 112, 2008
[14] OHNO, T. Toyota Production System, Beyond Large-Scale Production, Productivity Press, 1988
[15] OSER, R.L. *et al.* Training team problem solving skills: an event based approach, Computer in Human Behavior 15: pp. 441-462, 1999
[16] PETERS, J. Total Strategy, Management Decision, Vol. 30 No. 8, 1992
[17] RAHMAN, SU. Theory of constraints A review of the philosophy and its applications, International Journal of Operations & Production Management 18: pp. 336- , 1998
[18] SCHRAGENHEIM, E. Management dilemmas: the theory of constraints approach to problem identification and solution, St. Lucie Press, APICS series on constraints management, 1998
[19] SIHA, S. A classified model for applying the theory of constraints to service organizations**,** Managing Service Quality 9: pp. 255- , 1999
[20] SUMMERS, G. J.; SCHERPEREEL, C. M. Decision making in product development: are you outside-in or inside-out?, Management Decision, Volume: 46 Issue: 9 Page: 1299 – 1312, 2008
[21] SUTHERLAND, L. Developing problem solving expertise: the impact of instruction in a question analysis strategy, Learning and Instruction 12: pp. 155-187, 2002
[22] TAYLOR III, L. J.; MOERSCH, B. J.; FRANKLIN, G.M. Applying the theory of constraints to a public safety hiring process, Public Personnel Management 32: pp. 367- ,2003

The Business Data Integrity Risk Management Model: A Data Center Case of Information Service Firm

M. K. Chen[a], Shih-Ching Wang[b1]

[a] Graduate Institute of Industrial and Business Management, National Taipei University of Technology.
[b] Graduate Institute of Commerce Automation and Management, National Taipei University of Technology.

Abstract: In recent years, some major earthquakes, such as those in Indonesia and Sichuan, China, have brought out huge losses to enterprises, and have made many enterprises disable to recover their transport business information system and transaction data. Therefore, many enterprises have started to outsource their operative information to the internet data center in order to monitor critical business data operations. Although there are literatures proposing some information safety related subjects, it is still immature to conclude with a suitable data integrity risk & data center operation research.

Based on literature review, case study, deeply interview with 11 top CEOs & experts, and Fuzzy Delphi method, we came up with two frameworks: business data integrity risk management framework and data center service operation reference framework. The former is from the enterprise risk viewpoint and the later one is from IT service firm's viewpoint.

Keywords: Data Integrity, risk management, fuzzy Delphi, data center, case study, security, information service

1 Introduction

Natural disasters, such as earthquake, generally occur without any early warning, and are often difficult to estimate the losses to the related enterprises. The earthquakes in Indonesia and Sichuan, China have also affected several units' business operation risks, causing huge belongings losses. Therefore, if an enterprise or organization only uses an inner disaster recovery strategy, it would be unable to remedy the losses from disaster. When an enterprise more and more relies on information science and technology technology to maintain business

[1] Corresponding Author E-mail: shihching.wang@gmail.com, Graduate Institute of Industrial and Business Management, National Taipei University of Technology, or Industrial Development Bureau, Ministry of Economic Affairs, Taiwan. Address: 3, Sec. 5F, 41-3, Sinyi Rd., Taipei 106, Taiwan, fax: +886-2-27020406. 2.

operations and competition, IT's damage probability from natural disasters also increases. Therefore, in order to reduce the risk of enterprise operations, firms should build an appropriate security policy and strategy [13]. Nevertheless, many organizations monitor the security vulnerabilities introduced through back-up software, procedures and the location of the data. In order to reduce business data risk, outsourcing has become one of the primary ways that large firms deliver information technology (IT) services. Increasingly, firms that seriously consider outsourcing are taking a second look. However, IT outsourcing is no longer simply a strategic move that can be done without a thorough business strategy. Nerveless, data center is a popular way to back-up data for enterprise. Data centers are facilities that house large information systems and related equipments, including telecommunications and storage systems for firms information by a variety of industries. According to IDC research, by 2010, the global requirement for data storage by then will be one zetabyte or one billion terabytes. Bishop pointed the data center design of the future will only be successful if it enables them to adapt to specific business operational characteristics for high-velocity electronic and algorithmic trading, risk management, and complex synthetic hedging strategies [1]. Based on literatur review, we analyze several key operative points of data integrity risk through case-study with a Taiwan data center. In addition, due to this issue is highily involves business policy and decision-making. Therefore, we used fuzzy - Delphi method to find those elements. Finally, we propose a referance framework of business data integrity risk management model and data center service content.

2 Business Information Security

IT security is a key policy for firms, and it involves technology, processes, and people. It has a variety of measures, such as passwords, biometrics, firewalls, etc. As the business transactions become more complex, IT security events become more frequent. With this, enterprise should follow a strict procedure or mechanism to help business to be managed under stable operation. Many organizations have defined the components of information security [10] [8]. BS 7799 is the international standard of information security. BS 7799 gives guidance and recommendations to assist firm with a reference solution for most situations that IT systems are applied. Glazer claimed "information is the firm's primary strategic asset", that IT is the key element in strategic planning and decision making for business. Consequently, enterprise should do their best effort to hold that their information resources keep their confidentiality, integrity and availability [6]. Therefore, it has become a highly complex and challenging task to be holding the security of corporate information assets. On one hand, Garg et al. points out that the growing value of IT resources and the increased levels of connectivity among information systems both within and among organizations [5]. Distinctly, the weak sensitivity of security means that enterprises are weak to plan their information resources effectively [12]. One increasingly important model for protecting company IT system, and for reducing times of security breaches, is through the formulation and application of a formal information security policy (InSPy) [7]. A research conducted by Eloff and Eloff introduces a comprehensive approach towards information security, namely PROTECT [4]. This is an acronym for Policies, Risks, Objectives, Technology, Execute, Compliance, and Team. PROTECT focuses on addressing all aspects of IT security. It involves an approach that considers various and well-integrated controls in order to reduce risk and keep effectiveness and efficiency in the organization. The seven control components of PROTECT focus on implementing and managing an effective information security program from a technology viewpoint and human aspect. Lindenmayer also points out 10 key components to ensure the security of network data such as layered approach, encrypt, etc. [9]. Besides, the risk framework of security is important issue. Generally, there are four main principles that underpin a risk management strategy: co-ordination, credibility, effectiveness and transparency.

As above, establishing information security is a more complex, time & cost consuming, operational risk. Therefore, enterprise should apply a strict procedure to reduce risk of business. ISO standards play a more significant role in security management, such as the code of practice for information security management (ISO,17799) and the requirements for information security management systems (ISO 27001). Brenner points out the differences between these standards [2]. It has been accepted that there are very close ties between information security and risk management, and these standards contribute to this relationship. These standards are based on the PDCA cycle rule: Plan, Do, Check and Act to maintain complete security process as shown in Figure 1. The "PLAN" focuses on design ISMS content, the "DO" focuses on implement & uses ISMS, the "CHECK" focuses on monitor & review ISMS and the "ACT" focuses on Maintain & improve ISMS.

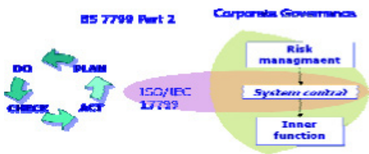


Figure 1. ISMS – PDCA cycle

3 Data Center

As far as business operation concerned, build better IT service management and governance task will consume huge cost and time. Therefore offshore IT outsourcing model, such as data center, is a trend development, solution for enterprise [2]. According to Ovum Research pointed data center service content two types [11]: telecomm related vendor and IT related vendor with them service items as shown Table 3. Application services providers (ASPs) service contents including: wave one/ wave two/applications, wireless ASP (WASP), messaging applications, desktop applications, and multi-channel centers. Managed service providers (MSPs) service content including: managed security, storage virtual IT department, performance monitoring, business continuity, disaster recovery. Web hosting service contents including: dedicated hosting, shared hosting, co-location, application management, application integration, and infrastructure solution. It was quite obvious that company enterprise more care the data risk strategy, they have more value added from data center service (Table 1).

Table 1. Data center service type and items

Hosting service(Data Center)			Value Added
IT related vendors	Application services providers(ASPs)	Wave one/ Wave two/applications	High
		Wireless ASP(WASP)	
		Messaging applications	
		Desktop applications	
		Multi-channel centers	
	Managed service providers(MSPs)	Managed security	
		Storage	
		Virtual IT department	
		Performance monitoring	
		Business continuity	
		Disaster recovery	
Telecomm. Related venders	Web hosting	Dedicated hosting	
		Shared hosting	
		Co-location	
		Application management	
		Application integration	
		Infrastructure solution	Low

Source: Ovum (2007)

In recent, data center service model industry is growing up. Data center servers should operate 7days and 24 hours with no failures for the servers are sensitive to keep their business operation. Data center assist enterprises to obtain their business goals and provides faster service to business. Carter and Pultorak suggested that one can classify IT frameworks with following six dimensions [3]: structure and roles, metrics, processes & practices, technology, controls and people. Chang et. al also pointed the concept of pervasive internet data center (P-IDC) for preservation of high security on confidential information for large enterprises [3]. However, there are several researches the issue of IT securities and data center service, but still a development stage.

4 Research Methodology

According to the literature review in business data integrity (information) risk issue, we found only a few of qualitative research reports related to enterprise data integrity risk management model. Hence, the issue is still in a premature stage. Based on the preceding literature, we use Delphi approach & case study method [14], deeply figure out several key operation factors of a representative data center case in Taiwan. According to these analyses and observations, we collected identical judgments by experts and consultants through several iterative procedures (called Delphi method). The research key problems in this case study are identified as follow:

1. How do organizations keep their information infrastructure absolutely safe?

2. *How do organizations keep their data disaster recovery or remedy system absolutely safe?*
3. *How do organizations assure their security/safety system for managed hosts?*

Base on literature review in business data integrity (information) risk issue, and interviewed the CEOs of dater center's firms (such as ACER etc.), we sum up these key elements of operation in business data integrity risk management for enterprise and data center service firm. As regards to the industrial (firms) domain, these elements are: facility location & structure, host servers management level, information security management level and disaster recovery level. In the other hand, as regards to the IT service (such as dater center' firms), these elements are: customized service modular domain, service modular/specification, IT infrastructure and security modular/specification. Due to these elements involves much strategic policy and higher order decision-making. Therefore, we used fuzzy - Delphi method to find those elements. The Delphi method is better as a research instrument when there is incomplete knowledge about a issue. The Delphi method is an iterative process to collect and distill the anonymous judgments of experts. Fuzzy Delphi method collects a serious of data and analysis techniques interspersed with feedback. A fuzzy set is a class of objects with a continuum of grades of membership [15]. In Fuzzy set theory, membership function is one of the basic concepts, which assigns to each object a grade of membership ranging between zero and one. Researchers always consider different issue with different membership function; the most used membership function includes Triangles membership functions, trapezoid membership functions, and Gauss membership functions. This research adopted Triangles membership functions for our primary membership functions. The equation of triangles membership functions include three parameters, i.e., LR_A, LM_A, LU_A, The parameters LR, LM, LU, respectively, denote the smallest possible value of factor "A", the most promising value and the largest possible value that describe a fuzzy event. For each number in the pair-wise comparison matrix represents the subjective opinion of decision makers and is an ambiguous concept, fuzzy numbers work best to consolidate fragmented expert opinions. The triangular fuzzy numbers are established as follows:

$$\tilde{v}_A = (LR_A, LM_A, LU_A) \quad (1)$$
$$LR_A \le LM_A \le LU_A) \quad (2)$$
$$LR_A = Min(X_{Ai}), i = 1,2,3,...,n \quad (3)$$
$$LM_A = (X_{A1} * X_{A2} * ... X_{An})^{1/n}, i = 1,2,3,...,n \quad (4)$$
$$LU_A = Max(X_{Ai}), i = 1,2,3,...,n \quad (5)$$

Where A is problem (factor); $\tilde{v}_A$ is Triangles membership functions for "A"; i is expert; X_{Ai} is the evaluated number of "A" from the ith expert; LR_A is the smallest number of "A" from experts; LM_A is the mean number of "A" from experts; LU_A is the largest number of "A" from experts. Based on the Delphi method and fuzzy theory as above, the authors set a series of Delphi — questionnaires which used Likert Five-Point Scale (1: Strongly disagree, 2: disagree, 3: slightly agree, 4: agree, 5: strongly agree). This research would use 15 questionnaires, and the respondents are professional views of experts in several domains (such as CEOs & experts on such area). Totally there are 11 people replied the questionnaire which made the rate of returns and effectiveness as high as 73.3%. The authors totally operate the questionnaire by 3 times back and forth (since May 15, 2008). After the fuzzy theoretical calculation, selects suitable threshold "S", and the value of LM_A is larger than "S" (based on expert's opinion, set S=3). The final results totally eight elements are selected from 13 elements (others, such as heat transfer problem, are less than "S" value). Therefore, there are eight strategic elements as listed in table 2.

Table 2. The results of elements - triangular fuzzy numbers

Experts / Elements	1	2	3	4	5	6	7	8	9	10	11	LR_A	LM_A	LU_A
Enterprise														
The degree of equipment management	4	5	4	5	5	5	5	5	4	5	3	3	**4.5**	5
The degree of servers management	3	4	3	5	5	4	3	4	3	5	3	3	**3.7**	5
The degree of information security management	5	4	3	5	5	4	5	4	3	5	3	3	**4.1**	5
The degree of disaster recovery	5	4	5	5	5	5	3	4	5	5	3	3	**4.4**	5
Information service firms														
Quick Response mechanism	3	4	3	5	5	4	3	4	3	5	3	3	**3.7**	5
Service model	4	3	3	3	5	3	3	4	3	4	3	3	**3.4**	5
IT infrastructure	4	4	5	3	5	3	3	4	5	3	3	3	**3.7**	5
Security mechanism	5	5	5	3	5	4	3	4	5	3	5	4	**4.2**	5

5 Reference Frameworks

According to formerly analysis, this research proposed elements already might cover requirements for planing the data integrity. In other hand, using case study and fuzzy Delphi method, we point out four critical elements of business data integrity risk management framework as shown in Figure 2: the degree of equipment management, the degree of servers management, the degree of information security management, and the degree of disaster recovery. In order to enhance business data integrity, industry should focus on these elements to evaluate IT system risk level, such as IT infrastructure, disaster recovery or remedy system, security or safety system. Based on these factors and the factor relations, the enterprise may carry on the business data integrity risk management and evaluated their outsourcing plan.

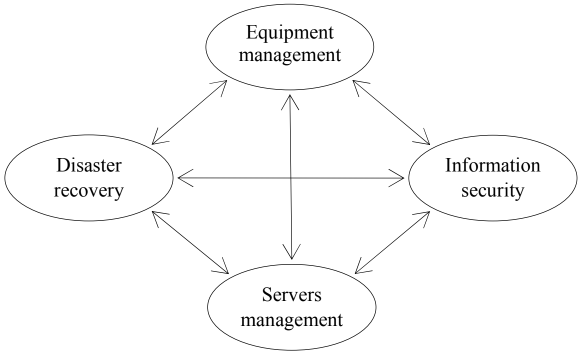


Figure 2. Business data integrity risk management framework

On the other hand, to keep a high service level, the data center firm should be concerned about inner and outer effect elements of information flow for each customer's business operation. Based on the analyses of business information security, case study and interview results, we suggest a data center service operative reference framework. It should draw out five dimensions as shown in Figure 3: quick response mechanism, service model, IT infrastructure, security mechanism , and knowledge database. The IT infrastructure and platform of data center should include strong network demanding equipments, facility and circumstances. The service module and specification of data center facility should construct flexible customer-oriented system functions. The security module and specification of data center facility have to set a sensitive monitoring system. All information of these should be tightly linked to knowledge base (or entities) operation kernel.

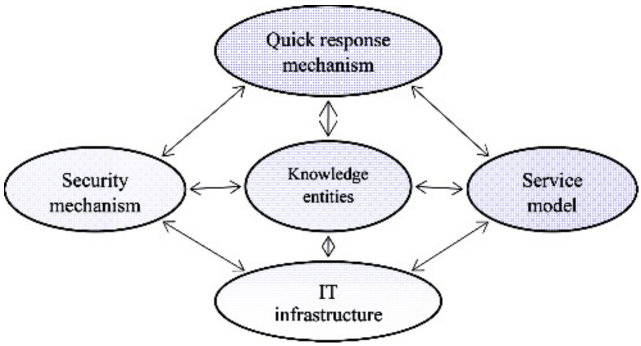


Figure 3. Data center service operation reference framework

5 Case Study

In order to verify the implementation of these elements for firms in business data integrity risk management. We selected a typical data center firm case in Taiwan, to analyze as below.

5.1 Background

The case company A is located in Asia Pacific Zone. Company A's PC brand is already a major worldwide business and it's headquarter set in Taiwan. Company A own worldwide revenue over US$10B in 2006. They focus on in-direct model that emphasizes the importance of selling through partners rather than selling directly to end-customers. Company A started to promote e-service since 2002. Their data center founded in 2000, business launched in early 2002. Because Company A is the largest data center in Asia-Pacific region, and company A's have a integrated facilities and high service quality in Asia. Their IT service's annual growth more then 30% in 2004 and 2005, and their revenue is US$20M (in Taiwan, 2005). Company A's data center has launched its international business in 2006. At that time, its IT service will maintain the "in-direct" business model of PCs and leverage the partners in local markets of countries. They also-seek opportunities in Japan and China by first identifying business partners in Japan and China. Company A is already the top choice of service providers for most international companies in Taiwan (such as Toyota, Ford, DeutscheBank, AstraZeneca, etc.). Besides, they also serve companies outside Taiwan through its data center facility in Taiwan. For example, Tyc-Healthcare, disaster recovery (a Japanese company), search engine, etc. Company A's data center own a robust system (including hardware and software entity). Their infrastructure of service is followed ITIL (ISO20000) and BS7799 (ISO17799) standards. Based on these IT standards, company A plan complete inner and outer IT service system. Additionally, Company A's own several better certification.

5.2 Data center management

The company A's data center is belong to its security operation center key strategy global market. By the data center operation, they also enhance market expend zone through IT service project. Base on this project, they hope to achieve the target (2006~2008): over NT$30B revenues produced more than two bigger national customer move operation center to Taiwan, etc. The data center operation goal has two targets: 1. to assist operation center of large enterprise through local dater center. Firstly, they focus market on Japan, Taiwan' enterprise in Asian market; 2. to serve Southeast Asia market in IT security business. The "safety operation" is the most important service kernel for customer. The framework (Figure 4) is consisting of six dimensions: operative on parts Its contain operation service platform, security infrastructure, technical operation team, IT service management specification, security management specification and service market target. There are several key business operations of company A as next section.

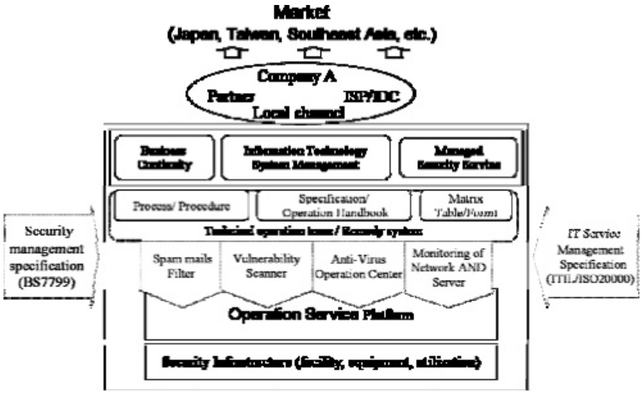


Figure 4. Company A's data center service operation framework

4.3 Infrastructure

Company A very near (less than 45km) from downtown Taipei, 40km from CKS international airport, elevation of 300m plateau precludes flooding potential, away from known faults, proximity to major highways and freeways, on-site hotel, dormitory, and condominium (with gymnasium and swimming pool, aerobic and exercise facilities), isolated rural surroundings for safety and access security, nearby several golf courses and resorts. Company A own a robust building, such as seismic design, pre-cast concrete construction, EMI-proof design(e.g., metal-enclosed data room, Metal deck allows efficient piping layout), power & cooling capacities, multi-tiered power design, redundant UPS design, fire suppression, cooling, CCTV monitoring and multi-tiered security screening, muti-communication features and auxiliary facilities. The company A's data center have 3-5 times of offices/factories in seismic strength, it can resist up to "Richter-scale 7" earthquakes and its collapse acceleration is 0.495g. Company A's data center is located The data center also have high standard of floor loading (the value: 1,000-kg/m^2 is larger than office:300-kg/m^2 or factory: 500-kg/m^2) Additionally, company A's layouts is: 1. five-floors, plus 6th-floor for facilities, 2. water reservoir over 6,500-ton capacity (for cooling), and 3.underground diesel fuel tank over 400m^3 capacity. There are several high level IT equipments: network operation center, security operation center, security surveillance center, and facility monitoring center.

4.4 IT Service Management

Company A followed ISO20000, implemented ITIL since 2002, IT service provider in Taiwan with ITIL practice only, 98% of data center staff passed external/internal fundamental certification, with three ITIL advanced managers certified. Company A's ITIL framework is shown in Figure 5.

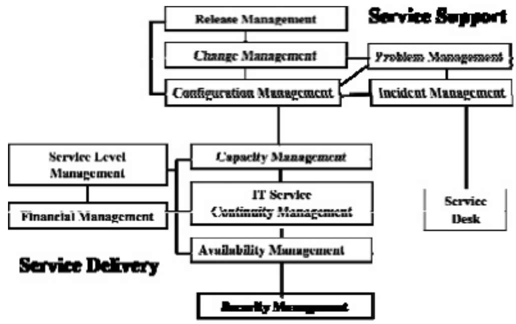


Figure 5. The ITIL framework of company A's data center

Company A's data center service operation platform include: Periodic data-mining on operation statistics to identify potential improvements; Runs matured incident/problem management; change/release/configuration management all in-place; CMDB (configuration management database) in production; remedy (BMC), the most comprehensive helpdesk system on the market shown in Figure 6.

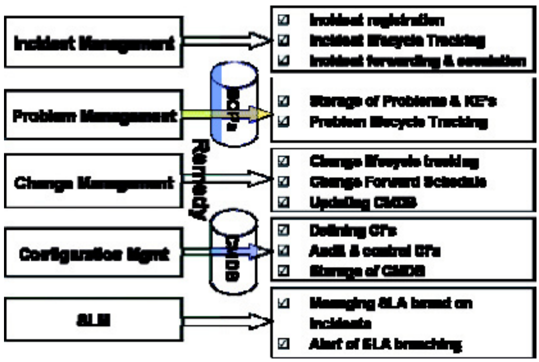


Figure 6. Company A's data center remedy system (compare to ITIL contents)

4.5 IT Security Management

Company A' security management followed ISO27001 standard is a well and large IT security service provider in Taiwan: 80% of government market, 60% of (emerging) commercial market, hardware/software-neutral, can work with most major models/brands. It's As regards to the monitoring & management service domain, company A has three strategies: security operation center (SOC), firewall, IDS/IDP (intrusion detection/protection system) and Web-site monitoring. The SOC solution adopted ArcSight and TeraGuard / VeriSign, called Managed Security Service (MSS), and the MSS has alerting, escalation, and response upon security incidents and attacks. The firewall and IDS/IDP play a role in monitoring and management for network and hosts. The Web-site monitoring can against defacing by hackers, with automatic web restoration. As regards to the auditing service, there are three function: vulnerability scanning, system hardening and penetration tests. Additionally, company A has consulting & professional service, including: consulting for the framework of information security, periodic healthy check and doing the log analysis.

4.6 Disaster Recovery Management and Host Servers

As regards to the disaster recovery management, company A signed technology transfer agreement with SunGard for disaster recovery (DR) (2002).SunGard is the leading service provider of DR in North America, with more then 50% of market share, as wall as Company A has over 200 man-days trained at SunGard facility in US. Company A is the partner and co-brand with SunGard in the Asia Pacific region. Company A is a large scale service provider in Taiwan, they provide: data center space, command center space (space for performing DR operations), customers and delegated DR operations to company A' data center (no need to present at data center) and leasing for shared/dedicated IT equipments.

The procedure of company A's security/safety for managed hosts also followed ISO20000. Their customer can delegate to data center or through remote control model. Their management is performed via secured lines and jump servers with logging of all actions for auditing. This system is 7x24x365 no-stopping servicing.

6 Conclusion and Disscusion

This research provides a useful framework to plan and analyze the business data integrity risk Management in IT security & service policy for a company. It has pointed a performed architecture from data center's firm and user (company), respectively. Their detail operation items includes operation service platform, security infrastructure, technical operation team, IT service management specification, security management specification & service market target and performance, etc.

Based on literature review, case study, deeply interview with 11 top CEOs & experts, and Fuzzy Delphi method, we came up with two frameworks: business data integrity risk management framework (from enterprise risk viewpoint) and data center service operation reference framework (from data center firm's viewpoint). Although these two frameworks are different in Figure 7&8, they are in the same aspects of business data integrity risk management base; in other words, these two frameworks have mapping effects. However, in order to reduce risk from the disaster, firms could utilize the research results and the reference framework to develop a data integrity risk management policy, or to evaluate data center's security policy and service effectiveness.

We also use this framework to point out several critical contents by the reference framework. In other words, this framework is suitable for firms. On the other hand, base on this analysis, we suggested company A had better develop a more flexible service mechanism such as "management security service", "information system management outsourcing", "business continuity" and "facility Outsourcing", to fulfill other new market needs (such as US, Europe…,etc.).

Based on this paper, for future research, other research can be further improved by including more others factor and through statistics, decision making analysis, optimal method, etc. We believe that the further research can constitute a more complete and solid research in this issue.

7 Acknowledgements

We are so grateful to get many helpful suggestions from CEOs and experts which work for the famous information service & risk management corporations and the ministries of government in Taiwan.

8 References

[1] BISHOP, T. Building a Better Data Center -- The dos and don'ts of state-of-the-art data center design. Wall Street & Technology. Vol. 26, No.10, pp.36, 2008.
[2] CARTER, K., PULTORAK, D. (2003) ITIL, COBIT, CMMI, Balanced Scorecard, IS-9000, Six Sigma: Putting Them All Together.
[3] CHANG, P.T., HUANG, L.C., AND LIN, H.J. The fuzzy Delphi via fuzzy statistics and membership function fitting and an application to human resources. Fuzzy Sets and Systems. vol.112, pp.511-520, 2000.
[4] ELOFF, J. H. P., ELOFF, M. Integrated Information security Architecture. Computer Fraud and security. Vol.11,pp.10-16, 2005.
[5] GARG, A., CURTIS, J., & HALPER, H. Quantifying the financial impact of information security breaches. Information Management and Computer Security. Vol.77, No.2, pp.74-83, 2003.
[6] GLAZER, R. Measuring the value of information: The information intensive organization. IBM Systems Journa. Vol.32, No1, pp.99-110, 1993.
[7] HINDE, S. Security surveys spring crop. Computers and Security. Vol.27, No.4, pp.310-321, 2002.
[8] ISO/IEC 17799. BS 7799-1. (2005). Information technology/security techniques/Code of practice for information security management. Britain.
[9] LINDENMAYER, G. Information security standards: The 10 Keys to protecting your network. Risk Management. Vol. 54, No. 12, pp11, 2007.
[10] McCARTHY, M. P. and Campbell, S. (2001). Security transformation. McGraw-Hill: New York.
[11] Ovum research(2007). http://www.ovum.com/expertise/itservices/
[12] STRAUB, D.W. & WELKE, RJ. Coping with systems risk: Security planning models for management decision making. MIS Quarterly, Vol.22, No.4, pp.441-470, 1998.
[13] VEIGA, A. D. and Eloff, J.H.P. An Information Security Governance Framework. Information Systems Management, Boston: Fall, Vol.24, No.4, pp. 361-73, 2007.
[14] von SOLMS, B. & von SOLMS, R. The ten deadly sins of information security management. Computers & Security. Vol.25, pp.371-376. 2004.
[15] YIN, R.K. (1994). Case Study Research: Design and Methods, California: SAGE Publications.
[16] ZADEH, L.A., Fuzzy Sets. Information Control , 8, 338–353, 1965.

The Key Dimensions for Information Service Industry in Entering Global Market: a Fuzzy-Delphi & AHP Approach

M. K. Chen[a], Shih-Ching Wang[b,1]

[a] Graduate Institute of Commerce Automation and Management, National Taipei University of Technology.
[b] Graduate Institute of Industrial and Business Management, National Taipei University of Technology.

Abstract: Due to the trend to globalization and multi-country companies penetrating into domestic market, Taiwan medium-sized IT service firm (as it grows to a certain scale) had better to develop international market in order to maintain competetive advantage. Rajala (2003, 2006) et al. applied qualitative research and case-study to explore four major dimensions: profit model, product strategy, marketing channels and services. The four dimensions focused mostly on individual firm's operation. But in order to catch the globalization trend of information services industry, we need to have a comprehensive model to enter into the international market. This research, through Delphi method theory and AHP approach, would extract the key business dimensions provided by the experts and the enterprise decision-makers in Taiwan. We proposed six dimensions and 14 sub-dimensions and pointed out the weights of each through AHP appraoch. We believe our porposal will be a better model for developing the international market than the four dimensions suggested by Rajala et al.

Finally, we hope that the results and recommendations of this research will help the information services industry develop the international market more successfully.

Keywords: information service industry; cross-case study; fuzzy Delphi method; analytic hierarchy process (AHP)

1. Introduction

Since business operations are changed and moved to other markets rapidly, internationalization should be an imperative strategy for information service firms, especially, for the place where there are very limited domestic markets for information service firms [2, 3] such as Taiwan. Over the past few years, to promote global target market, several Taiwan firms have developed various

[1] Corresponding Author E-mail: shihching.wang@gmail.com, Graduate Institute of Industrial and Business Management, National Taipei University of Technology, or, Industrial Development Bureau, Ministry of Economic Affairs, Taiwan. Address: 3, Sec. 5F, 41-3, Sinyi Rd., Taipei 106, Taiwan, fax: +886-2-27020406. 2.

software products. However, Taiwan is located in a special area of the Asia's major economic region. Therefore, Taiwan should assess the potential information service sectors and find the best software service solution to enter the global market. On the other hand, Rajala et al. [20, 21] proposed a model more focused on the internal operations of individual business. So it is still insufficient in providing the factors for considering how to develop global market. The key concern of the business operation model is to consider both the internal and external operations for making profits on selling [14]. Hence it is necessary to provide more sufficient operational dimensions for the industry. Overall, the research on operational issues of information service industry to develop global market is still in immature stage and most relevant studies are more inclined toward qualitative approaches. Based on these literature and our observations of the development of information service industry, this study aims to propose key business operative dimensions for information service industry to develop international markets. We extract several key dimensions through cross-case study and fuzzy Delphi method [26]. After the operational dimension was found, we could figure out the weight of each dimension through AHP process. Finally, we would propose a referential framework for information service industry to develop global market.

2 Business Operative Dimensions

Generally, business model means "how a firm works, who its customers are, what the customer value, and how a firm makes money" [6, 14, 20]. However, as comparing to physical product, software product is generally expensive to produce, but very cheap to reproduce [12, 17], and software includes working solutions (in pre-sales) and services [4, 9]. Software implementation always combines with the service contents. Therefore, the software industry is also called "information service industry". In Taiwan, the category of the information service industry includes three types: package, customization, and project. From reviewing the earlier literature, there are few research focused on identifying key dimensions for investors to enter global market [13]. Nevertheless, some researchers have studied different aspects on the influences of internationalization; for example, customization [4, 15], intangible assets [3], pricing [8] and product offerings [2]. Based on Rajala et al. [20], Ojala and Tyrväinen [19] proposed to analyze the entry model for software firm to go into Japanese market. And he pointed out that product strategy is a key for software internationalization. Moreover, some other researches that quote the framework Rajala et al. proposed [20]. Rajala et al. proposed 4 operative dimensions: product strategy, revenue logic, distribution model, and service and implementation model. "Product strategy" is a firm's product offering which has various types such as providing customers with specific, highly standardized products [11]. "Revenue logic" is the rule a firm gains profits from selling products and services. "Distribution model" is the channel a firm uses to get its products and services to the market [1]. "Services/implementation model described" is the product needs to be installed, implemented, maintained, and supported. Whatever, Rajala et al. [20, 21] emphasized the concerns of individual firm businesses, but short of considering other key dimensions for entering global market. For example, market size and culture [18], outsourcing [23, 24], service and consulting [10], domain knowledge [21, 23] and partnership [5]. Therefore, it is also need to more carefully consider other dimensions, such as target market, strategic cooperation, talented service people, fast duplication of product, etc.

3 Research Method

Due to the business model of information service industry to develop global market is still in an exploratory and premature stage, Based on the preceding literature, the research would use cross-case studies [9, 25] method to analyze several critical dimensions of success among Taiwan's benchmark information service firms. Because this issue involves much more firm's strategy by decision-maker (such as CEO). However, Delphi method is fit approach to this issue, and it is an expert opinion survey method with three features: anonymous response, iteration and controlled feedback and finally statistical group response. Unfortunately, some weaknesses have been exposed, it needs repetitive surveys to allow forecasting values to converge which requires several time and cost. Furthermore, in many real states, experts' judgments cannot be properly reflected in quantitative terms. Since people use linguistic terms, such as 'good' or 'very good' to reflect their preferences, the concept of combining fuzzy set theory and Delphi was proposed by Murray, Pipino, and Gigch (1985). To improve the shortcomings of vagueness and ambiguousness of Delphi, group membership functions were used to compute the expected values of fuzzy numbers, and then a forecasting value could be obtained. Therefore, we extract the most important business dimensions through Fuzzy Delphi method. Furthermore, we would build up hierarchy business model and weights of each dimension through the AHP approach. Finally, we would propose a referential framework for firms as they plan to develop global market.

3.1 Cross-Case Study

In order to find the key dimensions of information service industry in developing global market ,we study prior literatures and collate these dimensions through cross-case studies are joining the "information services flagship-program" (also called BEST project) in Taiwan. The program is encouraged by the Industrial Development Bureau of Ministry of Economic Affairs (MOEAIDB), Taiwan, and its main purpose is to push the backbone firms onto activating the cooperated firms to form a cooperation strategic way for opening international target markets. Totally there are over 8 bigger flagship companies and 42 SEM partners joining this program, the gross of all those firms' scales has surpassed more than half of the information service industry in Taiwan and has got the representative status. We use multi-case study to interviews over 30 CEOs, 15 experts and consultants from May 25, 2008 to May 15, 2009. Each CEO or expert is a key leader in BEST projects. Each project has four key men to handle the main & secondary program in their BEST project. The key issues, such as target market, customers, co-

competition and partner, are listed in Table 1. Then, content analysis principle is used on those data and the key observations are listed in Table 2.

Table 1. Summary of cases

Case	Employees	Customer type / Target market	Number of partner (Service contents)
Case 1	1044	Manufacturing (SME) / China, East-Southern Asia	5 (GSCM, Logistics, e-invoice, KM, e-Learning)
Case 2	414	TFT-LCD, Traditional / Manufacture,Government	5 (HR, Web report, ERP, PLM, e-Learning)
Case 3	290	Steel / Mainland China, east-southern Asia (Malaysia)	4(KM, marketing partner, marketing system, search engine)
Case 4	226	Petroleum &Chemical, Steel, Semiconductor /Mainland China	4(network communication solution, KM, expert system platform, MES)
Case 5	515	Bank & Securities / Mainland China, east-southern Asia	4(data mining, domain know-how, consultant, e-learning)
Case 6	1017	ICT,SI, Government, Traffic/Asia, Middle East, Europe	3(PKI,KM, search engine)
Case 7	1700	Bank & Securities, financial,manufacturing & logistics/ China, east-southern Asia	4(SI, KM,CRM,BI)
Case 8	295	Bank,& Clearing, 7-11 retailer, Expense finance / Thai, Hong Kong, China	3(SI, reader card machine, card)

Note: These benchmark firms are qualified by Taiwan, Industrial Development Bureau, Ministry of Economic Affairs (MOEAIDB),which are joining the Flagship-program.

Table 2. Cases's observations

Key observations	Obvious case
When the advance of core competition is clearer and the technique of the products is more powerful, it will be easier to promote global markets.	Case 8
The quality of the products and services influences the market response and market image, and also has a positive impact on international orders.	Case 5
Strategic alliance becomes a collaborative system for entering the global market. Mutual-support brings multi-effects for developing markets.	Case 1
Familiarity with international market scale, market culture, language and regional specialty and realizing news of competitors are helpful in making investment decisions.	Case 1,5
The efficiency of the distributive strategy and the capability of cost management will impact on extending the market. Pricing models for products and service impacts profits greatly.	Case 3
The knowledge and manpower for the product and service is one of the main pillars of globalization.	All
Others issue (such as knowledge, manpower, etc.)	Case 1,3,5,7

3.2 Fuzzy - Delphi Method

The purpose of this study is to find specific business operative dimensions for information service industry to develop international markets. In other words, we had listed the operating dimensions of the initial content as described in section 3.1, but the significance of each dimension still need to be quantified further. Nevertheless, the business operative dimension involves much strategic policy and higher order decision-making. Therefore, we used fuzzy - Delphi method to find those dimensions. Delphi method is an applicable research instrument when there is incomplete knowledge about an issue. Delphi method is an iterative process collecting and refining the expert judgments. Furthermore, in order to prevent the extreme value of experts' opinions from no convergence, the thesis has used fuzzy theory to substitute for traditional geometric average of Delphi method to search out more exact weight of the dimension. A fuzzy set is a class of objects with a continuum of grades of membership [22]. In Fuzzy set theory, membership function is one of the basic concepts, which assigns to each object a grade of membership ranging between zero and one. Researchers always consider different problems (or issue) with different membership function; the most used membership function includes Triangles membership functions, trapezoid membership functions and Gauss membership functions. This research would adopt Triangles membership functions for our primary membership functions. The equation of triangles membership functions include three parameters, i.e., LR_A, LM_A, LU_A, shown as Figure. 1. The parameters LR_A, LM_A, LU_A, respectively, denote the smallest possible value of factor "A", the most promising value and the largest possible value that describes a fuzzy event. For each number in the pair-wise comparison matrix represents the subjective opinion of decision makers and is an ambiguous concept, fuzzy numbers work best to consolidate fragmented expert opinions. The triangular fuzzy numbers are established as follows:

$$\tilde{v}_A = (LR_A, LM_A, LU_A) \quad (1)$$

$$LR_A \le LM_A \le LU_A) \quad (2)$$

$$LR_A = Min(X_{Ai}), i = 1,2,3,...,n \quad (3)$$

$$LM_A = (X_{A1} * X_{A2} * ... X_{An})^{1/n} \quad (4)$$

$$LU_A = Max(X_{Ai}), i = 1,2,3,...,n \quad (5)$$

Where A is problem (factor); $\tilde{v}_A$ is Triangles membership functions for "A"; i is expert; X_{Ai} is the evaluated number of "A" from the ith expert; LR_A is the smallest number of "A" from experts; LM_A is the mean number of "A" from experts; LU_A is the largest number of "A" from experts.

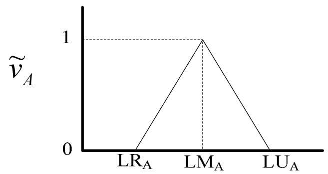


Figure 1. Tri-angle membership function

Based on the Delphi method and fuzzy theory as above, we set a series of Delphi – questionnaires, and used five point types (1: Strongly disagree, 2: disagree, 3: slightly agree, 4: agree, 5: strongly agree). This research would use 25 questionnaires, and the respondents are professional views of experts in several domains (such as CEOs & experts). Totally has 22 people replied to the questionnaire which made the rate of returns and effectiveness as high as 88.0%.

We totally operate the questionnaire 3 times back and forth (about 2 months). According to the data from the questionnaire, the author figured out the weight of fuzzy numbers-LM_A value of dimensions through software (Microsoft excel 2003). The final results of totally six dimensions are extracted from 13 dimensions (others as human resource, intellectual property, etc) as listed in Table 3.

Table 3. The final fuzzy Delphi experts reply results lists

Experts / Dimensions	1	2	3	4	5	6	7	8	9	...	22	LR_A	LM_A	LU_A
Market segment	4	5	4	5	5	5	5	5	4	...	3	3	4.491*	5
Strategic alliance	3	4	3	5	5	4	3	4	3	...	3	3	3.730*	5
Service model	5	4	3	5	5	4	5	4	3	...	3	3	4.093*	5
Product strategy	5	4	3	5	5	4	5	4	3	...	3	3	4.093*	5
Distribution/channel model	5	4	5	5	5	5	3	4	5	...	3	3	4.375*	5
Revenue strategy	3	4	3	5	5	4	3	4	3	...	3	3	3.730	5
New techonlogy application	2	2	3	2	1	3	3	2	3	...	3	3	2.258	3
Investment strategy	3	2	2	2	3	3	3	2	3	...	2	3	2.405	3
Intellectual property	3	3	2	1	2	3	2	1	3	...	2	3	2.043	3
Manpower training	2	1	3	2	2	3	1	3	3	...	3	3	2.120	3

Note: Mark "*" is selected dimensions (i.e., over the threshold "S" vale = 3.0)

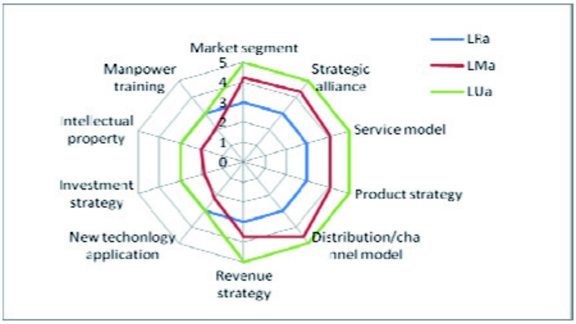


Figure 2. The fuzzy numbers-LMa value of elements

Based on Delphi method, we require set suitable threshold "S=3.0" (i.e., expert's viewpoint) and select the value of LM_A that is larger than "S" from Table 3. Obviously, these selected dimensions are the most influential factor for information service industry to open global market (Figure. 2). After obtained the main dimensions as above, we also run the same procedure through fuzzy Delphi method to propose 14 sub-dimensions of each dimension as shown in Figure 3. We found the largest LM_A - value is distribution /channel model, which three sub-dimensions' total LM_A - value is 12.895, also is the largest among 14 items. This means that the results of sub-dimension fit the main dimension computing results as above. In accordance with the above cross-case observation and fuzzy Delphi method, we have extracted six key dimensions and 14 sub-dimensions for the information service industry to open the global market. All those results are adequately integrated with the consensus of experts and business decision-makers; it has reached a higher experts and content validity.

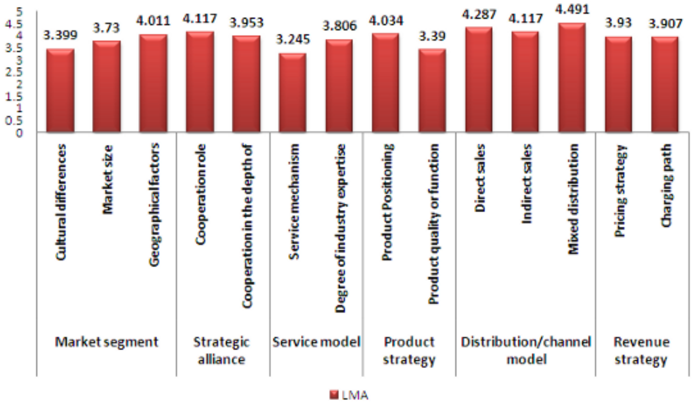


Figure 3. The fuzzy numbers-LM_A value of sub-dimensions

Compared to Rajala [22] who proposed 4 business operative dimensions, the results of this study not only add dimensions of the "market segmentation" and "strategic alliances", but also increase 14 sub-dimensions to form a more complete new business operative framework of information service industry in developing global market as shown in Figure 4. Apparently, it is more convenient for the information service industry in developing global market.

3.3 Dimensions' Weights: an AHP Procedure

According to the new operating architecture, information services sector in expanding the international market also need to consider each dimension in the input of resources available in order to exert the greatest effect of business operations. For this reason, we quantified weight of each dimension through the AHP method the proportion of value [27, 28]. AHP approach is a qualitative and quantitative method which is useful for evaluating the complex multiple criteria involving subjective judgment; AHP is an especially symmetric way that could transform complex problems into simple hierarchic structure. A decision-maker should determine the weights by conducting pair-wise comparisons between two criteria. Therefore, we propose to analyze the critical factors of success for developing international markets through the AHP approach. The procedures of AHP are:

(1) Drawing a pair-wise comparison matrix. An attribute compared with itself is always attributed to value 1(i.e., the main diagonal entries of the pair-wise comparison matrix are 1). Numbers 3, 5, 7, and 9 mean 'general importance', 'strong importance', 'very important', and 'absolutely important'; and 2, 4, 6, and 8 for compromise between 3, 5, 7, 9. If it has "m" attributes, the pair-wise comparisons would form a square matrix as matrix M:

$$M = \left[a_{ij}\right] = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1m} \\ a_{21} & 1 & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1m} \\ 1/a_{12} & 1 & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 1/a_{1m} & 1/a_{2m} & \cdots & 1 \end{bmatrix}$$

Where $a_{ij} = 1$ and $a_{ji} = 1/a_{ij}$; i,j = 1, 2, … m.

(2) Obtaining the relative normalized weight (wj) of each attribute by calculating the geometric mean (GM) of the i^{th} row normalizes the geometric means of rows in the comparison matrix. The GM method of AHP is used to find out the relative normalized weights of the attributes because of its simplicity and ease to find out the maximum eigenvalue and reduce the inconsistency in judgments.

- Matrix M, the problem involves assigning a set of numerical weights w_1; w_2; . . . ; w_m to the m criteria a_1; a_2; . . . ; a_m that ''reflects the recorded judgments".
- If A is a consistency matrix, then the relations between weights w_j and judgments a_{ij} would be simply given by $w_i / w_j = a_{ij}$ (for i, j= 1, 2, . . . ,m).

(3) Finding out the maximum eigenvalue λmax, Saaty (1988) suggested the largest eigenvalue is:

$$\lambda_{\max} = \sum_{j}^{n} a_{ij} \frac{w_i}{w_j} \qquad (1)$$

(4) Calculating the consistency index as equation (2). The smaller the value of *CI*, the smaller is the deviation from the consistency. The consistency in the judgments of relative importance of attributes reflects the cognition of the analyst.

$$CI = \frac{\lambda_{\max} - m}{m-1} \qquad (2)$$

(5) Obtaining the random index (RI) for the number of attributes used in decision-making. Calculating the consistency ratio CR= CI/RI. Usually, a CR of 0.1 or less is considered as acceptable and reflects an informed judgment that could be attributed to the knowledge of the analyst.

4 Reference Framework

Based on AHP approach as above, we use business operative framework of information service industry in developing global market as section 3.2, and propose the hierarchic structure layer in Figure 4. In Figure 4, the first layer is the issue goal, the second layer is the first criteria and the third layer is the sub-criteria. Under such framework, we summarize 35 questionnaires and calculate each criteria weight by AHP software. Based on this result, information service companies could adopt these weights of criteria to enter the global target market. Besides these three layers, whether the decision-maker should consider the fourth layer or not is depended on their firm's business operative needs, and we may also calculates the weights through AHP process again.

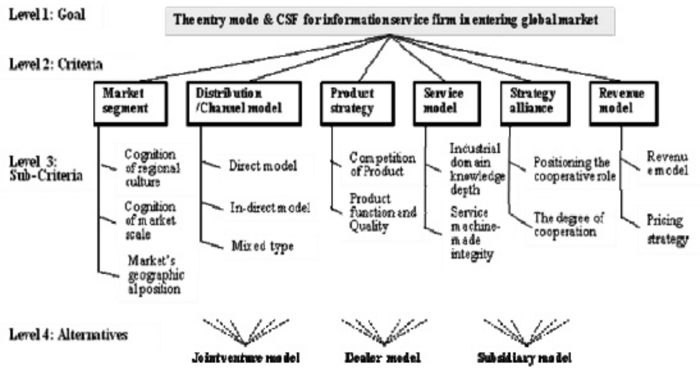


Figure 4. The AHP hierarchy framework

We use 30 AHP questionnaires in this research and all of the respondents are familiar with the domain of information service industry for over 10 years. There are 27 people replied to the questionnaire which made the rate of returns and effectiveness as high as 90.0%. According to the turnout from the questionnaire, the author figured out the weight of each item by AHP software (Choice-Maker 2002 and Microsoft excel 2003). After computing, I found that nearly all replies to the questionnaire reached a consistency ratio (i.e. CR =CI/RI) less than 0.1, hence this result is acceptable. We also listed overall weight of the criteria & sub-criteria as shown in Figure 5. In Figure 5, the ranking of the main criteria's weights are: product competition (0.2673), market segment (0.2543), service model (0.164), revenue efficiency (0.122), strategy alliance (0.102) and distribution/channel model (0.091).

These results pointed out that the most influential criterion for the information services industry to develop global market is product strategy. This means the company should strengthen the product strategy, including product positioning and product quality or function of the layout strategy. The second significant factor is market segment. Therefore, company need to understand more about the cultural differences, market size and geographical factors while comprehensively developing global market. The least influential factor listed is distribution/channel model. That is the optimal channel selected=based on others five dimensions (i.e., criteria). Among these criteria, there are three (i.e., the product competition, market segment and service model) whose weights get over 50% in all dimensions. The weights of sub-criteria are listed in Figure 6.

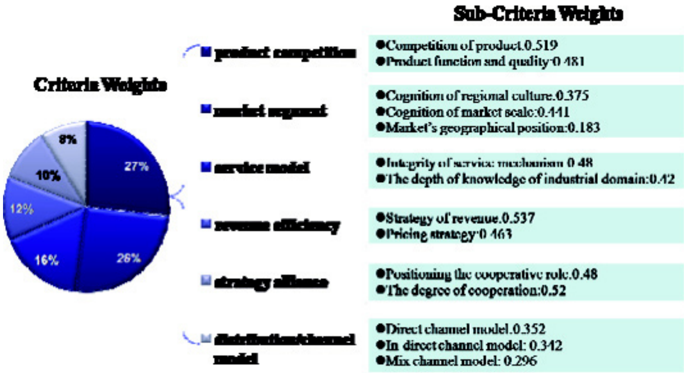


Figure 5. The AHP weights of criteria and sub-criteria (local value)

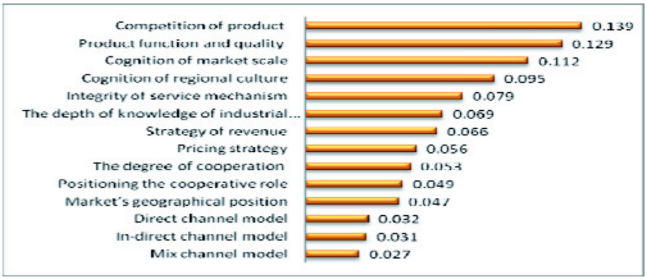


Figure 6. The AHP weights of sub-criteria (global value)

5 Conclusion and Future Work

The main purpose of this study is to identify the operating dimension for information services industry in developing global markets. We had proposed six main dimensions (by order): product competition, market segment, service model, revenue efficiency, strategy alliance and distribution/channel model.

We also pointed out 14 sub-dimensions: competition of product, product function and quality, cognition of regional culture, cognition of market scale, market's geographical position, integrity of service mechanism, the depth of knowledge of industrial domain, strategy of revenue, pricing strategy, positioning the cooperative role, the degree of cooperation, direct channel model, in-direct channel model and mix channel model. This result is based on the model proposed by Rajara et al. which includes the observations of industry, cross-case studies, deeply interview of CEOs and run fuzzy-Delphi procedure. With these dimensions, we built up a useful architecture. On the other hand, we also found weights of each dimension through AHP approach. For the future, based on this framework, we suggest to focus on correlative research, such as excogitating the types of information service industries or the domain of IT application.

6 Acknowledgements

We are so grateful to get helpful suggestions from CEOs and experts, which are famous for software corporations and the ministries of government in Taiwan.

7 References

[1] Barcus, A. and Montibeller, G. Supporting the allocation of software development work in distributed teams with multi-criteria decision analysis. Omega, 36(3), 464-475, 2008.

[2] Bell, J. A comparative study of the export problems of small computer software exporters in Finland. Ireland and Norway. Int. Bus Rev. 6(6),585–604, 1997.

[3] Bieberstein, N., Bose, S. Walker, L. and Lynch,A. Impact of service-oriented architecture on enterprise systems, organizational structures, and individuals. IBM Systems Journal. 44(4), 691-709, 2005.

[4] Burgel, O, and Murray, G. C. The international market entry choice of start-up companies in high-technology industries. J. Int Market. 8(2), 33–62, 2000.

[5] Coviello, N. and Munro, H. Network relationships and the internationalisation process of small software firms', International Business Review. 6(4), 361-386, 1997.

[6] Hedman, J. and Kalling, T. The business model concept: theoretical underpinnings and empirical illustrations. European Journal of Information Systems. 12. 49–59, 2003.

[7] Hoch D, Roeding C, Purkert G, Lindner S, M¨uller R. 2000. Secrets of software success: management insights from 100 software firms around the world. Harvard Business School Press, USA

[8] Jain, S. and Kannan, P.K. Pricing of information products on online servers: issues, models, and analysis. Management Science. 48(9), 1123–1142, 2002.

[9] Gable, G. G. Integrating Case Study and Survey Research Methods：An Example in Information Systems. European Journal of Information Systems. 3(2), 112–126, 1994.

[10] Ganek, A. and Kloeckner, K. An overview of IBM Service Management. IBM Systems Journal. 46(3), 375-386, 2007.

[11] Krishnan, M. S., Kriebel, C. H., Kekre, S. and Mukhopadhyay, T. An Empirical Analysis of Productivity and Quality in Software Products. Management Science, 46(6), 745-759, 2000.

[12] Krishnamurthy, S. 2003. E-Commerce Management, Text and Cases, Thompson Learning

[13] Leidecker, J. K. and Bruno, A.V. Identifying and Using Critical Success Factors. Long Range Planning. 17(1), 23, 1984.

[14] Magretta, J. Why business models matter, Harvard Business Review, 80(5), 86-92, 2002.

[15] McNaughton, R.B. Foreign market channel integration decisions of canadian computer software firms. Int Bus Rev. 5(1), 23–52, 1996.

[16] Morris, M., Schindehutte, M. and Allen, J. The entrepreneur's business model: toward aunified perspective. Journal of Business Research. 58, 726–735, 2005.

[17] Nambisan, S. Why Service Business Are Not Product Businesses. MIT Sloan Management Review. 42(4), 72-80, 2001.

[18] Ojala, A. and Tyrväinen, P. Business Models and Market Entry Mode Choice of Small Software Firms', Journal of International Entrepreneurship. 4(3), 69-81, 2006.

[19] Ojala, A. and Tyrväinen, P. Market Entry and Priority of Small and Medium-Sized Enterprises in the Software Industry: An Empirical Analysis of Cultural Distance,

Geographic Distance, and Market Size', Journal of International Marketing. 15(3), 23-149, 2007.
[20] Rajala, R., Rossi M, and Tuunainen V.K. Software vendor's business model dynamics case: TradeSys', Annals of Cases on Information Technology. 5(1), 538–548,2003.
[21] Rajala, R. and Westerlund, M. a business model perspective on knowledge-intensive services in the software industry. International Journal of Technoentrepreneurship. 1(1), 1-19, 2007.
[22] Saaty, T.L. How to make a decision: the analytic decision processes. Interfaces. 24(6), 19-43,1994.
[23] Tiwana, A. An empirical study of the effect of knowledge integration on software development performance', Information & Software Technology. 46(13), 899-906, 2004.
[24] Willcocks, L. and Choi, C. J. Co-operative partnership and 'total' IT outsourcing: From contractual obligation to strategic alliance? European Management Journal. 13(1), 67-78, 1995.
[25] Winkler, J. K, Dibbern, J., and Heinzl, A. The impact of cultural differences in offshore outsourcing: Case study results from German-Indian application development projects. Information Systems Frontiers. 10(2), 243-259, 2008
[26] Yin, R.K. Case Study Research: Design and Methods. California: SAGE Publications. 1994.
[27] Zadeh, L.A., Fuzzy Sets. Information Control , 8, 338–353, 1965.

Problem-Based Construction of Engineering Curricula for Multidisciplinary and Concurrent Engineering Practice

Gloria Lucía Giraldo[a,1] and German Urrego-Giraldo[b]

[a] Universidad Nacional de Colombia, Escuela de Sistemas, Colombia.
[b] Universidad de Antioquia, Dpto. Ingeniería de Sistemas, Colombia.

Abstract: This paper presents an approach to construct the study plans of engineering programs from problems identifed in social context. A model of four human training dimension: to be, to know, to do and to cooperate, is used, aiming to define the curricular thematic units. This model leads the multidisciplinary and concurrent engineering practice.

Keywords: problem, training dimension, curriculum, concurrent engineering

1 Introduction

Within every pedagogical conception many engineering curricula models are created. They reproduce, at higher or lower levels, the same shortcomings, some of them limiting the multidisciplinary and concurrent engineering practice, such as: lack of global and complex thinking, deficiency in the capacity of modelling, lack of skills for the management of technology; lack of contact points among different engineering fields; lack of some human dimensions in the training, etc.

The problem concept is used in techniques known as Problem Based Learning (PBL) as referred in [1]. PBL constitutes a training strategy in which students solve problems in collaborative groups. Some applications of this concept, in engineering curricula, are presented in [2]. Many approaches use the concept of problem as a learning strategy or to construct individual curricular units using the PBL concepts, but we did not find models which systematically obtain the whole program of studies from social context problems. A set of four training dimensions is used aiming to promote physical, intellectual, emotional, affective, social, and human features, and to benefit of pertinent and meaningful curricula knowledge. This knowledge train for good human and social beings able to apply multidisciplinary and concurrent engineering concepts, for intervening the realities with scientific, technical, and ethical criteria. The use of training dimensions: to be, to know, to do, and to cooperate, for structuring the thematic areas, is also a contribution of our

[1] Corresponding Author Email : glgiraldog@unal.edu.co

research. These dimensions, the concepts of problem, and training purposes are the core of the proposed curricula models. The concept of problem expresses the needs identified in social and organizational contexts. The human dimensions of training constitute the framework for entailing relationships among the engineer and culture, organizations, products and processes. The concept of training dimensions links the problems that gave rise to the program with the training purposes and the thematic concepts necessary to find the required solutions. Problems in engineering domain imply, at the same time, social, scientific, technological, and economic aspects, which claim for multidisciplinary and concurrent engineering interventions. Our proposal refers to the problems and the knowledge required for developing curricula, which enable multidisciplinary and concurrent engineering practice.

Multidisciplinary and concurrent engineering pose additional requirements to the social context problems that give birth to engineering programs. An extract of a general model for this type of engineering let enhance the questions about the social context problems. In this way the training purposes, which answer these questions, include more accurate elements focused on the features of multidisciplinary and concurrent engineering. These enriched learning purposes determine, in Thematic Units of the program, specific concepts related to the practice of the mentioned engineering.

The proposed problem-based model for engineering curricula construction, presented in this article, drives and verifies the correspondence between the requirements of the multidisciplinary and concurrent engineering and the pertinent contents within Thematic Units.

In Section 2, this article describes the dimensions of training for multidisciplinary and concurrent engineering practice. Section 3 explains the model for obtaining the thematic units of the curriculum from the problems, which the curriculum should solve. The determination of the thematic concepts and units is explained in Section 4. Conclusion and future research as well as bibliography are presented in sections 5 and 6.

2 Training Dimensions for Multidisciplinary and Concurrent Engineering Practice

The evolution of society and knowledge sets out new requirements for engineers in relation to their abilities to incorporate and integrate knowledge, models, methods and tools in the products and processes of their fields of activity. Every engineer requires new abilities for contributing to human, scientific, and technological progress and for establishing points of contact among different fields of science and technology to address complex problems in changing situations. The multidisciplinary and concurrent engineering pose additional training requirements to the problems that each engineering branch should solve. In order to comply with these requirements we have constructed a requirement model, which groups training requirements in three categories: *technical issues*, *human interventions*, and *product lifecycle*. Some specific requirements in these categories are: ubiquity (characterized by place and platform independence, real time, multi-communication media access, tele-communication, and tele-working facilities), interrelated tasks, technological solvency, systemic thinking, product lifecycle management, process integration, team spirit, communicability, etc.. The whole requirement model is not presented here. These requirements and the questions about the problems identified in the social context gave birth to the *training purposes* centered in four training human dimensions: *To be*, *to know*, *to do*, and *to cooperate*. These training dimensions, defined in Table 1, orient the definition of *thematic unit* contents which contribute for developing the intellectual, physical, emotional, social, and attitudinal human faculties to face the solution of social context problems and training requirements. The model, which embodies all these elements, is described in the following section.

The four *training dimensions* (to be, to know, to do, and to cooperate), Table 1, gather and enhace the core elements of cognitive approaches which involve competences, originally discussed in the cognitive framework and then extended to education and labor field [3].

Table 1. Training Dimensions

To be Aims to develop intellectual, emotional, and ethical capacities for acting independently and responsibly.
To know Seeks the growing up of empirical, technical, technological, or scientific knowledge to address the solution of problems from one or several disciplines, acting either on products or on processes.
To Do Favors the development of skills for passing from abstractions, ideas, and reflexions to concretions, assembly, and materialisation of images of an object or phenomenon. It is then the deployment of the potential for developing, using knowledge, experience, and intuition.
To cooperate It aims to help train for the interaction with society, nature, culture, organizations, and processes, with reason, dignity, solidarity and justice criteria. It is the development of the system thinking, the sprit of teamwork and the intellectual autonomy for intervening the whole processes.

These dimensions characterize the training purposes, which lead to pertinent contents of thematic areas and units. To cooperate, for example point to contents, which favour the teamwork, system thinking, the autonomy and the view of the whole process, etc., determinant features for the multidisciplinary and concurrent engineering practice

3 Model Description for the Acquisition of Thematic Concepts

In Figure 1, the first source of knowledge, consisting on philosophy, epistemology, psychology, and pedagogy provides human dimensions for training and contributes to elaborate the pedagogical model that establishes the students and teachers' relationships, among them, and with knowledge and resources needed for learning. This first source also contributes to the formation of *Curricular Organization Units*

(COUs), which are units of knowledge to satisfy the training purposes of a program.

The knowledge embodied in the COUs is required to solve some society's problems, a knowledge which mainly comes from the second source: Universal Knowledge Areas (see Figure 1).

Problems, *training purposes*, and dimensions of training are the central elements *of the model for constructing* thematic *Areas* and *Units*, the subject of this article.

Problems are identified from social context knowledge, described as extra-academic knowledge, which constitutes the third source of knowledge (see Figure 1). These problems lead to generate *training purposes* focused on *training dimensions*: to be, to know to do, and to cooperate.

The fourth source refers to internal and external academic knowledge, located on the fourth source of knowledge (see Figure 1), contains knowledge of the general thematic areas managed by domestic and foreign academic institutions, responsible for the development and control of programs and their curricula [4, 5], such as public and private universities, technological institutes and technical associations. This source also includes the current curricula and the knowledge and experience of the institution responsible for the program. General thematic areas are refined from the training purposes characterized by training dimensions, in order to determine the proper thematic areas of the program. These areas incorporate the concepts required to meet the training purposes of the program in the four above-mentioned training dimensions. The COUs are developed from concepts of the proper thematic areas, and are enriched and completed with the contents, structures, and methods offered by the Universal Knowledge Areas.

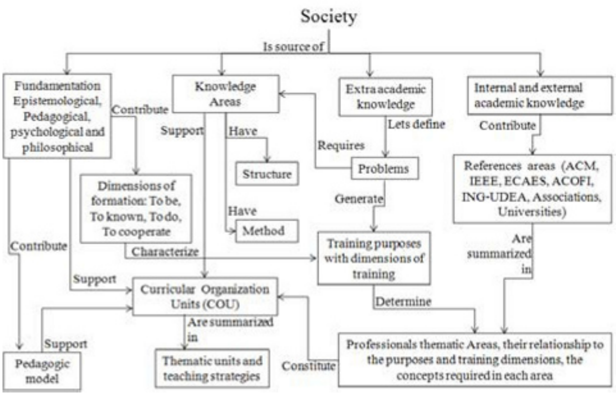


Figure 1. Structure of Contents Defining Process of Professional Component

Knowledge and competences of COUs are specified in the thematic units (courses, workshops, conferences, seminars, etc.) which adopt appropriate teaching strategies.

4 Identifying the Concepts of the Thematic Units Corresponding to the Professional Component

We identify three major components in an undergraduate engineering curriculum: basic science, basic engineering and professional development.

Exact and natural sciences, humanities, and social sciences nourish basic sciences component. These sciences contribute to the development of intellectual skills and support knowledge necessary for the identification and characterization of problems and for the conceptualization, design, and implementation of appropriate solutions. Basic engineering component brings together technical and socio-humanistic knowledge, which is an essential foundation for acquiring advanced knowledge and for the achievement of competencies related to the professional component.

Professional component includes advanced scientific and technologic knowledge, leading to achieve the competences, which characterize professionals from different branches of engineering.

This paper presents only the professional component.

4.1 Identification of Problems

The purpose of the program is determined from existing studies and results of researches aimed at characterizing the social, economic, scientific, and technological development, at local, national, and international levels [6]. The next step leads to identify the problems in the engineering field corresponding to the program for which the curricular model is built. These problems make it possible to define both general and some specific purposes. That is, general and specific objectives, which the engineering training program aims to achieve.

In our case study, a group of teachers, professors, and researchers, all of them engineers working for Universidad de Antioquia Computer Science Engineering Department identified problems which a professional holding a degree in this area should be able to solve, based on diagnostic and prospective studies on the development of the software industry [7], regional development plans [8], and general publications on the impact of Information Technology and Telecommunications in the economic and social development [9]. They focused the study object of this program on *physical and/or logic systems of symbolic processing*. A general statement of problems in terms of areas, in which intervention is required, is shown in Table 2. This way of presenting identified problems, approximated problems to the proper object of the program and guided the formulation of questions for identifying the specific purpose of training.

Table 2. Areas intervened by the Computer Sciences Engineering Program

1.Development of software solutions for the transaction (transmission, consulting, storage and processing) of information across large networks of computers.
2. Development of software systems that support modem organizational management.
3. Development of didactic software tools that support the teaching-educational process.
4. Automation and process control.

4.2 Defining the Training Purposes

Problems identified in social context require the intervention of an engineering program, which supports the training of skilled people to take on the search for solutions to these problems. To face this challenge the program takes a general purpose which is detailed in a set of specific training purposes. The general purpose expresses the global response that the curriculum gives to major problems identified in the social and organizational context.

In our example, for Computer Science program, the general purpose of training, defined according to the problems shown in Table 1 states the following:

"Computer Sciences Engineering Program articulates research, teaching, and service to society for the training of engineers, linking them to critical study of realities and projects, which promote social, economical, and cultural development, so engineers participate with scientific, and technological knowledge as well as with ethical values, intervene processes of the life cycle of products in domain of knowledge, information, and computing at a national and international context".

Specific purposes are the answers to what you need to know to address the solution to each problem in Table 2. The last two questions in Table 3 are related to specific training requirements posed by the multi-disciplinary and concurrent engineering, stated in section 2. Questions begin with the word "how" and are followed by a verb which expresses actions leading to the problem solution. The question is supplemented by a concept involved in the problem and/or in requirements of the multidisciplinary concurrent engineering. For example, in the case of Computer Sciences Engineering Program, the second problem and its related questions are presented in Table 3.

Table 3. Questions about Problem 2

Problem 2: The development of software systems that support modern organizational management
How to develop sequential and competing solutions?
How to design file systems and databases?
How to design queries and reports?
How to develop application-specific languages?
How to design communication protocols?
How to develop human-machine interfaces?
How to configure computer networks?
How to develop applications with multicomunication media access?
How to develop systems which support interrelated tasks and complete processes?

In the same way, questions for the other three problems listed in Table 2 were formulated. The answers are specific training purposes. Some answers may correspond to points raised in several questions. In this way, a group of answers which satisfy all submitted questions is formed. Table 4 shows 18 specific training purposes for the Computer Sciences Engineering Program. Each specific purpose is defined in one or more *training dimensions*, which correspond to the last four columns of Table 4, marked with "Y". The *training purposes* characterized by the training dimensions, specifically numbers 2, 3, 4, 6, 7, 9, 12, 13, 14, 16, 17 and 18, encourage effective multi-disciplinary and concurrent engineering practice, according to the training requirements formulated in section 2. This guidance sets out the thematic areas and units, which will be explained in the next section.

Table 4. Specific training purposes

Specific purposes of training	To know	To do	To be	To share
1. To understand and apply laws, concepts, principles, and basic theories in the field of computer science, as scientific and technological support in developing software and hardware.	Y			
2. To understand and apply concepts, principles, and theories of software engineering field in the develpment of systems that meet technical, economic, ethical, social, and legal conditions.	Y	Y		
3. To model phenomena in domains which may fall outside the scope of computing, appreciating the value of these intellectual skills in the engineering design process.		Y	Y	Y
4. To evaluate the solutions, as well as the impacts under standards accepted by the professional community within the area.		Y		
5. To identify, in the software lifecycle, potential breaches to security, privacy, and intellectual property; to know legal consequences of them and to their impact on the information and institutions.	Y	Y		
6. To understand and apply basic laws, concepts, principles, and theories, in the field of Computer Engineering, for taking into account restrictions imposed by software-hardware relationship in systems development.	Y	Y	Y	Y
7. To understand and apply basic laws, concepts, principles, and theories in the field of Computational Information Systems in order to develp projects according to organizational objectives and meeting legal, technical and economical criteria,such as: consistency, clarity, and workability.	Y	Y		
8. To manage IT departments in the organizations.		Y	Y	Y
9. To use the human creative potential in technologies and innovation development for satisfying socio-economic needs.		Y		
10. To act ethically in the professional and personal life.		Y	Y	Y
11 To provide technical support, with transparency and honesty, to ensure the functioning of information systems.			Y	Y
12. To act with respect, initiative, and effectiveness in work teams, for projects development.		Y	Y	Y
13. To present and communicate verbally, written and visual, in a clear, coherent and respectful way, proposals in domain of the information systems, and social and organizational impacts on the and on.			Y	Y
14. To manage projects, particularly those involving information technology, establishing acceptable compromises within the limits of cost, time, knowledge, and existing systems.		Y	Y	Y
15. To communicate ideas in a second language, with a good level of competition.	Y	Y	Y	Y
16. To manage the own learning and personal development, updating on the progress of disciplines, and learning new models, techniques and technologies awhen these emerge.		Y		
17. To think critically, in oreder to propose solutions to problems in a global context in intervention.			Y	Y
18. To understand concepts and theories about the cognitive process for applying them to systems	Y		Y	Y

4.3 Obtaining the Proper Thematic Areas and their Concepts

Models, proposals, knowledge and experience of national and foreign institutions, provide the general themes and concepts related to the under study engineering program. These broad themes are cross-checked with the specific purposes of the training program in order to obtain a summary of areas and their concepts, adjusted to program purposes, leading to establish the proper thematic areas. When defining each training purpose, its involved training dimensions were introduced (to be, to know, to do, and to cooperate) for multidisciplinary and concurrent engineering. This characterization, in terms of training dimensions for the exercise of concurrent and multidisciplinary engineering, is kept when we define proper thematic areas, thematic units, and the teaching strategies. For our case study, in Computer Sciences Engineering Program we obtained nine proper thematic areas: Algorithmic and Programming, Discrete Mathematics, Information Management, Social and Professional Elements, Software Engineering, Information Systems, Architecture of Machines and Operating Systems, Computational Sciences, and Data Communication.

The proper thematic areas are associated with a number of specific *training purposes*, each one characterized by its training dimensions.

4.4 Obtaining Curricular Organization Units (COUs)

For each proper thematic area, specific training purposes are synthesized in one or more general training purposes, settling in this way one or more COUs, each one with a defined general training purpose as its objective. These units belong to the professional component of the program and support its training purposes. The COUs of the professional component defines and develops professional competences of graduates. In our application case in Computer Sciences, each proper thematic area constitutes a COUs with the same name and contents. For example, for Computer Siences Engineering Program, the COU Software Engineering supports the *training purposes* 2, 3, 4, 6, 9, 12, 13, 17 and 18, Table 4, defined for satisfying the training requirements posed by the multidisciplinary and concurrent engieering. A synthesis of these trainig pruposes constitutes the objective of COU Software Engineering and leads the elaboration of the thematic units.

Thematic concepts of COUs are organized into *thematic units* designed for teaching and learning. For each of these, didactic strategies are adopted, which facilitate the teaching and the appropriation of knowledge by students. For example, the COU Software Engineering groups the following four thematic units: Requirements Engineering, Analysis and Design of Systems, Software Construction, and Software Quality. The four trainning dimensions: To be, to know, to do, and to cooperate, as well as the *training requirements*, introduced in section 2, determine contents and teaching strategies for each *thematic unit*. These units support the practice of the multidisciplinary and concurrent engineering, acording with the objetive of the COU Software Engineering, stated as follows: "Build high quality software using development environments, methodologies, and platforms, which let to solve complex problems, interacting with different experts, knowledge sources, technological resources, and systems; according to the international standards, for satisfying customer needs and changing situations".

4.5 Achievement of Training Purposes Supported on COUs

In our curriculum study in Computers Science Engineering Program, the COU Software Engineering addresses the achievement of above mentioned training purposes, 1, 2, 3, 4, 5, 6, 9, 10, 12, 13, 17, 18, in their proper training dimensions. The COU Software Engineering and their thematic units should include contents and didactic strategies belonging to training dimensions to be, to know, to do, and to cooperate in order to satisfy, for example, the training purpose 5, which covers the four training dimensions, expressed as: to Identify any breaches of security, privacy and intellectual property, to know the legal consequences of them and estimate their impact on information from the institutions during the development and use of software.

In this example, the COU Software Engineering found, in the four training dimensions to be, to know, to do, and to cooperate, contents, such as: multi-communication media simultaneous access requirements, information representation and retrieval, information traceability, systems and processes integration etc., for achieving the training purpose 5, in their four training dimensions. In this way, the Computer Science Engineering Program incorporates the knowledge for satisfying their training purposes, which face the solutions of social context, which gave birth to the program.

This correspondence verifies the support of thematic of COUs for achieving training purposes in the four training dimensions adopted in our model, looking for a social and personal responsible professional practice applying multidisciplinary and concurrent engineering concepts.

5 Conclusion and Future Work

The diagnostic and prospective knowledge about the software industry, the national and local planning, and the information related to technological aspects on society and economy leaded to the identification of problems and to the definition of a curriculum for a Computer Sciences Engineering Program in Colombia.

The introduction of a training dimensions framework (to be, to know, to do and to cooperate) and training purposes involving training requirements of multidisciplinary and concurrent engineering, address thematic units to the practice on this last engineering.

Applying the model referred in this article, the curriculum for Computer Sciences Engineering Program at the Antioquia University was implemented in October 2008. For sake of brevity only any examples of the whole curriculum were presented in precedent sections.

Any ongoing works apply the here presented curriculum construction model to others engineering programs. Equally, the model is applied for obtaining thematic units of basic sciences components of engineering programs. New works are required for gathering and managing of curricular knowledge provided by several sources. It is also important to compare the proposal model with other currently ones.

6 References

[1] Davies, T. (2006). Creative teaching and learning in Europe: Promoting a new paradigm. The Curriculum Journal, 17(1), 37-57.

[2] Rau, D.C., S.T. Chu, & Y.P. Lin (2004) “Strategies for Constructing Problem-Based Learning Curriculum in Engineering Education.”Proceedings of Int. Conference on Engineering Education (ICEE 2004). October 17-21, 2004. Florida, United States.

[3] Delors. J. (1996). La educación encierra un Tesoro. Madrid. Santillana. Unesco.

Crawley, E. et al. Rethinking Engineering Education, the CDIO Approach. Springer: New York, 2007.
[4] ACM, Computer Curricula, Information Tecnology Volume, Oct. 2005 http://www.acm.org/education/curric_vols/IT_October_2005.pdf
[5] IEEE Computer Society. Guide to the Software Engineering Body of Knowledge, SWEBOK™, 2004 version. IEEE Computer Society, Los Alamitos, 2004.
[6] Morell, Lueny et al, Engineering Capacity Building in Latin America, ASEE 2006 Conference Proceedings.
[7] Cátedra del software (1999) Propuestas para fortalecer la industria del software en Antioquia, Centro de Ciencia y Tecnología de Antioquia.
[8] Plan de Desarrollo de Antioquia. 2008-2011. Colombia http://www.antioquia.gov.co/plandesarrollo/ordenanza/3_2cienciaytecnologia.html
[9] Career Space. *Directrices para el desarrollo curricular. Nuevos currículos de TIC para el siglo XXI: el diseño de la educación del mañana*. Oficina de Publicaciones Oficiales de las Comunidades Europeas. Luxemburgo, 2001. www.carrer-space.com.

Competences Supported on Thematic Contents for Evaluation of Curricula Aiming to Concurrent Engineering

German Urrego-Giraldo[a1] and Gloria Lucía Giraldo[b]

[a] Universidad de Antioquia, Dpto. Ingeniería de Sistemas, Colombia..
[b] Universidad Nacional de Colombia, Escuela de Sistemas, Colombia

Abstract: This work adds the determination of competences to the model for constructing the curriculum of engineering programs, from problems identifed in social and organizational context. The competences achievment is oriented by a model of four training dimension: to be, to know, to do and to cooperate, aiming to support the multidisciplinary and concurrent engineering practice. The problem-based and competence-oriented model is illustrated in the curriculum of a Computer Science Engineering Program.

Keywords: Competence, Thematic Unit, Problem-based, Training Dimension, Curriculum, Concurrent Engineering

1 Introduction

In the scientific field, the concept of competence arises in the cognitive psychology, the psycholinguistic and the neuropsycholgy [1,2]. From these sciences, this concept is deployed into the educational and occupational fields [3,4]. The circulation of the competence concept in the mentioned fields leads from the theory to the practice [5, 6, 7]and spreads it in different human being faculties: mental, physical, social, and affective one. Generally, educational competences are considered as occupational ones and intent to focus all human faculties on the learning processes. The intentional character of competences, as an achievement metric of training programs, becomes concrete in the impact of the program on proposed human aspects. Four training dimensions are adopted in this work: to be, to know, to do and to cooperate. These training dimensions were conceived in a previous curriculum construction model, included in [8], to guide enginners' trainning and for adressing the satisfaction of requirements of multidisciplinary and concurrent engineering.

These requirements and training dimensions led to definition of training purposes of engineering programs.

[1] Corresponding Author E-mail: gaurrego@udea.edu.co

These training dimensions were checked with the learning outcomes defined in Dublin Descriptors [9, 10], finding coincidence of our training dimensions to *be*, to *know*, and to *do* with four of the five above mentioned outcomes: *knowledge and understanding*, *application of knowledge and understanding*, *judgement*, and *learning skills*. The remained outcome, *communication*, is exceeded by the training dimension to *cooperate*, which looks for higher human development, and majeure social responsibility and commitment.

In this article we add the concept of competence to the mentioned curriculum construction model, which is used for the definition of thematic concepts for engineering programs. Competence concept leads the program evaluation, verifying the correspondence between competences supported by the thematic areas and units of the engineering program, and the training purposes generated from social context problems. Equally, the correspondence between competences and social context problems is verified. The use of competences as a measure of program achievements is illustrated in a Computers Science Engineering Program oriented to multidisciplinary and concurrent engineering practice.

Some models as described in [11, 12, 13] have been proposed for introducing the concept of multidisciplinary and concurrent engineering in special courses, or industrial training cycles; but, the incorporation of this engineering concept in thematic areas and units for the whole curriculum is also a contribution of this paper.

Only the existence of these pertinent contents in programs does not assure that students reach the competences. The top competences achieved by students are developed step by step incorporating the competences supported by each one of thematic areas and units. Competences-oriented curricula claim the use of competences for courses evaluation, but we do not find competences models for evaluating the construction and application of a whole program. The verification of engineers' competences constitutes a way for program, namely, curriculum evaluation. Thus, the proposed model aids to elaborate models oriented to the evaluation of training's results supported by competences-oriented curricula.

The proposed model was evaluated for establishing the correspondence of competences with training purposes of the program, and with the social context problems. These correspondences were verified in the construction of a curriculum of a Computers Science Engineering Program.

In section 2 of this paper, the model for determining of the competences that the program students should obtain from thematic contents is described. The determination of the competences is explained in section 3. The aid for the verification of program competences achievement, as well, the correspondences of competences, with the problems, and the training purposes are explained in section 4. The conclusion and future Works are in included in section 5. The bibliography is in section 6.

2 Model for Obtaining the Competences from Thematic Contents

Traditional models for elaboration of curricula do not include either the social context problem concept or the competence concept. Proposes involving the competence concept, as a didactic strategy, have been developed in [14,15,16]. Other approaches use the competence concept as a guide for defining the thematic contents and the strategies for reaching them. The objective of this article is to complete a previous problem-based model, presented in [8], for including the determination of competences, which are supported by the engineering programs, from thematic contents. The competences model closes the cycle which begins in the context problems identification, progresses to constructing the program of studies, goes on to develop the competences and verifies the correspondence between the competences and the problems. In this paper the competences achievement, which should let to face responsibly and wisely the solution of referred problems are explained.

2.1 Extention of the Model in Order to include the Competences

The determination of competences from the Curricular Organization Units (COUs) is added, Figure 1, to the Model in [8], which is used for constructing the thematic contents. In the same way, the relationship among competences, training dimensions (To *be*, to *know*, to *do* and to *cooperate*), training purposes and context problems are added. Training purposes should be satisfied by the competences, and engineers, supported on their competences, may face the problem solution with intellectual and ethical solvency.

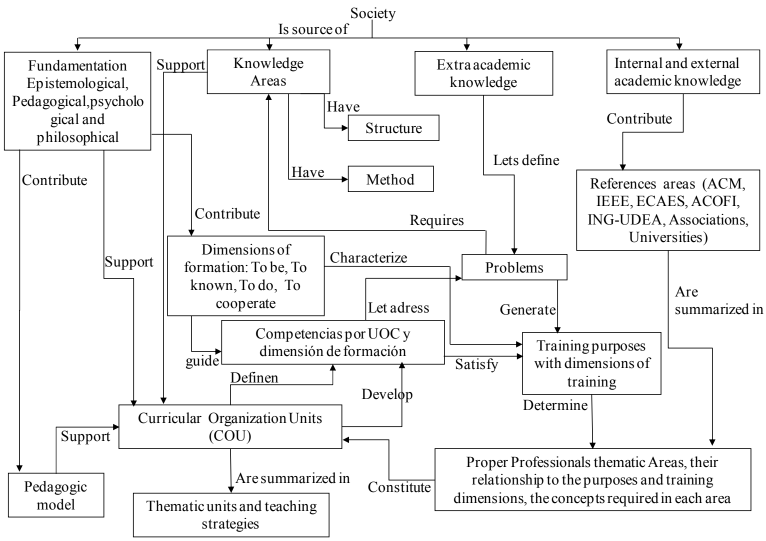


Figure 1. Process Structure for Obtaining Competences Supported on Thematic Contents

The original model for obtaining the COUs and their contents considers four major knowledge branches, Figure 1. The first one, the knowledge coming from the psychology, the pedagogy, the philosophy and its filial the epistemology. This knowledge supplies the pedagogical model, which supports the training program and the elaboration of the COUs.

The second branch points to the universal knowledge, in particular, for the engineering programs, to the natural and exact sciences, with their structures and methods.

The third one encloses the extra academic knowledge from which the problems are defined. These, in turn, generate the training purposes, characterized by training dimensions and oriented to determine Proper Thematic Areas of engineering program.

The fourth branch corresponds to the internal and external academic knowledge. In this category, general thematic areas of national and foreign institutions, the knowledge of the entity which elaborates the curriculum, the own experience and precedent curricula appear [17, 18, 19]. General Thematic Areas are synthesized in Proper Thematic Areas determined by the training purposes.

2.2 Curricular Organization Units (COUs) and Thematic Units

In a previous work, the model for obtaining the thematic units of an engineering program of studies has been presented and illustrated. The thematic units arise from a set of problems identified in social context and the requirements for the multidisciplinary and concurrent practice. The program of studies should address the solution of above mentioned problems and requirements.

In relation to a Computer Science Engineering Program, our study case, problems mentioned in terms of its solutions, are described in Table 1:

Table 1. Social and Organizational Context Problems

1-Development of informatics solutions for the transmission, the query, the storage, and the processing of information through a wide net of computers.
2-Development of informatics systems, which support the modern organizational management.
3-Developpñment of didactic informatics tools which support teaching and learning processes.
4-Automation and processes control.

Multidisciplinary and Concurrent Engineering Requirements are groped in [8] in three categories: *technical issues*, *human interventions*, and *product lifecycle*. An extract of specific requirements contents: ubiquity (characterized by place and platform independence, real time, multi-communication media access, tele-communication, and tele-working facilities), technological solvency, systemic thinking, product lifecycle management, process integration, team spirit, interrelated tasks, communicability, etc..

A set of questions about these requirements, and the problems, presented in Table 1, let to define a set of training purposes to be carried out in four training dimensions: to *be*, to *know*, to *do* and to *cooperate*.

For the case in Computer Science Engineering the 18 training purposes identified, correspond to problems found in the social context. For each purpose, training dimensions covered by it are pointed with Y. An extract of theses purposes is described in Table 2. These purposes, for example, take in account requirements of the multidisciplinary and concurrent engineering.

Table 2. Extract of Training specific purposes

Specific purposes of training	To know	To do	To be	To cooperate
1. To understand and apply laws, concepts, principles and basic theories in the field of computer science, as scientific and technological support in developing software and hardware.	Y			
2. To understand and apply concepts, principles and theories of engineering software field in the specification, design and implementation of systems that meet technical, economic, ethical, social, and legal conditions, for multiple contexts and domains.	Y	Y	Y	Y
3. To model phenomena in domains that may fall outside the scope of computing, appreciating the value of these intellectual skills in the engineering design process.		Y	Y	Y
4. To evaluate solutions, as well as their impacts under standards accepted by the professional community within the area.		Y		

The purposes are correlated with a set of general thematic areas obtained from the models and experiences of national and foreign institutions charged with the academic knowledge management, in order to determine the proper thematic areas with their corresponding concepts, for the envisioned engineering program.

For the Computer Science Engineering Program, thematic concepts are grouped in the following 9 proper thematic areas: Algorithms and Programming, Discrete Mathematics, Administration of Information, Social and Professional elements, Software Engineering, Information Systems, Machine Architecture and Operating Systems, Computational Science, Data Communication. For each one of last thematic areas the training purpose should be established.

For curricular effects referred areas are converted in COUs, each one, with an objective or own purpose which synthesizes the specific training purposes under its responsibility. The training purposes, since their definition, have aimed training dimensions associated (to be, to know, to do and to cooperate). This link is kept in proper thematic areas and in COUs. In our study case in Computer Science Engineering, each proper thematic area yields only one COU, with the same name, but it could have several COUs, Table 3. These areas are enriched with universal knowledge area, in particular natural and exact sciences.

Thematic concepts of the COUs are organized in Thematic Units (subject, course, class, etc.) adequated for teaching and learning. For each one, didactic strategies are adopted, such as: project, class, seminar, conference, workshops, exercise, etc., which develop the teaching processes and make the knowledge acquisition to the students easy.

The COUs and their Thematic Units constitute the professional component of the program, and support the training purposes. The determination of the COUs for the Basic Science, which supports the professional component, is the object of another work.

An example of one COU of Computer Science Engineering is depicted in Table 4. For achievement its objective, each COU is disaggregated in thematic units. In our study case, the COU Software Engineering yields four Thematic Units: Requirements Engineering, System Analysis and Design, Software Construction and Software Quality.

Table 3. COUs of Computer Science Engineering Program

COU of Computer Engineering Program	General objective of the UOC	Training purposes
Programming and algorithmic	To know, use and evaluate the main techniques of problem solving and the representation of information using the computer as a tool, and experiment with language to communicate with the machine from the point of view formal and user.	1,2,3,4,6
Discrete Mathematics	To understand and apply the concepts and properties of mathematical structures in the representation and study of discrete phenomena.	1
Information Administration	To capture, represent, organize, process and present information, based on effective and efficient algorithms for accessing, updating and storage of information (physic and logic), including aspects of security, privacy and integrity.	2, 5, 7
Social and Professional Elements	To perform the profession with the knowledge, skills and attitude in an ethical manner, with fluency in communication, coupled with multidisciplinary and social responsibility.	8,9,10,12,13,15,16,17
Software Engineering	To build software by applying methodologies for development, operation, maintenance and international standards, satisfying customer needs with high quality.	1, 2, 3, 4, 5, 6, 9, 17,18, 10, 12, 13
Information Systems	To managing information technology to support organizational strategies, emphasizing the process of negotiation, administration and control	5, 7, 8, 9, 11, 14, 17, 10, 12, 13
Machine architecture and operating systems	To understand and apply basic principles, methodologies and techniques for design of computer systems in a hierarchical manner from primary components to complex systems	6,2
Computational Science	To simulate different types of systems, building and using mathematical formulas, methods, computational tools and visual.	1,3,4,13
Data Communication	To understand and apply the basic knowledge required for the transmission of information, modeling, design and performance evaluation of data networks, using criteria of efficiency, safety, legality and interoperability	1, 5, 6, 11, 18

Section 3 describes how to obtain the professional competences from the knowledge contained in the COUs of the program.

3 Determination of Professional Competences

The COUs are responsible for achieving the competences. For each on, a general competence and multiple specific competences are set, which are organized by each of the four educational dimensions adopted for the training. Paragraph 3.1 explains the steps for reaching the competences.

3.1 General Competences of the COUs

In agreement with training purposes belonging to each COU, one general objective for each one is assigned. This objective expresses capacities that a student should develop through the engineering program. When objectives of the program and thematic areas are established in terms of student's achievements, it is clear that the general objective of a COU represents the general competence supported by this COU.

In our study case, Computer Science Engineering Program, one of the nine COUs, Software Engineering, is illustrated in Table 4.

Table 4. General Competence Supported by a COU and their Satisfied Training Purposes

COU of Computer Science Engineering Program	General objective of the UOC	Training purposes
Software Engineering	To build software by applying international standards, and methodologies for systems development, operation, and maintenance; satisfying with high quality level needs of customer.	1, 2, 3, 4, 5, 6, 9, 17,18, 10, 12, 13

General objective of a COU summaries training purposes which it covers. In terms of capacities to be reached by the students, the objective expresses the general competences supported by the COU, as it is showed in column 2 of the Table 4. The two next columns contain the thematic concepts considered by external and internal institutions, which manage academic knowledge. The last column contains the identification numbers, from [8], of the program training purposes supported by the COU named Software Engineering. These training purposes, some of them previously detailed in Table 2, show a clear orientation of this COU for supporting multidisciplinary and concurrent engineering.

3.2 Specific Competences of COUs

For each training purpose associated to a COU, in each adopted training dimension, what specific achievements students should reach, according to concepts mentioned in the general competences of the COU and to the thematic concepts previously defined are established. The detailed procedure for obtaining the competences is as follows:

1- To identify the training purposes and their training dimensions, supported by each COU.

2- To identify the training concepts mentioned by the general competence of the COU.

3- In each COU, for each training dimension of each training purpose the following question is formulated: What should students achieve, in this training dimension, in order to assure that the program satisfies the referred training dimension?. This reasoning is led by the concepts identified in the general competence and by the thematic concepts of the COU. Experts in these themes should conduct this reasoning.

The answers to each question constitute the specific competences that students should develop with the aid of the COU and in the training dimension where the question is formulated.

4- The answers take the student as tacit subject and begin with the expression “The student is able of …”, followed of a verb in infinitive and then the complete object of the answer.

Aiming to verify the correspondence of the competences with the training purposes and with the training dimensions, a table of competences, for each COU, is arranged. Table 5, shows only an extract of the specific competences for the professional COU named Software Engineering of the Computer Science Engineering Program.

The four columns of the table contain: the training dimension in which the competence is formulated, the identification number of the training purpose, the ordinal number of the competence inside the training dimension and finally the description of the competence.

In Table 5, only one specific competence for each training dimension is described, but in reality, in each of them many competences are identified. In this example, the training dimensions are obtained from training purpose number 2.

T**able 5.** Extract of specific competences by training dimension in a COU

Profesional COU: Software Engineering			
General Competence (Objective): To build software by applying international standards, and methodologies for systems development, operation, and maintenance; satisfying, with high quality level, needs of customer.			
Training Dimensions	**Training Purposes**	**Competence Ordinal Number**	**Descripción de la competencia específica**
To know	2	3	To understand, as they emerge, theories, methods, technologies, and techniques for the development, operation, and maintenance of software.
To do	2	5	To develop software in various domains using methods and standards internationally accepted.
To be	2	1	To be responsible of legal aspects and of the social impact of the projects.
To cooperate	2	1	knowledge to developing systems, which consider the complexity of the processes and the quality of results.

Specific competences are marked as essential or supplementary for the program. Thus, in each COU the non essential themes may constitute alternative, complementary or optional units in the curriculum.

In each COU, for each training dimension, the specific competences are synthesized in one objective, namely, one general competence to be supported by each mentioned dimension.

4 Competences Verification

In curricula construction and practice, the competences concept leads two verifications ways: Coherence of competences-oriented curricula, and achievement of competences. The coherence aims to establish the corrrespondence among problems, training purpouses and competeces. The achievement of competences is related to learning, and to the impact in facing social and organizaional context problems.

4.1 Coherence of Competences-Based Curricula

Competences model based on curricular contents, presented in this work, connects the problems, training purposes and the competences, accepting a double verification of competences. The first one sets the correspondence between competences and training purposes. The eventual mismatching demands an analysis of proper thematic areas of the program and of thematic contents.
In the verification of these correspondence for the Computer Science Engineering Program is clear that the training purpose number 2 in of the COU Software Engineering, expressed in Table 2 as: *To understand and apply concepts, principles and theories of engineering software field in the specification, design and implementation of systems that meet technical, economic, ethical, social, and legal conditions, in multiple contexts and domains*; is satisfied by the competences in training dimensions *to know* and *to do*, *to be*, and to *cooperate* stated in Table 5, respectively as:

- *To understand, as they emerge, theories, methods, technologies, and techniques for the development, operation, and maintenance of software.*
- *To develop software in various domains, using methods and standards internationally accepted.*
- *To be responsible of legal aspects and of the social impact of the projects.*
- *To contribute with scientific and technical knowledge to developing systems, which consider the complexity of the processes and the quality of products*.

The second verification is related to the problems, in order to establish which competences are, effectively, required for affronting the solution of the problems. The possible incongruence requires the verification of the correspondence between the problems and the training purposes.

For the Computers Science Engineering Program, the general competences supported by the COUs, in Table 3, constitute a set of pertinent answers to the social context problems, Table 1. This correspondence illustrates the efficacy of the model for establishing the competences for facing problems of social and organizational contexts.

4.2 Achievement of Competences

The coherence of the program-based and competence-oriented curricular model constitutes the base for creating models aiming to relate the competences with the learning of the concepts of thematic units, as well as, to compare the competences reached by the students with the quality of their interventions on the social context problems. These achievements models are not considered in this article.

5 Conclusion and Future Work

The proposed approach validates the correspondence between the social context problems and the competences of undergraduate students for facing the solution of problems. The determination of competences from knowledge involved in the UOCs of the program and the validation of those against the problems and training purposes show the consistence of the model and its possibilities of the evaluation and continuous improvement.

The model for obtaining the competences is applied in the ongoing studies for different engineering programs and for comparing its results with those coming from other models.

The training according to the dimensions to be, to know, to do, and to cooperate, based on social and organizational context problems, as well as, on the requirements of multidisciplinary and concurrent engineering, considered in [8, 11, 12, 13], supports the development of essential competences for the practice of this engineering. This may be appreciated in specific purposes 2 and 3 in table 2, and then, in the competences described in table 5. The approach of rationalization of problems and competences development highlights contact points of different engineering branches and it creates a reference for training an engineering thinking aiming to the development of the multidisciplinary and concurrent engineering.

6 References

[1] Gelman, R. y Greeno, J.G. On the nature of competence. Principies for understanding in a domain. In L.B. Resnick (Ed) "Knowing, learning, and instruction". Hillsdale, NJ, Eribaum, 1989.

[2] Hymes, D. Models of the interaction of language and social setting. Journal of social Issue.1967.

[3] Koschmann, T., Kelson, A.C., Feltovich, P.J., & Barrows, H.S. (1996). Computer-Supported Problem-Based Learning: A principled approach to the use of computers in collaborative learning. In T.D. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm*.(pp. 83–124). Hillsdale, NJ: Lawrence Erlbaum.

[4] EICE. Ingeniería Informática. Libro Blanco sobre las titulaciones universitarias de normalinformática en el nuevo espacio europeo de educación superior. 2004. http://www.sc.ehu.es/siwebso/Bolonia/textos/EICE/Libro%20Blanco%20v113.pdf.

[5] Brunner, JJ. Competencias de empleabilidad. Revisión bibliográfica. Internet, 2001.

[6] Puig J., Hartz B.(2005). Concepto de Competencia y Modelos de competencias de Empleabilidad. Primer Encuentro Internacional de Educación Superior: Formación por competencias. Medellìn.

[7] Carióla, M.L. y Quiroz, A.M. Competencias generales laborales y curricula. En M. Novick, M.A. Gallart (coordinadoras) "Competitividad, redes productivas y competencias laborales". Montevideo, CINTERFOR y Red Educación Y Trabajo, 1997.

[8] Giraldo, G. L., Urrego-Giraldo G. Problem-based Construction of Engineering Curricula for Multidisciplinary and Concurrent EngineeringPractice. Submitted to CE2009.

[9] Nick Clark (2007). Bologna: Curriculum Reform and other Considerations. Editor, World Education News & Reviews.

[10] Joint Quality Initiative (2004). Shared 'Dublin' descriptors for Short Cycle, First Cycle, Second Cycle and Third Cycle Awards.
[11] Balamuralikrishna R., Athinarayanan R., Song X. (2000). The Relevance of Concurrent Engineering in Industrial Technology Programs. Journal of Industrial Technology, Volume 16, Number 3, May 2000 to July 2000,www.nait.org.
[12] Smith R. P., Russell R. B., Craig A. N., Zayas-Castro J. L. (1996). Concurrent Engineering: A Partnership Approach. ASEE Annual Conference Proceedings.
[13] Zhou J., Carmi S., Lau A., Koulas S. (1996). Concurrent Engineering Concepts Applied to Concurrent Education and Research. Third ISPE International Conference on Concurrent Engineering:Research and Applications (CE96/ISPE).Toronto, Ontario, Canada, August 26-28, 1996.
[14] Mouthaan, T. J., Olthuis W., Vos Hen. Competence-Based EE-Learning:(How) Can We Implement It?.Proceedings of the 2003 IEEE International Conference on Microelectronic Systems Education (MSE'03)
[15] Tuning Project.Tuning Educational Structures in Europe. Available at : http://www.tunig;unideusto.org/tunigeu
[16] Proyecto Tuning.Reflexiones y Perspectivas de la educacion superior en America Latina. Informe final 2004. http://www.tunig;unideusto.org/tunigal
[17] ACM, Computer Curricula, Information Tecnology Volume, Oct. 2005 http://www.acm.org/education/curric_vols/IT_October_2005.pdf
[18] Crawley, E. F. Creating the CDIO Syllabus: A Universal Template for Engineering Education. ASEE/IEEE Frontiers in Education Conference, Boston, MA, USA, 06-09 November 2002.
[19] Crawley, E. et al. Rethinking Engineering Education, the CDIO Approach. Springer: New York, 2007.

Predicting the Yield Rate of DRAM Modules by Support Vector Regression

Shih-Wei Lin[a,1] and Shih-Chieh Chen[b]

[a] Department of Information Management, Chang Gung University, Taiwan.
[b] Department of Industrial Management, National Taiwan University of Science and Technology, Taiwan.

Abstract: Dynamic random access memory (DRAM) module is one of the principal components of electronic equipment, which impacts the quality, performance and price of the final products singinifcantly. Typically, DRAM module is composed of DRAM ICs (integrated circuit). DRAM ICs with higher quality can be used to produce DRAM modules with higher quality. Generlly speaking, high quality DRAM ICs are more costly. Due the the cost down and material saving reason, some DRAM module manufacturers purchase batches DRAM ICs containing defective units, and then have the batch tested in order to select DRAM ICs for production of DRAM modules. Thus, this kind of DRAM module is suitable only for products not intended for work in harsh environments being sold in lower price markets. Due to the lower quality of the DRAM ICs, the actual quality of the DRAM module is not easily predicted. Predicting the yield rate of the DRAM module is thus an important issue for DRAM module manufacturers who purchase DRAM ICs with lower quality at lower prices.
This study used support vector regression (SVR) to predict the yield rate of the DRAM modules produced using defective DRAM ICs. SVR is a very capable method and has been successfully applied across many fields. However, the parameters and input features differ depending on the application. Thus, a scatter search (SS) approach is proposed to obtain the suitable parameters for the SVR and to select the beneficial subset of features which result in a better prediction of the DRAM module yield rate. The experimental results showed that the performance is better than that of traditional stepwise regression analysis.

Keywords: DRAM yield rate, support vector regression, scatter search, feature selection, parameter determination.

1 Introduction

The semiconductor industry is one of the core industries of Taiwan, with the top five companies in Taiwan accounting for 65% of global annual revenues. The

[1] Corresponding Author : Associate Professor, Chang Gung University, No. 259, Wen-Hwa 1st Road, Kwei-Shan, Taoyuan, Taiwan; Tel: +886 (3) 211-8800 Ext. 5811; Fax: +886 (3) 211-8020; Email: swlin@mail.cgu.edu.tw

Dynamic random access memory (DRAM) is ubiquitous in electronic products, such as computers, automobiles, personal digital assistant, notebook computers, robots, and household appliances. The memory industry market often has large fluctuations. If the yield rate of the finished good (DRAM modules) can be accurately estimated when purchasing the raw material (DRAM ICs), DRAM module manufacturer can more precisely determine the allowable purchase price of the DRAM ICs and deliver the goods on time.

Some DRAM module manufacturers purchase batches of DRAM ICs that contain defective DRAM ICs, and then arrange testing to select DRAM ICs from this batch to produce DRAM modules. Such IC sorting services are common in Southeast Asia. This type of DRAM module is suitable for products not intended for harsh environments and is sold in lower price markets. Due to the lower quality of the DRAM ICs, the quality of the DRAM module is not easily predicted. For example, a DRAM module manufacturer plans to purchase 800,000 units of DRAM ICs at the price of US$1.20 each to assemble 100,000 DRAM modules, each with 8 DRAM ICs. Suppose the warranted yield rate of the DRAM IC supplier is 95%. The actual amount for purchase should therefore be 842,106 (800,000/0.95), for a total expenditure of US$1,010,526. If the actual yield rate achieves 97%, then there will be an extra 16,843 (842,106×0.97-800,000=16843) DRAM ICs, which amounts to an extra profit of US$20,212. But if the actual yield only reaches 94%, a shortage of 8,421 (842,106×0.94−800,000 = -8,421) DRAM ICs will result. It is likely that the DRAM module manufacturer will either suffer a loss or be forced to make urgent purchases of additional DRAM ICs.

At present, the estimated value, empirical value, and sample testing are the major measures used by manufacturers to predict the yield rates of DRAM modules. However, in practice, there is a considerable deviation between an actual yield rate and its estimated value, which may result in a shortage of DRAM modules for delivery to customers, leading in turn to delivery delays or second orders of DRAM ICs. The support vector regression (SVR) is a viable, reliable, and attractive approach for prediction problem. However, before applying a SVR to problem solving, the parameter settings of the SVR must be determined. The parameter settings for the SVR must be determined carefully to avoid constructing a suboptimal model, on the other hand the input features also a big problem these may significantly increase computational costs and produce inferior results.

This study proposes a scatter search based approach, denoted as SSSVR, to obtain the appropriate parameter settings for SVR, and to select the beneficial subset of features that will result in the best prediction of the yield rate of DRAM modules. We found that the proposed SSSVR approach is useful in predicting the yield rate of DRAM modules, and the performance of the SSSVR is better than that of traditional stepwise regression analysis (SRA).

The remainder of this paper is organized as follows. Section 2 reviews the studies on the DRAM IC, SVR and feature selection. Section 3 elaborates on the proposed SSSVR approach to determine the appropriate parameter settings for the SVR, and the beneficial subset of features, for predicting the yield rate of DRAM modules produced by defective ICs. Section 4 describes the experimental results. Conclusions and future research are offered in the last section.

2 Research background

2.1 DRAM IC

DRAM is a type of random access memory that stores each bit of data in a separate capacitor within an IC. Generally speaking, the IC manufacturing process consists of five steps: (1) wafer fabrication, (2) wafer probe IC, (3) device assembly, (4) raw test, and (5) final test. Among these steps, wafer fabrication and wafer probe IC are classified as forward processes, while device assembly, raw test and final test are backward processes [1].

DRAM ICs may be divided into different levels by the testing. IC testing determines the quality, and thus affects the price of the DRAM IC. A defective DRAM IC is defined as an IC which did not pass FT1, FT2, or FT3. Moreover, the quality of an FT3 defective IC (failed FT3) is better than that of an FT2 defective IC (failed in FT2), while the quality of FT2 defective IC is better than that of FT1 defective IC (failed in FT1). The presence of different grades of DRAM IC will affect the quality and price of the DRAM module. Furthermore, differing grades of DRAM IC will result in different DRAM module yield rates. The presence of defective DRAM ICs in a batch does not mean that all ICs cannot be used; non-defective ICs can still be selected from the batch of defective DRAM ICs through testing.

In this study we focus on DRAM ICs that have not passed the final test in predicting the yield rate of DRAM modules. DRAM ICs that have not yet passed the final test can be sent to the DRAM module manufacture for further IC sorting service. The prediction accuracy of the yield rate of DRAM modules using defective DRAM ICs may greatly affect the profitability of DRAM module manufacture. Therefore, if the prediction error can be reduced, inventory costs may also be reduced, or more revenue than expected obtained.

2.2 Relative works

Many studies have applied statistical approaches to the semiconductor industry. Wong [2] conducted a statistical correlation analysis to determine the significant electrical parameters that impacted line yield, and built a multiple regression model to predict the line yield using the electrical test parameters. Collica [3] presented a yield model for static random access memory based on die level bit fail counts on a wafer through the use of a logistic regression model. The model uses a binary response for when a chip does or does not have bit failures recorded. Jun *et al.* [4] proposed a semiconductor yield regression model including a new cluster index, whose performance seems to dominate the negative binomial model. The statistical approach has limitations in that the large amount of data in current semiconductor databases makes it impractical to efficiently analyze the data for valuable decision-making information.

In recent years, data mining approaches have been widely applied to semiconductor problems. Chen [5] proposed a hybrid SOM-BPN mechanism for lot output time prediction and achievability evaluation in a wafer fabrication plant (wafer fab). Gardner *et al.* [6] proposed three tough semiconductor manufacturing problems that can be solved using data mining methods. A combination of self-organizing neural networks and rule induction was used to identify the critical low yield factors from normally collected wafer manufacturing data. Last and Kandel [7] observed that the accurate planning of produced quantities is a challenging task in the semiconductor industry where the percentage of good parts (yield) is affected by multiple factors. They proposed a perception-based method, called automated perceptions network, for automated construction of compact and interpretable models from highly noisy semiconductor products data. Chen *et al.* [7] applied decision tree algorithm and other classification models to forecast dynamic random access memory prices. Chen *et al.* [8] use a decision tree and neural networks to establish a quality analysis system for the defects found in the production processes of package factories. Wang [9] proposed using a spatial filter to denoise the noisy wafer map and to extract meaningful defect clusters. To isolate various types of defect patterns, a hybrid scheme combining entropy fuzzy c-means with spectral clustering is applied to the denoised output. Chien *et al.* [10] develop a framework for data mining and knowledge discovery from database that consists of a Kruskal-Wallis test, K-means clustering, and the variance reduction splitting criterion to investigate the huge amount of semiconductor manufacturing data. To the best of the author's knowledge, no research has yet been carried out to the yield rate prediction of DRAM modules produced using defective DRAM ICs.

2.3 Support vector regression

Support vector machine (SVM) attempts to identify a hyper-plane, which functions as a separating plane for classification of data, in a multi-dimensional space [11]. In SVM for regression models (SVR), the aim is to find a pair, such that the output of the training vector *x* can be predicted according to the real-valued function.

The goal is to find a function that has at most ε deviation to from the actually obtained targets with a small risk (or test error). This function is called the ε-insensitive loss it can be expressed as:

$$l_\varepsilon = \left|y_i - f(x_i)\right|_\varepsilon = \left\{ {}^{0,\ \text{if}\, y-f(x_i,w)\le\varepsilon}_{\left|y-f(x_i,w)\right|-\varepsilon\ ,\ \text{otherwise}} \right\} = \max(0, \left|y_i - f(x_i)\right| - \varepsilon) \tag{1}$$

where y_i is the predict value, w is the weight vector with d dimension, X is some nonempty set from which the patterns x_i are taken, ε is a parameter of controlling tube size, and l represents the total number of training data samples with d input variables. Usedε-insensitive function we need a variable ξ (slack variable) to control the margin as equation (2):

$$\left|y - f(x,w)\right| - \varepsilon = \zeta,\ \text{for data above an } \varepsilon \text{ margin}$$
$$\left|f(x,w) - y\right| - \varepsilon = \zeta^*,\ \text{for data below an } \varepsilon \text{ margin} \tag{2}$$

The goal of learning is to find a function with a small risk on the test samples.

Minimize：$$\tau(w,\xi_i^*) = \frac{1}{2}\|w\|^2 + C\Box\left(\frac{1}{l}\sum_{i=1}^{l}(\xi_i + \xi_i^*)\right) \tag{3}$$

Subjected to：
$$\begin{aligned} &\langle w \cdot x_i + b\rangle - y_i \le \varepsilon + \xi_i, \\ &y_i - \langle w \cdot x_i + b\rangle \le \varepsilon + \xi_i^*, \\ &\xi_i^* \ge 0, \varepsilon \ge 0. \end{aligned} \tag{4}$$

where $\xi_i^{(*)} = \left(\xi_1, \xi_1^*, \ldots \xi_i, \xi_i^*, \ldots \xi_l, \xi_l^*\right)^T$ that are slack variables guarantee the meet of constraint condition.

The constrained optimization (in Eq. 3 and 4) can be reformulated into dual problem formalism by using Lagrange multipliers, which leads to the solution:

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x) + b \tag{5}$$

where b is the bias, α_i , and α_i^* are the Lagrange multipliers and $K(x_i,x)$ the so called kernel function. Several kernel functions help the SVM to obtain the optimal solution. In this study, the radial basis kernel function (RBF) is applied in the SVR. Three major parameters applied in SVR, C, Υ, and ε must be set appropriately. Parameter C control the trade-off with the complexity penalizer $\|w\|^2$. Parameter Υ affects the partitioning outcome in the feature space. The ε parameter controls the desire accuracy of the approximation to a specific level. Parameters setting for support vector regression (SVR) are important step in SVR design, which strongly affects the performance of SVR [12]. Suggesting a good setting is thus an important problem. In this study we used scatter search algorithm to determined parameters setting and input feature simultaneously in order to build the appropriate model to precise predict DRAM IC yield.

2.4 Feature selection

The SVR requires a dataset to construct a model. A dataset may consist of many features, but not all features are helpful for classification. If the dataset has tremendous noise and complex dimensionality, the SVR may face limitations in learning the classification patterns. Feature selection may be considered part of the data cleaning and/or pre-processing step where the actual extraction and learning of knowledge or patterns is done after a suitable set of features is extracted.

The primary benefits of feature selection are as follows: (1) computational cost and storage requirements are reduced; (2) degradation of classification efficiency due to the irrelevant or redundant features used in the training samples is overcome; (3) training and prediction time are reduced; and, (4) data understanding and visualization are facilitated [13].

In feature selection, as whether each feature is useful must be determined, the task of finding an optimal subset of features is inherently combinatory. Therefore, feature selection becomes an optimization problem. An optimal approach is then needed to examine all possible subsets.

3 The Proposed Approach

The scatter search (SS) is a population-based approach that starts with a collection of reference solutions obtained by applying preliminary heuristic processes. The SSSVR approach follows the steps of the SS template [13] and is described as follows.

(1) Diversification generation method: Population P with P_{size} solutions is generated randomly; that is, each variable in the solution is uniformly generated from 0-1.

(2) Reference set update method: The size of the reference set is $b = b_1 + b_2 = |RefSet|$ Construction of the initial reference set starts with selecting b_1 best solutions (solutions with the best classification result) from P. These solutions are added to *RefSet* and deleted from P. For each solution in the *P-RefSet*, the minimum Euclidean distance to the solutions in *RefSet* is calculated. The solution with the greatest minimum distance is selected. This solution is then added to *RefSet* and deleted from P, and the minimum distances are updated accordingly. This process is repeated b_2 times, where $b_2 = b - b_1$. Thus, the resulting reference set has b_1 high-quality solutions and b_2 diverse solutions.

(3) Subset generation method: The size of subsets is set to 2; that is, only subsets consisting of all pair-wise combinations of solutions in *RefSet* are considered. Therefore, at maximum, $b(b-1)/2$ subsets exist.

(4) Solution combination method: The method employed consists of finding linear combinations of reference solutions. Each combination of two reference solutions, denoted as $X^{'}$ and $X^{''}$, are employed to create three trial solutions. These three trial solutions are (a) $X = X^{'} - d$, (b) $X = X^{'} + d$, and (c) $X = X^{''} + d$, where $d = u(X^{''} - X^{'})/2$ and u is a random number generated from 0-1.

This study presents an approach which uses SS for parameter determination and feature selection of SVR, called the SSSVR. For the solution representation, if n features are required to decide which features are chosen, then 3+n variables must be adopted (3 variables are used for C, Υ and ε). The value of these n variables ranges between 0 and 1. If the value of a variable is less than or equal to 0.5, then its corresponding feature is not chosen. Conversely, if the value of a variable is greater than 0.5, then its corresponding feature is chosen. The values of C, Υ and ε are also ranges from 0–1 in the solution representation for the easy implementation. However, their real values can be scaled to the specific value (given the lower and the upper bound for their values). Take the case of a scaled solution as an example, if the data set has 6 attributes and the SVR requires 3 parameters, there are 9 variables used as shown in Figure 1. This solution can be decoded as follows. The C is 2500, the, Υ is 0.02, the, ε is 0.001, and the selected features are 1, 2, 3, and 6.

1(C)	2(r)	3(ε)	4(F1)	5(F2)	6(F3)	7(F4)	8(F5)	9(F6)
2500	0.02	0.001	0.96	0.73	0.90	0.49	0.36	0.65

Figure 1. An example of solution representation for the proposed approach

The parameter settings and selected features are then adopted to build a SVR model. Finally, the Root Mean Squared Error (RMSE) of the current parameter values and the selected features for SVR can be computed. To evaluate the robustness of the proposed SSSVR approach, the k-fold approach [15] is used. The average RMSE of this experiment was calculated by summing the individual RMSE for each fold of testing, and then dividing the total by 3. Figure 2 describes the architecture of the proposed SSSVR approach to establish the SSSVR prediction model.

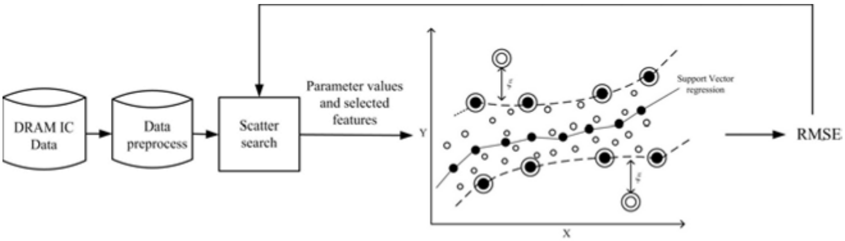


Figure 2. The architecture of the proposed SSSVR approach

4 Experimental Results

The proposed approach was implemented using the C language and the Windows XP operating system and run on a personal computer with Pentium IV-3.0GHz CPU and 512 MB of RAM. In order to verify the proposed SSSVR approach, a DRAM module dataset was used for evaluation. This study used 3-fold cross-vaildation to evalute prediction performance, the data were collected from a DRAM module manufacturer in Taiwan, and consisted of 500 records without missing values. Each record represented the data of a DRAM IC order and have 36 features to predict DRAM IC yield.

In this study RMSE is used as the objective function value in SSSVR. After the termination of SSSVR, two commonly used performance criteria, RMSE and MAPE (mean absolute percentage error), are applied to evaluate the prediction performance of SVR. The smaller the RMSE and MAPE of the SVR is, the better the prediction performance of the SVR. The best the RMSE and MAPE obtained by the proposed SSSVR for each fold (both the training data and testing data), are compared with those of the SRA as shown in Tables 1. It is noted that because the difference in RMSE between the training data and testing data is not excessive, the proposed approach appears to avoid the problem of under-fitting and over-fitting, and achieve better prediction.

Table 1. Comparison of SSSVR with feature selection and stepwise regression analysis (SRA) for RMSE and MAPE

Fold	Training data (SSSVR)	Testing data (SSSVR)		Training data (SRA)	Testing data (SRA)	
	RMSE	RMSE	MAPE	RMSE	RMSE	MAPE
1	0.03339	0.03397	0.02889	0.06486	0.07408	0.05858
2	0.01856	0.02883	0.02629	0.06877	0.06566	0.05900
3	0.02111	0.02237	0.01931	0.06335	0.06092	0.05077
Average	0.02435	0.02839	0.02483	0.06566	0.06689	0.05612

The average RMSE and MAPE of the SSSVR with feature selection approach are 0.02839 and 0.02483, respectively. The average RMSE and MAPE of the SRA approach are 0.06689 and 0.05612, respectively. It is clear that the RMSE and MAPE obtained by the proposed approach were better than those of the SRA.

5 Conclusions and Future Research

The yield rate of DRAM module is one of the most important elements of the competitiveness of DRAM module manufacturers. This study applied the scatter search-based approach for SVR parameter determination and feature selection. The proposed approach is then applied to predict the yield rate of DRAM modules which are produced using defective DRAM ICs. The main contributions of this study include:

(1) The trial-and-error method traditionally used with SVR to determine the parameters is time-consuming and cannot guarantee the best results. The proposed approach can be used to automatically determine parameters and feature selection for the SVR using DRAM IC data.

(2) From the viewpoint of prediction error, the prediction ability of the SSSVR was superior to that of the SRA approach;

(3) This approach provides the DRAM module manufacturer the advantage of estimating yield rate of DRAM modules in advance, which can be further used to determine whether the cost of DRAM ICs is acceptable.

Comparison of the SRA and the proposed SSSVR approach with feature selection shows that the SSSVR obtains better result. The proposed approach with feature selection effectively deletes some moderating or non-affecting features and provides good prediction result.

6 Acknowledgements

The authors would like to thank the Chang Gung University, Taiwan, for financially supporting this research under Contract No. UARPD380011 and the National Science Council of the Republic of China, Taiwan, for financially supporting this research under Contract No. NSC 97-2410-H-211-001-MY2.

7 References

[1] Uzsoy, R. C.; Lee, Y.; Martin-Vega, L. A. A Review of Production Planning and Scheduling Models in Semiconductor Industry, IIE Transactions, vol. 24, p. 47-60, 1992.
[2] Wong, A. Y. A Statistical Parametric and Probe Yield Analysis Methodology, In Proceedings of the IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems, Boston, p. 131-139, October 15-18. 1996.
[3] Collica, R. S. A Logistic Regression Yield Model for SRAM Bit Fail Patterns, In Proceedings of the IEEE International Workshop on Defect and Fault Tolerance in VLSI Systems, NY, p. 127-135, March 1-3. 1993.
[4] Jun, C. -H.; Hong, Y.; Kim, S. Y.; Park, K. -S.; Park, H. A Simulation-based Semiconductor Chip Yield Model Incorporating a New Defect Cluster Index, Microelectronics Reliability, vol. 39, p. 451-456, 1999.
[5] Chen, T. An Intelligent Mechanism for Lot Output Time Prediction and Achievability Evaluation in a Wafer Fab, Computers & Industrial Engineering, vol. 54, p. 77-94, 2008.
[6] Gardner, R. M.; Bieker, J.; Elwell, S.; Thalman, R.; Rivera, E. Solving Tough Semiconductor Manufacturing Problems Using Data Mining, In Proceedings of IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop, NY, p. 46-55, February 21-24. 2002.
[7] Last, M.; Kandel, A. Discovering Useful and Understandable Patterns in Manufacturing Data, Robotics and Autonomous Systems, vol. 49, p. 137-152, 2004.
[8] Chen, R. -S.; Wu, R. -C.; Chang, C. -C. Using Data Mining Technology to Design an Intelligent CIM System for IC Manufacturing, In Proceedings of IEEE the Sixth International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, DC, p. 70-75, December 10-23. 2005.
[9] Chen, R. -S.; Yeh, K. -C.; Chang, C. -C.; Chien, H. H. Using Data Mining Technology to Improve Manufacturing Quality a Case Study of LCD Driver IC Packaging Industry, In Proceedings of IEEE the Seventh ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, NY, p. 115-119, July 7-9. 2006.
[10] Wang, C. -H. Recognition of Semiconductor Defect Patterns Using Spatial Filtering and Spectral Clustering, Expert Systems with Applications, vol. 34, p. 1914-1923, 2008.
[11] SchÖlkopf, B.; Smola, A. J. Learning with kernels, Mapua Institute of Technology, London, 2002.
[12] Wang, X.; Yang; C. -H.; Qin, B.; Gui, W. -H. Parameter Selection of Support Vector Regression Based on hybrid optimization algorithm and its application, Journal of Control Theory and Applications, vol. 3, p. 371-376, 2005.
[13] Abe N.; Kudo M. Non-parametric classifier-independent feature selection. Pattern Recognition, vol. 39, p. 737-746, 2005.
[14] Laguna, M.; Martí, R. Scatter Search: Methodology and Implementations in C, Boston: Springer, 2003.
[15] Salzberg, S. L. On Comparing Classifiers: Pitfalls to Avoid and a Recommended Approach, Data Mining and Knowledge Discovery, vol. 1, p. 317-328, 1997.

Reflective Concurrent Engineering - 3^{rd} Generation CE

Shuichi Fukuda[a,1]

[a] Department of Mechanical Engineering – Stanford University, USA.

Abstract: As the changes in our society becomes more frequent and extensive and our customers' requirements get more diversified, a new framework of concurrent engineering is called for. If we pay our attention to hardware products, we realize that we have not paid enough attention to the fact that our customers would like to use our products more to their needs and to their own tastes. We have been not fully aware that our customers are very active and creative and they enjoy customizing our products. We have attached too much importance to our final product and have not considered enough how we can create life time value by supporting their customization. Software development affords a good example and we can use their continual prototyping approach in hardware development, too. But to make it possible, we have to develop another way of product realization,, where our design starts from elements and combine them into a product. Such a cell- structured or network-structured product realization will provide a basis for the 3^{rd} concurrent engineering.

Keywords: 3^{rd} generation concurrent engineering, reflective maintenance, continual prototyping, network-structured or cell-structured manufacturing.

1 Introduction

The frequent and extensive changes in our environments and situations brought about increasing diversification of our customers' requirements. In order to respond to such a change, a new framework of concurrent engineering is called for.

Hardware product producers are apt to consider their customers only as consumers or users. They often forget that their customers are very active and creative and that they enjoy customizing their products to make them adapt more to their environments and situations and to their own tastes.

Hardware engineers have attached too much importance to their final products and have not considered enough how they can create life time value by supporting customers at the operating stage. Software development affords a good example and we can use their continual prototyping approach in hardware development, too. But to make it possible, we have to develop new product realization approach. Then, our design starts from elements and combine them into a

[1] Corresponding Author Email : shufukuda@cdr.stanford.edu

product. This is very much reversed from the current one. Such a cell-structured or network-structured product realization will provide a basis for the 3rd concurrent engineering.

2 Concurrent Engineering; Yesterday, Today and Tomorrow

Our traditional concurrent engineering (CE) has been too much producer-oriented. As our society changes more rapidly and extensively, customers' requirements get more and more diversified. What the 1st generation CE achieved was to bring the downstream information as much upstream as possible because as we go downstream, the combination increases so that if we make appropriate decisions as early as possible, it will prevent the repetition of coming back and forth between different stages and to reduce time to market. But these conditions hold only if the producer can foresee the operating conditions. In the 20th century, conditions did not change appreciably from case to case so that the producer can foresee the operating conditions and once the operating conditions are known, it is the one way from design to manufacturing.

This product development style can be compared to a railroad. In the case of railroads, we can understand easily where a train is going because there are rails. So the decision we have to make is just which line we should take. Once the line is determined, the next problem is how fast and how accurately we can get to our destinations. Thus, the decision can be made at an early stage and once the decision is made, it is very straightforward so the problem is very much tactical.

The 21st century, however, is the age of voyage. On a voyage, our navigation course changes, depending upon the weather. If we come across a typhoon, we may have to make a port we have not planned. Or in the worst case, we may have to abandon our initial port of destination. In such a case we have to ask ourselves what was the primary purpose of this voyage? This is very much of a strategic problem. Therefore the 2nd generation CE focused more attention to collaboration, because to solve a strategic issue, we need many heads from different fields. In fact, CE sometimes stood for Collaborative Engineering.

But as environments and situations come to change more frequently and more extensively and our customers become more and more diversified, it is now time for us to develop the 3rd generation CE.

3 Creative Customers

In an attempt to develop the 3rd generation CE, what we have to remember is that we, product developers, often forget that our customers are not merely consumers or users. They are very active and creative. They would like to "customize" our products as the word "customer" literally means. We have assumed that they are very passive and our products will satisfy them, if they are good.

But we have to ask ourselves again whether these products are really "good" for our customers. This may be true when the changes were small and when there was an asymmetry of information between the producer and the customer (Figure 1) .

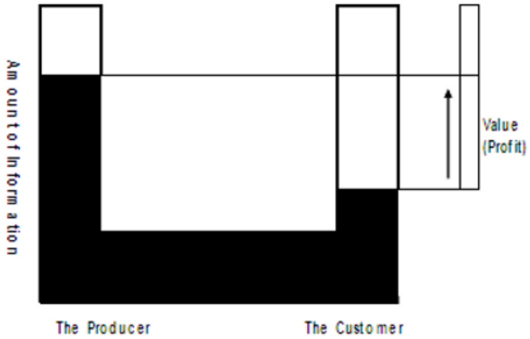


Figure 1. Asymmetry of Information

This is just like observing everything along the river. If it is a river, our customers downstream have no other choice than receiving the products produced upstream. Our traditional customers may be compared to river fishes. If we can produce good bait, these river fishes will certainly take the bait. But now our customers are salmons. Salmons are born and grow in a river but they migrate to the sea and get richer. In fact, our customers often have richer experience than us. Thus, asymmetry of information is quickly disappearing. We have to reexamine how we can develop products which will satisfy our customers. Our new direction would be to co-create value together with our customers (Figure 2).

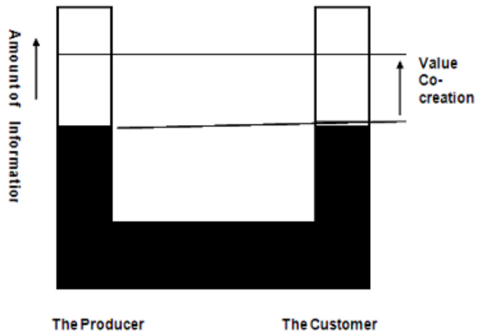


Figure 2. Value Co-creation

The concept of value co-creation was proposed by Prahalad and Ramawamy [1]. But what they discussed were the production stage and the value of a final product, as we have discussed traditionally. So let us consider again what the real value of hardware products is.

4 Hardware Product Development with Fixed Functions

Our hardware products are developed with fixed design functions (Figure 3). Their values have been discussed only with respect to the final products at the time of delivery. How much high level and how wide varieties of functions final products have are considered most important. In other words, hardware product values have been discussed only in terms of market economy. We have never stepped outside of Adam Smiths' world, where the law of one price dominates.

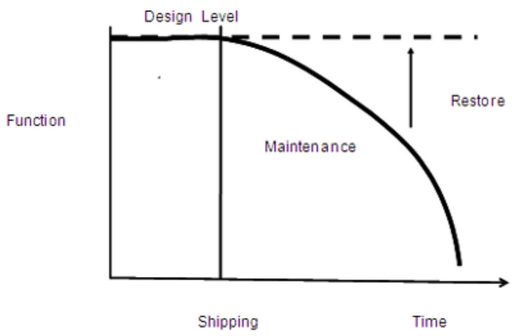


Figure 3. Hardware product development

But hardware products are physical entities so immediately after delivery, they degrade and although the whole operational period is occupied by this degrading stage, there has been not too much discussion as to maintenance in the 1st and 2nd generation CE. In fact, value engineering discusses value only in terms of final products. Degrading has been interpreted as nothing other than devaluing. We completely forgot that it is not us, but our customers who evaluate the products, and that they evaluate them when they are using, not at the time of delivery. Whether they can use them as they like in their environments and situations is the matter of their primary concern.

As hardware engineers have regarded degradation simply as devaluation, they have interpreted maintenance as an activity of restoring the degrading functions back to their original design level.

But let us take cars for example. You will be happy when you get a brand new car, but you would feel very awkward and feel uneasy when you drive such a new car. Yet after you drive the car for a long time, you would feel it has adapted to you and you can drive it more relaxed and easier. So longer usage does not necessarily mean degradation and renewing does not necessarily mean value recovery.

This was true even in the 20th century when environments and situations did not change appreciably. But today, it is no exaggeration to say that our environments and situations change from moment to moment and further these changes are very extensive. So what our customers expects is not hardware products with exactly verified design functions, but rather those which would adapt to their current needs and to their own tastes.

The 1st and 2nd generation CE's were based upon a linear model. The greatest problem is that these CE models did not pay much attention to this operational period when hardware products are really working. Traditional CE discussion did not extend beyond delivery stage. But if we consider product "development", we should cast our attention to this working stage. Our traditional CE discussion has been limited to product "realization".

To put it another way, we have been trying to produce a pair of better shoes. Whether a pair of shoes is better or worse can be evaluated after we start walking. Only the walkers can tell if the shoes are better. Our 3rd generation CE must be more customer-centric.

5 Reflective Practice / Thinking in Action

Donald Schon pointed out that in management, design and medical fields, rationality does not hold and the professionals in these fields are reflective practitioners or they are thinking in action [2].

Charles Sander Peirce proposed abduction [3]. In a closed world where boundaries can be clearly defined, the logics of induction and deduction are effective, but as the boundaries disappear and the world becomes open, these logics are no more effective. In fact, the time when Peirce proposed abduction is when Americans explored the west. The world tomorrow may be another one.

Abduction may be very roughly described in such a way: Build up a hypothesis, apply it to the current situation and if it does not work, build up another one that would work. Repeat the process until the problem is solved.

Thus this approach is very much cyclic and reflective. Schon's and Peirce's ideas have very much in common and are deeply connected to pragmatism. Pragmatism originated in UK. She is a sea-faring country and dominated the seven seas. Such a philosophy is expected to work effectively in the 21st century, i.e., the age of voyage.

Our 1st and 2nd generation CE's were based upon a linear model (Figure 4). But if we note the similarity between maintenance and remedial treatment, we will understand we should introduce cyclic or reflective approach to our hardware product development. And that will be the 3rd generation CE, which focuses more attention to the operational stage. Then, how can we make our hardware product development cyclic or reflective?

Figure 4. Linear development

6 Software Development With Evolving Functions

Although the main focus of this paper is on hardware products, software development affords an excellent example. Software used to be produced in the same way as hardware. But it soon turned out that software cannot be developed in such a way. Hardware is physical and software is non-physical. Although hardware maintenance is very difficult, there are physical indications of degrading. But software does not deteriorate and it is impossible to remove all the bugs before and during operation.

AI removed the wall between development and execution software and introduced the concept of continual prototyping. Software development does not end at the time of delivery but it extends beyond that point and keeps going on in the operational stage. Thus, software development is seamless all through its product life cycle. Software development now goes on in cycles (Figure 4).

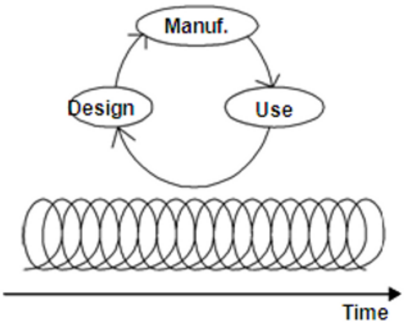


Figure 5. Cyclic development

This makes it possible for software engineers to provide a very basic level of functions at first and when their customers get familiar with this level of functions and they call for higher level ones, software engineers can develop a little bit higher level of function and provide this new version to them. Thus, software functions grows or evolves with time and with customers (Figure 6).

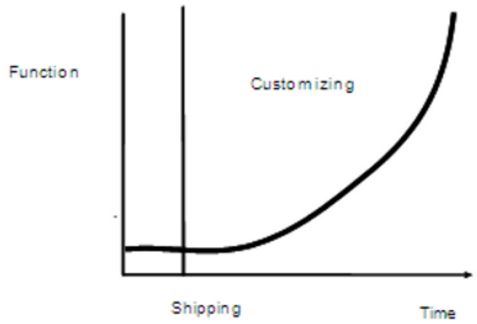


Figure 6. Software development

The advantages of continual prototyping are,

(1) Software engineers do not have to worry about debugging and about the complexity of a system. They can develop a simple program each time so they can “see” and “feel” far better. Besides they can understand how much time and cost will be necessary
(2) Customers are provided with a narrower and better perspective so that they can learn a system more easily and they will know what they should expect. Systems work as they expect because each added program is simple. They can also “see” and “feel” the system. As they learn, their confidence grows so they put trust in the system.
(3) Once the customers get familiar with the system, they know what they really need to solve their own problems at hand. So they feed back their requirements to software engineers. When the new version arrives, all of their problems may not be solved but the customers feel happy because their voices are heard. No matter how small the solution may be, it is more important for them that that there is a feed back loop and that their voices were heard.

Thus, software development changed its style from a linear model to a cyclic or reflective model. Linear model is a feed forward system (Figure 7).

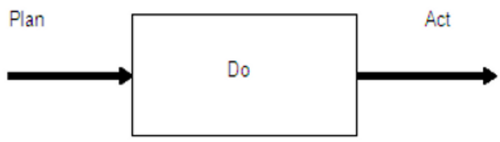


Figure 7. Linear Model

A cyclic model may be represented as Figure 8. This is Shewhart Cycle used in quality management. If we interpret P (Plan) as Build up a hypothesis, D (Do) as Apply it, S (Study) as Study if it works, and A (Act) as Solve the problem using this hypothesis, then we would know Shewhart Cycle and Peirce's Abduction have much in common.

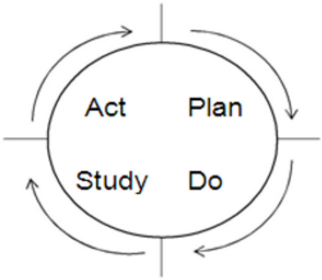


Figure 8. Cycle Model

Further, if we represent PDSA cycle as Figure 9, we would know that these concepts may be modeled as a system with a feedback loop

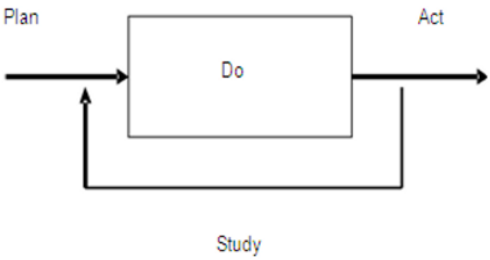


Figure 9. System with a feedback loop

7 Customization Magnifies Value

When engineers talk about "creativity", they discuss how creative they can be in designing a new product. They are apt to think that their customers are passive consumers or users. They are not fully aware that their customers are very active and creative and they invent new ways of using their products. In other words, they are very creative in customizing their products. They are literally "customers". We should keep this in mind and develop a new creative way of maintenance. If we note that hardware products are physical entities, can we really get them back to the original state? Or even if we could, will our customers be happy if we restore their product functions to the designed state? We have to remember that physical entities adapt to situations. After hardware products are delivered, they adapt to environments and situations which vary from customer to customer. The more adapted a product is, the better a customer would feel. This is already described in the above, taking a car as an example.

To put it differently, human and machine work together as a team. If machines are renewed, then humans have to be renewed, too. That is, customers have to tune up their machines all over again from the start and they have to accustom them.

Further, it must be added that our customers will enrich their experiences through customizing their products, which will make up stories. So they are ready to pay more for customization, because they can share their experience and tell their stories to their friends. Harley-Davidson is a good example of such a case.

8 How Can We Make Hardware Maintenance Reflective?

Our hardware product development was fundamentally very linear and tree-structured. But the concept of continual prototyping would change the whole picture. Current approach starts from the whole and breaks it down to elements. But continual prototyping enables us to go the other way around and start from designing functional elements and combine them into a whole product. We provide a basic combination first and as our customers get accustomed to it, we will make this combination more complex with higher levels and wider varieties step by step. Then, we could make our hardware product flexible and adaptive enough to cope with customers' expectations which vary with time and with situations. A tree structure has only one output node, but in a network structure, any node can be an output. While current development considers only one time value at the time of delivery, such hardware products will create a life time value. Therefore, our customers will turn out to be more loyal to us. So we could sell less, but could gain more in the long run.

This is very much opposite to the 1st and 2nd generation CE's, which aimed to shorten time to delivery and to reduce cost. If we can establish such kind of reflective CE, we can get our customers involved in our product development so that time to delivery would not be too much of a problem. Besides the initial version is so simple that it is easy to estimate time and cost for its production anyway.

It should be noted that this way of development is what micro- or nano-technology is practicing now. Micro- or nano-technologies start from examining manufacturability of elements or parts and then study the possibility of putting them together into a product. The flow of reasoning is reversed. Currently, our product development starts from the whole and breaks it down to elements. Network-structured manufacturing is where micro- or nano-technology is heading for.

Such network-structured or cell-structured development is very much similar to electronics design. There, they produce functional elements and constitute a whole product by assembling them together. The concept of electronics design is discrete. But most of our hardware products, especially mechanical products, have been produced based upon the idea of continuum. Their concept is in essence analog.

9 Emotional Satisfaction

The 1st and 2nd generation CE's framework is basically analog and paid most of their attention and efforts to the realization of a final product with a very wide variety of full fledged high level functions.

Donald Norman pointed out [4] that in the old age, machines were simple. People knew what to expect from them and they operated as they instructed. So people put trust in them. But today, they become too much complex. People don't know what to expect and when they give instructions, machines do not operate as they expect. So they are losing trust very quickly.

The reflective or cyclic maintenance will provide our customers with a sense of safety and security. We have been discussing value only in terms of functions. But these emotional aspects should also be considered in the 3rd generation CE. In short, the 3rd generation CE should be directed more toward how we can validate our customers' expectations rather than how we can verify the requirements.

The continual prototyping or step by step approach would provide our customers with the sense of confidence and trust. Today our customers have a sense of uneasiness and concern because many new and never-experienced products are put into market and what makes the matter worse, their life cycles are getting shorter and shorter. Thus, our customers don't have enough time to really convince themselves that these products are really good to them.

If we can introduce network-structured or cell-structured and cyclic or reflective product development in our hardware products, we can make our customers happy by removing these concerns and make their customization more enjoyable to them.

This is in a way we are co-creating product value with customers. But this value co-creation is different from the one proposed by Prahalad and Swamy who focused their attentions to a final product realization.

The product development approach described here is in essence selling "services" or "engineering" to our customers. We are no more just selling products. This approach has another great advantage. Hardware products are physical so that they cannot be divided. Therefore, if we sell products, it will bring about "win-lose relationship". That is why hardware producers are competing so hard to win the larger market share. But if we adopt this approach, then we could possibly create "win-win relationship" just as Web 2.0 demonstrates.

10 Summary

The framework for the new concurrent engineering, the 3rd generation CE, is proposed in order to respond to the frequent and extensive changes in our environments and situations and therefore increasingly diversfiying customers' requirements.

Main focus is placed on hardware product development. Although the 1st and 2nd generation CE focused their attention to product realization and made efforts to produce a good quality final product, the 3rd generation CE will pay major focus upon the operational stage and how we can create life time value by supporting our customers to customize our products to their current needs and to their own taste.

To realize such a goal, our product realization has to be changed from the current one way and tree-structured production to the cyclic and network-structured or cell-structured one.

11 References

[1] PRAHALAD, C. K.; SWAMY, V. The Future of Competition: Co-creating Unique Value with Customers, Harvard Business School Press, 2004

[2] SCHON, D. A. The Reflective Practitioner: How Professionals Think in Action, Ashgate, 1995.

[3] http://plato.stanford.edu/entries/peirce/

[4] NORMAN, D. A. Emotional Design: Why We Love (or Hate) Everyday Things, Basic Books, 2004.

The Study of Autonomous Negotiation System Based on Auction Enabled Intelligent Agent – Using Parking Tower Asset Maintenance as Case Example

Yu-Lin Liu[a, 1], David W. Hsiao[a] and Amy J.C. Trappey[a, b]

[a]Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan
[b]Department of Industrial Engineering and Management, National Taipei University of Technology, Taiwan

Abstract: Satisfying customers' needs is one of the most critical points in the customer-oriented market space, especially for the service industry. Therefore, the issues consumers care most when using the automatic parking tower is that if it can provide efficient service at right time. In the traditional maintenance practice of automatic parking tower, the maintenance demanders (i.e., parking towers) and maintenance providers (i.e., system providers) have different viewpoints of maintenance actions, especially on the time for maintenance activities. For maintenance demanders, they would like to minimize the influence of maintenance jobs and maximize the profit from parking service provision. However, for the maintenance providers, they would like to minimize the cost spent on maintenance jobs, especially minimizing the overtime works. To provide solutions for this conflict for automatic parking tower industry, this research proposes an integrated maintenance chain. In the proposed maintenance chain, the maintenance demanders and maintenance providers are coordinated by the service center with expertise in diagnosis and prognosis to help transform the traditional maintenance price and time negotiation as a non-cooperative game into a cooperative game. Moreover, to increase the proposed maintenance chain efficiency, multi-agent system with auction authorities is introduced to realize all-day cooperation.

Keywords: Parking Tower, Multi-Agent System, Game Theory, Auction, Goodwill.

[1] Corresponding Author : Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu (30013), Taiwan, E-mail: u940810@oz.nthu.edu.tw

1 Introduction

Engineering assets contain public facilities, manufacturing machineries and service facilities (e.g., automatic parking tower, the case example of this research) which are high-valued and require efforts from multiple organizations to construct, install and maintain. When speaking of engineering asset management, extending the asset operational time is always one of the most concerned issues. To extend operational life of engineering asset, multiple functions and resources are required, including monitoring techniques, data transformation and transmission techniques, database technology, diagnosis and prognosis expertise, maintenance human resources allocation and maintenance part preparation [4] [9] [3] [5]. However, above functions and resources required to accomplish engineering asset maintenance are distributed among different organizations. Therefore, the ability to integrate above mentioned functions and resources is the critical point determining maintenance efficiency.

As mentioned above, key elements for preventive maintenance, containing monitoring techniques, data transformation and transmission techniques, database technology, diagnosis and prognosis expertise, maintenance human resources allocation and maintenance part preparation should be integrated for the purpose of extending the asset life span. Consequently, this research aims to propose an integrated maintenance chain to jointly combines automatic parking tower (i.e., maintenance demanders), service center (i.e., a fair third party with diagnosis and prognosis expertise) and system provider (i.e., maintenance providers), and maintenance part suppliers to integrate the critical elements for preventive maintenances. In the proposed maintenance chain, maintenance demanders operate monitored engineering assets (automatic parking towers) and share transformed data with maintenance coordinator. The maintenance coordinator is devoted to the researches of engineering asset diagnosis and prognosis to provide good references for maintenance and integrates maintenance information as the basis of future design improvement. Moreover, the maintenance coordinator takes charge of integrating and coordinating maintenance resources to provide timely and reliable maintenance to consolidate customer relationships. With the introduction of service center, the original non-cooperative game for maintenance time and cost negotiation can be transformed into a cooperative game by the fair third party as the coordinator to increase the maintenance chain negotiation efficiency [1].

In the network of maintenance chain, a requested maintenance job may have numbers of choices of maintenance providers to fulfill the maintenance job. Consequently, the maintenance demanders can have the negotiation power with these maintenance providers. As a result, the auction mechanism is arisen to provide solutions for this situation. However, there are lots of auction strategies, and how to choose the most suitable auction for the maintenance bidding is a critical issue. The category of auction can be separated into two ways, independent private value [12] and common value [13]. The bidder in this case is the maintenance provider, who is capable to judge the value of commodity in the cost of different technology and parts. Furthermore, value can only be known by itself, so we classify it into independent private value. Consequently, we consider the situation in finding a proper maintenance provider as follows (Figure 1):

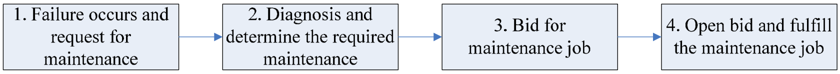


Figure 1. The process of maintenance provider determination

In step two, after receiving the parking tower's failure signal, the maintenance provider will quote its own price after estimating its own technique and inventory. In this case, no maintenance provider knows the quotation of the others, in other words, the process resembles a sealed auction. Allocating the maintenance requirement must be fair and efficient. Among these types of auction, the second price sealed auction [8], which is efficient, private, and appeal to bidders to write down their real price that they held in mind, is the most appropriate auction in this case.

The bid of maintenance construction, also called individual value, is different from normal product transaction, but similar to professional technique provision. Therefore, except price, the difference among various elements is also important, such as the quality of maintenance, the efficiency of maintenance, and the cooperation of maintenance providers. Other values, which is not be able to show in accounting, we call it intangible values. The intangible asset we are going to discuss is goodwill, which cannot be classified as tangible asset but clearly has the ability of earning profit. The common way we use to evaluate the goodwill is to generate from the inner industry standard method. The basis of the industrial standard method is the related techniques, which have been explained by the industrial economics; they have influence on the market and will produce the goodwill eventually [2]. In this case, the related techniques can be divided into three parts, maintenance technique, maintenance efficiency, and market sharing rate. We take these factors as our base and consider the willing fee of each maintenance provider. Thus, we can use the industrial standard method to estimate the value of goodwill.

During negotiation, decision-making will change with the restriction and needs of each other. It also needs to exchange the information frequently to make mutual decisions. The intelligent agent has the features of reactivity, pro-activeness, and social ability and overturns traditional ways. Traditionally, relying on phone calls or fax may lead to artificial mistakes in systems generating decisions and transmitting information immediately. Other features, adaptability and mobility enable the intelligent agent to interact with the environment through the Internet [10] [7].

By illustrating the maintenance of parking tower, we focus on the maintenance fees of parking tower and maintenance provider, and discuss the negotiation of different strategies between each side. Finally, a platform with which is capable to negotiate and integrate both sides is developed.

2 As-is Model of Parking Tower Maintenance

Conventionally, the parking tower's maintenance requires human judgments and communication between maintenance demander and maintenance demander by telephone or fax. After the maintenance provider receives the maintenance requests, it still relies on human judgments and personal technique and experiences. To compromise maintenance time and price which both sides are willing to accomplish the maintenance that also need a great deal of negotiation. However, the complicit decision tree will make the common view harder, it will also take a lot of time which maintenance requirements cannot be tolerated since the maintenance needs to shut down the whole parking tower. The maintenance time must be decided carefully, but not every maintenance time can be satisfied by the maintenance provider, in other words, it may need more expense to maintain the equipment in some particular time. In conclusion, we can describe the maintenance problems that parking tower may face as the following, and the current flow-chart of emergent maintenance processing is shown in Figure 2.

1. The diverse types of parking facility need multiple maintenance providers to fulfil different types of maintenance. This issue potentially increases the difficulty of choosing proper maintenance in timely matters.
2. The maintenance of automatic parking tower is constricted by human resource, spare part inventory, maintenance time and price, and so on. Therefore, the distribution of these maintenance resources makes maintenance a complicated and time-consuming job.
3. The huge system and database can't reflect abrupt events elastically. Owing that lots of systems and databases are involved in the maintenance chain, the linkage of these systems and databases decreases the flexibility of emergent maintenances handling.
4. The condition of assets can't be known well by the parking tower, so it can only maintain the assets passively or periodically to procure the good function of assets.
5. The repetitive failure is considered as a totally new failure, so it can't get rid of the failure efficiently.

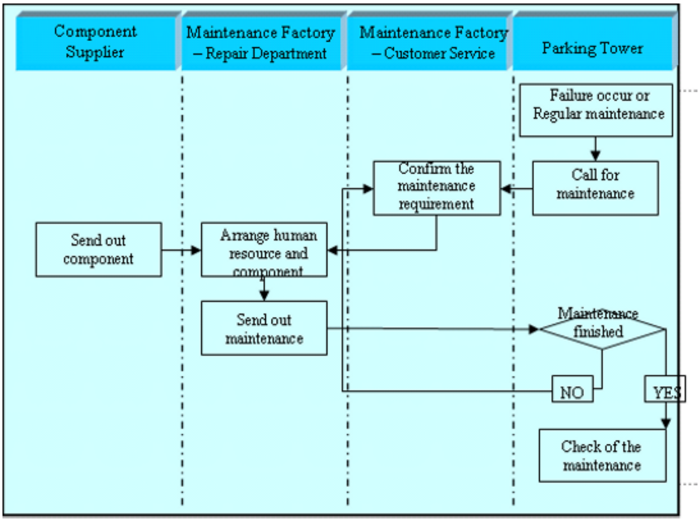


Figure 2. The flow chart of as-is parking tower maintenance processes

To solve the concerns depicted above, this research proposes the architecture of maintenance chain integrated by the service center. With the assistance of this service center, the demand and supply of maintenance can be bridged to fulfill maintenance request by most suitable maintenance providers. Meantime, the service center provides professional consultants of prognosis and diagnosis, which potentially increase the efficiency of maintenance owing that the situation of engineering assets are more well controlled. Moreover, to realize the proposed maintenance chain integration, multi-agent system (MAS) is applied to ensure all-day operation and cooperation. In next chapter, the to-be maintenance chain integration architecture will be depicted in detail.

3 To-be Model of Parking Tower Maintenance

According to the sub-problem, we bring up an integration of maintenance chain, which can be described as shown in Figure 3.

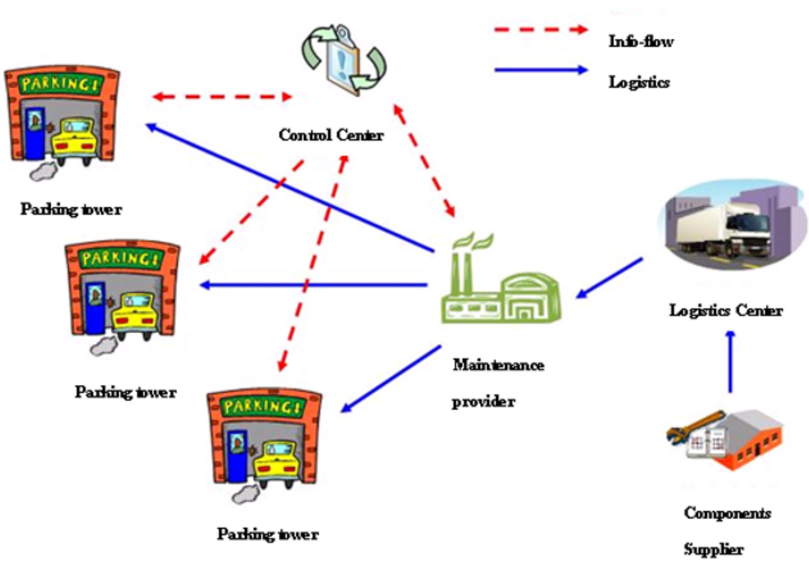


Figure 3. To-be model of parking tower maintenance framework

The maintenance chain integration has brought up a new model which provides an independent and fair third party, which is also known as the control center (service center) separated from suppliers and demanders. The maintenance system integrations also use intelligent agent to communicate and develop the strategy. The main function of control center is to deal with all the information shown from the failure until the maintenance is satisfied. The control center also has an advantage which is able to integrate the condition monitoring data of all parking towers. So it can prognoses errors accurately and determine the life cycle and the status of facilities. Moreover, it also establishes the goodwill database to avoid malicious competition and reflect the connection between price and quality more sufficiently. As we have mentioned above, the proposed maintenance chain integration have following four major functions.

1. Condition monitoring: Condition monitoring module continuously monitor the asset critical parameters to have more clear view of engineering assets. To monitor the engineering assets.
2. Prognosis: The service center has dominant expertise in prognosis, which is very helpful to tell the participants the upcoming symptoms of engineering assets. With this assistance, maintenance resources, including human resources and spare parts can be well prepared in advance.
3. Life cycle prediction: With precise life cycle prediction, the maintenance schedule can be arranged more efficiently.
4. Resource allocation: For high level maintenance jobs, they require the cooperation of different maintenance providers. With the integration by service center, the cooperation of maintenance can carried out more efficiently.

4 System Analysis, Design and Implementation

4.1 System analysis

In order to analyze the integrated maintenance chain, the first step is to establish the maintenance environment. After the control center diagnoses the failure, it will transmit the error type and failure rate to all maintenance providers and parking towers. Once a parking tower receives complete information, including all maintenance providers' fee and their best maintenance time, which means the provider could have a lower cost of labor and parts at that time, the parking tower will decide both maintenance time and fee and will send that information to the control center. After the maintenance providers send the truly bid price, a price re-calculated by the given time of parking tower, back to the control center, the negotiation begins. When the auction begins, the control center will check whether the bid price of the maintenance providers is above or under the maintenance fee which is acceptable for the parking tower. If the bid price is under the fee, the system will process goodwill evaluation. In the goodwill evaluation, with goodwill ratio and weight given by the parking tower and the historical database, we are able to evaluate individual qualified provider's goodwill. When the negotiation failed, the control center will send the result to each participator and force them to adjust their prices in order to reach a cooperative game which can maximize the profit of both sides. However, if there is still no match exists, the control center will send the second favor maintenance time to all providers and continues the route until there is a match appears between a parking tower and a maintenance provider, and both of them will schedule this new maintenance task into their original schedule. Once the maintenance is done, the parking tower will send back the goodwill score of the maintenance provider to the control center as a feedback.

4.2 System design

The development of cost calculating for parking towers and maintenance providers and the calculating of goodwill will be discussed in this section.

- **Estimation of the maintenance cost for the parking tower**

The cost of parking tower can be distinguished into two different costs, containing the internal cost and the external cost.

The internal cost is the expected lost earning due to the shut-down, including both the unexpected shut-down before a designated maintenance schedule and the shut-down during maintenance hours, which can be written as the failure rate multiplied by the accumulated expected earning added up the expected earning during the maintenance period.

The external cost is the direct maintenance fee paid for the maintenance provider, and it can be describe in three circumstances as the following:

1. The cost of direct fee at off-peak time.
2. The cost of direct fee at normal time.
3. The cost of direct fee at maintenance provider's suggestive time.

The objected function can be written as follow:

Min $\sum_{m=1}^{i-1} L_m \sum_{n=m}^{i-1} I_n + \sum_{l=i}^{i+t-1} I_l + C_i$ s.t. $a \equiv i (\mathrm{mod}\, 24)$ $C = S_k y_1 + F_1 y_2 + F_2 y_3$ $y_1 + y_2 + y_3 = 1$ $y_1 = \begin{cases} 0, hk = 0 \\ 1, hk \neq 0 \end{cases}$ $y_2 = \begin{cases} 0, hk \neq 0 \\ 1, 8 \leq a \leq 18 \end{cases}$ $y_3 = \begin{cases} 0, hk \neq 0 \\ 1, a > 18, or, a < 8 \end{cases}$ $\begin{cases} 0 < x < 24, x = x \\ x > 24, x \equiv x (\mathrm{mod}\, 24) \end{cases}$	Where $0 < t, i$ $0 \leq C_i, \forall i$ $0 \leq I_x, \forall x$ $0 \leq L_m, \forall m$ $0 \leq S_k, \forall k$ $0 \leq H_k, \forall k$ $0 \leq F_1, F_2$ T: Estimated time need for maintenance. I: Execution time of the maintenance after diagnosis. L_m: The accumulated failure rate after diagnosis m hours. I_x: The expected income for parking tower at x hour. H_k: Maintenance provider k's suggestive maintenance time at k hour. S_k: Maintenance provider k's needed maintenance fee at H_k hour. F_1: The cost of direct fee at the normal time. F_2: The cost of direct fee at the off-peak time.

- **Estimation of the maintenance cost for the maintenance provider**

The maintenance provider's cost can be classified into the cost of labor and that of parts. Both of these factors will have different costs in different situations. For example, the human resource can be classified into inside-provider or outside-provider, and the parts also can be classified into inside-warehouse and outside-warehouse. The equation of minimum cost is showed as below:

Min $\sum_{m=1}^{24}\sum_{n=1}^{2} P_{mn}N_nX_m + \sum_{l=1}^{K}\sum_{n=1}^{2} C_{ln}M_{ln}$ s.t. $\sum_{m=1}^{i+t-1} X_m = t$ $\sum_{m=1}^{24} X_m = t$ $0 < t, k$ $0 \le P_{mn}, \forall m,n$ $0 \le C_{ln}, \forall l,n$ $0 \le M_{ln}, \forall l,n$ $0 \le N_n, n=1,2$	Where t: Estimated time need for repair. k: The kind of parts. X_m: X_m is a dichotomous variable which is either one or zero. N_n: The required number of people for maintenance in situation n. While n=1, inside-provider; n=2, outside-provider. P_{mn}: The time rate needed to pay in situation n. C_{ln}: The cost of l part in situation n. n=1, inside-warehouse; n=2, outside-warehouse. M_{ln}: The needed number of l part in the situation n.

- **Second evaluation – goodwill evaluation**

In this case we use the diverse analysis method to quantify Goodwill. This is a synthetic analysis method which can take care of both the monetary and non-monetary factors. Its original equation is showed as follows:

$$\frac{A}{B} = \left(\frac{M_A}{M_B}\right)^a \left(\frac{R_A}{R_B}\right)^b \left(\frac{L_A}{L_B}\right)^c \cdots$$

This assay method can compare the goodwill of firm A with that of firm B. The variables M, R, and L are factors which parking tower are considered to be goodwill factors. The exponents a, b, c is the weight of each factor. This method provides a rational-basis evaluation which compares two firms by their ratios. Moreover, we make a transformation from the original equation. In this research, e don't compare one firm to another firm, instead we compare one firm with market average. While we consider the market average value as a numerator and the firm's goodwill value as a denominator, we can get the conclusion that if the value is less then one which means the performance of one firm is above average. As a result, we can compare more than two firms at the same time to figure out different maintenance providers' performances. In this research, we firstly choose four factors for consideration, containing (1) the efficiency of maintenance, (2) the quality of service, (3) The quality of maintenance and (4) The cooperation of maintenance provider. In the future, the proposed system will aim to look for more suitable factors and extend the evaluation scope to provide more precise evaluation.

The quality of service and the cooperation of the provider cannot be quantified easily and objectively, so we apply Fuzzy Set Theory [14] to quantify those factors. The first step is to choose the type of membership function, and two of those factors have the features of "the larger the better," so we use S-Type standard membership function [11]. And we set an interval value between one and zero as the longitudinal axis's unit, and set a one-to-ten score as the lateral axis's unit. Furthermore, we assume the grade follows the normal standard distribution.

Finally, the parking tower will have management and finance phases according to its own business strategy to assign the ratio between these two phases. For example, if the parking tower set 80% for finance phase and 20% for management phase, we will take the 80% of the maintenance provider's bid price and take the rest of 20% multiplied by the equation which we mentioned above, and add up these two numbers to get a new number which is the basis of second evaluation.

4.3 System implementation

This research uses JADE (Java Agent DEvelopment Framework) [6] to simulate the system implementation. However, before implementing agent system on JADE agent development platform, we need to define agents' mutual behaviors. In the parking tower maintenance case, identified agents' behaviors can be divided into eight types, including "listening queue inquiry," "inquiry delivery," "quotation received," "bid-open," "order delivery," "inquiry received," "report quotation" and "order received." The next step is to develop identified agents, there are three organizations running agents, containing parking tower, maintenance provider and control center. The detailed relationship of agents and organizations is depicted in Table 1. Furthermore, the agents' communication linkages with initial communication protocols are shown in Figure 4.

Table 1. The relationship of organizations and agents

Organization	Agent
Automatic Parking Tower Site (Maintenance Demander Site)	Monitoring Agent
	Production Scheduling Agent
	Set Maintenance Price Agent
	Adjusted Agent
	Scheduled Agent
Service Center Site (Maintenance Coordinator Site)	Monitor Agent
	Diagnosis Agent
	Auction Agent
	Goodwill Analysis Agent
System Provider Site (Maintenance Provider Site)	Scheduled Agent
	Human Resource Agent
	Inventory Control Agent
	Maintenance Time Advised Agent
	Bid Agent
	Adjusted Price Agent

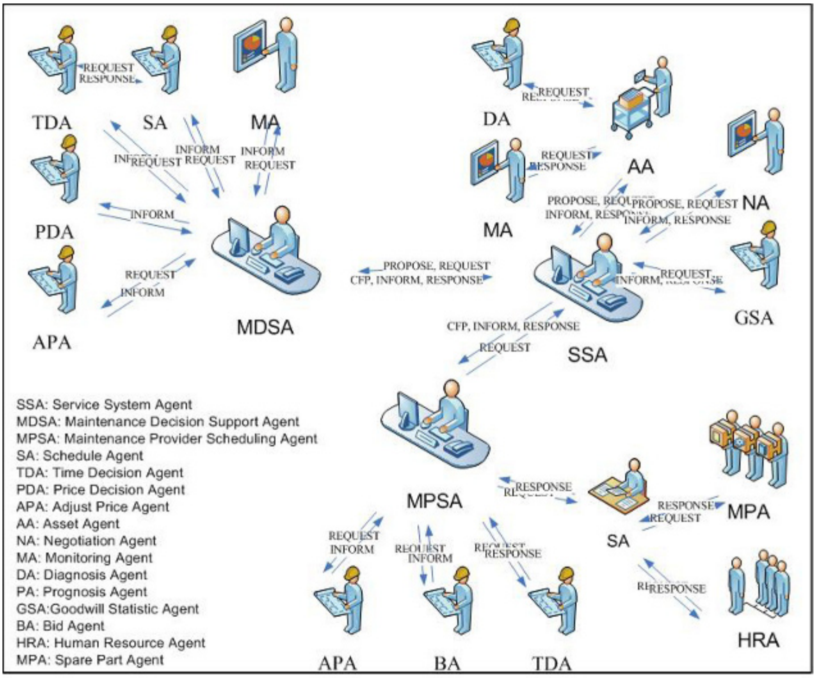


Figure 4. The communiate net of multiagent system with communication protocols

5 Conclusion

With the introduction of the proposed maintenance chain integration, the maintenance demanders and suppliers are well bridged to promote the chain efficiency. Moreover, the service center acting as the fair third party ensures the chain operating as a cooperative negotiation game. Finally, with the assistance multi-agent system with negotiation authorities, the information systems and databases are solidly interlinked and contribute to better efficiency. This research has following contributions. First, this research develops a proper game theory or auction theory to achieve the goals of maximizing the whole profit and getting the best equilibrium solution. Second, this research builds up an agent structure according to the decision model of each side, interchanging information by agent system immediately. Third, this research develops a multi-agents platform allowing every stage of agent system to negotiate and coordinate. Fourth, this research establishes goodwill evaluation considering not only the maintenance fee but also the bidder’s goodwill, quality of service, and efficiency of maintenance. Finally, this research applies fuzzy membership to quantify the numerical values.

6 References

[1] Bian. Z.-A., and Luo., J., 2006, “A Cooperative Game Model for Agent Negotiation in Network Service,” Proceedings of the 10th International Conference on Computer Supported Cooperative Work in Design.

[2] Chen, D, 2001, ”The Empirical study on the determinant of good will” Chung Yuan Christian University.

[3] Davis, D.N., 2000, “Agent-based decision-support framework for water supply infrastructure rehabilitation and development,” Computers, Environment and Urban Systems, Vol. 24, pp. 173-190.

[4] Fu, C., Ye, L., Liu, Y., Yu, R., Iung, B., Cheng, Y., and Zeng, Y., 2004, “Predictive Maintenance in Intelligent Control Maintenance Management System for Hydroelectric Generating Unit,” IEEE transactions on energy conversion, Vol. 19, No. 1, pp.1-8.

[5] Han, T., and Yang, B.S., 2006, “Development of an e-maintenance system integrating advanced techniques,” Computers in Industry, Vol. 57, Issue 6, pp. 569-580.

[6] JADE (Java Agent DEvelopment framework). Available at: <http://jade.tilab.com>. Accessed on: Jan. 12th 2009.

[7] Jennings, N. R., Sycara, K., and Wooldridge, M., 1998, “A Roadmap of Agent Research and Development,” Autonomous Agents and Multi-Agent Systems, 1, pp.7-38.

[8] Klein, S., 1997, “Introduction to Electronic Auctions,” Electronic Markets, Vol. 7, No. 4, pp.3-6.

[9] Li, Y., Chun, L., and Ching, A.N.Y., 2005, “An Agent-based Platform for Web-enabled Equipment Predictive Maintenance,”Proceedings of the 2005 IEEE/WIC/ACM international Conference on Intelligent Agent Technology (IAT’05), pp.132-135.

[10] Lu, H.-H., 2003, “Applying Intelligent Agents in Knowledge Management System,” National Sun Yat-sen University.

[11] Mendel, J.M., 1995, “Fuzzy Logic Systems for Engineering: A Tutorial”, Proceedings of the IEEE, Vol.83, No.3.

[12] Vikrey, W., 1961, “Counterspeculation, Auctions and Competitvie Sealed Tenders”, Journal of Finance, Vol. 16, No.1, 8-37.

[13] Wilson R., 1977, “A Bidding Model of Perfect Competition,” Review of Economic Studies. 44:511—518

[14] Zadeh, L. A., 1965, “Fuzzy Sets”, Informat. Control, Vol.8, pp.338-353.

Value Engineering

KBE and Manufacturing Constraints Management

Richard Curran[a,1], Wim J.C. Verhagen[b], Ton H. van der Laan[d], MJT van Torren[e]

[a] Chairholder, Aerospace Management & Operations, TUD
[b] PhD. Candidate, Aerospace Management & Operations, TUD
[d] Stork Fokker AESP BV
[e] Chairholder DAR, TUD

Abstract: The paper is associated with Manufacturing Knowledge Management and so that stream of the work has been formalised into the KNOMAD methodology. The Knowledge Optimized Manufacture And Design (KNOMAD) is a methodology for the analytical utilisation of manufacturing knowledge within design. The main contribution is to highlight the validity of the KNOMAD methodology and to show how that can be incorporated within a KBE approach in order to show how manufacturing knowedlge management can be used to aid decision making in analysing the implications of various solution choices.

Keywords: Knowledge Management, Knowledge-based Engineering

1 Introduction

This paper is primarily associated with Manufacturing Knowledge Management and so that stream of the work has been formalised into the KNOMAD methodology. Consequently, Knowledge Optimized Manufacture And Design (KNOMAD) is a methodology for the analytical utilisation of manufacturing knowledge within design. The main contribution of the work here-in is to highlight the validity of the KNOMAD methodology and to show how that can be incorporated within a KBE approach in order to show how manufacturing knowedlge management can be used to aid decision making in analysing the implications of various solution choices.

Many projects within the aerospace sector and in general across automotive etc have the challenging objectives of meeting difficult production cost and rate targets for the future, as well as addressing environmental economic and environmental performance requirements for future products. An important enabler in this repect is to be able to carry out full optimisation and integration, for exmple in application to the design and manufacture of aerospace structures as presented later in the

[1] Corresponding Author : Delft University of Technology, Faculty of Aerospace Engineering, Chair of Aerospace Management & Operations, Kluyverweg 1, 2629 HS Delft. Tel: +31 (0)15 27 88288. E-mail: R.Curran@tudelft.nl; www.lr.tudelft.nl/amo

paper - including rapid and automated assembly of structures and systems. Full optimisation and integration will require the early and continuous use of Knowledge-Based Engineering (KBE) solutions that must be integrated over the various design domains. Currently, design processes and commercially available solutions (e.g. Dassault CATIA) feature some integrated optimisation over various disciplines, but these processes are still essentially sequential; in particular, manufacturing engineering objectives and constraints for optimisation purposes are not (fully) included into current KBE applications. Besides this shortcoming, existing KBE solutions are not very well suited to represent and re-use their embedded knowledge, while existing Knowledge Management (KM) solutions fail to properly interface with KBE applications to iteratively transfer essential knowledge for use in the KBE routines. To overcome this, research programmes are being carried out to investigate and develop technologies necessary for design, development and validation in the production environment. A critical part of this is to research and develop a suitable Knowledge Management (KM) solution for use in an aerospace manufacturing environment. Such a manufacturing KM solution should at least fulfil the following major objectives:

- Transfer manufacturing knowledge to the upstream design disciplines
- Provide the knowledge base & rationale(s) from which manufacturing rules/constraints/algorithms for use in engineering applications can be extracted
- Enable early estimation and optimization of aerospace design based on manufacturing parameters and knowledge

In this paper, initial research into this area will be presented. First, the research objective will be briefly discussed. After this, the theoretical framework will be elaborated. This consists of short concept definitions and a discussion of available KM methodologies, which ends with a selection of the most appropriate KM methodology to use for aerospace manufacturing knowledge management. Subsequently, a case study is identified and performed; it is comprised of an implementation of manufacturing knowledge into a knowledge management application. Finally, some conclusions and recommendations for further research are discussed.

2 Research Objective

The research objective of this paper is to explore and develop a Proof-of-Concept Knowledge Management (KM) solution for use in an aerspace manufacturing environment. The Proof-of-Concept solution must use the right KM methodology for the manufacturing environment, in conjunction with a suitable Knowledge Management software application for actual knowledge base construction. The Proof-of-Concept is an initial deliverable on the road towards the final KM solution, which must fulfil the objectives as stated in the introduction. The Proof-of-Concept will be further explored in the case study. A first and essential step is to form a theoretical framework to be able to properly conduct the case study.

3 Theoretical Framework

Fiure 1 introduces the general classification of knowedlge types. There are in addition 3 key areas for interest. Knowledge acquisition (KA) is the transformation of knowledge from the forms in which it exists (be it tacit or explicit) into forms that can be used in a knowledge-based system. Or alternatively: KA includes the elicitation, collection, analysis, modeling and validation of knowledge for knowledge engineering and knowledge management projects. Knowledge management (KM) [1] 'is a systematic, organized, explicit and deliberate ongoing process of creating, disseminating, applying, renewing and updating intellectual and knowledge-based assets of the organization to achieve organizational objectives'. Knowledge-Based Engineering (KBE) [5] 'is a technology based on the used of dedicated software tools that are able to capture and re-use product and process engineering knowledge. The main objective of KBE is reducing time and cost of product development by means of: Automation of repetitive, non-creative design tasks & Support of multi-disciplinary integration starting from the conceptual phase of the design process'.

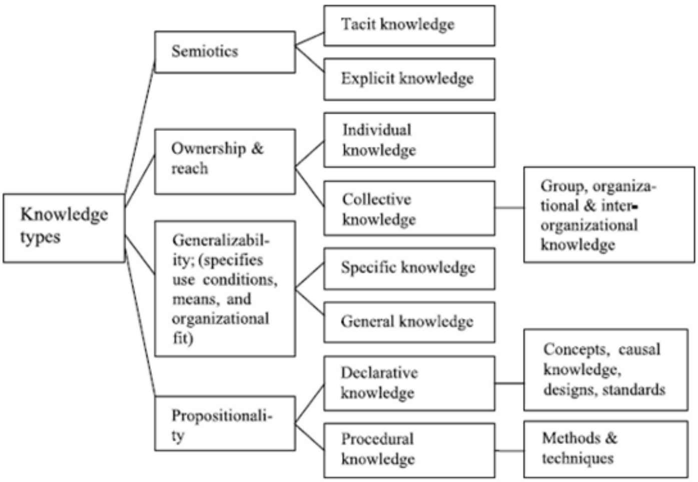


Figure 1: classification of knowledge types [6]

Knowledge acquisition is the most narrowly defined concept. Knowledge management goes a step further than KA by having a wider organizational focus and including an iterative perspective. Knowledge-based Engineering retains some aspects of KM, but is more focused on the implementation of knowledge into engineering applications in order to automate design tasks and support multi-disciplinary conceptual design. In other words, KM & KA deal with defining and registering the rationale, context and background of the knowledge. KBE is more concerned with the application of knowledge; vital principles in the form of algorithms, constraints, rules etc. are extracted from the knowledge base to be applied in engineering applications. The applications to be reviewed can all be categorised as being either KA, KM or KBE. Commonly, the KA & KM solutions lack working interfaces with engineering applications, which has evident drawbacks in terms of user involvement, user feedback and end-user worth in design involvement. KBE applications lack a supporting knowledge base detailing the rationale and context of the knowledge, which may lead to misinterpretations and frustrations on the part of the end-user ('why am I restricted in my design options?'), while failed heritage of knowledge over multiple applications is a huge risk – if knowledge is represented in detailed, KBE application-specific language, the knowledge is difficult to extract and transfer to new engineering solutions.

KA, KM and KBE can be linked with the MOKA KBE lifecycle through combination with the relevant cycle elements. KA is concerned with the capture phase. KM is primarily concerned with the capture and formalise phases, but also includes aspects of the identification, justification and use phases. Finally, KBE is connected to all phases, but is of primary interest in the formalise, package, distribute and use phases.

Some well known Knowledge Management Methodologies are compated comparison in Table 1 while Figure 2 presents the MOKA KBE Lifecycle & ontology

Table 1: Knowledge Management Methodologies (adapted from Skarka, 2007)

Methodology	**MOKA**	**CommonKADS**	**SPEDE**[2]
KB system specialization	KBE – designing	KBS – no specialization	KBE – no specialization
Range of use	CAx, automotive, aeronautical industry	Banking, food, car manufacture	Aeronautical industry
Modelling languages	MML, UML	UML	MML, UML
Models	Informal model (ICARE – Illustration, Constraint, Activity, Rules, Entity forms) and Formal model (product model, design process model & rationale)	Organization, task, agent, knowledge, communication, design models	Procedural model (based on gating criteria), process model

The main requirements to such a framework shown in Figure 2 are (Colledani et al., 2008):

- *Flexibility:* model must be easily adaptable to describe many different production system architectures, processes and product features
- *Extendibility:* model must guarantee the possibility to rapidly extend the level of detail
- *Scalability:* model must be able to support the description of product, process and production system at different levels of detail
- *Integration:* interrelationships must be taken into the framework, as all is part of the manufacturing environment. Also, integration with other disciplines (Verhagen).

---

[2] Source: http://www.epistemics.co.uk/Notes/125-0-0.htm

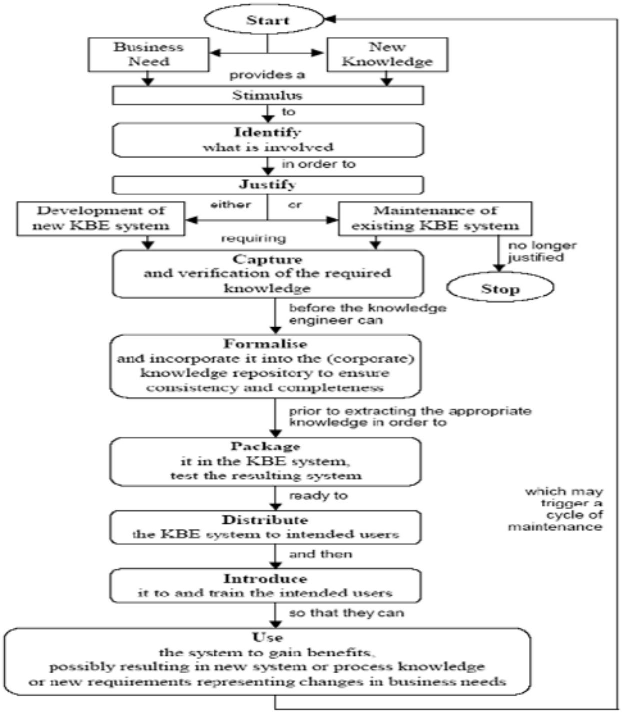


Figure 2. MOKA KBE Lifecycle [2]

The theoretical approach to Knowledge Management has been has now been established but still that is only effective if properly used with the business process, whether for general governance or analytically with the engineering process. An example of a KBE framework for implementation of the design cycle is presented in Figure 3. The domain knowledge is structured into a modelling framework that outputs Report Files that can then be used to identify an optimal design solution according to an Evaluation module, which tests against preset evaluation criteria in determining a good solution. This is first facilitated by certain analysis models, e.g. manufacturing cost, which use the data generated in the Report Files to provide new Data Files with information that links directly to the established Requirements.

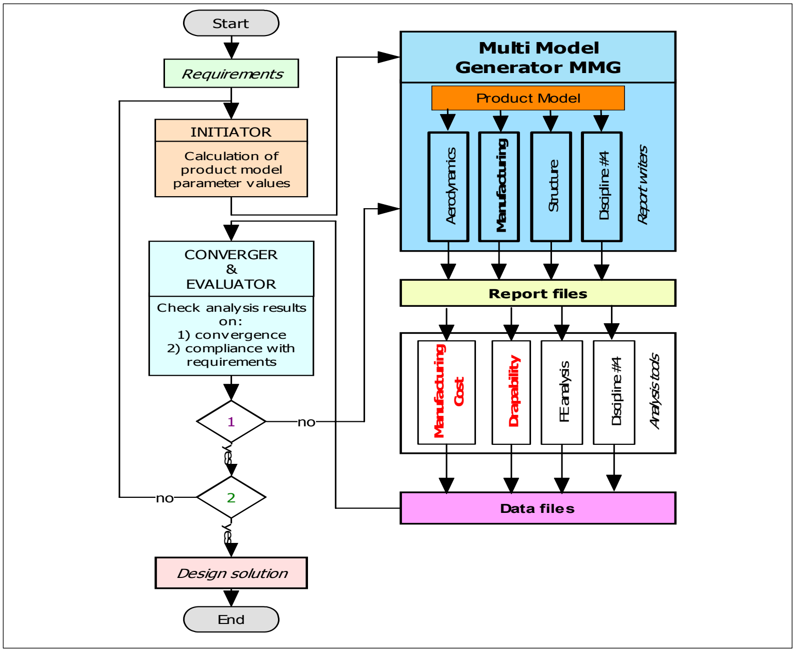


Figure 3. KBE framework for design cycle hierarchy

This paper is primarily associated with Manufacturing Knowledge Management and so that stream of the work has been formalised into the KNOMAD methodology. Consequently, Knowledge Optimized Manufacture And Design (KNOMAD) is a methodology for the analytical utilisation of manufacturing knowledge within design. The acronim has also been developed to aid in the implimentation of the specific formalised proces step, in order to simplify the basic implimentation and make it more user-friendly:

(K)nowledge identification, capture, and formalisation (for manufacture).
(N)ormalisation of data/information according to agreed standards.
(O)ptimisation strategy defined for design integration of manufacture knowledge.
(M)anufacturing verification of operational issues within application.
(A)uthentication of new (manufacturing) knowledge for sustainable improvement.
(D)esign implimentation.

The KNOMAD methodology was then used in a case study to look at the use of manufacturing knowledge in the assessment of manufacturing cost and that will be addressed in the following section which essentially deals with the validation of the theoretical material presented.

4 Case Study: Implementing manufacturing knowledge into a knowledge management system

The following case study provides a practical example of the implementation of the KNOMAD manufacturing knowledge management approach. This was carried out on a 'moveable' control surface used on aircraft, e.g. a flap, aeleron, rudder etc. This use case was then developed to do a trade-off study between to fastening processes that could be utlised in a manufactiring design solution: traditional reviting and Friction Stir Welding (FSW). Consequently, the goal of the study was to highlight the validity of the KNOMAD methodology and to show how that could be incorporated within a KBE approach in order to show how manufacturing knowedlge management can be used to aid decision making in analysing the implications of various solution choices.

Figure 4 illustrates in some more detail the actual analysis process taken in implementing manufacturing knowledge management within a facilitating KBE framework, such as that in Figure 4. The ultimate output was to be manufacturing cost, being one of the requirements to be tested in the solution evaluation. In this case the geometry is assessed for interfaces that are to be joined or fastened together, along with any general manufacturing knowledge and data that needs to be used. Figure 5 highlights that the product can be structured in the knowledge base according to the product architecture, so that part geometry and associated manufacturing knowledge is made available at a range of more global levels, as well as for the particular localized analytical model. The cost estimation process then investigates each of the interfaces by inputting the local geometry to cost estimating functions that are stored in libraries that can be viewed as specific process knowledge repositories. As stated, the analysis was run for two fastening processes, traditional riveting and FSW.

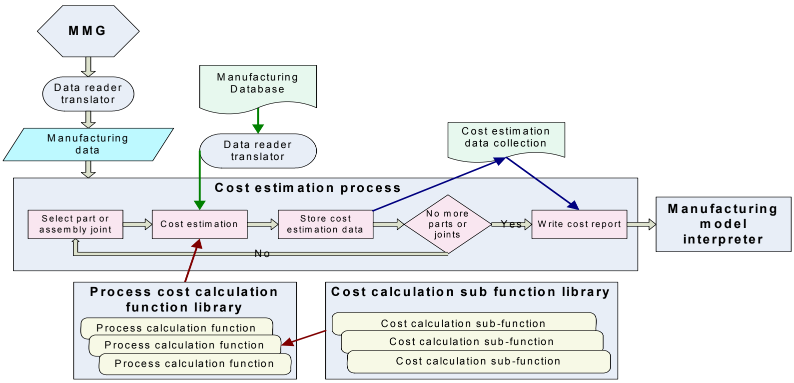


Figure 4. Modelling process for the KBE assessment of manufacturing cost

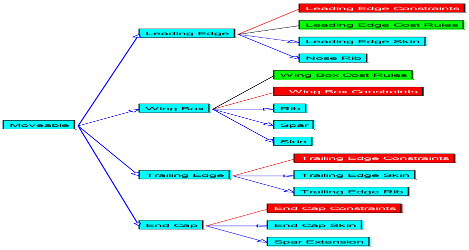


Figure 5. Knowledge base archetecture for moveables

As an example of the specific localized process knowledge exploitation, Figure 6 presents the FSW manufacturing characteristic that was stored numerically in the Function Library shown in Figure 5. The Initiator shown in Figure 3 generates the geometrical definition to be considered in the manufacturing modelling process (Figure 4), which is then used as an input for the FSW process (Figure 6).

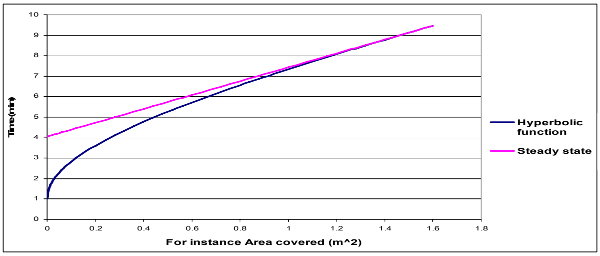


Figure 6. Model for KBE characterisation of FSW

Figure 7 presents the results of the trade-off analysis. It has been stated that the main goal of the study was to highlight the validity of the KNOMAD methodology and to show how that could be incorporated within a KBE approach in order to show how manufacturing knowledge management can be used to aid decision making in analyzing the implications of various solution choices. Therefore, the case study looked at the manufacturing cost trade-off between to processes. Figure 7 is explicit validation of the power of manufacturing knowledge management as specific manufacturing knowledge has been exploited using the KNOMAD methodology to provide crucial information in terms of the implication of design decisions in terms of both cost and process time.

The results in Figure 7 are very supportive of the KNOMAD approach as it is apparent that FSW is estimated to be almost three times faster and 20% cheaper. That result of both production rate and cost improvement would probably not be captured at the same time (concurrently) by industry, or at least not on a repeatable basis. This is an excellent example of the fundamental role of KM in capturing and formalizing manufacturing understanding to be used in an exploitive KBE framework so that manufacturing implications can be integrated into design.

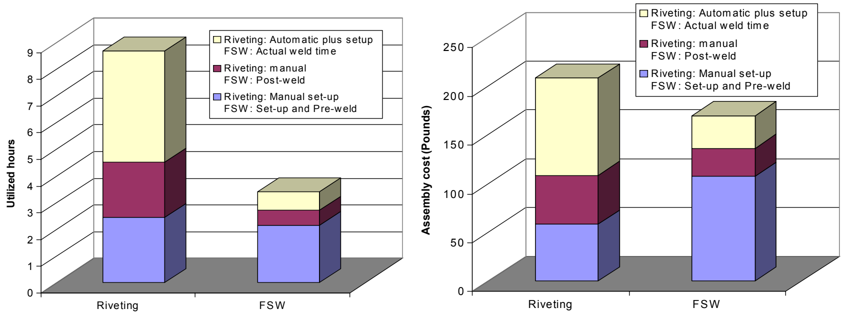


Figure 7. Comparision of KBE analysis of process 1) cost & 2) time

5 Conclusions & Recommendations

The paper is associated with Manufacturing Knowledge Management and so that stream of the work has been formalised into the KNOMAD methodology. The Knowledge Optimized Manufacture And Design (KNOMAD) is a methodology for the analytical utilisation of manufacturing knowledge within design. The main contribution is to highlight the validity of the KNOMAD methodology and to show how that can be incorporated within a KBE approach in order to show how manufacturing knowedlge management can be used to aid decision making in analysing the implications of various solution choices.

6 References

[1] Ammar-Khodja S., Bernard A. An Overview on Knowledge Management in Bernard A., Tichkiewitch S. (eds.). Methods and Tools for Effective Knowledge Life-Cycle Management, Berlin: Springer, 2008

[2] Brimble K., Callot M., Kneebone S., Murton A., Oldham K. MOKA – A Methodology for developing KBE applications, Proceedings of the 8th European Conference on Product Data Technology, Stavanger, Norway, 1999.

[3] Colledani M., Terkaj W., Tollio T., Tomasella M. Development of a Conceptual Reference Framework to Manage Manufacturing Knowledge related to Products, Processes and Production Systems, in Bernard, A., Tichkiewitch (eds.), Methods and Tools for Effective Life-Cycle Management, Berlin: Springer, 2008.

[4] Laan, van der A.H., Tooren, van M.L.J. Parametric modeling of moveables for structural analysis, 45th AIAA Structures, Structural Dynamics & Materials Conference, AIAA 2004 – 1849.

[5] La Rocca G. KBE Techniques to support Aircraft Multi-Disciplinary Design and Optimization, PhD. Thesis, to be published, 2009.

[6] Skarka W. Application of MOKA methodology in generative model creation using CATIA, Engineering Applications of Artificial Intelligence 20 (2007), pp. 677-690.

[7] Wijnhoven F. Manufacturing Knowledge Work: The European Perspective, in Bernard, A., Tichkiewitch (eds.), Methods and Tools for Effective Life-Cycle Management, Berlin: Springer, 2008.

Manufacturing Cost Contingency Management: Part a) Methodology Development

Richard Curran[a,1], Marc Gilmour[b], C. McAlleenan[d], P. Kelly[e]

[a] Chairholder, Aerospace Management & Operations, TUD
[b] Queens University Belfast, NI,
[d] UK & Bombardier Aerospace Belfast,
[e] Airport Road, Belfast, NI, UK

Abstract. The key aim of the presented work is to provide validated evidence of a manufacturing cost contingency management approach. However, the work is split into two papers due to the length with Part a) addressing methodology and Part b) addressing the application and validation. A truly concurrent engineering process is established by capturing a mixture of knowledge from the design, manufacturing and procurement functional areas, and putting it in a form that can be automatically accessed with no specialist manufacturing knowledge, thus enabling the company to quickly estimate their designs based on actual company manufacturing capability and associated cost contingency. A use-case study is ultimately presented in Part b) which estimated the rolled-up assembly time for a thin-walled stiffened structure to be within 7% of the actual recorded value. A rolled-up accuracy of ± 10% was indicative in terms of total part and assembly cost, although it was evident that some of the individual deltas deviated more significantly, tending to cancel each other out in the roll-up.

Keywords.

1 Introduction

Cost estimation in the early stages of production is challenging due to: (1) the partial amount of product and process data available, (2) the limited amount of design data that is available being highly volatile and likely to change, (3) a lack of manufacturing knowledge by those responsible for either designing the product or those responsible for putting together the estimate, (4) a lack of clarity of the causal cost drivers of a design, (5) the estimation process is resource intensive and (6) it is

[1] Prof. Dr. Richard Curran, Delft University of Technology, Faculty of Aerospace Engineering, Chair of Aerospace Management & Operations, Kluyverweg 1, 2629 HS Delft. Tel: +31 (0)15 27 88288. E-mail: R.Curran@tudelft.nl; www.lr.tudelft.nl/amo

highly dependant on subjective expert judgment. A key driver in establishing this project was the recognized difficulty in evaluating and subsequently tracking direct recurring costs that are associated with a complex product as it evolves through its development process to the stage of being manufactured.

The work presented aims to illustrate a more concurrent engineering process by capturing a mixture of knowledge from the design, manufacturing and procurement functional areas and putting that into a form that can be automatically accessed with no specialist manufacturing knowledge, thus enabling the company to quickly estimate their designs based on actual company manufacturing capability. It will be shown that this ability can have a profound affect on product development, when cost is integrated as an analytical performance metric.

2 Research Problem

Often within aerospace, the selling price is the target price for the contract, with the supplier being allowed a certain profit margin. The target cost is determined as: Target Cost = Target Price – Target Profit. The selling or target price is often specified in the Request for Proposal (RFP) and analysis this industry will embark on a top-down costing exercise using the overall market value as the starting point. Internal programmes are expected to create a bottom-up estimate based on a Bill of Material (BOM) that is used to verify that they can meet the target cost. Once the value of the work package has been agreed then the target cost becomes the sold level cost of the piece of work, which the company is contractually obliged to meet through the absorption of any deviations. It is for this reason that initial BOM based estimate is so crucial and why so much effort is currently expended in its generation. The Estimate at Completion (EAC) evaluation is a rolling cost estimate that is evaluated as the design activity progresses through detailed design and manufacture, as illustrated in Figure 1. The EAC is compared to the Sold Level in order to evaluate whether the design is on target to meet the target cost objective.

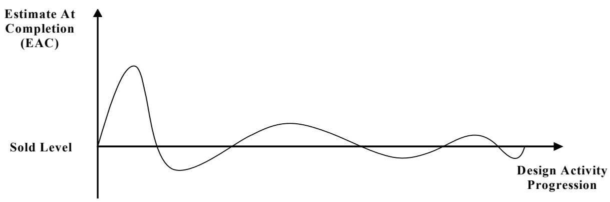


Figure 1. Comparison of EAC to Sold Level to ensure design activity meets target cost

3 Theoretical Proposition and Methodology

Manufacturing companies have been embracing Lean theory following the success of the Toyota Company, in the attempt to improve the efficiency of their operations. A product and process that is developed or reengineered along lean principles should be optimised for minimum effort production. Work in excess of a theoretical minimum can result for a number of different reasons, each of which can be linked to one or more of the seven wastes that lean theory acts to eliminate. Such reasons are: 1) Work added due to product design defects –which results in the waste of over processing. This can be avoided by good communication between the methods and design functions to ensure products are 'designed for manufacture'; 2) Work added due to ineffective production & operation methods – such as the transportation of parts arising from bad facility layout, the excess motion of operators due to poorly designed work stations, defects arising from poor selection of processes and tools, waiting arising from poor scheduling; 3) Work added due to poor planning by management such as excess inventory and over production; 4) Ineffective time due to workers such as time taken off above the allocated rest periods, i.e. lateness, deliberately working slowly.

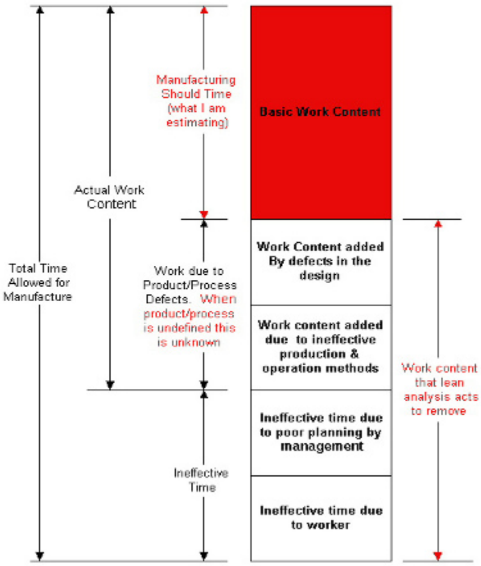


Figure 2. Breakdown of Product Work Content (Adapted from ILO, 1978)

The breakdown of the total work content experienced by an organisation during production is displayed in Figure 2, where the actual work content of a product is composed of the theoretical minimum work content plus that work which arises from defects in the design of the product and process. The work content associated with problems with the design of the product and process can only become quantifiable once the product and process has been completely specified. Ineffective time arising from worker inefficiency and poor managerial planning is only experienced, and hence accurately quantifiable, once production begins. Thus the total time that will be required for production will be the sum of basic work content, the work due to product/process defects and the ineffective time. Lean practices adopted post product and process definition act to reduce the amount of excess work due to ineffectiveness and process defects, with only minimal scope for product change.

Consequently, work content estimated at the concept stage (when neither product nor process is fully defined) must only capable of accurately quantifying the basic work content. The use of standard hours negates the need to take into consideration operator learning within the proposed methodology. The standard hour measurement assumes that work is carried out by experienced competent personnel and that the process has attained steady state conditions. In reality it takes time for an operator to become experienced and competent with the processes relating to a particular product, especially in situations where there is an extensive amount of manual work such as that experienced in aerospace assembly operations. Initial work times will therefore be much greater than those experienced after a number of products have been manufactured where lessons are have been learnt, the principle is displayed in Figure 3 for a typical learning curve (Butterfield et al, 2007). Learning is dependant on the nature of the product (complexity) and process (high automation versus low automation). Identification of suitable learning rates to be used on the programme is the responsibility of senior management and these are established by the Manufacturing and Commercial business functions prior to project start up. These rates are based on a mixture of standardised industry rates, such as those offered by the RAND Corporation (Boren & Camphill, 1970) and through organisational past experience on previous contracts. Figure also displays the potential effect that the introduction of lean principles post production has on the standard hours for the process.

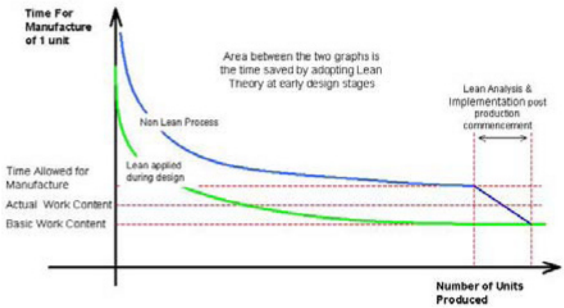


Figure 3. Work Breakdown, and the Effect of Lean in the Context of the Product Learning Curve

The methodology for the process mapping illustrated in Figure 4. It is readily evident that there is an enforced integration of the Manufacturing and Procurement functions with only one feedback loop for design revision. In addition, data management tool and knowledge repositories have been utilized to facilitate the automated retrieval of knowledge and automatic generation of data, with clear demarcation of responsibility in terms of the data categorization. The overall product estimate is the sum of the estimates associated with each of the line items within the BOM (***Product Std Hours Estimate*** $= \sum$***Line Item Std Hour Estimate***). The resultant estimate is therefore a sum of parts and it is useful to be able to quantify this range that can be considered as estimate assurance or uncertainty analysis. Uncertainty analysis concerns the evaluation of the 'goodness' or confidence associated with a generated estimate and indicates the credibility of the study. Uncertainty can arise due to: 1) Uncertainty with regard the product design and 2) Uncertainty in the cost estimation approaches

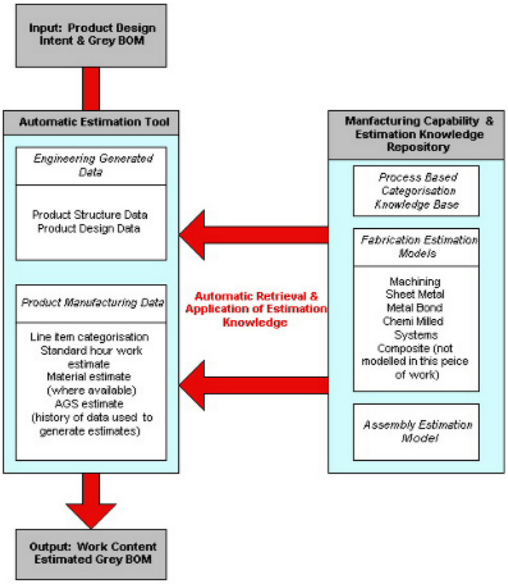


Figure 4. Future-State approach to work content and material planning

Product design inputs into the cost model are considered to be fixed (i.e. single point estimates) and so sensitivity analysis considers the effect of changes to design parameters on the total estimate. Therefore, parts of the product can be ranked according to the overall effect that changes to their design have on the total product estimate. There is then the opportunity to identify parts that require more design attention if manufacturing time is to be minimized. Each line item estimate has an associated range of values, or a confidence interval, expressed as a mean, minimum and maximum value. The upper and lower bounds for the range can easily be discerned from the upper and lower bounds of the estimates from each of the line items : ***(Product Std Hours Estimate)*** $_{Lower\ Bound}$ ***= ∑(Line Item Std Hour Estimate)*** $_{Lower\ Bound}$ and *(Product Std Hours Estimate)* $_{Upper\ bound}$ *= ∑(Line Item Std Hour Estimate)* $_{Upper\ Bound.}$. Likewise the mean of this range can also be gauged from the mean of each of the individual estimates: ***(Product Std Hours)*** $_{Mean}$***= ∑(Line Item Std Hour)*** $_{Mean}$. These values are important as they specify what the limits of the estimate are, i.e. they add credibility and confidence to referencing estimates by stating that the actual value will likely not fall below (*Product Std Hours)* $_{Lower\ Bound}$, nor will it likely exceed *(Product Std Hours)* $_{Upper\ bound}$, and it is most likely to be *(Product Std Hours)* $_{Mean}$.

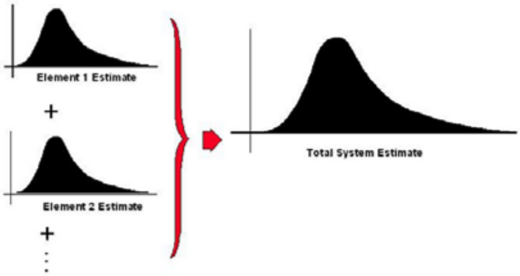


Figure 5. Generation of the Total System estimate

To quantify the likeliness of experiencing an actual value within this range of estimates every possible instance of variable needs to be considered and summed against every possible instance of the other variables within the system. The frequency of each potential output obtained then must be converted into a percentage probability as illustrated in Figure 5. Using Monte Carlo analysis, the variables can be randomly sampled and the resulting outputs considered being representative for the whole system. However, it is primarily the distribution of output from the system that is of interest as this aids in the understanding of the behavior of the output of the system. Figure 5 displays the individual Probability Density Function (PDFs) for the components of a particular system and illustrates how these can be added to obtain the PDF for the global system. Each variable is assigned a random value from its known or assumed probability distribution and these are input into the system model, with the process repeated for each of the variables many times over. The results of the simulation are also displayed in the form of Cumulative Distribution Function (CDF) graphs. The central limit theorem states that the sum of a large number of variables can be approximated to a normal distribution. The standard deviation can be used as a measure of confidence in the results space, whilst the CDF, facilitates the quantification of certainty associated with a particular point estimate.

Sensitivity analysis involves studying how the outputs of a system change with variations in the assumptions or inputs of the system. It is useful for identifying those system variables that are important in controlling system behavior. Within the context of the prototype tool, sensitivity analysis is conducted in order to identify those parts or assemblies within the product that will alter the total product manufacturing time most, should their individual design attributes change. All product attributes are subject to change, especially in the early product development stage. To apply sensitivity analysis for the purpose of estimate control, product design attributes can be changed within likely ranges. The attributes are increased or decreased by a specified amount and then the estimation activity is rerun. These new estimates can be compared to the original estimated values and relative differences, at a sub system, and at a total system level, calculated. These relative differences in the estimates, both at a component level, and at an all up level, are then considered the sensitivities of the product estimate to the changes made in the product design attributes. Two different types of sensitivity can be calculated, namely, (1) the sensitivity of each individual part estimate compared to the original estimate:

$$\% \,\Delta \text{ Part Estimate} = \frac{\text{Part Estimate}_{+/-\%} - \text{Part Estimate}_{\text{Original}}}{\text{Part Estimate}_{\text{Original}}} \times 100$$

and (2) the total product estimate sensitivity due to a change in one individual part:

$$\% \,\Delta \text{ Total Product Estimate} = \frac{\text{Part Estimate}_{+/-20\%} - \text{Part Estimate}_{\text{Original}}}{\text{Total Product Estimate}_{\text{Original}}} \times 100$$

It is the second sensitivity, that of the total system sensitivity to the change in part cost, that is of most interest. Each part or assembly within the product can be ranked according to the relative effect on the cost of the overall system. This ranking defines the priority given to the part during design changes in order to best control the manufacturing estimate. Due to the non-continuous method of estimation achieved through use of the categorization technique, it is appropriate to consider the maximum potential variation in input parameters, and determine the effect this has on the generated estimate. This is different to a more typically employed sensitivity technique, determined by calculating the change in total system output per unit input change. Figure 6 contrasts a discontinuous estimating space with that of a continuous linear estimating space, and indicates why, in the case of discontinuous ranges, it becomes necessary to consider the upper and lower attribute input bounds. It is evident that neglecting to include a sufficiently large range may result in failure to capture a step change in an estimate, affected as a result of a change in part category, resulting in erroneously ranked system components.

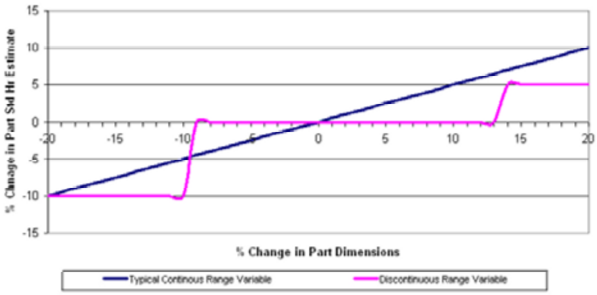


Figure 6. Contrast continuous and discontinuous estimation space

5 Discussion and Conclusions

The presented work has presented the development of a manufacturing cost contingency management methodology. A truly concurrent engineering process has been established by capturing a mixture of knowledge from the design, manufacturing and procurement functional areas, and putting it in a form that can be automatically accessed with no specialist manufacturing knowledge, thus enabling the company to quickly estimate their designs based on actual company manufacturing capability. The application and validation of the methodology is address in part b) of this paper where use-case study was presented which estimated the rolled-up assembly time for a thin-walled stiffened structure to be within 7% of the actual recorded value. A rolled-up accuracy of ± 10% was indicative for all the case studies investigated, in terms of total part and assembly cost; although it was evident that some of the individual deltas deviated more significantly, tending to cancel each other out in the roll-up.

6 References

[1] Bode J. Decision support with neural networks in the management of research and development: concepts and application to cost estimation, Information & Management 1998,34: 33-40.

[2] Brinke, ET. 2002 Costing Support and Cost Control in Manufacturing: A Cost Estimation Tool applied in the sheet metal domain, PhD Thesis, Printed by PrintPartners Ipskamp, Enschede, The Netherlands, ISBN 90-365-1726-5.

[3] Cooper R, Kaplan RS. The design of cost management systems; text, cases and readings, New Jersey, Prentice-Hall Inc, 1991

[4] Curran R, Kundu AK, Wright JM, Crosby S, Price M, Raghunathan S, et al. Modeling of aircraft manufacturing cost at the concept stage, The International Journal of Advanced Manufacturing Technology, 2006; 1 – 14.

[5] Curran R, Price M, Raghunathan S, Benard E, Crosby S, Castagne S, et al. Integrating Aircraft Cost Modeling into Conceptual Design, Journal of Concurrent Engineering: Research and applications, Sage Publications, 2005, ISSN: 1063-293X, 13:4, 321-330.

[6] Curran R, Kundu A, Raghunathan S, Eakin D. Costing Tools for Decision Making within Integrated Aerospace Design, Journal of Concurrent Engineering Research, 9(4), 2002, 327-338.
[7] Curran R, Raghunathan S, and Price M. A Review of Aircraft Cost Modeling and Genetic Causal Approach, Progress in Aerospace Sciences Journal, 40:8, 2004, 487-534.
[8] Curran, R, Rush, C, Roy, R and S Raghunathan (2002). Current Cost Estimating Practice in Aerospace, Proceedings of Concurrent Engineering: Research and Applications (CE2002), 894-902.
[9] Fiegenbaum AV. *Total Quality Control*, McGraw-Hill, New York, 1993 Meyer, C. *Fast Cycle Time: How to Align Purpose, Strategy and Structure for Speed*, Free Press, New York 1993.
[10] Hammer M, Champy J. *Reengineering the Corporation*, Nicholas Brealey Publishing:London 1993.
[11] Jiao J, Tseng M. A pragmatic approach to product costing based on standard time estimation. International Journal of Operations and Management, 1999, 19:7, 738-755.
[12] Olds L. Integration of Cost Modeling and Business Simulation into Conceptual Launch Vehicle Design AIAA 97-3911.
[13] Singh V. Systems Integration – Coping with legacy systems"; 8/1 24-28; 1997.
[14] Suri R.*Quick Response Manufacturing: A Companywide Approach to Reducing Leadtimes*, Productivity Press:Portland, Oregon 1998
[15] Towill DR. Management theory: Is it of any practical use? Or, how does a fad become a paradigm, *Engineering Management Journal,* 1999*,* 111-122.
[16] Watson P, Curran R, Murphy A, Cowan S, Cost Estimation of Machined Parts within an Aerospace Supply Chain, Journal Concurrent Engineering, 2006 14:1, 17-26.
[17] Zandin, KB. MOST® Work Measurement Systems, Marcel Dekker 1998

Manufacturing Cost Contingency Management: Part b) Application and Validation

Richard Curran[a,1], Marc Gilmour[b], C. McAlleenan[d], P. Kelly[e]

[a] Chairholder, Aerospace Management & Operations, TUD
[b] Queens University Belfast, NI,
[d] UK & Bombardier Aerospace Belfast,
[e] Airport Road, Belfast, NI, UK

Abstract. The key aim of the presented work is to provide validated evidence of a manufacturing cost contingency management approach. However, the work is split into two papers due to the length with Part a) addressing methodology and Part b) addressing the application and validation. It was shown in Part a) that a truly concurrent engineering approach can be established by capturing a mixture of knowledge from the design, manufacturing and procurement functional areas, and putting it in a form that can be automatically accessed with no specialist manufacturing knowledge, thus enabling the company to quickly estimate their designs based on actual company manufacturing capability and associated cost contingency. This Part b) of the paper presents the use-case study which estimated the rolled-up assembly time for a thin-walled stiffened structure to be within 7% of the actual recorded value. A rolled-up accuracy of ± 10% was indicative in terms of total part and assembly cost, although it was evident that some of the individual deltas deviated more significantly, tending to cancel each other out in the roll-up.

Keywords.

1 Introduction

This paper forms Part b) of the Manufacturing Cost Contingency Management work, where Part a) already presented the basic methodology to be follow, with this Part b) presenting the application of the work and the validation. The reader is therefore encouraged to read Part a) first and then this Part b) as the papers need to be read simultaneously. As established in Part A), the work aims to illustrate a more concurrent engineering process by capturing a mixture of knowledge from

[1] Corresponding Author : Delft University of Technology, Faculty of Aerospace Engineering, Chair of Aerospace Management & Operations, Kluyverweg 1, 2629 HS Delft. Tel: +31 (0)15 27 88288. E-mail: R.Curran@tudelft.nl; www.lr.tudelft.nl/amo

the design, manufacturing and procurement functional areas and putting that into a form that can be automatically accessed with no specialist manufacturing knowledge, thus enabling the company to quickly estimate their designs based on actual company manufacturing capability.

2 Use-Case Study and Basic Estimation Results

Like with the approaches developed for dealing with unknowns in the development of the fabrication models, Watson's enhanced multiple linear regression statistical analysis tool again will be utilised in order to generate statistical relations that link known assembly attributes to fastener counts. The development of such relations are only semi causal, exploiting observed correlation between such attributes and the fastener counts. They do not fully model the real causal relationships.

The primary difficulty with relying on a regression-based approach is the need for a sufficient number of similar past or current products on which to base the analysis. This is particularly problematic with assemblies, as the further up the product structure the SSC goes, then the fewer instances of these there become, and therefore the reduced certainty with which statistical models can be constructed.

Fuselage SSCs have been chosen in this instance to present the potential, and the limitations, of the approach. Fuselage assemblies have been chosen primarily due to the availability of suitable data at BAB. In all, 67 sub assemblies where considered, all of which relate specifically to regional jets. The 67 assemblies are from 3 derivative regional aircraft programs and are made up of the following SSCs: Skin panels – 19 sub assemblies; Floor sections – 4 floor sub assemblies; Longitudinal joints – 17 longitudinal joint assemblies; Circumferential joints – 5 circumferential joint assemblies; Door surrounds – 7 door surrounds; Major load bearing structures -13 different load bearing structures; Door structures – 2. As skin panels over the greatest quantity of data points they will be used to demonstrate the regression models. All panels considered are metallic and mechanically fastened.

Canvassing the observations of the domain experts it was observed that the total number of fasteners in a panel would potentially correlate with (1) the total surface area of the panel and (2) the total part count in the assembly. Both of these attributes will be known at the concept design stage and both can be determined for the BOM.

Table 1. Correlation matrix for skin panel fastener counts and known assembly attributes

	Surface Area (sq inches)	*Rivets*	*Bolts*	*Total Part Count in Assembly*	*Total Fasteners*
Surface Area (sq inches)	1.00				
Rivets	0.78	1.00			
Bolts	0.10	-0.12	1.00		
Total Part Count in Assembly	0.64	0.81	-0.28	1.00	
Total Fasteners	0.81	0.98	0.06	0.76	1.00

To validate the observations a correlation analysis was carried out the results of which are summarised in Table 1. The high degree of co dependence between surface area and part count precludes the use of both these attributes in the same relationship; therefore relationships for each must be developed independently. Carrying out linear regression analysis, in the first instance based on the surface area of the skin panels, and secondly on the number of parts, gives the graphs displayed in Figure 1a and Figure 1b.

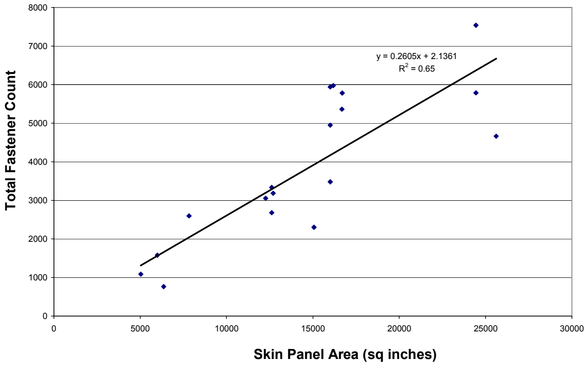


Figure 1a. Total fastener count vs. skin panel area for BAB regional jet skin panel assemblies

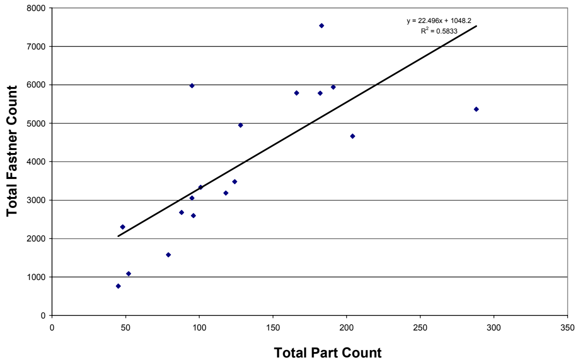


Figure 1b. Total fastener count vs. total part count for BAB regional jet skin panel assemblies

The R squared values for the relation of total fastener values and skin panel area and total part count respectively are 0.65 and 0.58 respectively. It is noted that rivets are much more prevalent than the lightweight structural bolts, and both these figures could be improved marginally by considering the relationship of both the variables with rivets alone.

A more detailed breakdown of the composition of skin panels could potentially offer better results. For example consider the constituent parts when grouped by their fabrication centre. The observation is made that machined parts, being thicker than sheet metal and extruded parts, are more likely to require the use of bolts as opposed rivets. Multiple linear regression on this basis resulted in an Adjusted R square value of 0.62, which is no improvement over that already considered. Indeed the difficulty with using these particular set of variables is the limited number of data points, and the larger number of variables, thus calling into question the validity of the derived relationship.

Also consider the break down of parts by the size categories defined previously, where the assumption in this instance is that larger parts are more likely to require a greater quantity of fasteners. The R squared value has been improved to 0.82, the problem again being one with regard the credibility of the relationship since there is 5 independent variables but only 19 data points. **Figure 2** displays the error associated with the fastener estimates for each of the skin panels.

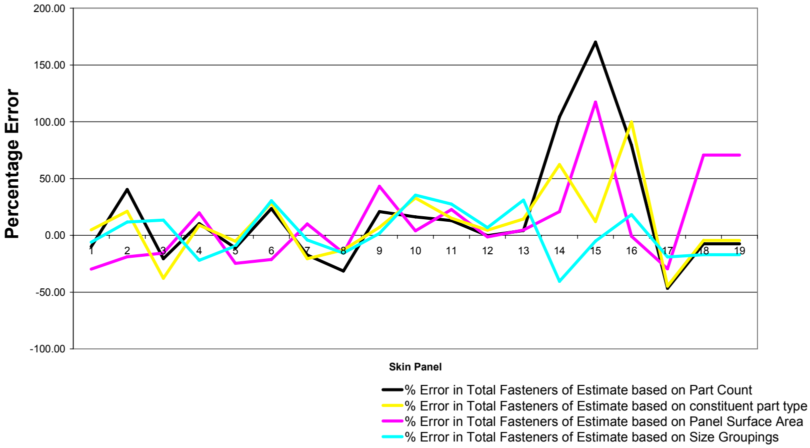
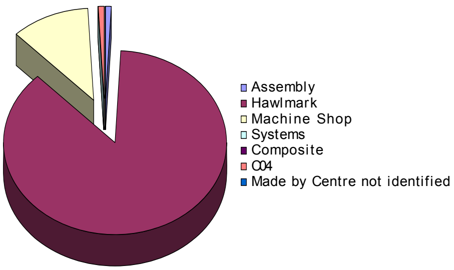


Figure 2. Percentage Error for the skin panel assemblies of the various statistical estimating techniques.

The fundamental issue with regard using a statistical based approach is that it is not fully representative of the product being developed. The assumption is made that the new product is similar to all current products. For example consider the data set used for the development of the previous relations; included within the dataset are three skin panels that are significantly different from the other panels, namely panels 14, 15 and 16. These panels are located in the fuselage centre section just above the wing pick up points. They have significantly large cut outs for emergency exit doors, and included in their structure is three heavy G frames that are used to attach the fuselage to the wings. As these entities carry high loads they are fastened to the skin panel by means of lightweight bolts as opposed rivets. Consequently for these centre panels the lightweight bolts equates to about 30% of the total fastener count whilst for the other panels the proportion is less than 2%. Their inclusion therefore within the skin panel family of parts skews the relationships developed for that family, consequently reducing the accuracy of the estimating relation. It is observed on **Figure 2** that consistently the most accurate technique is that of the regression based on the size groupings. As expected, larger parts will require a larger number of rivets. This relationship is only appropriate though where the size category is based on the same dimension on which parts are to be joined, in this instance the part length. Consider the assembly of a skin panel, where the majority of fasteners are consumed by the attachment of the frames and stringers to the skins. Both these parts are fastened along their length. On another type of SSC such as a floor structure where parts are not joined along their length, i.e. at their end points, the relationship is likely to be less obvious and therefore more difficult to generate statistical based relations.

Similarly, the issue regarding a purely statistical based approach is that although it offers the ability to develop fastener estimation models in a relative quick manner, the accuracy and representative-ness of these models may be questionable, and that they are really only applicable when sufficient data exist with regard past products, which, for the majority of SSCs, is seldom the case.



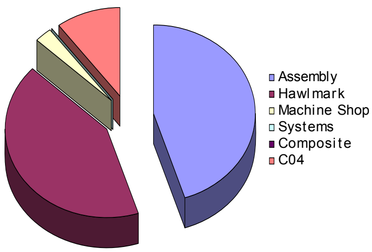


Figure 3a and 3b. Total parts broken down for new skin panel assembly & Mean Manufacturing hours broken up by Manufacturing Centre

3 Cost Contingency Management Application

A validation exercise was carried out for a new fuselage skin panel assembly, which originated from a new regional aircraft. The skin panel BOM, as provided by the design engineers, contained 51 different line items. Again the majority of parts are sheet metal details, being fabricated in the sheet metal manufacturing facility. Figure 3a and 3b shows a breakdown of the part number and mean manufacturing hours respectively by manufacturing centre. Again the large amount of manufacturing time (45%) is consumed by the assembling of the product, whilst although the product is made of 87% of sheet metal parts these account for 42% of the manufacturing time. The one chemi-milled part, the fuselage panel skin, contributes to 10% of the total manufacturing time. Figure 4 displays the spread of the total manufacturing estimate and the risk associated with each value within the range. The 10%, 50% and 90% figures for the total manufacturing time estimate are 100, 110 and 119 hours. Due to the large number of parts to be manufactured in the sheet metal facility is appropriate to carry out a Monte Carlo simulation to determine the spread of the total hours for the facility.

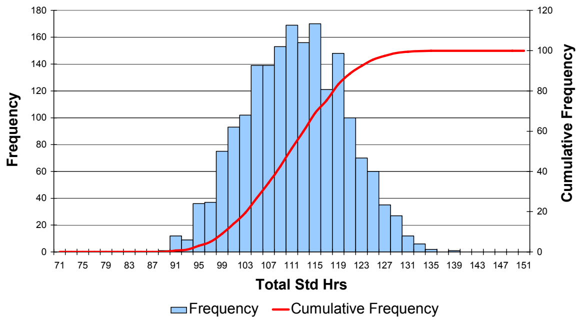


Figure 4. Uncertainty Analysis for new skin panel for all up manufacturing time

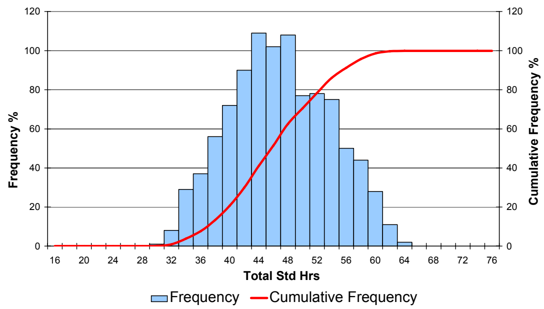


Figure 5. Sheet metal manufacturing hours for new skin panel assembly

This is displayed in Figure 5. For the sheet metal MC the 10%, 50%, and 90% figures for the amount of work for which they will be responsible for within this product are 37, 45, and 55 hours respectively. If the sheet metal fabrication manager was going to be prudent when providing a quote on the amount of work for which they would be responsible for with regard they product manufacture he could for example quote a figure as high as 54 hours, which would be associated with an 80% probability that it would not be exceeded. This illustrates how the approach can provide greater estimate understanding and indicates how estimates can be managed, not just with regard the product design but also for manufacturing operations.

Figure 6 displays a sensitivity analysis for the assembly, again based on a generalised +/-20% change in the product dimensional attributes. It is noted that the parts that require control during the design activity with regard their overall dimensions will be assembly, the chemi-milled skin, the re-enforcing frame angles and two machined fixtures.

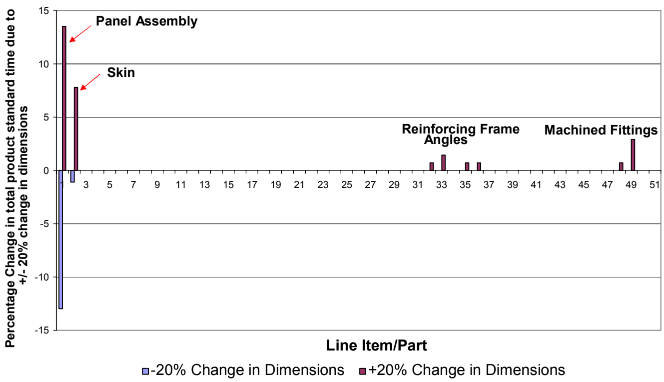


Figure 6. Sensitivity Analysis for new skin panel assembly

In this instance it was requested that the assembly estimate be broken and displayed graphically based on the skill/resource required, as shown in Figure 7, as each of these resources have different cost rates for their time. Consequently, this breakdown offers greater insight with regard the cost of the product, which in turn acts to facilitate greater cost control.

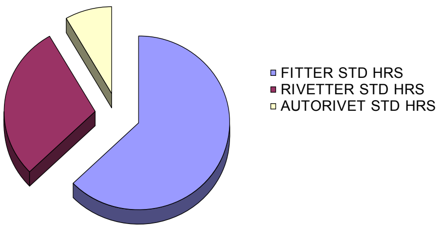


Figure 7. Breakdown of mean assembly hours by resource for new skin panels

5 Discussion and Conclusions

Following on from Part a) of this paper, the presented work has provided a validated evidence of robust value management applied to the manufacturing cost constraints analysis. A truly concurrent engineering process has been established by capturing a mixture of knowledge from the design, manufacturing and procurement functional areas, and putting it in a form that can be automatically accessed with no specialist manufacturing knowledge, thus enabling the company to quickly estimate their designs based on actual company manufacturing capability. A use-case study was presented which estimated the rolled-up assembly time for a thin-walled stiffened structure to be within 7% of the actual recorded value. A rolled-up accuracy of ± 10% was indicative for all the case studies investigated, in terms of total part and assembly cost; although it was evident that some of the individual deltas deviated more significantly, tending to cancel each other out in the roll-up.

6 References

[1]. Bode J. Decision support with neural networks in the management of research and development: concepts and application to cost estimation, Information & Management 1998,34: 33-40.

[2]. Brinke, ET. 2002 Costing Support and Cost Control in Manufacturing: A Cost Estimation Tool applied in the sheet metal domain, PhD Thesis, Printed by PrintPartners Ipskamp, Enschede, The Netherlands, ISBN 90-365-1726-5.

[3]. Cooper R, Kaplan RS. The design of cost management systems; text, cases and readings, New Jersey, Prentice-Hall Inc, 1991

[4]. Curran R, Kundu AK, Wright JM, Crosby S, Price M, Raghunathan S, et al. Modeling of aircraft manufacturing cost at the concept stage, The International Journal of Advanced Manufacturing Technology, 2006; 1 – 14.

[5]. Curran R, Price M, Raghunathan S, Benard E, Crosby S, Castagne S, et al. Integrating Aircraft Cost Modeling into Conceptual Design, Journal of Concurrent Engineering: Research and applications, Sage Publications, 2005, ISSN: 1063-293X, 13:4, 321-330.

[6]. Curran R, Kundu A, Raghunathan S, Eakin D. Costing Tools for Decision Making within Integrated Aerospace Design, Journal of Concurrent Engineering Research, 9(4), 2002, 327-338.

[7]. Curran R, Raghunathan S, and Price M. A Review of Aircraft Cost Modeling and Genetic Causal Approach, Progress in Aerospace Sciences Journal, 40:8, 2004, 487-534.

[8]. Curran, R, Rush, C, Roy, R and S Raghunathan (2002). Current Cost Estimating Practice in Aerospace, Proceedings of Concurrent Engineering: Research and Applications (CE2002), 894-902.

[9]. Fiegenbaum AV. *Total Quality Control*, McGraw-Hill, New York, 1993

[10]. Meyer, C. *Fast Cycle Time: How to Align Purpose, Strategy and Structure for Speed*, Free Press, New York 1993.

[11]. Hammer M, Champy J. *Reengineering the Corporation*, Nicholas Brealey Publishing:London 1993.

[12]. Jiao J, Tseng M. A pragmatic approach to product costing based on standard time estimation. International Journal of Operations and Management, 1999, 19:7, 738-755.

[13]. Olds L. Integration of Cost Modeling and Business Simulation into Conceptual Launch Vehicle Design AIAA 97-3911.

[14]. Singh V. Systems Integration – Coping with legacy systems"; 8/1 24-28; 1997.

[15]. Suri R. *Quick Response Manufacturing: A Companywide Approach to Reducing Leadtimes*, Productivity Press:Portland, Oregon 1998

[16]. Towill DR. Management theory: Is it of any practical use? Or, how does a fad become a paradigm, *Engineering Management Journal,* 1999*,* 111-122.
[17]. Watson P, Curran R, Murphy A, Cowan S, Cost Estimation of Machined Parts within an Aerospace Supply Chain, Journal Concurrent Engineering, 2006 14:1, 17-26.
[18]. Zandin, KB. MOST® Work Measurement Systems, Marcel Dekker 1998

Systems Engineering Methodology for Concurrent Engineering Education

Richard Curran[a] , Michel van Tooren[b], and Liza van Dijk[d]

[a] Chairholder, Aerospace Management & Operations, TUD
[b] Chairholder, Design of Aircraft and Rotercraft, TUD
[c] Director of Education, Faculty of Aerosapce Engineering, TUD
[d] Management Assistent, Aerospace Management and Operations, TUD

Abstract. The main contribution of the paper is to present and review the main proposition of this paper, that: 'Systems Engineering provides a beneficial educational framework and process for helping students to understand the aerospace design process more quickly'. The authors 1) illustrate the Systems Engineering components, 2) the integration into the actual final year design project and 3) present evidence that the SE educational framework is an effective way of teaching aerospace design. Finally, an example from the design project is presented to highlight the technical quality of work ensured through the use of the SE process.

Keywords.

1. Introduction

The main contribution of the paper is to present and review the effectiveness of a Systems Engineering approach in the aerospace design education at the Technical University of Delft (TUD). The basic approach will be outlined as well as presenting some of the main components that make up the Project Management aspect and the more technical Systems Engineering aspect. The integration of these into the actual design project will also be addressed moving to the main proposition of this paper, that: '*Systems Engineering provides a beneficial educational framework and process for helping students to understand the aerospace design process more quickly*'. In order to help validate this proposal, the paper will also provide evidence from historical records to show that a good assessment mark in the Systems Engineering component (1/8th of the overall mark allocated) is highly coupled to a good assessment mark in the actual technical design excellence. Consequently, it is the intention of the authors through this paper to 1) illustrate the Systems Engineering components, 2) the integration into the actual final year design project and 3) present evidence that the SE educational framework is an effective way of teaching aerospace design. Therefore, the relevance of the paper addresses the implementation and understanding of SE methods, the relevance and potential exploitation within design education. Finally, as an alternative formulation of the main proposition, the paper will also assess whether SE tends to limit students and their creativity by formalising and de-skilling design, or whether it provides a framework that allows them to more quickly carry out the procedural tasks that are also a part of design so that they have more time to be creative and achieve more outstanding design synthesis.

2. The Systems Engineering Methodology For Concurrent Engineering at TUD

A M.Sc. from TUD takes 5 years to complete in principle and includes a full-time 10-week design project called the Design Synthesis Exercise. Each student is required to complete this in the final year of their Bachelor Degree part of the studies and each student forms part of a team of 10 persons working on their own distinct project. Brief proposals are presented to the students by a Principle Supervisor as part of the early selection process so that they have some choice in selecting a particular project that they will be more highly motivated by. The Principle Supervisor is supported by 2 other members of staff, typically from different Chairs so that there is a range of competency, while the students are also encouraged to consult other relevant staff members during the project as required.

The general outline of the DSE is illustrated in Figure 1 which shows a first facilitating Systems Engineering phase, with associated Project Management tasks, that leads up to the Requirements Review. There then follows two design loops at a conceptual and preliminary level respectively, with an associated mid-term review and final review respectively.

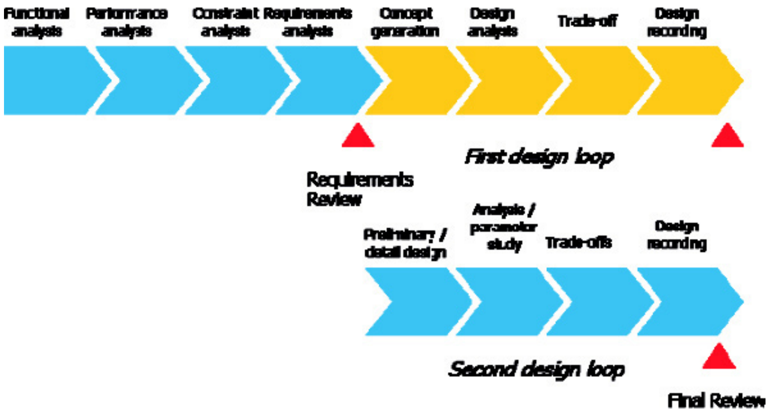


Figure 1. General SE process flow for the DSE project

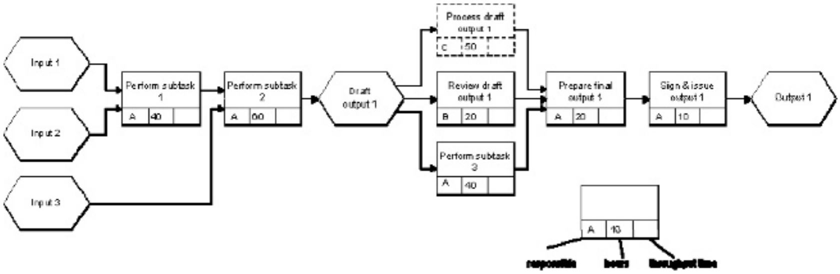


Figure 2. Generic Work Flow Diagram describing the sequence of activities, inputs and outputs, responsibilities and times

The Requirements Review entails the formation and submission of a Project Plan followed by a Baseline Report. The Project Plan is associated with project management deliverables such as a Functional Flow Diagram, Function Breakdown Structure, Gantt Chart, Organogram, and project description while, subsequently, the Baseline Report includes the technical systems engineering deliverables. The Functional Flow Diagram and Functional Breakdown Structure in the Baseline Report relate specifically to the technical design, rather than project management. In addition, this includes a Requirements Discovery Tree, Technical Budget Breakdown, Technical Risk Assessment and Maps, and a Design Option Tree.

Figure 2 shows the generic form of the Work Flow Diagram (WFD) describing the sequence of activities, inputs and outputs, responsibilities and times relating either to the project management or the technical design for the Project Plan and Baseline Report respectively. This is developed in detail to take the project up to the Mid-Term Review point, at which stage all these deliverable are then redone for the second stage of the design loop shown in Figure 1. An example of the generic Work Breakdown Structure (WBS) is shown in Figure 3, again used as part of the project management definition (in the Project Plan) or technical systems engineering definition (in the Baseline Report). The WFD describes a chronological view of the functionality to be captured while the WBS is hierarchical in nature, presenting the user with two view points in order to help capture all of the relevant functionality. It is often helpful to do the WFD first and then to develop a consistent WBS that can be used to capture addition aspects that are less obvious when thinking of the design's functionality, such as maintenance, and of course the WBS offers the chance to develop a much more detailed view of the required functionality.

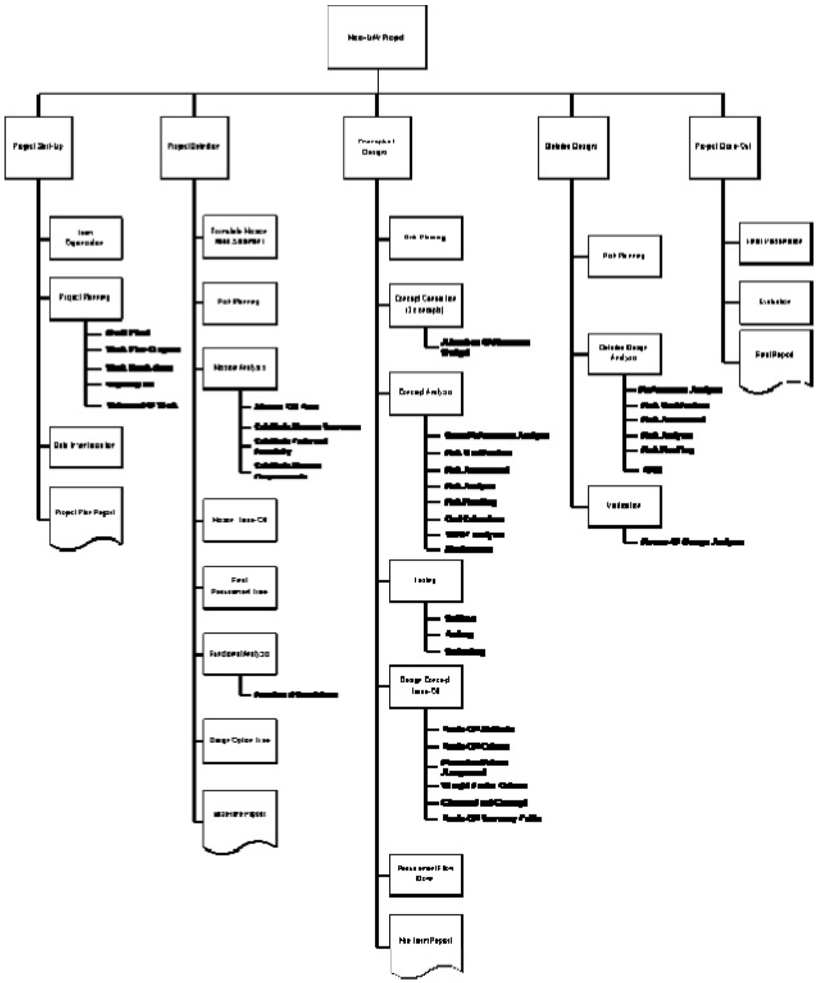


Figure 3. An example of a Work Breakdown Structure

3. Intregration of the System Engineering Framework into the Design Project

The true integration of the Systems Engineering and Program Management elements into the design project itself is of critical and fundamental importance. It must be stressed to the students that the SE aspects are actually there to help them do the design project in a more effective manner, rather than just jumping straight into the designing. Figure 4 presents a full check list of the deliverables associated with the four major submissions: Project Plan, Baseline Report, Mid-Term Report, and Final Report.

No	Product	Project Plan	Baseline Report	Mid-term Report	Final Report
1	DSE work-flow diagrams	•		•	
2	DSE work break-down structure	•		•	
3	DSE project-approach description	•		•	
4	DSE Gantt chart	•		•	
5	DSE team organization/organogram & HR allocation	•		•	
6	Functional flow diagram(s)		•		•
7	Functional breakdown		•		•
8	Requirements discovery tree		•		
9	Resource allocation/ Budget breakdown		•		•
10	Technical risk assessment/risk map		•	•	•
11	Design option structuring (tree)		•	•	
12	Interface definition/N2 charts			•	
13	Trade method, rationale and organization			•	
14	Trade criteria			•	
15	Criteria weight factors			•	
16	Trade summary table			•	
17	Operations and logistic concept description			•	•
18	Project design & development logic (post DSE)				•
19	Project Gantt chart (post DSE)				•
20	Cost break-down structure				•
21	H/W, S/W block diagrams (interactions, flows)				•
22	Electrical block diagram				•
23	Data handling block diagram				•
24	Sustainable development strategy		•	•	•
25	Compliance Matrix (table with tick marks for all requirements that are met)				•
26	Communication flow diagram		√	√	
27	Manufacturing, Assembly, Integration plan (production plan)				√
28	Return on investment, operational profit			√	√
29	Market needs estimate		√	√	√
30	Reliability, Availability, Maintainability, and Safety (RAMS) characteristics		√	√	√
31	Performance analysis (e.g. flight profile diagrams, payload-range diagrams, climb performance, noise characteristics, emissions, etc.)			√	√
32	Configuration / layout (internal /external)		√	√	√
33	Spacecraft system characteristics (e.g. communications link budget, memory size, etc.)	N/A	N/A	N/A	N/A
34	Aircraft system characteristics (e.g. fuel and hydraulic system lay-out, auxiliary power estimate, environmental control, etc.)			√	√
35	Aerodynamic characteristics estimate (e.g. lift, drag, aerodynamic moments, drag polar, etc.)			√	√
36	Structural characteristics (e.g. loading diagrams, stresses, bending, flutter, etc.)				√
37	Stability and control characteristics (e.g. control forces, c.g. limits, etc.)				√
38	Material characteristics (e.g. yield, ultimate, fatigue, etc.)				
39	Astrodynamic characteristics (e.g. orbit, trajectory, decay rate, ΔV-budget)	N/A	N/A	N/A	N/A

Figure 4. Full listing of deliverables in the DSE

However, this check list can be counter-productive if the students only think that each of the items needs to be included and that it correlates well to the marking structure (although even with it they still tend to forget items!). In fact, the deliverables for any one Report are highly interdependent and are better understood as relating to a process highlighted in Figure 1. For example, it has already been mentioned that for the Baseline Report, the FFD leads nicely into the FBS and that the FBS should be consistent with the FFD as both offer two views in the inclusive capture of function requirements, both chronological and hierarchical. Furthermore, these then facilitate the generation of the Requirements Discovery Tree (RDT), the basic functional requirements from the WFD and WBS being augmented further by constraining requirements to develop a more comprehensive RDT that is still consistent with the WFD and WBS. This then acts as the precursor to doing a technical budget breakdown where a very high level breakdown of technical parameters such as weight, drag and even cost for example. However, this needs to be consistent with the Design Option Tree so that the product architecture considered is realistic and relevant. The technical risk assessment and mapping then gives an overview of all of the relevant risk, whether technical, scheduling or cost.

The integration into the design project is an aspect that is specifically picked up in the assessment process. In the assessment process each of the SE elements identified in Figure 4 are assessed relative to 1) being included, 2) being of sufficient quality and 3) being successfully integrated into the DSE design project. In addition, marks can be gained or lost for consistency, version tracking and traceability, report quality, etc. The grading structure for the non-SE elements of the DSE is shown in Figure 5, being the other 7/8^{ths} of the total mark, but the effective SE process (1/8th of the mark) is specifically designed to directly improve the effective design process.

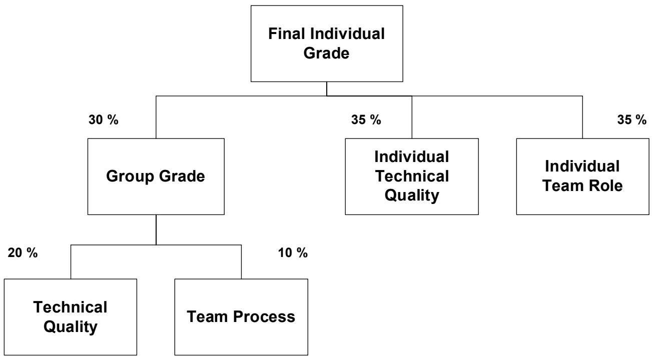


Figure 5. Grading structure for the non-SE elements

4. Systems Engineering as an Effective Framework for Teaching Aerospace Design

The objectives of the Design Synthesis Exercise at TUD are to enhance the students' skills in:

a) Designing
b) Applying knowledge
c) Communicating (discussion, presentation, report)
d) Working as a team
e) Sustainable development

As the Bachelor DSE design exercise at TUD is part of the MSc. undergraduate program the aim is not to attain a flawless final result but as stated at the beginning of this paper, the main proposition is that: '*Systems Engineering provides a beneficial educational framework and process for helping students to understand the aerospace design process more quickly*'. Consequently, the students are encouraged to harness their natural drive to create "the best design", but to realise that certain design choices will be made which will not be optimal ultimately as they are at a very early stage of the design process. However, this is entirely in sync with the real design experience when choices and decision making can only be partially iterated within the limited timeframe and budget of any design exercise. The simulation of this aspect within ISDE of such an iterative process is an invaluable element in obtaining design experience. That brings us back to the main proposition of this paper, which is to consider SE as an effective framework for achieving the five objectives listed above. Each objective is well associated with the goal of SE, sustainable development fitting within the SE goal of achieving a life-cycle balanced solution that integrates product, process and people requirements.
The 3rd objective of this paper was 3) to present evidence that the SE educational framework is an affective way of teaching aerospace design education, as well as 1) illustrating the Systems Engineering components and 2) showing the integration into the actual final year design project. Figures 5 and 7 illustrate some of the established benefits of SE within the industrial setting where it can be seen that a 7-10% cost allocation to early SE effort can reduce cost and schedule overrun from 200% to under an acceptable 20%, with little additional cost-benefit evident beyond this.

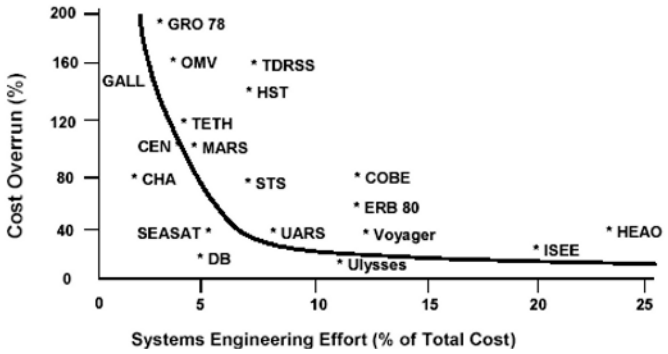


Figure 6. Impact of SE of cost overrun (Moody et al)

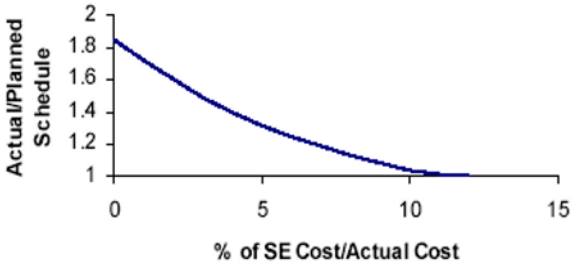


Figure 7. Impact of SE on schedule overrun (Mar & Honour)

However, how do we interpret this relative to design projects being carried out within the educational sector? Figure 6 would suggest that within an educational design project, the students will 1) have to put in double the effort (time) that they anticipated in order to do what they want (with all of the associated issues within a team context and varying levels of expectation and commitment) and that projects tend never to finish unless bounded by time or cost! Figure 7 would suggest that the students 2) need to put 1/10th of their effort (time) into SE, especially at the beginning, if they are to achieve what they want within the allocated time scales, but encouragingly, that this is possible!

5. Examplar Case Study of the Technical Educational Benefits

An example of the trade-off process formalized within the SE approach is provided in Figure 8. The SE process requires the students to spend approximately 1/3 of their time in developing a strong engineering rationale (through the FBS, RDT, and DOT) for their assessment of a broad range of design synthesis solution, as illustrated by Figure 8. The exemplar project entailed the design of a micro-air-vehicle that could perform complex aerodynamic maneuvers to enable it to be an unobtrusive surveillance agent, with obvious energy challenges in terms of effectiveness.

	Monoplane	Biplane	Tandem
			
Dimensions (l x w x h)	0.36 x 0.45 x 0.06	0.43 x 0.45 x 0.06	0.47 x 0.45 x 0.05
Mass	16.48 g	17.48 g	18.18 g
Aspect Ratio	3.90	3.90	3.90
Wing Span	0.450 m	0.450 m	0.450 m
verage flight speed	2.35 m/s	1.40 m/s	1.36 m/s
Centre of gravity	51% of chord from leading edge	41% of chord from leading edge	51% of chord from leading edge
ower consumption	1551.2 mW	1440.8 mW	2149.5 mW
lapping frequency	3.69 Hz	6.17 Hz	7.90 Hz
Tail configuration	Bird tail	Inverted V-tail	Conventional tail
Rocking amplitude	8 mm	± 0 mm	Rotation
Comments	The monoplane that is considered in this phase of the design process is based on the Freebird model that is available at ornithopter.org.	The biplane that is considered is based on the Luna, which is also available at ornithoper.org.	The tandem is based on an existing model that has only been built on a very small scale. In order to build the test model for this configuration, this small model first had to be scaled up.

Figure 8. Exemplar trade-off considerations formalized through the SE approach

Figure 9 illustrates that the SE approach helps greatly in developing early solutions to some of the wider design requirements captured through the FBS, RDT and trade-off study.

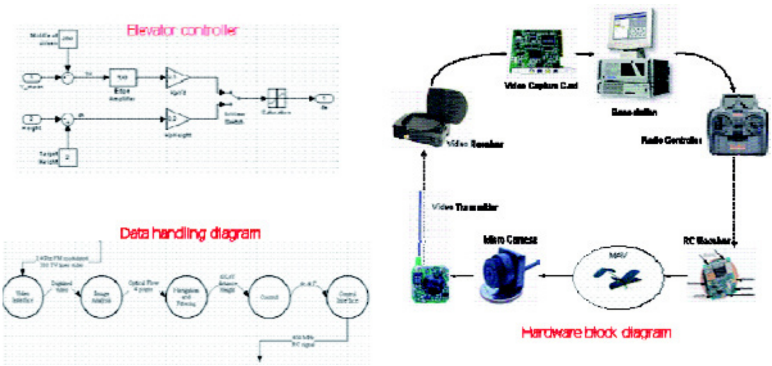


Figure 9. Example of truly systems oriented approach required through SE

Finally, Figure 10 presents a solid model representation of the result of the design project. The Figure is included to highlight that the students are still expected to deliver a solid engineering design, although the result should be even better given the SE approach.

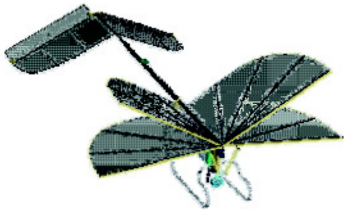


Figure 10. Solid model of final design to illustrate engineering delivery within a SE approach

6. Conclusion

Ultimately, the paper triangulates all of the components mentioned in order to address the main proposition that: '*Systems Engineering provides a beneficial educational framework and process for helping students to understand the aerospace design process more quickly*'. This is achieved by: 1) illustrating the Systems Engineering components utilised at TUD, 2) looking at the integration of SE into the final year DSE design project at TUD and 3) presenting evidence that the SE educational framework is an effective way of teaching aerospace design. Therefore, the relevance of the paper addresses the implementation and understanding of SE methods within concurrent thinking related to design education.

7. References and Bibliography

[1] Synthesis of Subsonic Aircraft Design, E. Torenbeek,
[2] Airplane design, Roskam
[3] Aircraft Design, A conceptual Approach, D.P. Raymer
[4] JAR-21, subpart JA, JB and appropriate appendices
[5] ISO-9000-1 and 2
[6] Guide to Microsoft Projects
[7] Product Design: Fundamentals and Methods, N.F.M. Roozenboom, J. Eekels
[8] Jane's All the Worlds Aircraft
[9] World Aircraft & System's directory – Flight International
[10] Airframe Structural Design, M.C.Y. Niu
[11] Airframe Stress Analysis And Sizing, M.C.Y. Niu
[12] Analysis and Design of Flight Vehicles Structures, Bruhn
[13] Aircraft Landing Gear Design: Principles and Practices, Currey, N.S.
[14] Aircraft Performance and Design, J.D. Anderson
[15] Fluid-dynamic drag, S.F. Hoerner
[16] Fluid-dynamic lift, S.F. Hoerner and H.V. Borst
[17] Personal aircraft drag reduction, B. Carmichael
[18] Theory of wing sections, Abbott and von Doenhoff
[19] Stuttgarter Profilkatalog, D. Althaus and F. Wortmann
[20] Niedrig-geschwindigkeits-profile, D. Althaus

[21] Elements of airplane performance, G. Ruijgrok
[22] JAR-25
[23] Aerodynamics, Aeronautics and Flight mechanics, B.W. McCormick
[24] Aircraft Structures for Engineering Students, Megson

Creating Value by Measuring Collaboration Alignment of Strategic Business Processes

Frank van der Zwan[a,1], Sicco Santema[b], Richard Curran[b]

a Assistant professor Aviation, Aerospace Management & Operations, Delft University of Technology (Delft), The Netherlands
b Professor Aerospace Management & Operations, Delft University of Technology (Delft), The Netherlands

Abstract. In today's competitive aviation landscape, it is imperative for a hub-airport and its network carrier to create sufficient added value in order to ensure survival. One way of doing this is to improve the collaboration on interdependent business processes. In this research, four of such processes have been studied, by using a model which contains thirty-one factors that determine the level of alignment within the collaboration. Studying the characteristics of the collaboration has proved a useful tool for identifying opportunities for an airport and an airline increasing the value that is to be created.

Keywords. Interorganisational relationship, airport, airline, alignment

1 Introduction

This paper focuses on the aviation industry and more specifically on the relationship between airports and airlines. In the last few decades, the aviation landscape has changed dramatically: we have seen the privatization and deregulation of airports and airlines. This facilitated the proliferation of new airlines, such as Southwest Airlines, Easyjet or Ryanair. With the introduction of these low cost carriers (LCC's) competition in the aviation industry increased, ticket fares were drastically cut, making air transport more affordable and thereby commoditizing air travel.

Since the airline business is a fixed-cost business [8], traditional flag carriers, like British Airways, have to adapt to this altered competitive landscape if they are to survive. As if one compares the unit costs of different European airlines, expressed in cost per available seat kilometer, it becomes clear that the costs of the traditional flag carriers, like Lufthansa, British Airways, Air France/KLM are on average 73% higher than these of the LCC's like Ryanair and Easyjet [12].

[1] Corresponding Author: Aerospace Management & Operations, Delft University of Technology, The Netherlands; Tel: +31 (0) 15 2781483; Email: f.m.vanderzwan@tudelft.nl

These traditional airlines are generally network or hub-and-spoke carriers, which means that they operate a network of connections, in which all traffic moves along spokes connected to the hub at the center. In producing air transport services, airports act as providers of the on-ground infrastructure for flight operations while airlines offer the transportation services *per se* [1]. After decades of largely independent, solitaire development of strategies, both airports and airlines have started to rethink their traditional customer-supplier relationship [2]. So understandably, there is a great dependency between the network carrier and its hub-airport for revenue and profit generation. One could say that their destinies are inter-twined and that their existence as viable economic entities depends upon their joint performance [10]. Therefore it is imperative for the airport and airline to improve their collaboration, with the aim to improve the function over cost ratio.

Thus, this research focuses on how collaboration on interdependent business processes between airports and airlines can be better aligned, thereby reducing costs, increasing efficiency and added value, and diversifying their businesses. This research mainly focuses on the strategic level; the level at which decisions are made, policies are established, rules are created, and financial levels are set [7]. Based on previous research [9][11] and input from airport and airline, the following interdependent business processes have been chosen:

Environmental capacity	the process during which the capacity of the airport is determined in order to cope with environmental impact of air transport
Network planning	the process during which airport and airline are determining which type of network would be most beneficial for the airport
Infrastructure planning	the process during which airport infrastructure is developed in a effective and efficient way
Aircraft stand allocation	the process during which gates are allocated to aircraft on a seasonal and daily basis

The structure of this paper is as follows: In the next section, the research methodology is outlined. Next, the results of this research is presented and discussed, and finally, recommendations, for further research and application within both firms are made.

2 Method

This research is performed by four researchers each studying two interdependent business processes from a perspective at either the airport or the airline. In order to get a useful comparison, the researchers used a common research framework.

First, a literature study has been carried out in order to establish a framework about aspects that influence collaboration between organizations at a general level [4]. In this study, thirty-one factors were found that give an indication of the level of alignment of business processes between two or more organizations. The factors were then grouped into ten clusters, which are divided into two phases: the pre-alignment, or partner selection phase, and the alignment phase. The factors, clusters and phases are structured in the Factors for Alignment model (the FA model), as shown in figure 1.

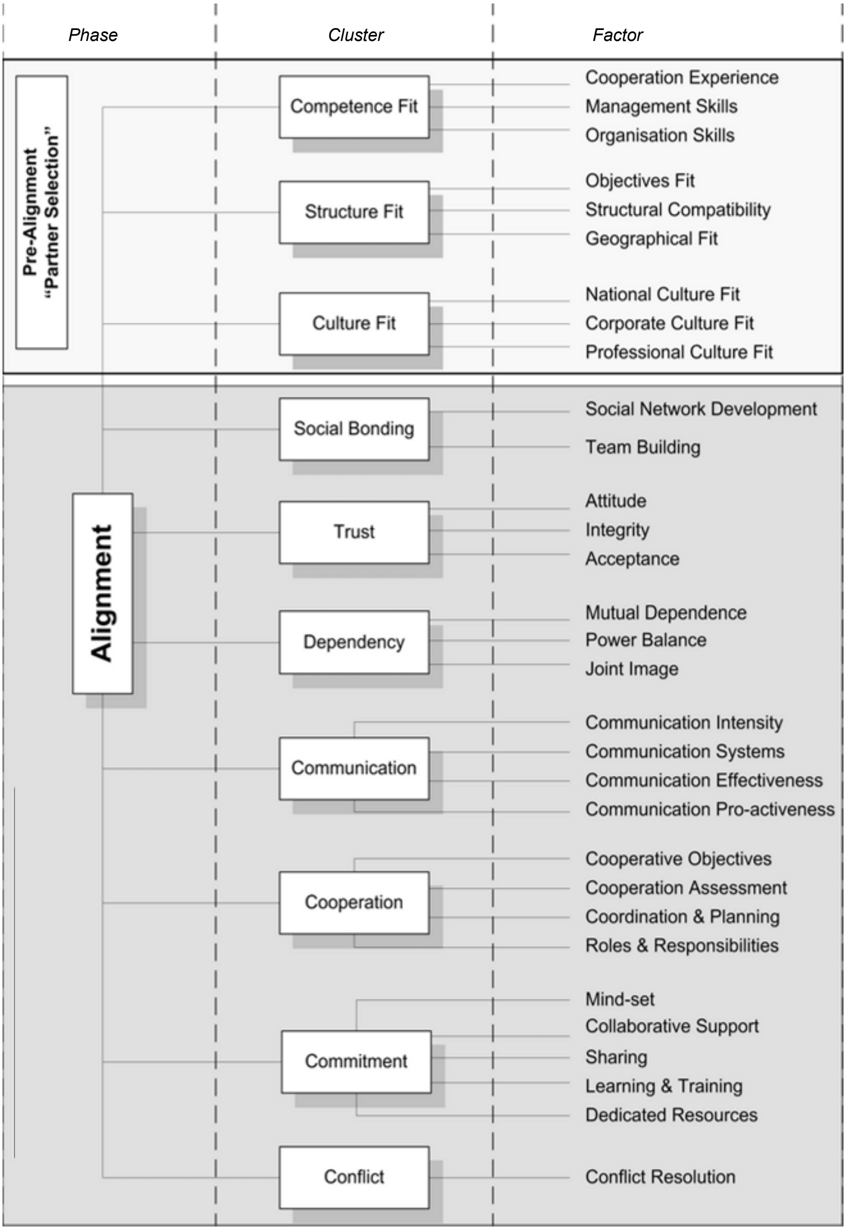


Figure 1. Factor for Alignment Model

The FA model formed the basis of questionnaires that were developed and interviews that were held amongst all employees of the airport and the airline involved in the specific strategic business processes. The questionnaire has been applied in order to measure the employee's perception concerning the collaboration, with respect to the factors of alignment. By design, the questionnaire is set up to measure how far off the current collaboration (as perceived by employees involved) is from 'good', *i.e.*, a delta value. This score is regarded as the implicit potential for improvement. By using a seven-point Likert scale, the scores of the respondents can be combined and average scores are produced.

From the interviews held among both the airline and airport employees, issues can be extracted that negatively influence the airline-airport collaboration. Each of these issues can be coupled to one or more factors for alignment. The total number of times that each factor is coupled to one of the issues is called the occurrence and symbolizes the importance for improvement on that specific factor. This score is regarded as the explicit potential for improvement. The multiplication of both implicit and explicit potential for improvement; *i.e.*, the combination of the data streams from both applied research methods, results in the definition of the Potential for Improvement on alignment for each factor. In order to be able to compare the scores of the airport and the airline, scores are normalized per firm by dividing the multiple score of the individual factors by the highest multiple score among the thirty-one factors.

This potential for improvement can be ranked top-down for all factors thereby constructing a prioritization of factors for alignment on which airline and airport should focus to improve collaboration. Alternatively, the two underlying variables can be set out in a plot with X-axis being the Factor Occurrence and the Y-axis being the Factor Delta.

In the next step the results of the airline and airport are compared in order to find the factors for alignment representing the areas for improvement within the processes that have been investigated. Factors are clustered into so called 'factor families' With the data obtained, a prioritization is made in the factors for alignment that are significant in aligning the collaboration. Next also areas of improvement, so called 'problem areas', are described so as to give both organizations a clear and practical view of the challenges within the collaboration that could be solved by focusing on the identified factors for alignment.

3 Results

Figure 2 below shows Factor Delta plotted against the Factor Occurrence for each of the thirty-one factors of airport and airline respectively for the airline-airport interdependent business process of *Infrastructure Planning*. In order to be able to better compare the scores of the airport and airline, the two previous plots can be superimposed, and their relative position can be compared.

In order to prioritize the factors on which the airport and airline could focus on in order to improve their collaboration, Figure 3 below shows the sums of the normalized multiples for the airline-airport interdependent business process of Network Planning. Factors with a normalized sum greater than 1 are the primary factors for this process (Group I) and factors with a normalized sum between 1 and 0.3 are secondary factors for this process (Group II).

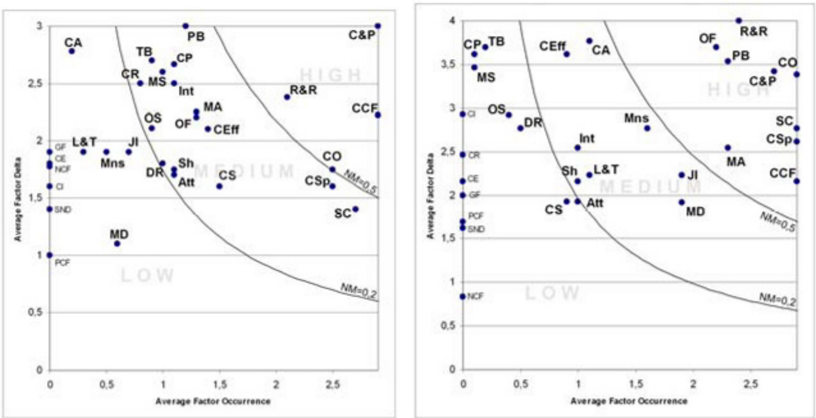


Figure 2. Factor Delta vs. Occurrence Plot Airport and Airline, *Infrastructure Planning* [10]

Gr.	Factor	Airline	Airport	Σ Norm multiples
I	Coordination & Planning (C&P)			1.57
	Communication Intensity (CI)			1.30
	Objectives Fit (OF)			1.43
	Cooperation Objectives (CO)			1.36
	Integrity (Int)			1.05
II	Communication Proactiveness (CP)			0.72
	Collaborative Support (CSp)			0.66
	Roles &Responsibilities (R&R)			0.65
	Mindset (Mns)			0.65
	Mutual Acceptance (MA)			0.51
	Mutual Dependence (MD)			0.46
	Communication Effectiveness (CEff)			0.40
	Joint Image (JI)			0.33
	Structural Compatibility (SC)			0.32
III	Attitude (Att)			0.21
	Sharing (Sh)			0.18
	Communication Systems (CS)			0.12
	Corporate Culture Fit (CCF)			0.12
	Power Balance (PB)			0.11
	Cooperation Assessment (CA)			0.09
	Social Network Development (SND)			0.04
	Conflict Resolution (CR)			0.03

Figure 3. Combined multiple rank, *Network analysis*

During the analyses of all the gathered issues during which the researchers at the airport and airline worked in close consultation, certain common topics have been distinguished. For *Network Analysis* and *Environmental Capacity* six recurring topics have been defined as problem areas. Each deals with a part of the collaborative effort in one of the interdependent business processes. The six problem areas have been determined by forming clusters of issues for each company in both of the processes. The definitions of the six identifiable problem areas are [13]:

Vision & interests	differences between the individual visions and interests of each collaborating party, compared to the joint vision and interests
Process design & execution	differences in the joint set up by the collaborating parties, as well as its implementation
Internal alignment	differences in the internal setup of the process as well as the communication between those involved in the process, within each of the collaborating parties
Interorganisational relationship	differences between the collaborating parties on both a department and individual level
Corporate structure	differences between collaborating parties in the organizational structure or culture
Third-party relationship	differences in the way the collaborating parties interact with third-parties, i.e. other stakeholders

In the case of *Infrastructure planning* and *Aircraft stand allocation* the issues were clustered into factor families based on similarities between the factors they are linked to. The making operational of the factor families has resulted in coherent problem areas that are characterized by a specific cluster of issues. These problem areas were discussed in focus group discussions by the employees involved from both airport and airline. There problem areas were translated into so called 'joint challenges', *e.g.*, " How can both firms achieve a higher level of trust and willingness to share, improve the insight in each other's processes and developments for the benefit of joint initiatives?" [5][6]

During the focus group discussion initial options for these challenges were generated. It is acknowledged by all participants of the focus group discussion that such sessions are a valuable source of providing insights in the issues present within the partner firms [5].

4 Discussion and conclusions

The application of the FA model with its questionnaires and interviews has proved useful for studying interdependent business processes within the airport and airline. It facilitated the process with both airport and airline of becoming more sensitive for each other's perspectives.

The factors with the highest priority to the alignment of both collaborating parties are recognizable to airport and airline alike through the issues mentioned by their own representatives.

Due to the research methods that have been used, comparing the results across the four business processes has its limitations; *e.g.*, different people were involved, and the processes have their own characteristics and dynamics. But when the primary (p) and secondary (s) factors of the four processes are added up, useful insights for future research can be gained, as is shown in figure 4 below.

		Business Process			
		Environm. capacity	Network planning	Infrastr. planning	Aircraft Allocation
Factor name	**Cluster**				
Objectives Fit	*Structure fit*	s	p	p	p
Cooperation Objectives	*Cooperation*	p	p	p	
Coordination & Planning	*Cooperation*	p	p	p	
Mutual Acceptance	*Trust*	s	s	p	p
Roles & Responsibilities	*Cooperation*	p	s	p	
Integrity	*Trust*	s	p	s	s
Power Balance	*Dependency*			p	p
Communication Intensity	*Communication*	p	p		
Structural Compatibility	*Structure fit*	s	s	p	
Communication Pro-activeness	*Communication*	p	s	s	
Corporate Culture Fit	*Culture fit*	s		p	
Collaborative Support	*Commitment*		s	p	

Figure 4. Most important factors for alignment across the four business processes

The figure shows that the factors *Objectives Fit*, *Cooperation Objectives* and *Coordination & Planning* have been listed as primary factors three times, thus making these among the most important factors to further investigate when aiming to improve collaboration between the airport and airline. Therefore, future research will focus on looking at how one stakeholder can be made more sensitive for the key performance indicators of the other stakeholders, with the aim to looking at how processes can be redesigned so as to optimize the value created.

5 Acknowledgements

This research is performed by four researchers each studying two processes from a perspective at either airport or airline [3][5][6][13] as part of a PhD research project [10].

6 References

[1] Albers A, Koch B, Ruff C, Strategic alliances between airlines and airports – theoretical assessment and practical evidence, Journal of Air Transport Management, No. 11, pp. 49-58, 2005.

[2] Auerbach S, Koch B, Cooperative approaches to managing air traffic efficiently – the airline perspective, Journal of Air Transport Management, No. 13, pp. 37-44, 2007.

[3] Driessen EK, Schiphol Mainport Alignment, an analysis of interdependent business processes at AAS and KLM, MSc thesis, Delft University of Technology, The Netherlands, 2008.

[4] Driessen EK, Drijgers M, Fiksinski MA, The Search for factors for alignment, Delft University of Technology, Faculty of Aerospace Engineering, January 2008.

[5] Drijgers M, Great Firms think aligned, a mirrored perspective on alignment of KLM-AAS inter-organisational business processes, MSc thesis, Delft University of Technology, The Netherlands, 2008.

[6] Fiksinski MA, Great Firms think aligned, a mirrored perspective on alignment of KLM-AAS inter-organisational business processes, MSc thesis, Delft University of Technology, The Netherlands, 2008.

[7] Fry M, The Key Levels in an Organisation, Enterprise Leadership.org, 2007; 1-6.

[8] Gillen D, Lall A, Competitive advantage of low-cost carriers: some implications for airports, Journal of Air Transport Management, 10, pp. 41-50, 2004.

[9] Oei HY, Strategic management for inter-organizational processes in aviation building for the future of Schiphol Mainport, MSc thesis, Delft University of Technology, The Netherlands, 2007.

[10] Perié RP, All for One: factors for alignment of inter-dependent business processes at KLM and Schiphol, PhD thesis, Delft University of Technology, The Netherlands, 2008.

[11] Rennestraum R, Survival through sustainable synergy, MSc thesis, Delft University of Technology, The Netherlands, 2007.

[12] Ringbeck J, Schneiderbauer D, Mostajo M, Fauser B, Time for take-off: Setting new strategic directions for mid-sized European Airlines, Booz Allen Hamilton, 2007.

[13] Wever CJC, Alignment of AAS-KLM collaboration on specific strategic business processes, MSc thesis, Delft University of Technology, The Netherlands, 2008.

Drivers of Customer Satisfaction in a Project-Oriented, Business-to-Business Market Environment: an Empirical Study

Wim J.C. Verhagen[a,1], Wouter W.A. Beelaerts van Blokland[b], Richard Curran[b]

[a] PhD. Candidate, Aerospace Management & Operations, Delft University of Technology
[b] Profile Coordinator, Aerospace Management & Operations, Delft University of Technology
[b] Chairholder, Aerospace Management & Operations, Delft University of Technology

Abstract: Quantitative analysis with respect drivers of customer satisfaction in a project-oriented, business-to-business environment is very scarce. The aim of this paper is to make headway into this terrain by statistical analysis of quantitative empirical research. Five drivers of satisfaction are posited. Analysis of 41 observations shows that the overall model is highly reliable, valid and generalizable. The overall model fit is excellent. Bi-variate correlation shows the promise of the individual drivers of satisfaction. However, multiple regression analysis suffers from a large degree of multicollinearity. Based on this regression, 'Solution' is the only driver of satisfaction that can unequivocally be accepted, being both significant and meeting effect size requirements. A larger dataset (N > 80) is recommended for further research to combat multicollinearity and test smaller effect sizes.

Keywords: Customer satisfaction, Business-to-Business, project environment

1 Introduction

A functioning and up-to-date customer satisfaction measurement system is critical for any company that wants to remain competitive. Research has shown that losing customers is much more expensive than retaining customers [5, 8, 20]. A critical part of customer retention is to fulfil customer expectations. Customer satisfaction measurement fills the information pipeline driving company decision-making, which is aimed at better fulfilling the aforementioned customer expectations. Customer satisfaction measurement has been practiced in business-to-customer markets for years; in this field, scientific research and associated theories are well

[1] Corresponding Author : PhD. Candidate, Delft University of Technology, Faculty of Aerospace Engineering, Chair of Aerospace Management & Operations, Kluyverweg 1, 2629 HS Delft. Tel: +31 (0) 6 4845 5291. E-mail: w.j.c.verhagen@tudelft.nl; www.lr.tudelft.nl/amo

developed. Customer satisfaction measurement is adopted into business-to-business markets as well, spurred by developments such as increasing adoption of quality control in the form of ISO certification. However, scientific research in B2B market environments is relatively scant, particularly when additional restrictions apply.

One such example is a project-oriented, B2B market. Companies operating in such markets fulfil project demand rather than product demand. Each project has its own specific requirements; the supplying company delivers a complete solution for a business problem, rather than a single product. A successful project requires a long-term commitment from both supplier and buyer, involving many steps to bridge the gap between initial system definition to having an operational system with associated service activities on the work floor. Measuring customer satisfaction in these circumstances is generally more involved than measuring in B2C markets, while the variety of projects makes it difficult to identify general drivers for customer satisfaction. To remedy this situation, this research paper is aimed at discovering drivers of customer satisfaction in project-oriented, business-to-business markets. To uncover these drivers, a model is proposed that is subsequently statistically tested using data from a company operating in a project-oriented, B2B market.

The structure of this paper reflects this approach. First, the research question and research method underlying this paper will be outlined. This will be followed by a brief review of relevant theoretical concepts related to customer satisfaction measurement and analysis. Subsequently, relevant research is assessed to reveal previously identified drivers of customer satisfaction. This acts as a starting point for the construction of a model of customer satisfaction in a project-oriented, B2B market environment. This model is tested using research data from a company operating under the aforementioned conditions. The practical validity of the results is briefly considered before concluding with an evaluation of the model, along with recommendations to improve the research.

2 Research setup

The objective of this paper is to derive, test and validate a model of customer satisfaction in a business-to-business, project-oriented market environment. This model will reveal satisfaction drivers that are specific to a project-based organisation in a B2B environment, an area in which very little research has been performed. This thesis will address this knowledge gap. The resulting research question is as follows:

What are drivers for customer satisfaction in a project-oriented, business-to-business market environment?

The research method consists of two main branches. First, a literature review is performed to establish the theoretical background and to inspect scientific research with regard to customer satisfaction. The latter evaluation serves a basis for model construction. The second branch of the research method consists of empirical research: using interviews and an e-survey, data is collected with which the model will be analysed.

3 Theoretical framework

The research question contains a number of elements that have to be defined: customer satisfaction, project-orientation, and B2B. Alongside these definitions, the reasons behind customer satisfaction measurement are evaluated, thus giving some theoretical background. After this, existing research will be evaluated to find previously proven drivers of customer satisfaction, which will serve as a basis for model construction.

3.1 Definitions and theoretical background

According to Kotler [17], customer satisfaction is ‘a measure of product performance with respect to expectations’. Hill [15] offers a similar but subtly different definition: customer satisfaction is ‘a measure of how your organization’s total product performs in relation to a set of customer requirements’. Synthesizing leads to the following definition: *customer satisfaction is a measure of the organization’s total product performance with respect to customer expectations (consisting of both formal and informal requirements)*. Formal requirements are laid down in a contract (e.g. price, quality), whereas informal requirements deal with soft factors (e.g. supplier attitude, professionalism).

Why would measuring customer satisfaction be profitable? Hill [15] shows the links between customer satisfaction and bottom-line results in his satisfaction-profit chain model (see Figure 1). The (individual) links in this chain have been validated by various studies [5, 8, 15, 17, 20]. The causative mechanism (the presence of occurrence, antecedence and contiguity) between Customer Value Package and Customer Satisfaction is validated by value generation theories such as Porter’s Value Chain [19] and Beelaerts’ 3C model [1].

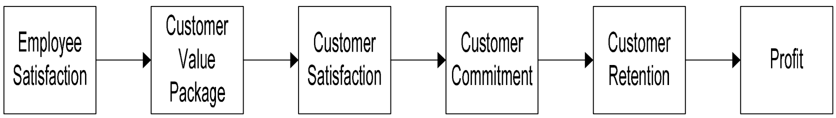


Figure 1: The Satisfaction-Profit chain (Source: [15], figure 4.1)

The second element in the research question is the project orientation. This simply indicates that the company in question is focused on supplying projects rather than products. A typical project consists of delivering a total system solution. This starts at system definition in the sales phase (commonly in multiple iterations), runs through system installation, testing and acceptance in the project phase, and continues throughout operational life with service and maintenance. In such an environment, transactions between companies are far more limited. Sometimes, there are replacements of system parts. Also, system extensions occur. After a certain period of time (commonly >10 years), the system will be replaced. Under such conditions, forming and maintaining relationships – which logically includes delivering to satisfaction - with customers over years of time is critical in gaining follow-up business. The resulting repeat business can be a major driver for revenue and profitability. This mechanism is very comparable to the one described in the Satisfaction-Profit chain, indicating that this concept is also viable for project environments.

Finally, a B2B setting implies that a transaction occurs between two companies, as opposed to between a company and a customer. The term also encompasses the transfer of goods and/or services between companies.

3.2 Previously identified drivers of customer satisfaction

Research regarding customer satisfaction drivers in B2B markets is relatively limited. The following drivers were identified from literature. Characteristics of the product (e.g. price, product quality) significantly influence an industrial relationship [13]. Besides this, services delivered along with the product are also drivers of satisfaction. Homburg and Rudolph [16] show that product-related information is frequently a driver of customer satisfaction in a B2B context. Chakraborty et al. [5] have researched and validated reliability, product-related information and commercial aspects as significant drivers for customer satisfaction in a B2B environment. Gonzalez and Fredericks [11] particularly stress the importance of including emotional (or 'soft') factors in a B2B environment into survey efforts, something that is often forgotten, 'yet the most committed relationships are bases on creating both rational and emotional bonds with customers' [11]. A notable characteristic of existing research is the wide scope of the drivers of satisfaction.

In B2C settings, more research is available. In the next section, the model will be constructed using identified drivers from both the B2B and B2C market environments; the drivers from the B2C sector will be briefly summarized in the categorization table (Table 1).

The previously mentioned drivers of satisfaction apply to B2B and/or B2C markets. However, the influence of a project orientation on customer satisfaction has not yet been taken into account. Research that takes into account this aspect is even scarcer than B2B-related customer satisfaction research. The study of Chakraborty et al. [5] is the only B2B study to take into account project environment considerations. The identified drivers of satisfaction (reliability, product-related information and commercial aspects) are based upon the method of distilling general factors that apply to each project. Hill [15] shortly advises the same when considering project environments: try to find factors that apply to each project. If one does so, these factors can be measured and compared irrespective of specific project conditions. This approach will be followed when constructing the model.

4 Model definition

The identified drivers of satisfaction can be used to validate the theoretical strength of the drivers of customer satisfaction posited in this paper. Before performing such a validation, it is of course necessary to declare these drivers of satisfaction. A Needs-Benefits-Features analysis has been performed to identify possible drivers of satisfaction in a project-oriented, B2B environment. The NBF analysis uses needs (a customer want or desire), benefits (the value of a feature to a customer) and features (characteristics of the product) to describe what a supplier must deliver to its customer. As the needs of a customer are the needs of the supplier, the resulting needs and benefits can be used to define drivers of satisfaction. The following drivers are identified:

- **Reputation:** perception of image; the way in which the market appreciates and describes performance.
- **Industry know-how:** knowledge of customer processes and expectations.
- **Capabilities & Competences:** the various organizational and associated functional capabilities.
- **Solution:** the spectrum of systems & products offered to customers
- **Innovation & Improvement:** efforts at continuous improvement & introduction of new products/product variants.

Besides the NBF analysis, the concept of customer total value [17] and its contributing elements (value aspects of product, service, staff and reputation) served as inspiration for these drivers. The drivers have been validated in practice by interviewing a selected number of customers (N = 5), in which the operationalizations of the constructs were also deemed satisfactory. Having obtained the drivers, the model can be configured. An overview of the model is given in Figure 2.

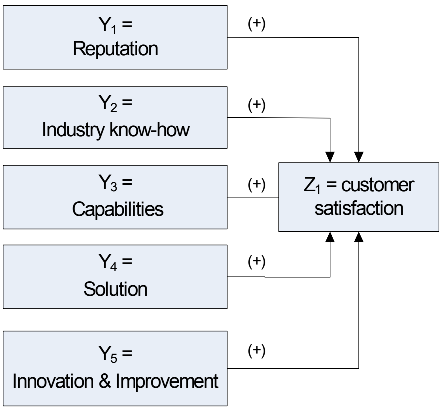


Figure 2: Model of satisfaction drivers

The key dependent variable is customer satisfaction. It is variable Z_I at the end of the causal chain depicted in the conceptual model. There are five drivers of satisfaction (Y_I up to Y_5), which are adopted as model constructs.

The hypotheses associated with this model are given below:

- H_1: reputation is positively related with customer satisfaction.
- H_2: industry know-how is positively related with customer satisfaction.
- H_3: capabilities are positively related with customer satisfaction.
- H_4: the provided solution is positively related with customer satisfaction.
- H_5: innovation & improvement is positively related with customer satisfaction.

All hypotheses are one-tailed; each driver is assumed to have a positive influence on customer satisfaction. This assumption is based on the aforementioned causative mechanism between the Customer Value Package and customer satisfaction as observed in the Satisfaction-Profit chain. The associated null hypotheses can be summarized as

H_0 = (satisfaction driver)$_n$ is not or negatively related with customer satisfaction,

where the satisfaction drivers are the five familiar drivers given above. The associated main multiple regression equation (based on Field, 2005) of this model is

$$Z_i = (b_0 + b_1Y_1 + b_2Y_2 + b_3Y_3 + b_4Y_4 + b_5Y_5) + \varepsilon_i \quad (1)$$

Does the previously discussed literature support these assumed relations? As mentioned before, the Needs-Benefits-Features analysis and the value theorem of Kotler [17] are the sources of the five constructs. Both are untested analyses and therefore a weak rationale for the relations between Z_I and constructs Y. However, categorising the identified and validated drivers of satisfaction (by means of detailed construct and operationalization analysis) shows that the assumption of positive relationships between the constructs Y and Z is sufficiently founded to warrant analysis. This categorisation is performed in Table 1. The identified drivers of satisfaction are grouped along the five constructs.

Table 1: Grouping the satisfaction drivers along constructs

Construct	Associated drivers of satisfaction
Reputation	Quality [2, 8, 13, 21, 22] Value [2, 5]
Industry Know-How	Product-related information [5, 16] Environment [2]
Capabilities	Availability [3] Responsiveness [3]

	Reliability [3, 5] Completeness & Professionalism of service [3] Value [2, 5] Timeliness, Efficiency, Self-management, Teamwork, Commitment [2]
Solution	Quality [2, 8, 13, 21, 22] Value [2, 5] Availability [3] Reliability [3, 5] Product characteristics [13, 21, 22] Product-related information [5, 16] Environment [2]
Innovation & Improvement	Product characteristics [13, 21, 22] Innovation [2]

The next step is to validate the model by analyzing the quantitative results.

5 Analysis of results

Research data was obtained by performing 15 interviews with customers of a company operating in a project-oriented, B2B environment within The Netherlands and Belgium. Furthermore, 26 observations were attained through distribution of an e-questionnaire (termed the 'standardized survey') to international customers. The overall sample size is 41 (N = 41), with an overall response rate of 32%.

Model validation requires the use of parametric tests. The data has to be checked to see whether it conforms to the four assumptions underlying parametric testing [9]. The use of interval data and independence of results are met through the survey and research design. To assess normality, the Shapiro-Wilk test has been used (since N < 50 [12]). Testing of the original dataset has revealed that the data is highly positively skewed, preventing conformance to a normal distribution. To remedy this, the data has been transformed, using methods such as data reversal and square root transformation. The output of the Shapiro-Wilk test for the transformed data is represented below (
Table 2).

Table 2: Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
Satisfaction	,958	41	,131
Reputation	,956	41	,110
Industry know-how	,954	41	,100
Capabilities	,956	41	,114
Solution	,946	41	,051
Innovation & Improvement	,951	41	,074

All variables possess normally distributed data, as significance values consistently exceed .05. Homogeneity of variance is tested using Levene's statistic.

Table 3: Test of Homogeneity of Variance

		Levene Statistic	Sig.
Satisfaction	Based on Mean	,465	,708
Reputation	Based on Mean	1,787	,167
Industry know-how	Based on Mean	,703	,556
Capabilities	Based on Mean	1,464	,240
Solution	Based on Mean	1,879	,150
Innovation &Improvement	Based on Mean	1,339	,277

Significance of all variables exceeds .05; each variable possesses homogeneity of variance.

Both the model and measurement instrument validity are satisfactory (assessed using checklists from [10] and various statistical tests). Reliability is assessed using Cronbach's Alpha. Assessment of the model gives an Alpha of .933 for 5 items, which easily exceeds the .7 threshold value for acceptance of reliability [9].

A check for influencing cases has been performed using a generation of Mahalanobis' and Cook's distances. No undue influencing cases have been discerned (Cook's distance = .657 < concern threshold of 1 [9], Mahalanobis = 13,510 < threshold of 15).

The dataset is tested using multiple regression. Unfortunately, the sample size is rather limited. With N = 41, only effects above r ≈ .55 may be accepted at a power level of .8 and significance level of .05 [9]. Generalizability is still warranted, since the minimal value of 5 observations per independent variable [12] is exceeded. A multiple regression test has been performed on the normalized data to test the model. Table 4 shows the overall model characteristics.

Table 4: Model statistics summary

Model	**r**	**R^2**	**Adjusted R^2**	**Std. Error of the Estimate**
1	,853	,727	,634	,13650

The overall model fit is very good (R = .853) and exceeds the effect size limit of r = .55. Generalizability is assessed using the adjusted R-square characteristic, which is calculated using Stein's formula [9]:

$$\text{Adjusted } R^2 = 1 - \left[\left(\frac{41-1}{41-5-1}\right)\left(\frac{41-2}{41-5-2}\right)\left(\frac{41+1}{41}\right)\right](1-.853^2) = .634$$

The F-characteristic [9] value is 18,681 (F >> 1), and is highly significant (p < .001). This indicates that the model successfully improves the prediction of the outcome compared to the level of inaccuracy in the model [12]. Table 5 indicates the significance and effect sizes of the individual predictor variables.

Table 5: Regression coefficients

	Unstandardized Coefficients		Standardized Coefficients	
Predictor variable	B	Std. Error	Beta	Sig.
Reputation	,228	,142	,260	,117
Industry Know-How	-,382	,167	-,418	,028
Capabilities	,231	,178	,280	,204
Solution	,552	,172	,521	,003
Innovation & Improvement	,194	,125	,236	,129

These results have implications for the hypotheses:

- H_1: reputation is rejected as driver of satisfaction (not significant)
- H_2: industry know-how is rejected as driver of satisfaction, since its effect size does not exceed the $r = .55$ threshold value.
- H_3: capabilities is rejected as driver of satisfaction (not significant)
- H_4: solution is a highly significant ($p < .01$) driver of customer satisfaction, with $R = .552$.
- H_5: innovation & improvement is rejected (not significant)

The hypotheses results are shown in Figure 3.

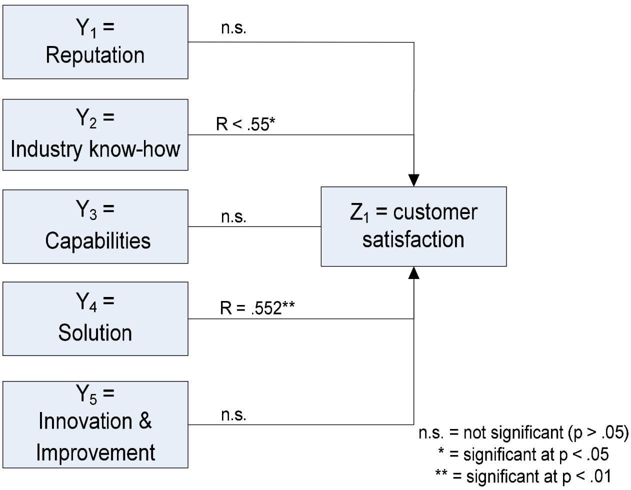


Figure 3: Predictor results using multiple regression

The results do not fully conform to expectations. The negative influence of industry know-how is particularly noteworthy, as is the non-significance of 3 out of 5 drivers. This has two main causes. The first cause is the limited sample size. With 41 cases, significance testing is likely to come up with non-significant results.

The second cause is multicollinearity: the various drivers are highly interrelated to each other - the model suffers from data noise. Table 6 shows the collinearity statistics. Tolerance values should not fall below .2, whereas the VIF statistic should not exceed 5 [9]. Clearly, the Capabilities construct violates these thresholds. Industry know-how is in the ‘danger zone’. Due to high multicollinearity and a limited sample size, the individual predictors of the model are difficult to estimate reliably.

Table 6: Bi-variate correlations and Multicollinearity statistics

	Correlations	Collinearity Statistics	
Model construct	Zero-order	Tolerance	VIF
Reputation	,678	,298	3,353
Industry Know-How	,556	,233	4,286
Capabilities	,732	,167	5,998
Solution	,769	,294	3,397
Innovation & Improvement	,753	,336	2,973

From this perspective, it is a useful idea to see whether the individual predictors are related with customer satisfaction, when viewed on their own. In other words, a bi-variate analysis is performed for each relation. The resulting correlations are shown in Table 6, under ‘zero-order correlation’. The effect size r exceeds .55 for each construct and all results are highly significant ($p < .001$ for every construct). The lower correlation between industry know-how and customer satisfaction explains why it turns out as negative in the multiple regression analysis.

The bi-variate results show that the reasoning behind the five drivers of satisfaction is sound; each construct shows a high correlation with satisfaction. It would be worthwhile to investigate the model with a larger dataset ($N > 80$). This allows inspection of smaller effect sizes, and a larger dataset also is the best remedy to combat multicollinearity and improve significance test results.

6 Practical validation

The contribution and use of the results have been tested in two ways. First, the operationalizations of the constructs enabled a detailed analysis of company performance. The involved company has been advised to improve its performance with respect to knowledge of customer processes, spare parts management, quotation transparency and technical reporting by means of various specific actions. These actions are part of a comprehensive customer satisfaction system which has been set up to guarantee continuity in measurement and follow-up at the involved company. The questionnaire used in obtaining the data for the model testing has been adopted as the primary tool for this customer satisfaction system. During an ISO audit performed by KEMA in September 2008, the system was particularly praised as being one of the strongest improvements for 2007-2008.

The second validation arrived in the form of an independent customer satisfaction research effort performed by the Dutch branch of the involved company. This concerned a study in which customers from the Netherlands and Belgium were sampled by means of an e-questionnaire (N = 54, response rate = 31%). There was no overlap in addressed customers between this survey and the previously described interviews (see section ‘analysis of results’). Unfortunately, the model constructs were not tested in the Dutch branch survey. However, many of the operationalizations used in the Dutch survey were highly comparable to those for the model constructs. The resulting output of the Dutch survey shows a high degree of resemblance with the results from the standardized survey (discussed previously in the analysis of results). This is a good indicator of the ability of the standardized survey to uncover ‘truthful’ results that can be used to indicate priorities for improvement for use in a business environment.

7 Conclusions & Recommendations

The overall model is highly reliable and valid. The predictive power of the overall model is high (r = .853). The individual predictors (or constructs) of the model are more difficult to judge, as the dataset suffers from multicollinearity and is too small to generate positive significance results, unless the relation between predictor and customer satisfaction is very pronounced. The research question of this paper is as follows:

What are drivers for customer satisfaction in a project-oriented, business-to-business market environment?

Based on the statistical analysis of research results from a company in a project-oriented, B2B sector, the research question can be answered. The sole predictor that can be accepted unequivocally as a driver for customer satisfaction in a project-oriented, business-to-business environment is Solution.

The quality of the overall model (with r = .853) and the high bi-variate correlation between the individual drivers and customer satisfaction in a project-oriented, B2B environment shows the promise of the model.

With a higher sample size, the problems with multicollinearity and unfavourable significance of most of the predictors can likely be resolved. Further research is recommended to test the model; a sample size of N > 80 is recommended to detect smaller effect sizes and combat multicollinearity.

8 References

[1] Beelaerts WWA, Verhagen WJC, Santema SC. The Effects of Co-Innovation on the Value-time Curve: Quantitative Study on Product Level, Journal of Business Market Management, 2008, 2:1, pp. 5-24.

[2] Berry LL. Discovering the Soul of Service: the Nine Drivers of Sustainable Business Success, Palm Reader E-books, 1998.
[3] Bingguang L, Riley MW, et al., A comparison study of customer satisfaction between the UPS and Fed-Ex: an empirical study among university customers, Industrial Management & Data Systems, 2006, 106: 2, pp. 182-199.
[4] Bruhn M, Grund MA. Theory, development and implementation of national customer satisfaction indices: the Swiss Index of Customer Satisfaction (SWICS), Total Quality Management & Business Excellence, 2000, 11:7, pp. 1017-1028
[5] Chakraborty G, Shrivastava P, Marshall F. Are drivers of customer satisfaction different from buyers/users from different functional areas?, Journal of Business and Industrial Marketing, 2000, 22:1, pp. 20-28.
[6] Cronin JJ jr., Taylor SA. Measuring Service Quality: A re-examination and Extension, Journal of Marketing, 1992, 56:3, pp. 55-68.
[7] Eskildsen J, Kristensen K, Juhl HJ and Østergaard P. The Drivers of Customer Satisfaction and Loyalty. The Case of Denmark 2000-2002, Total Quality Management & Business Excellence, 2004, 15:5, pp. 859 -868
[8] Fornell C. A national customer satisfaction barometer: the Swedish experience, Journal of Marketing, 1992, 56:1, pp. 6-21.
[9] Field A. Discovering Statistics using SPSS, 3rd ed., London: Sage Publications, Inc., 2005.
[10] Fink A. Conducting Research Literature Reviews, 1st ed., London: Sage Publications, Inc., 1998.
[11] Gonzalez H, Fredericks J. Guidelines to Co-operation, Quirks Marketing Research Review, 2007, 21:3.
[12] Hair JF, Anderson RE, Tatham RL, William C. Multivariate Data Analysis, 5th Ed., Upper Saddle River: Prentice-Hall, 1998.
[13] Hakonsson H. Industrial Marketing and Purchasing of Industrial Goods – An Interaction Approach, John Wiley & Sons, London, 1982.
[14] Heskett JL, Sasser WE, Schlesinger LA. The Service Profit Chain: How Leading Companies Link Profit and Growth to Loyalty, Satisfaction, and Value, Free Press, 1997.
[15] Hill N. Customer satisfaction measurement for ISO 9000, 1st ed., Elsevier Science & Technology, 2001.
[16] Homburg C, Rudolph B. Customer satisfaction in industrial markets: dimensional and multiple role issues, Journal of Business Research, 2001, 52:1, pp. 15-33.
[17] Kotler P, Keller K. Marketing Management, 13th ed., Upper Saddle River: Prentice-Hall, 2008.
[18] Oshagbemi T. Overall job satisfaction: how good are single versus multiple-item measures?, Journal of Managerial Psychology, 1999, 14:5, pp. 388–403.
[19] Porter, ME. Competitive advantage: creating and sustaining superior performance, London: Simon & Schuster Ltd., 1998.
[20] Reichheld FF. The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value, Boston: Harvard Business Press, 1996.
[21] Rossomme J. Customer satisfaction measurement in a business-to-business context: a conceptual framework, Journal of Business and Industrial Marketing, 2003: 18:2, pp. 179-95.
[22] Sharma S, Niedrich RW, Dobbins G. A framework for monitoring customer satisfaction: an empirical illustration, Industrial Marketing Management, 1999, 28:1, pp. 231-43.

Web Technologies

Development of a Web-Based Mass Customization Platform for Bicycle Customization Services

Tung-Hung Lu[a,b,1] , Amy J.C. Trappey [a, c]

[a] Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan
[b] Identification and Security Technology Center, Industrial Technology Research Institute, Taiwan
[b] Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei, Taiwan

Abstract: Bicycle manufacturing is a mature industry in Taiwan and several companies already own the world-known brands such as Giant, Merida and KHS. These companies based on not only their core design capabilities but also well established supply chain to delivery high quality products to their customers. However, none of them focus on providing bicycle customization services to a single customer with affordable price. In contrast with these companies, some small workshops provide full customize services, but the price is not suitable for general public. This paper develops a web based mass-customization platform to aid the bicycle customization services through cyberspace. The online platform composes of two major parts, which are the bicycle requirement extraction service and the custom bicycle specification management service. These two modules can assist customers define their needs intuitively (and often at home) while maintain their spending within budgetary constraints. By using these two modules, the bicycle industry can establish an easy customization process for the mass general market. Further, with the mass customization function online, all customer orders and specifications are collected in electronic forms in realtime in order to optimize production planning and execution for quick market response and order fulfillment.

Keywords: Mass customization, information technology, web based product platform, bicycle industry.

1 Introduction

It is only a human nature that people prefers and often is willing to pay extra to be different from others, e.g., one-of-the-find designer outfits and customized home

[1] Corresponding Author: Researcher, Identification and Security Technology Center, Industrial Technology Research Institut.; Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan; Tel: +886-3-5913408; Fax:+886-3-5826464; Email: tonylu@itri.org.tw

decoration. The need for distinction is often used as a product design principle, particularly for the high-end market. From the product design perspectives, the needs for differentiation can be categorized into two folds. One is the appearance and the other is the functionality. The appearance means that the outlook style of a product can be built based on customer's preferences. In functionality aspect, the physiological differences cause a product functioning differently from expected, e.g., a bicycle, used by two people with two different weights and heights, will affect its acceleration rate. Hence, employing human factors into product design becomes a well known approach for mass customization. From the endusers perspectives, customers may very likely to choose the type of product that does not fit their physiological condition due to the lack of human factor knowledge during the product design. In order to close the gap between design and customers needs, this research focuses on developing a system with guidance-based interface to extract customer's preferred functionality and appearances. We use the bicycle customization management service to demonstrate the concept and the IT implementation.

The remaining of this paper is organized as follows. Section 2 is the literature review of mass customization on product design. Section 3 depicts the customization process model and discusses the drawbacks we found in this model. Section 4 describes the web based bicycle mass customization process model and discuss the adventage of the process model. Section 5 presents the system architecture and implementation details. The final section concludes the results and contributions of the research efforts.

2 Literature Review

The definition of mass customization (MC) is to produce a product based on consumers' requirements while the cost is still affordable for general public [1]. The concept of mass customization quickly becomes an emerging research topic. The growing interest in MC led researchers to suggest companies to shift from mass production to mass customization [2]. However, the successful story in real MC adoption is rare. MC itself is viewed being neither a methodology nor as a technology, but a service strategy [3, 7-9]. It involves not only production management but also direct customer requirement response, modular product design, supply chain management and after-service management. The difference between mass production and MC can be easily identified. Mass production is product centre and focuses on controling the cost of product. MC is customer centre and focuses on the customer value creation [7]. The customer-oriented means the product is made based on the requirements set by a customer. In other words, retrieving customer requirements is the crucial first step for customization. Although the traditional face-to-face interview is an effective way to collect customer's first hand requirements, the cost is very high and the efficiency is very low. Further, the requirement extraction process is often an iterative process. Customers may need to review and modify their previous decisions to ensure their needs are recognized and fulfilled. The skill of the interviewer may effects customer's decision and causes unsatisfactory. Thus, using web as an interactive media to collect customer's needs becomes an important option for implementing a system to mass customization [4-6]. Companies can establish an interactive web site to provide the process to customize a product and collect customer's requirements online[11-12]. However, we found most of web sites are designed for people who are familiar with the products at the first place [6]. This situation forms a gap for novice people to participate in the customization process. Thus, this paper presents a process to establish interactive web environment for both novice and experienced customers. No matter the product knowledge level of the customer, a customized bicycle can be configured through the interactive process.

3 Processes of Bicycle Customization

The bicycle manufacturing is a mature industry. Companies in the industry can be divided into two catalogues based on their service attributes. The component provider provides the fundamental parts such as stems, pedal, et al. The finished bicycle companies design their own frame and then assemble the frame and other components. A bicycle construct and its bill of materials (BOM) generally follow a standard design. Thus, it is possible to assemble a bicycle complying with the customer's requirements. If we look closer, a bicycle is composed of 5 sub-systems including frame, transmission, control, wheel and accessories as shown in Figure 1. Each sub-system can be split into further detailed sub-assembly and components, such as fork assembly and head parts assembly. In mass production, the whole assembly process is frame centre. They put the bicycle frame on a stand and mount other parts to the frame. After running through three or four assembly stations, a bicycle is produced completely.

Although the finished bicycle companies already provided four or five different sizes of frames to fit the different heights of bicycle riders, there still are high demands to build customized bicycles. Figure 2 shows the bicycle customized process model. The model is constructed by using petri-net [10]. The model describes not only tasks definition but also the input, output and role in charge the activity in the customize process. As can be seen in Figure 2, the customize process can be split into three phase. The first phase is interview and key dimensions collect. Customer who wants to build a customize bicycle has to interview with bicycle store employee. In the interview, the employee will reveal customer's needs such as riding style and preference, painting style, and some preference regarding to accessories parts selection. Customer's pelvis, thigh, shank will be measure as key dimensions for building the frame. The second phase of the customization process is pre-engineering phase. In this phase, the design employee starts to modify the template frame design with customer's specific dimension that collected from first phase and the purchase engineer starts to send purchase orders to related component vendors by customer's requirements. The third phase is the construction phase. In this phase, the technician employee is responsible to assemble the customized frame and parts into a bicycle. The final adjustment is based on the customer's riding preference.

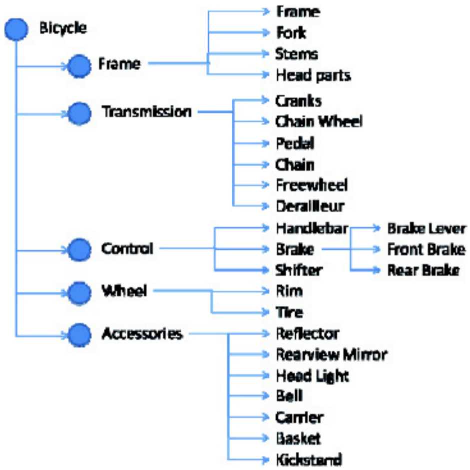


Figure 1. Bicycle bill of material (BOM) structure

Although the customization process model seems be an executable procedure, we found the actual customization is by far a very difficult task to accomplish. Things may go wrong in many stages, e.g., misunderstand customer's riding preference, purchase wrong parts, and assemble the incorrect components. The overall efficiency of MC is too low and the cost of MC is too high. Therefore, the customization is hard to be adopted by general public due to the higher prices comparing to the mass-produced bikes. Hence, we introduce a new process model and information system to improve the MC process.

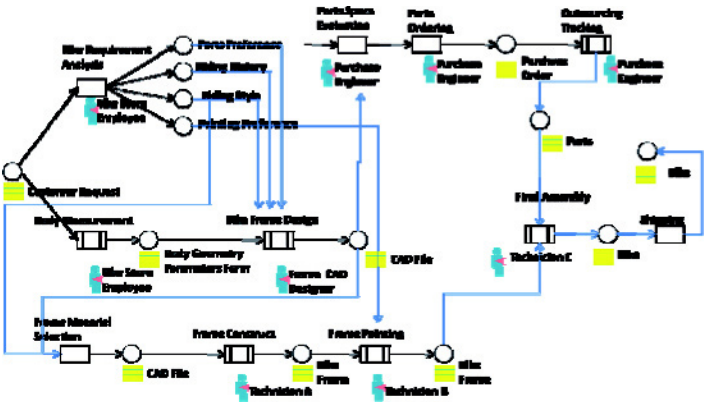


Figure 2. Bicycle customization process model

4 Web-based Bicycle Mass Customization

How to increase the efficiency of the customization process and reach the economic scale is the main consideration of this research. The ideas to improve the overall efficiency of the customization process are in three aspects. First task is introducing a web-based self service system for customers. Second task is establishing a centralized customization order management system to increase the bargaining power to the component vendors and saving the component transporting fee. Third task is establishing a strategic partnership with local bicycle shops for final adjustment and providing after-services for the MC customers. Figure 3 shows the improved bicycle customization process model. The to-be process starts when customer logs onto the site. The self service function collects customer's body geometry, riding style preferences, and key parts preferences step-by-step and interactively. For novice people, the website provides online help video for self education. Besides, customers can review the decision he or she had specified at any time during process. The requirements collection process can be iterated for many times until the customer feels comfortable to submit the final MC design. Once customer submits the customization requirements, the "Bike B2C Service" will be triggered. Customer's requirements will be transformed into a custom bike specification containing the original customer data and some basic parts information, and then be submitted to "Custom-made Bike Engineering." In "Custom-made Bike Engineering," the system will rout the specification to an available engineer. The engineer will analyze the specs and configure the BOM structure regarding to the customer's preference. After the BOM structure is completed, then "Purchase and Build" process starts. In order to reduce the cost, components are not purchase directly. Each part or sub-module has its own minimum purchase quantity limit setting. The purchase order is not released until the minimum purchase quantity is reached. We call this function as the joint purchase approach, which may make a purchase order wait for a period of time before official order being released for final production. Customers may be impatient in waiting if they do not understand the joint purchase policy. Thus, the system provides a progress report to customer as shown in Figure 4. It reveals the status of each part, e.g., "on waiting", "purchased", "checked in", "processing" and "finished." With the status report, customers are able to learn the progress of their MC order in real time. After the assembly process complete, the bicycle will be delivered to the local bicycle shop for final adjustment. Then, the customer will be notified to receive the custom-made bicycle.

There are several advantages of web based bicycle MC. From customer aspect, the web based platform provides an easy to use interface to build a custom bicycle, a transparent progress tracking interface for catching the progress status and cost saving due to the grouping parts purchase mechanism. From engineering and manufacturing aspect, the system provides an integrating environment ensure the original customer requirements can be consistence from the beginning to the final assembly. This can prevent any errors happened during stage transition.

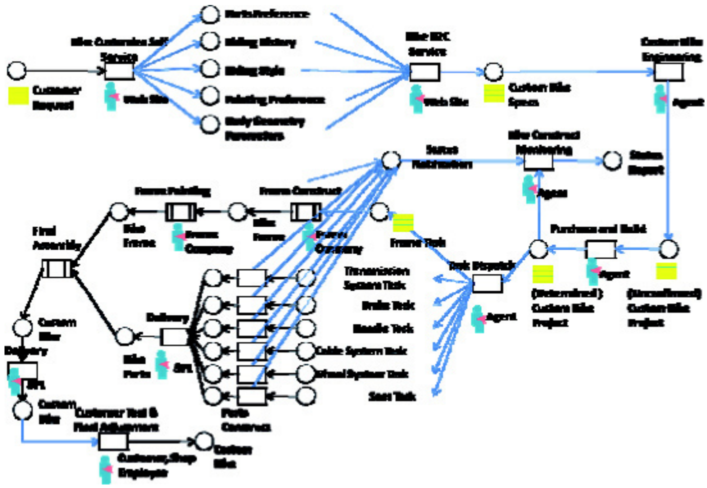


Figure 3. Process model of web-based bicycle mass customization

5 System Architecture and Implementation

The system architecture of this research is shown in Figure 5. The system can be divided into three major modules, i.e., customer requirements collection, order manager, and object persistent module. The customer requirements collection includes customer geometry collection, customer riding preferences, and customer preferred parts selection. The customer preferred parts selection is for professional customers. They can specify their preferred parts or sub-system with specific component or product brands. The order manager module is the core module of the system. For example, the custom specification editing module provides functions to process customer's requirements into production-able specification. The status tracking module collects the information of the actions taken by the engineer for customer to reveal the progress status. Finally, the object persistent module provides a common logic in handling the data storage issues of the entire system.

In addition, the customer preferred parts selection module and the custom specification editing module use the same configuration rule to validate the parts and sub-modules selection. The configuration rule is based on the assembly relationship between parts as shown in Figure 6. For example, a fork is selected only if the fork's head tube outer diameter matches the frame's head tube inner diameter. The part specification data are stored in the database table. Different type of parts has its own table field layout. The information will be loaded into a Java object when a rule is triggered. The rule is expressed as a Java compare clause, e.g., fork.headTube.outterDiameter == frame.headTube.innerDiameter. The rule is stored in a XML file and loaded at run time. The entire system is implemented using Java Enterprise Edition 6 with Apache Tomcat 6 as the web server and MySQL database as the backend data storage.

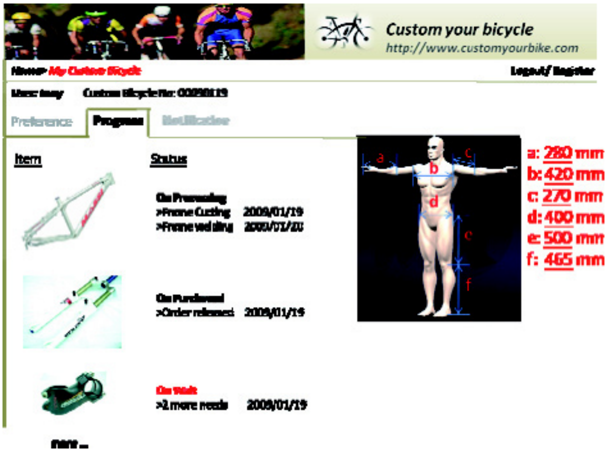


Figure 4. Status report

6 Conclusions

This paper presents a bicycle mass customization system that utilizes process modeling, Internet-enabled and object-orient web technologies. In particular, the improved process model reveals an efficient way of customizing a bicycle from both perspectives of customer order and manufacturing assembly. In addition, two important features are employed in this research, which make the customization more acceptable than the traditional face-to-face approach. They are the joint purchase strategy (for cost control) and constraint based configuration (for manufacturability and assemblability). The former can effectively reduce the overall cost of a customized bicycle, while the later provides an automatic validation of part selections for both customers and engineers. Finally, the web-based system provides a transparent information environment for both customers and manufacturers to track progresses and exchange messages on the fly. The system can keep customers informed in the real-time progress of the MC order and is also the key to keep a good cutomization experiences to the mass market.

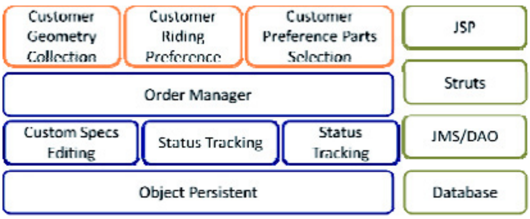


Figure 5. Web based bicycle MC system architecture

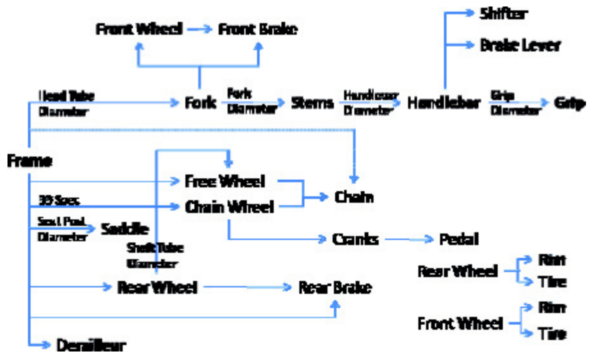


Figure 6. Bicycle parts assemble relationship

7 References

[1] Davis, S. M., Future Perfect. Reading, MA: Addison-Wesley, 1987.

[2] Silveira, D. G.; Borenstein, D.; Fogliattoc, F. S. Mass customization: Literature review and research directions, International Journal of Production Economics, vol. 72, no. 1, pp.1-13, 2001.

[3] Pine, B. J. Making mass customization happen: strategic for the new competitive realities, Planning Review, vol. 21, no. 5, pp. 23-24, 1993.

[4] Chen Y. H.; Wang Y. Z.; Wong M. H. A web-based fuzzy mass customization system, Journal of manufacturing systems, vol. 20, no. 4, pp. 280-287, 2001.

[5] Janda, S.; Trocchia, P. J.; Gwinner, P. J. Consumer perceptions of Internet retail service quality, International Journal of Service Industry Management, vol. 13, no. 5, pp. 412-431, 2002.

[6] Simpson, T. W.; Umapathy, K.; Nanda, J.; Halbe, S.; Hodge, B. Development of a Framework for Web-Based Product Platform Customization, Journal of Computing and Information Science in Engineering, vol. 3, no. 2, pp. 119-130, 2003.

[7] Kotha, S. Mass Customization: Implementing the Emerging Paradigm for Competitive Advantage, Strategic Management Journal, vol. 6, no.5, pp.21-42, 1995.

[8] Kotha, S., From Mass Production to Mass Customization: The Case of the National Industrial Bicycle Company of Japan, European Management Journal, vol. 14, no. 5, pp. 442-450, 1996.

[9] Gilmore, J.; Pine, J. The four faces of mass customization, Harvard Business Review, vol. 75, no.1, pp. 91-101, 1997.

[10] PROMATIS., Income Suite: Modeling, simulation and visualization of business processes, [online]. Available from: http://www.promatis.com/cbusiness/products/income/index.htm. [cited 20 Jan 2009].

[11] Frutos, J.D.; Borenstein, D. A framework to support customer–company interaction in mass customization environments, Computers in Industry, vol. 54, no. 2, pp. 115-135, 2004.

[12] Jiaoa, J.; Tsenga, M.M. Customizability analysis in design for mass customization, Computer-Aided Design, vol. 36, no. 8, pp. 745-757, 2004.

A Manufacturing Grid Architecture Based on Jini and SORCER

Wensheng Xu[1] , Jianzhong Cha

School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing, China

Abstract. A manufacturing grid is composed of distributed heterogeneous manufacturing resources to be shared and federated as services within a virtual organization. But because of some unique characteristics of manufacturing resources and the dynamic nature of the underlying Internet, a static service-oriented architecture based on Web services is not suitable for constructing a manufacturing grid in some cases. A dynamic service-oriented manufacturing grid architecture (DySOMGA) is presented in this paper based on the Jini and SORCER infrastructure. Based on DySOMGA, distributed manufacturing resources can be provisioned as Jini/SORCER services and join a collaborative engineering application dynamically when requested, therefore the dynamic nature of manufacturing resources and Internet can be handled properly.

Keywords. Service-oriented architecture, manufacturing grid, computing grid.

1 Introduction

A manufacturing grid is a special type of utility computing grids in which distributed heterogeneous manufacturing resources may be shared and federated within a virtual organization. Manufacturing grids play an important role in the modern manufacturing industry in which collaboration from multiple parties is essential. It is a key enabling technology to support virtual organizations. Engineering applications including multidisciplinary product developments often need to span across various departments to utilize various resources located in different places in a firm or in multiple firms. These resources may include various business databases, software tools, computing facilities, and human experts and engineers, etc., and they may be distributed across a country or even globally. A manufacturing grid needs to connect these resources together through Internet and provide transparent, dependable and consistent support to a wide range of engineering applications. A lot of research and development efforts in the computing grid field are based on static Service-Oriented Architecture (SOA) [1][2], for example the IBM Rational SOA solution [13] and the Globus Toolkit 4 [3] are based on Web services [4], and Web services are basically static. But since

[1] Corresponding Author E-mail: wshxu@bjtu.edu.cn

manufacturing grids are based on the Internet and the network is inherently a dynamic environment, without fully considering the inherent changing nature of the network, the grid systems will lack of scalability, flexibility and reliability.

In recent years, dynamic SOA has evolved rapidly from static SOA. Dynamic SOA can provide a flexible infrastructure for grid computing over the dynamic network. The Federated Intelligent Product Environment (FIPER) was developed under the sponsorship of the National Institute for Standards and Technology (NIST), which can provide a collaborative product design environment based on dynamic SOA [14]. Jini is one of the network technologies that are suitable for building the middleware for grid computing. The Jini technology aims to support the rapid configuration of devices and software within a distributed computing environment [5]. These devices and software are made available to remote clients as Jini services. Based on Jini, the SORCER (Service ORiented Computing EnviRonment) system which is developed by the SORCER Lab, Texas Tech University defines a dynamic SOA meta-computing infrastructure [6][7]. Based on Jini and SORCER, we have proposed a flexible manufacturing grid architecture—a Dynamic Service-Oriented Manufacturing Grid Architecture (DySOMGA) in this paper. Based on this architecture, human experts and various software tools and utilities can be conveniently and reliably integrated as loosely coupled services that can form a dynamic service federation for engineering applications when requested. The DySOMGA can provide for reusability, scalability, reliability and efficiency for a manufacturing grid in the dynamic Internet environment.

In this paper, the features of manufacturing grids and dynamic SOA are first analyzed (Section 2), then a dynamic service-oriented manufacturing grid architecture—DySOMGA, is proposed based on Jini and SORCER (Section 3), and an application of DySOMGA is introduced (Section 4), and finally the conclusions are given (Section 5).

2 Analysis on manufacturing grids and dynamic SOA

One widely accepted approach to construct service-oriented computing grids is to adopt the Open Grid Services Architecture (OGSA) [8] developed through the Global Grid Forum. OGSA is a specification that aims at defining a standard and open architecture for grid systems based on Web services protocols. The Globus Toolkit 4 (GT4) is an open source software package that implements OGSA, which can be used to develop grid applications. GT4 provides services implemented on top of the Web Services Resource Framework [9], a specification that extends Web services with stateful services and other features. But since Web services are static, grid systems based on OGSA/GT4 will lack of reliability and flexibility in the real dynamic Internet environment.

In OGSA/GT4, the network environment is assumed constant and reliable, and no latency exists. Therefore in reality, computing grids based on OGSA/GT4 could make the grid applications less reliable in the Internet environment, since Eight Fallacies of distributed computing are neglected in the grid architecture [10]. When the network topology is changed accidentally or partial system failures occur, some grid resources integrated in the computing grids could no longer provide reliable services. Another disadvantage for OGSA/GT4-based grids is related with the underlying Web services. Due to the limitations of Web services, service discovery and service registration are static, not dynamic, and the communication protocol between requestors and the service providers should be known to both sides beforehand—so this is a type of static SOA. As a result, the structure of the computing grids is mostly static, where the location of all integrated services should be pre-defined in the grid systems. It is not flexible for service requestors, and a failure of one component in the grid application could lead to failure of the whole system. Another obvious disadvantage for Web services in some situations is that transmitting all the data in the XML format in Web services protocols is not as efficient as using a proprietary protocol in the binary code format.

The varieties and complexity of the manufacturing resources to be integrated in a manufacturing grid is another concern for constructing a manufacturing grid. There are various manufacturing resources in a manufacturing grid: various engineering software tools such as Computer-Aided Design and Computer-Aided Engineering, etc., engineering data and databases, engineering standards and specifications for product development and manufacturing, manufacturing equipment, high performance computing facilities, and even human experts and engineers when human expertise is needed. These resources are distributed, dynamic, heterogeneous, and may not necessarily be working online continuously all the time as the computing resources do in a regular compute grid. For example, manufacturing equipment may be working and down, and human experts and engineers may be working and leaving in a day. Thus, on top of the service-oriented architecture, several basic problems need to be solved to cope with the dynamic nature of manufacturing resources and the Internet when constructing a manufacturing grid: (1) a proper service access interface for accessing various manufacturing resources; (2) a dynamic service-oriented architecture which allows for dynamic service discovery and service registration, lease and transaction management for services; (3) service coordination and federation services for a manufacturing grid.

Since the inherent weaknesses of the static SOA infrastructure exist in the dynamic Internet environment, dynamic SOA evolves from static SOA, where requestors do not need to know locations of providers and their communication protocols, as shown in Figure 1. In dynamic SOA, service providers and requestors can dynamically discover registries on the network. The provider is responsible for deploying the service on the network and publishing its service proxy for accessing the service to one or more registries. The proxy is created and published by the service provider and it is configured by the provider to provide the efficient communication protocol for the requestor. The requestor looks up a service by sending queries to registries, and queries generally contain search criteria related to the type and quality of service. When the requestor makes a selection from the available services, it downloads the service proxy from the registry and then invokes the service via the ready proxy, and it does not need to know location of the service and the details of the communication protocol between the proxy and the service. Because of features of dynamic SOA, it can adapt to dynamic network environment very well and is suitable for constructing a manufacturing grid.

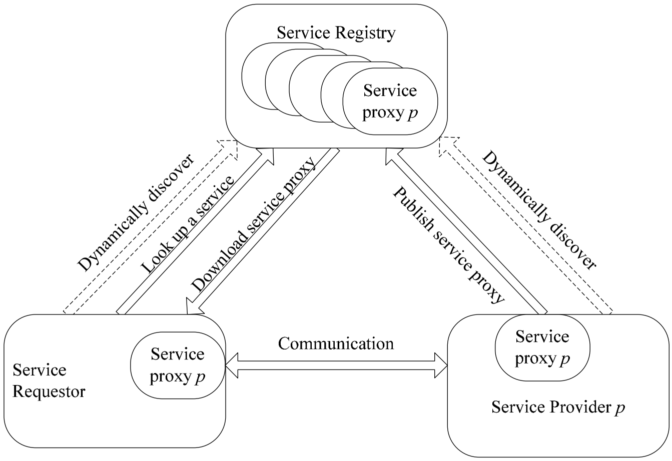


Figure 1. Dynamic service-oriented architecture

3 The architecture of DySOMGA

3.1 The layered structure of DySOMGA

Jini is an implementation of dynamic SOA and it focuses on individual service management in a networked environment [5]. In Jini not only the location of the service providers can be dynamic, also the communication protocol is neutral, and the location of the registries can be dynamic as well. Jini has a programming model with explicit leases, distributed events, transactions, and dynamic discovery/join protocols, so the dynamic nature of the network environment can be well handled, and services can dynamically join and leave and rejoin the grid environment without causing breakdown of the whole system. Based on Jini, the SORCER environment is developed to provide a federated service-to-service (S2S) metacomputing programming model [6][7]. Based on SORCER, Jini services can dynamically form federations through service-oriented programs initiated by service requestors. SORCER is focused on service-oriented programming and the execution environment, so service coordination and federation can be achieved. The features of Jini and SORCER make them a suitable infrastructure for constructing a manufacturing grid in the Internet environment, where scalability, reliability and flexibility are critical for distributed applications while dynamic nature exists for manufacturing resources and the underlying Internet environment.

Therefore, DySOMGA—a Dynamic Service-Oriented Manufacturing Grid Architecture, is proposed here based on the Jini and SORCER infrastructure. Following the principles of the hourglass model for computing grids [11], DySOMGA can be divided into five layers, as shown in Figure 2.

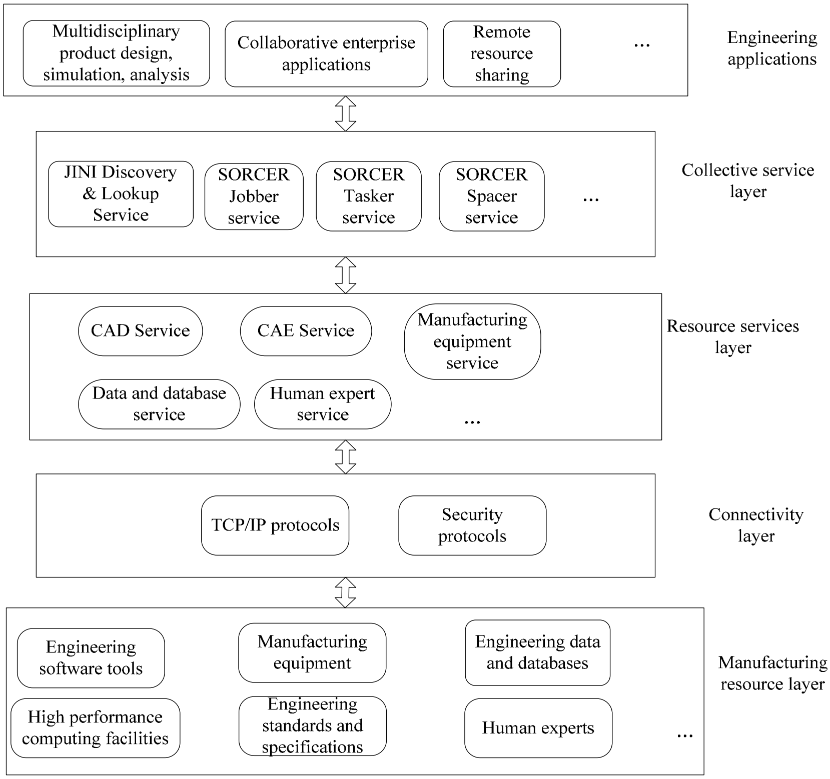


Figure 2. The layered structure of DySOMGA

The manufacturing resource layer contains all types of manufacturing resources to be shared in the grid, including engineering software tools, computing hardware, manufacturing equipment, engineering data and databases, and human experts, etc. These resources will be connected through the connectivity layer in the Internet. The connectivity layer includes communication and authentication protocols required for grid-specific network transactions, so secure data communication between manufacturing resources can be achieved.

The resource services layer provides access to manufacturing resources for engineering applications over the connectivity layer, and a primitive manufacturing resource can be encapsulated and provisioned in this layer as an individual standard Jini/SORCER service for remote requesters. Resource services in this layer are built on top of the Jini/SORCER service provider interfaces and can interact with the collective service layer through Jini discovery, join and invocation protocols. The resource services layer is the basis for implementing a grid system.

The collective service layer contains various management and coordination services for the manufacturing grid, including Jini discovery & lookup service, which is a service registry, and SORCER Jobber service, SORCER Tasker service, and SORCER Spacer service, etc. which are responsible for managing and coordinating multiple resource services in grid applications.

The application layer includes various engineering applications for specific collaborative tasks or jobs in a virtual organization. These engineering applications are created through exertion-oriented programming [6] by requesters and will be interpreted and executed through the collective service layer. Subsequently the resource services layer and the manufacturing resource layer will be involved in the collaborative applications.

3.2 The functional structure of DySOMGA

The functional structure of DySOMGA is shown in Figure 3. Based on Jini basic service interfaces, the SORCER core elements, namely core classes, have been defined to support service definition and development in SORCER, including Servicer, ServiceTasker, ServiceProvider, etc. [6] With these core elements, various manufacturing resources can be encapsulated as resource services by extending these classes, so these resource services can be exposed to requestors in the grid. Basically there could be two types of resource services in DySOMGA: one is automatic service providers which can run automatically and need no human involvement, such as some software services; the other is human-involved services, which need human involvement, such as human expert services or some manufacturing equipment services. Some resource services could have interfaces for both types of services and be working in either mode in different applications. Actually there is no fundamental difference for these two types of services from the perspective of service interfaces.

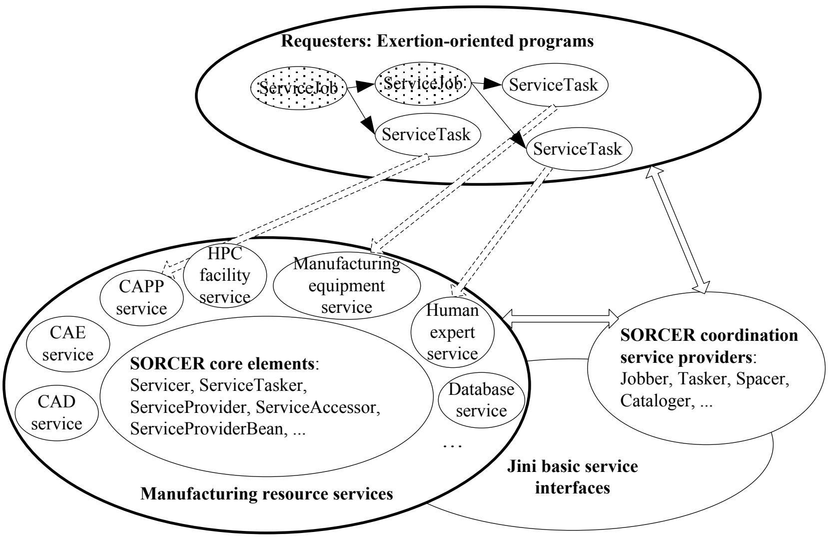


Figure 3. The functional structure of DySOMGA

In DySOMGA, an engineering application to utilize and federate multiple manufacturing resource services is achieved by requesters' exertion-oriented programs. Exertion, or namely service exertion, is a remote message in the SORCER infrastructure, which is the basic means for service requesters to communicate with service providers and pass controls to service providers [6]. An exertion encapsulates both a service signature and a service context. The service signature denotes the operation that needs to be performed, and the service context contains a tree-like data structure which can store multiple data for the operation. There are basically two types of exertions in SORCER: ServiceTask and ServiceJob. A ServiceTask is an elementary exertion which can correspond to a single service provider, while a ServiceJob is a composite exertion which can contain multiple ServiceTasks and even other nested ServiceJobs, as shown in Figure 3. The relationships between the ServiceTasks and ServiceJobs contained in a composite ServiceJob can be defined by their input-output relationships. With the support of SORCER coordination service providers, a composite exertion can discovery and invoke multiple needed service providers in the required sequence, thus a federation of services can be formed dynamically. So writing an engineering application for a service requester in DySOMGA is mainly about writing exertion-oriented programs.

The SORCER coordination service providers including Jobber, Tasker and Spacer, etc. are Jini services themselves based on Jini basic service interfaces, so they can be dynamically discovered and looked up by service requesters in DySOMGA. The SORCER coordination services can interpret exertion-oriented programs in terms of the control contexts contained in the exertions and create a federation of required service providers for application jobs in runtime, and manage a shared context for the job federation. The coordination services are a broker between the requesters' applications and the manufacturing resource services.

Therefore logically there are two spaces in DySOMGA: the requesters' application program space and the manufacturing resource service space. An exertion-oriented program in the program space represents a collaborative engineering application and contains multiple ServiceTasks and ServiceJobs, and these ServiceTasks and ServiceJobs will be mapped to the service space, as illustrated in Figure 3. So the corresponding manufacturing resource services in the service space will be selected and executed accordingly. The mapping process is coordinated by the SORCER coordination services.

4 An application of DySOMGA

Based on DySOMGA proposed in this paper, we have developed a service-oriented collaborative design platform—SCoD, in which three design and analysis engineering tools and a human expert service have been integrated in the platform [12]. Pro/Engineer, Ansys, HyperMesh, and a human Finite Element Analysis (FEA) expert service are encapsulated as Jini/SORCER services by extending ServiceTasker in SCoD. The Pro/Engineer modeling service is located in a railway vehicle manufacturing firm, while the Ansy service, the HyperMesh service and the human FEA expert service are located in a vehicle research institute, as shown in Figure 4. These four services are in the manufacturing resource service space and can join a service federation for a collaborative design job in concurrent engineering dynamically when requested by exertion-oriented programs from the requesters' application program space, and they can be coordinated by the SORCER coordination services in SCoD.

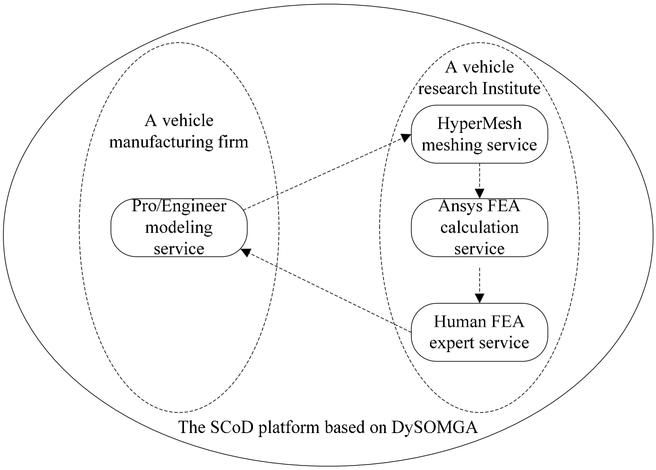


Figure 4. The SCoD platform based on DySOMGA

For example, a static stress analysis job can be composed of four service tasks—a Pro/E service task, a HyperMesh service task, an Ansys service task and a human expert service task, and each service task is defined by a service signature and a service context, and the output and input of the service contexts are linked by the mapping relations, so the four tasks can be linked in sequence in this case. According to the exertion-oriented programs, when the static stress analysis job is executed in the SCoD application layer, the four service tasks will dynamically bind to the four corresponding distributed services through Jobber's coordination and be executed by the four services iteratively until some application-specific requirements are satisfied. Based on DySOMGA, services can join and leave SCoD dynamically and the service federation for a collaborative design application can form dynamically.

Configurations are important issues in the implementation of DySOMGA. When the distributed environment spans across enterprise borders and several Jini registries are involved, the Jini registries in each network should be configured and linked with a service registry bridge such as Lincoln [15], thus remote services can be exposed to other local services or requesters as though they are local. Or alternatively, if only one Jini registry is being used for the whole environment, the underlying SORCER infrastructure can be configured explicitly with the host name or IP address of the Jini registry host, but of course in this approach, flexibility and reliability for this single registry could be compromised.

5 Conclusions

Since the Jini and SORCER infrastructure provides effective mechanisms to cope with dynamic environment, the manufacturing grid architecture—DySOMGA based on Jini and SORCER can integrate various distributed manufacturing resources as services and adapt well to the Internet environment. Multiple service providers can be redundant and dynamically provisioned in the grid for reliability and flexibility to compensate for network failures and service shutdown. Joining and leaving of an individual manufacturing service in the manufacturing grid will not affect the configurations of the whole system. A needed service can be discovered dynamically in runtime by the service type it implements. Based on SORCER exertions, a requestor can write exertion-oriented programs to invoke manufacturing resource services in a dynamic federation with no need to know the exact locations of the service providers and the communication protocols beforehand, so flexibility can be well achieved, and a virtual organization can be well supported in the distributed dynamic environment.

6 Acknowledgement

This paper is supported by Beijing Jiaotong University Research Fund (Project No.2007RC035, No.2006XZ011) and a collaborative project between the Intelligent Engineering Lab, Beijing Jiaotong University and the SORCER Lab, Texas Tech University. We would like to thank Prof. Michael Sobolewski of the SORCER Lab for his tremendous help in this joint project.

7 References

[1] I. Foster, C. Kesselman and S. Tuecke. The anatomy of the grid: Enabling scalable virtual organizations, International Journal of Supercomputer Applications, vol.15, no.3, pp.200-222, 2001.
[2] J. Darlington, J. Cohen and W. Lee. An architecture for a next-generation Internet based on Web services and utility computing, Proceedings of 15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, pp.169–174, 2006.
[3] http://http://www.w3.org/2002/ws/
[4] http://www.globus.org
[5] http://www.jini.org
[6] M. Sobolewski. SORCER: Computing and Metacomputing Intergrid, Proceedings of 10th International Conference on Enterprise Information Systems, pp.74-85, 2008.
[7] M. Sobolewski, R. Kolonay. Federated Grid Computing with Interactive Service-oriented Programming, International Journal of Concurrent Engineering: Research & Applications, vol. 14, no. 1, pp.55-66, 2006.
[8] Towards Open Grid Services Architecture. Available at http://www.globus.org/ogsa. Accessed on: Jan. 29, 2009.
[9] The Web Services Resource Framework. Available at http://www.globus.org/wsrf. Accessed on: Jan. 29, 2009.

[10] Fallacies of Distributed Computing. Available at: http://en.wikipedia.org/wiki/Fallacies_of_Distributed_Computing. Accessed on: Jan. 29, 2009.
[11] Computer Science and Telecommunications Board. Realizing the Information Future: The Internet and Beyond, National Academy Press, Washington DC, 1994.
[12] W. Xu, J. Cha and M. Sobolewski. A Service-oriented collaborative design platform for concurrent engineering, Materials and Product Technologies/Proceedings of 2008 International Conference on Advances in Product Development and Reliability, vol.44-46, pp.717-724, 2008.
[13] http://www-01.ibm.com/software/info/developer/solutions/soadev/index.jsp
[14] M. Sobolewski. Federated P2P Services in CE Environments, Advances in Concurrent Engineering, A.A. Balkema Publishers, pp.13-22, 2002.
[15] https://rio-utilities.dev.java.net/

Minding the Gap between First and Continued Usage: An Empirical Study of the Implementation of a Corporate E-Learning English-Language Program at a Financial Firm in Taiwan

Tainyi (Ted) Luor[a,1], Hsi-Peng Lu[b], Robert E. Johanson[c], Ling-ling Wu[d]

[a] Graduate School of Managenent, National Taiwan University of Science & Technology- Taiwan, R.O.C.
[b] Department of Information Managenent, National Taiwan University of Science & Technology- Taiwan, R.O.C.
[c] Department of Department of Applied Foreign Languages, National Taiwan University of Science & Technology- Taiwan, R.O.C.
[d] Department of Information Managenent, National Taiwan University - Taiwan, R.O.C.

Abstract. While extensive scholarly attention has been paid to various aspects of on-line language programs, the implementation of on-line English-language learning (COEL) programs in business contexts remains an unexplored area in the CALL (computer assisted language learning) literature. Moreover, while many studies have focused on learners' first usage intentions in TAM (technical acceptance model), few have explored their intentions to continue using them. To address this lacuna, a framework of COEL in order to explore learners' perceptions of COEL had been developed first. An examination of the proposed two models revealed gaps between learners' intention of first usage and re-usage of the COEL. Further analyses revealed a significant difference between groups of high intention learners to re-use the COEL and low intention learners to re-use the COEL. A second investigation determined eight factors that contributed to the differences between these two diverse groups of learners. This study's findings shed light on the relationship between TAM model and the perceived enjoyment, learners' attitudes, intentions and overall reactions to COEL programs. Limitations and suggestions for future studies are discussed as well.

Keywords. Corporate on-line English learning, computer assisted language, TAM, perceived enjoyment, users' reactions

[1] Corresponding Author E-mail: a384@ibfc.com.tw

1 Introduction

As English continues to solidify its position as the lingua franca of the Internet, academics, and international commerce [1, 2], it is not an overstatement to say that a firm's chances of survival in future years will be directly linked to the aggregate English proficiency of its workforce [3,4]. Therefore, it is not surprising that companies the world over are scrambling to implement English-language teaching programs to improve their employees' English ability [5]. This wealth of studies conducted on CALL reveals much about the increasing availability and increasing complexity of CALL program design. However, to the authors' knowledge, this ever-increasing body of CALL literature has failed to address the potential role that CALL English-language learning (COEL) programs play in the corporate context.

The purpose of the present study is to address the aforementioned lacunae in the literature by investigating the factors that influence business COEL learners' intentions of utilizing corporate e-learning English-language learning and teaching programs. Our study comprises of two components. First, we develop and test a COEL framework on the basis of reasoned action theory [6] by integrating the Attitude towards Corporate E-learning (ATT), Intention to Use Corporate E-learning (INT), PEOU and PU in the TAM model, perceived enjoyment, Satisfaction (SAT) and Affective Reaction to E-learning (AR). This framework was designed to help explain how learners' perceptions affect their intentions to re-use COELs and to ascertain stronger causal rationales behind the observed relationships. Second, the attempt to identify potential gaps between users' INT of first use (INT-FU) and re-use (INT-RU) was conducted with quantitative analyses to determine the critical factors that may lead to gaps.

Taiwan was considered to be an ideal site for the present study because it is a rapidly-developing country whose future economic and political survival depends largely on its ability to improving the English-proficiency of its workforce [7]. Despite recent efforts to improve the island's aggregate English proficiency levels, however, the education system continues to teach English according to the grammar-translation methods and place a tremendous emphasis on testing rather than communicative competence [8]. Hence, it is believe that in Taiwan furnishes us with an appropriate sample from with which we can rigorously assess employees' perceptions of COEL and employees' UGE behaviors.

2 Hypothesis Development

2.1 Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) and Attitude toward COEL (ATT)

TAM [9,10] has received major attention in the IT/IS acceptance literature. Deriving from TRA and theory of planned behavior [11], TAM posits that behavioral intention to use a new technology is determined by perceived ease of use, the user's perception of the quantity of effort needed to use the system and perceived usefulness, and the user's perception of the extent to which using the system will advance his or her performance in the workplace [9,10]. TAM also suggested that the interactions of these two core components in which perceived usefulness is determined by perceived ease of use; that is, the easier a system is to use, the more assistance it can provide to those who choose to use it [12]. In this study, *perceived ease of use* and *perceived usefulness* were adopted, these two core TAM components were used to measure learners' attitudes as expressed in the following three hypotheses in order to ascertain the relationships among learners' PEOU, PU and ATT. The following hypotheses proposed.

H1: An individual's PEOU relates positively to the individual's PU so that the higher the individual's PEOU, the higher the individual's PU.

H2: An individual's PEOU relates positively to the individual's ATT so that the higher the individual's PEOU, the higher the individual's ATT.

H3: An individual's PU relates positively to the individual's ATT so that the higher the individual's PU, the higher the individual's ATT.

2.2 Perceived Enjoyment (PEOJ)

Results of other studies have revealed that the effect of perceived enjoyment was pivotal in influencing an individual's attitude [13]. We proposed the following hypotheses.

H4: An individual's PEOU relates positively to the individual's PENJ so that the higher the individual's PEOU, the higher the individual's PENJ.

H5: An individual's PENJ relates positively to the individual's ATT so that the higher the individual's PENJ, the higher the individual's ATT.

2.3 Intention to Use Corporate E-learning (INT)

According to the theory of reasoned action [6] and the theory of planned behavior [11], an individual's intention to engage in a specific action derives from his or her individual's attitude toward the behavior. Many empirical studies have confirmed the relationship between attitude, intention, and behavior [14]. The two theories (TRA and TBA) assume that practical behavior such as the intention to use a system can draw strength from this intention to act [15, 16, 17]. On the basis of TRA and TPB, the following hypotheses regarding ATT and INT was proposed.

H6: An individual's ATT relates positively to the individual's INT so that the higher the individual's ATT, the higher the individual's INT.

H7: An individual's PU relates positively to the individual's INT so that the higher the individual's PU, the higher the individual's INT.

In the first stage, learner's intentions to first use of COEL (INT-FU) must be explored; Figure 1 shows the conceptual model and the relationships existing among variables of PEOU, PU, ATT and INT-FU.

2.4 Reaction to COEL

Kirkpatrick's [18, 19] four general criteria for training evaluations: reactions, learning, behaviors and results are widely considered appropriate and useful for evaluating training outcomes [20]. Among the four criteria, trainee reaction is the most frequently measured training outcome in training practices and studies [21, 22, 23]. Reaction may best be defined as how positively learners perceive particular a training program. It can also be defined as the measure of how participants feel about the various aspects of a training program, including the topic, speaker, and schedule [23]. In this sense, evaluating a training program in terms of trainees' reaction is the same as measuring trainees' feelings [18, 19]. In measuring trainers' reactions, researchers have identified a number of other types of reactions [24]. We adopted Brown's [11] two types of reactions: technology satisfaction (SAT) and affective reaction (AR). On the basis of the above rationale, the following hypotheses were proposed.

H8: An individual's PENJ relates positively to the individual's SAT so that the higher the individual's PENJ, the higher the individual's SAT.

H9: An individual's SAT relates positively to the individual's INT so that the higher the individual's SAT, the higher the individual's INT.

H10: An individual's AR relates positively to the individual's INT so that the higher the individual's AR, the higher the individual's INT.

Figure 1 shows the conceptual model employed in this study to survey learner's intentions to re-use COEL.

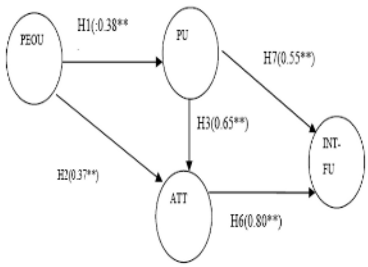


Figure 1. Research model: INT-FU

Note. Coefficients are marked as ** if significant at the .01 level. Coefficients are marked as * if significant at the .05 level

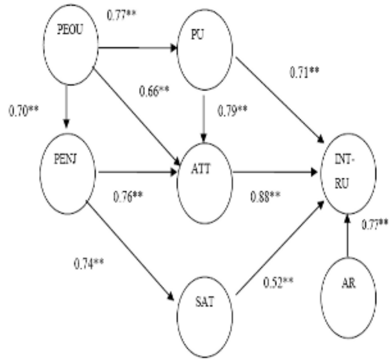


Figure 2. Research model: INT-RU

Note: Coefficients are marked as ** if significant at the .01 level. Coefficients are marked as * if significant at the .05 level

Figure 1. shows the conceptual model.

2.5 Minding the Gap between intentions to first use and re-use COELs

Studies on general e-learning initiatives have identified many critical factors that may affect learners' decisions to withdraw from e-learning programs, such as retention, motivation, and satisfaction [25, 26, 27, 28]. By identifying critical factors that contribute to a program's success, scholars and practitioners can better understand the dynamics of training processes and take steps to improve them.

To bolster the contributions of the current study to the existing literature on COELs, this study also identifies and examines COELs' critical success factors (CSFs). Although scant research has been conducting on CSFs, and it is generally acknowledged that there is very little research on the CSFs of students' perception-based acceptance of e-learning [29]. Several scholars have investigated the important role that trainees play in e-learning CSFs. Selim [30], for example, discovered that e-learning CSFs within a university environment can be grouped into four categories: IT, instructor, student and university support, while Masoumi [31] argued that student factors may affect the CSFs. In order to identify potential

CSFs in COELs, we explore the potential critical factors that contribute to the success of COELs by aiming to identify the factors that lead to a difference between two groups of users after their having used a COEL program for three months: Employees who have high INT continue to use a COEL, and employees who have low INT continue to use a COEL. By identifying these CSFs, organizations can improve the effectiveness of COELs when dealing with non-program-related CSFs.

2.6 Research Questions

(1) Does a gap exist between learner's INT-FU and INT-RU?
(2) What are the critical factors that may contribute to any gaps between high INT-RU learners and low INT-RU learners toward a COEL program?

3 Methods

3.1 Sample and Procedures

The participants in the present study were employees at a financial institution with a total of 180 employees that had a high market share of the Taiwanese money market. The company's CEO decided to implement the COEL examined in this study as a means of both improving employees' English abilities and simultaneously evaluating their English proficiency levels. The English training company that was hired to create the COEL designed it to comprise of three levels (Beginner, Intermediate and Advanced). The COEL program also provided three alternative learning modes to personalize their COEL experience and enhance their ownership of their own language learning: *The Dynamic Mode, The Guided Mode*, and the *Free-to Roam Mode*.

One month prior to the installation of the entire COEL program, the COEL was installed on selected computers at the research site for one month and users' experiences with the COEL were assessed via questionnaires to ascertain their reactions to - and general attitudes regarding - the COEL. The web-based questionnaires were distributed to all the 180 employees via the Intranet at the first stage, 59 usable responses were returned with a response rate of 32.78% in the pre-installation survey. Findings revealed that COEL users had average values of 3.8, 3.7, 3.7 and 3.8 (on a scale: from 1-the lowest to 5-the highest) of PU, PEOU, ATT and INT-FU, respectively from the response of fifty-nine employees. The relationships among the four variables are shown in Figure 1. After three months of using the COEL, a second survey was conducted to assess employees' reactions. The questionnaires were distributed to all the 180 employees again, this time fifty seven participants responded. Table 1 presents the demographic background of the second survey participants.

Table 1. Study Participants Demographic Backgrounds

Sex	Female	19	33.3%
	Male	38	66.7%
Age	20~29	7	12.3%
	30~39	30	52.6%
	40~49	17	29.8%
	50~59	3	5.3%

The relationships among the variables after three months of COEL usage are shown in Figure 2. The relationship among INT-RU and other variables is still positive; however, the average value of INT-RU is only 2.98 (with a standard deviation of 1.1) which is significantly lower (the F-value of Levene's Test for Equality of Variances is 21.8.) than the value of INT-FU (average value of 3.8 with a standard deviation of 0.5). Therefore, in the second stage, we sought to identify any differences of intention to re-use the COEL that might exist between high intention and low intention learners to re-use a COEL. To do this, we studied the high-intention (H-INT) participants and the low-intention (L-INT) participants who had completed the second survey and who had used a COEL. A measure of the 10 identified potential factors derived from both our interviews and the literature regarding e-learning critical factors [31, 32, 33] was assessed via this survey. Fifty seven participants participated in the COEL program were assessed to do the survey. We defined H-INT participants as those whose mean score on the intention scale was larger than or equal to 3 and L-INT to those whose score was lower then 3. Consequently, 20 participants were identified as H-INT participants and 37 participants were identified as L-INT participants.

3.2 Measures and Analysis

Unless specified, all items were measured on a 5-point Likert-type scale ranging from 1 (strongly disagree) to 5 (strongly agree). All items were adopted from well-developed scales that are commonly used in behavioral science and information management research. Corresponding reliability coefficients the of measured items are between .87 and .96 [34, 35, 36]. The data was analyzed by using path analysis with a regression approach, a multivariate analytical method for empirically examining sets of relationships in the form of linear causal models [37]. Figure 2 exhibits the path coefficients of our research model- regarding users' intentions to re-use the COEL. As shown in Table 2, most correlations among study variables were less than .8, indicating that there was no threat of multicollinearity inherent in the variables used in this study [38].

4 Results

4.1 Regression results

Table 2 presents the descriptive statistics of study variables. As Table 2 shows, the PEOU, PU and the SAT statistics have higher mean values, which imply that the participants generally had positive perceptions of *easy to use*, *useful* and *positive satisfaction* in using the COEL. The bivariate relationships show that all significant correlations were less than .80, save the correlation between PU to e-learning and AR to use e-learning ($r = .84$, $p < .01$), AR to ATT ($r = .85$, $p < .01$) and ATT to INT ($r = .88$, $p < .01$).

For the relationships among study variables, the regression results supported the proposed relationships from Hypotheses 1 through Hypotheses 10. However, we found that the average value of INT for learners who participated in the COEL was only 3.0 (as shown in Table 2), which was the lowest value among of all of the variables in our model. This finding suggests that other unaccounted for factors might have been present that affected either the learners' continued use of the COEL that caused them to give up and have low intentions to re-use the COEL as a viable English learning tool. To identify the potential factors that might contribute to learner's intention to use COEL, a further study was conducted.

4.2 Mind the Gap: A formative evaluation of intentions to re-use COEL

Igbaria et al., [39] used the TAM to investigate the impact that external factors (i.e., individual, organizational, and system characteristics) may have on users' acceptance of microcomputer technology. Their study confirms the effects of

Table 2. Means, Standard Deviations, and Correlation among Study Variables

	1	2	3	4	5	6	7
1 PEOU-AVG	**3.6(.8)**	.77**	.70**	.76**	.66**	.66**	.64**
2 PU-AVG		**3.5(.87)**	.69**	.64**	.84**	.79**	.74**
3 PENJ-AVG			**3.2(.87)**	.74**	.69**	.76**	.71**
4 SAT-AVG				**3.5(.81)**	.57**	.59**	.52**
5 AR-AVG					**3.3(.87)**	.85**	.76**
6 ATT-AVG						**3.2(1.04)**	.88**
7 INT-AVG							**3.0(1.13)**

** Correlation is significant at the .01 level (2-tailed).
* Correlation is significant at the .05 level (2-tailed).
Means (standard deviations) are shown diagonally.

individual, organizational, and system characteristics regarding users' perceived ease of use and perceived usefulness. Moreover, Luor et al. [40] identified ten potential factors contributing to the differences between two low usage learners and high usage learners in a corporate e-learning program to discover that time management and technical problems were the two critical factors that led to the usage differences. In the present study, we adopted most of the potential factors used in Luor et al. [40], in the second investigation and tried to tease out further information regarding the difference between levels of high and low intention to re-use the COEL between these two groups of learners. The ten factors that contributed to the differences of intention to re-use COEL were shown in Table 3.

In this study, a 10-item scale was developed to measure whether or not the *High Intention Group* (*Group H:* 20 observations with average INT-RU values greater than 3) and the *Low Intention Group* (*Group L: 37* observations with average INT-RU values equal or less than 3) differed from each other in their perceptions of the importance of the aforementioned ten factors. Table 3 presents both the items and the results of the independent sample *t*-tests. According to Table 3, the two groups differed from each other regarding individual characteristics of *Enjoyment*, *Intimidation*, *Time Management*, *Expectations of Efficiency and Effectiveness*, *and Improvement of English Abilities*. These two groups also differed from each other regarding organizational characteristics of *Motivation*, *Number of Right Courses* and *Management Support*.

5 Discussion

On the basis of information system literature and training literature [9, 11], we proposed and examined a COEL framework to explore the relationships among the factors involved in implementing an English-language COEL program at a financial firm in Taiwan. The results confirmed all of the above proposed hypotheses (as shown in Figure 1).

Our findings show that a user's PEOU has a positive relationship with the user's PU, ATT and PENJ. An individual's PENJ, has a positive relationship with the user's ATT and SAT. An individual's PU has a positive relationship with the user's ATT and INT, while an individual's ATT, SAT and AR all have a positive relationship with the user's INT. Our study also confirmed that after three months use of the COEL, the direct effects of PU on PEOU, PEOU on ATT, PU on ATT, PU on INT and ATT on INT are higher than the direct effects that measured users' first stage-intentions to use the COEL. These results may imply that increases of the practical usage of computer technology can better explain the impact of TAM model. The finding that PU and ATT are related positively to INT-FU and INT-RU suggests that a company that wants to sustain learner's intention to use COEL should provide more useful materials which can generate more positive ATT toward COEL to encourage these employees to participate in the COEL. The findings demonstrate that one's experience with a COEL relates positively to AR and SAT, so when employees try to use a COEL, they generally have positive reactions and satisfaction to it.

After conducting a t-test, the following eight factors belong to individual and organizational characteristics that may have led to this difference was identified: *Motivation*, *Enjoyment*, *Intimidation*, *Time Management* [31, 41], *Expectations of Efficiency and Effectiveness* [28], *Improvement of English abilities*, *Number of Right Courses* and *Management Support*. The abovementioned findings echo previous research on online learning environments that states that if they are designed in a way that enables learners to be self-directed to the point that they can take control of their own learning, learners will be highly motivated to participate in them [41].

Table 3. Descriptive Data: Mean (std. deviation) and T-test between High Intention (20 observations) and Low Intention (37 observations)

Factors (Questionnaire Items)	*High intention (group-H)*	*Low intention (group-L)*	*t*-value (*p*-value)
Intention	4.22(.47)	2.31(.74)	10.46 (0)**
Motivation (Our company encourages us to use the COEL.)	4.2(.7)	3.51(.9)	2.96 (0.01)**
Enjoyment (I enjoy using the COEL.)	4.0(.79)	2.86(.63)	5.91 (0)**
Intimidation (I feel intimidated to use the COEL.)	2.3(1.22)	2.81(.7)	-2.02 (0.05)*
Time management (In office, I have time to use the COEL.)	4.2(.52)	3.62(.59)	3.65 (0)**
Expectations of efficiency and effectiveness (I expect myself to use the COEL.)	4.15(.59)	3.0(.75)	5.96 (0)**
Improvement of English abilities (Using the COEL can improve my problem-solving abilities.)	4.2(.77)	3.43(.73)	3.73 (0)**
Number of Right Courses (So far, the COEL provides the right number of courses and exercises.)	3.75(1.02)	3.19(.58)	2.61 (0)**
Technical problems (I have technical problems using the COEL.)	2.55(1.19)	3.05(0.78)	–1.93
Promotion (The COEL provides better chances for promotion.)	3.6(.94)	3.59(.8)	0.02
Supervisor's encouragement (My supervisor encourages me to use the COEL.)	3.7(.86)	3.14(.71)	2.65 (0.01)**

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed). **Equal variances are not assumed.**

6 Conclusion

It is not an overstatement to say that the Internet has transformed English language teaching; however, there is a dearth of studies that address online English-language learning programs in business contexts. To address this lacuna in the literature, a framework that may help explain and allow one to predict users' intentions of continuing to use COELs was presented and tested. Our study revealed, for example, that ATT is significantly and positively correlated with PU, PEOU and PENJ; moreover, INT is significantly and positively correlated with PU, ATT, SAT and AR. The relationship between ATT and INT confirm Ajzen's TRA. Our findings also echo the results that the effect of perceived enjoyment is crucial in influencing individual's attitudes toward the COEL [13].

In summary, our findings underscore quite commonsensical findings in the general online learning literature that learners' perceptions of a program's usefulness, ease of use, and the amount of enjoyment derived from using it will generate positive attitudes of - and positive intentions to use the COEL. Furthermore, learners who perceive of the COEL with positive, satisfactory and affective reactions will be more likely to continue using it in the future. We also suggest that organizations - especially companies that encourage the learning of foreign languages as part of the employee-training programs by installing COEL programs choose a system that contains a variety of learning activities in order to produce positive learning dynamics because companies that use requirements or incentives to ensure that employees complete their training do not have significant problems with completion rates.

7 References

[1] Flowerdew, L. (2003). 'A combined corpus and Systemic Functional analysis of the problem-solution pattern in a student and professional corpus of technical writing'. TESOL Quarterly, 37, 489-511.

[2] Swales, J. (1997). English as Tyrannosaurus rex. World Englishes, 16, 373-382.

[3] Vollmer, H. (2001). Englisch und mehrsprachigkeit: Interkulturelles lernen durchEnglisch als lingua franca? In: D. Abendroth, Timmer and G. Bach (eds.) Mehrsprachiges Europa. Festschrift für Michael Wendt zum 60. Geburtstag (pp. 91-109).Tübingen: Narr.

[4] Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a lingua franca. International Journal of Applied Linguistics, 11, 133-158.

[5] Kaufmann, S. (2006). Business English: New approaches are needed to increase the ROI in corporate English learning. Published: 9/27/2006 http://www.buzzle.com/articles/business-english-new-approaches-increase-roi-corporate-learning.html. Uploaded 6 June, 2008.

[6] Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, Mass.: Addison-Wesley Pub. Co.

[7] Ministry of Education, Republic of China (Taiwan). (2006b, June). MOE Foreign English Teachers Project for Taiwan's Public Schools, the Republic of China (Taiwan). Press Release.

[8] Chen, C. H. The teaching behavior differences between native and non-native English teachers in early childhood education. Unpublished M.A. thesis, Providence University, Taiwan. 2005.

[9] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 318–339.

[10] Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. Management Science, 35(8), 982–1003.

[11] Ajzen, I. The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50, 179-211. 1991.

[12] Ndubisi, N. O. (2006). Factors of online learning adoption: A comparative juxtaposition of the theory of planned behavior and the technology acceptance model. International Jl. on E-Learning, 5(4), 571-591.

[13] Ha, I. Yoon, Y. Choi, M., Determinants of adoption of mobile games under mobile broadband wireless access environment. Information & Management, 44(3), 276-286.

[14] Kim, M.-s., & Hunter, J. E. (1993). Relationships among attitudes, behavioral intentions, and behavior: A meta-analysis of past research: II. Communication Research, 20(3), 331-364.

[15] Bagozzi, R. P. Attitudes, intentions, and behavior: A test of some key hypotheses. Journal of Personality and Social Psychology, 41(4), 607-627. 1981.

[16] Bagozzi, R. P., & Phillips, L. W. Representing and testing organizational theories: a holistic construal. Administrative Science Quarterly, 27(3), 459-489. 1982.

[17] Taylor, S., & Todd, P. (1995). Understanding information technology usage: a test of competing models. Information Systems Research, 6(2), 144-176.

[18] Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. Journal of the American Society for Training and Development, 13(11), 3-9.

[19] Kirkpatrick, D. L. (1996). Great ideas revisited: Revisiting Kirkpatrick's Four Level Model. Training and Development 50(1), 54-57.

[20] Arthur, W. J., Bennett, W. J., Edens, P. S., & Bell, S. T. Effectiveness of training in organizations: A meta-analysis of design and evaluation features. Journal of Applied Psychology, 88(2), 234-245. (2003).

[21] Brown, K. G.. An examination of the structure and nomological network of trainee reactions: A closer look at "smile sheets". Journal of Applied Psychology, 90(5), 991-1001. 2005.

[22] Brown, K. G., & Gerhardt, M. W. Formative evaluation: An integrated practice model and case study. Personnel Psychology, 55, 951-983. 2002.

[23] Noe, R. A. (2002). Employee training and development (2nd ed.). Boston, MA: McGraw-Hill/Irwin.

[24] Kraiger, K. (2002). Decision-based evaluation. In K. Kraiger (Ed.), Creating, Implementing, and Managing Effective Training and Development: State-of-the-art Lessons for Practice (pp. 192-233). San Francisco,CA: Jossey-Bass.

[25] Alexander, S. E-learning developments and experiences. Education + Training, 43, 240-248. 2001.

[26] Bonk, C. J. Online teaching in an online world. Bloomington, IN: CourseShare.com. 2001.

[27] Mason, R., & Weller, M. (2000). Factors affecting students' satisfaction on a web course. Australian Journal of Educational Technology, 16(2), 173-200.

[28] Packham, G., Jones, P., Miller, C., & Thomas, B. (2004). E-learning and retention: key factors influencing student withdrawal. Education & Training 46(6/7), 335-342.

[29] Lee, J., Hong, N. L., & Ling, N. L. (2002). An analysis of student's preparation for the virtual learning environment. The Internet and Higher Education, 4, 231-242.

[30] Selim, H. M. (2003). E-Learning critical success factors: An exploratory investigation of student perceptions. Paper presented at the 4th Annual Research Conference at the United Arab Emirates University.

[31] Masoumi, D. (2006). Factors for effective e-learning. The e-quality final seminar Retrieved April 30, 2007, from www.e-quality-eu.org/pdf/seminar/e-Quality_WS3_DMasoumi.pdf

[32] Féraud, G. (2005). Transferring successful corporate training into e-learning in a large international organization: The case of early warning in conflict prevention Systèmes d'Information et Management, 10(2), 89-106.

[33] Soong, M. H. B., Chan, H. C., Chua, B. C., & Loh, K. F. (2001). Critical success factors for on-line course resources. Computers & Education, 36, 101-120.

[34] Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology, 78(1), 98-104.

[35] Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. Psychological Methods, 1(2), 199-223.

[36] Schwab, D. P. (2005). Research methods for organizational studies (2nd ed.). Mahwah, N.J.: L. Erlbaum Associates.

[37] Klem, L. (1995). Path analysis. In L. G. Grimm & P. R. Yarnold (Eds.), Reading and understanding multivariate statistics (pp. 65-97). Washington, D.C: American Psychological Association.

[38] Emory, C. W., & Cooper, D. R. (1991). Business Research Methods. Homewood, IL: R.D. Irwin.

[39] Igbaria, M., Guimaraes, T., & Davis, G. B. (1995). Testing the determinants of microcomputer usage via a structural equation model. Journal of Management Information Systems, 11(4), 87-110.

[40] Luor, T., Hu, C., & Lu, H. (In press.). Mind the gap': An empirical study of the intention and actual usage of corporate e-learning programs in the financial industry (2008), British Journal of Educational Technology. (in-press).

[41] Dringus, L. P. (1999). Resource requirements for online learning environments. Journal of Instruction Delivery Systems, 13(4), 20-25.

Web-Based Mechanism Design of 3C Plastic Parts with Knowledge Management

Wen-Ren Jong[1], Chun-Cheng Lin, Yu-Hong Ting, Chun-Hsien Wu, Tai-Chih Li

Department of Mechanical Engineering and R&D Center for Mold and Molding Technology, Chung Yuan Christian University, Chung-Li, 32023, Taiwan, ROC

Abstract. The mechanism design of plastic parts provides means for assembly and strength for applications. The common mechanism designs include component separation, snap fit, boss, hook, lip, rib, gusset, etc., which are traditionally accomplished by the accumulated knowledge of well-trained designers. This paper presents a navigating system to accomplish the mechanism design of 3C products in a systematic, consistent, and scientific approach, making the design procedure transparent. This not only greatly saves the time in training and designing, but also reduces some preventable man-made errors.

Keywords. Top-Down Design, Collaborative Design, Knowledge Management, Parametric Design, Re-development.

1 Introduction

Due to the rapid development of the Internet in the 21st century, using the Internet integrates different platforms and designs related resources, which allow the traditional industrial manufacture increase efficiency immensely and bring out new competitiveness.

In order to systematize and speed up product development, many design methodologies are also starting to be applied, for example, Collaborative Design and Top-Down Design, etc. The so-called Collaborative Design is actually the realization of the concurrent engineering concept[1]. With the concept of concurrent engineering being applied in many fields, Computer Supported Collaborative Work (CSCW) is gradually replacing the traditional segmented sequence way of designing. CSCW uses computer technology and the internet to proceed to Computer Supported Collaborative Design (CSCD) [2]; in other words, CSCD offers designers to accomplish allocated assignments in different places.

[1] Corresponding Author Email : wenren@cycu.edu.tw

The most commonly-used designing strategy in Collaborative Design is the Top-Down Design. According to documents, Yang [3] and Aleixos [4] had already published related essays on the Top-Down Design method. The difference between Top-Down Design and Bottom-Up Design is that designers first proceeds the design at top assembly level, and then splits into sub-assembly and sub-components in order to design in more detail. This method is also capable of putting into practice the concept of Concurrent Engineering in CAD.

In the CAD system, Tang [5] also mentioned that although the CAD system already has inner features, which includes apertures, holes, ribs, etc. Even though these feature functions are able to satisfy the needs of the user in building and modeling, they don't apply to all different designs and cannot be modified in the fastest way possible. Therefore, when it comes to actual design and convenience, it cannot fully satisfy the needs in designing; users usually develop specific functions applied on customized designs by themselves. Thus, the establishments of these particular customerized functions are able to assist designers in accomplishing similar drawing at the fastest possible.

The research of navigating system in this study for Collaborative Mechanism Design of 3C product with knowledge management is mainly structured on the internet platform of Pro/Engineer. Moreover, it uses internet techniques and the operation of managing methods to structure the network system and information system of inner enterprises on the open system of sharing information. At the same time, using the three-level structure as a basis to apply the common web technologies and interface, simplifying the development and maintenance of the system. Moreover, this web-based collaborative system can increase communication of inner enterprises, shorten the time of product development, and strengthen competitiveness advantages of enterprises.

2 Collaborative Mechanism Design

Collaborative Mechanism Design Navigating System in this study utilizes the system structure by Jong [6], as Figure 1 shows; it uses project management to proceed, clearly structuring the whole designing procedure, maintaining the continuous process and managing the designing authority of engineers effectively. As 0 2 shows, when designing groups receive product information from clients, they aim at the contents of product information to proceed the design discussion. Then, they use the five major design modules of Collaborative Mechanism Design Navigating System to assist in the design development of products. The major five modules includes the establishment of product skeleton, the assigning of collaborative works, dividing of designing works, detailed design of mechanism, and lastly, examination of assembled product. Moreover, the main contents of the job are shown in Figure 3. During the process of designing, designers need to take the related manufacturing issues into consideration, because the design of plastic products will directly affect the mold design, mold manufacturing, and lastly, molding process. Designers obtain these cross-module designing information from technical knowledge databases, the proven design-process method is used inside every module, navigating designers to execute design instructions and connecting to designated design databases, which enhances the development power of modules and the ability of information circulation.

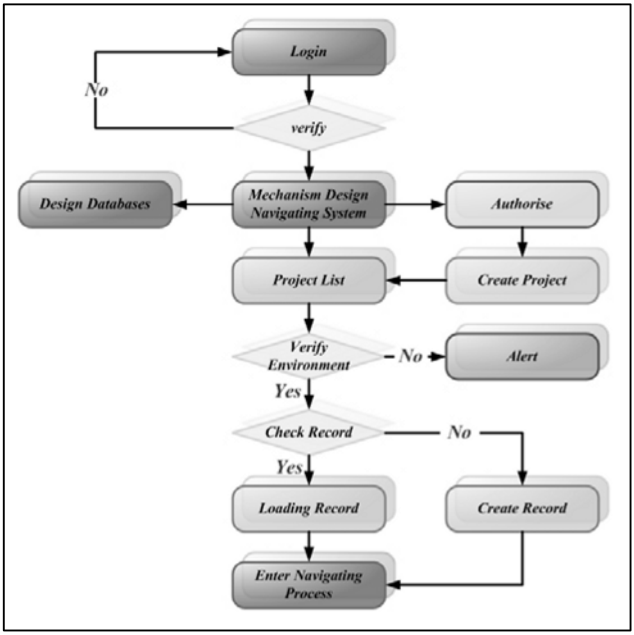


Figure 1. Project Procedure of System

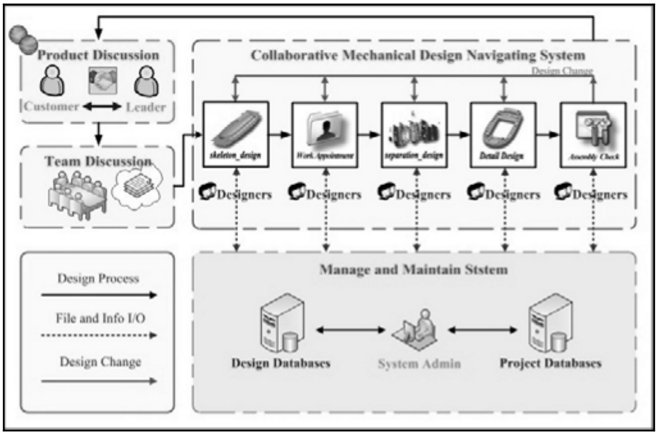


Figure 2. Collaborative Mechanism Design Navigating System

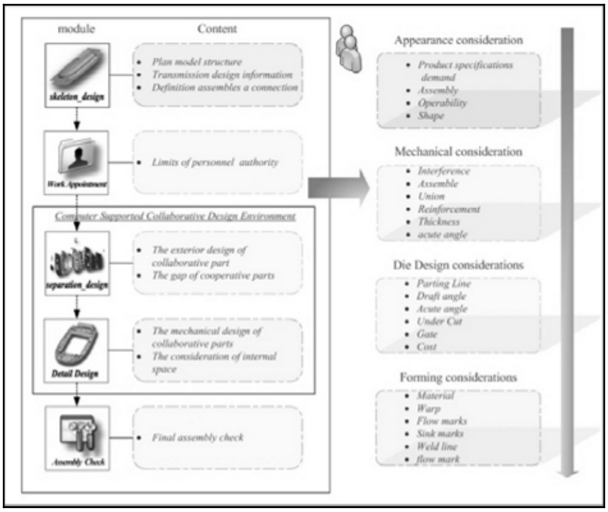


Figure 3. Mechanism Design Module

2.1 Top-Down Design

Nowadays, products are mainly composed of components with complex appearances; each design feature often implicates many techniques. If we apply the Top-Down Design onto our design, then product assembling will be connected, shared, and restricted. Thus, the design will become systematized. Under these circumstances, designers are able to keep the key design information, so that they can perform Concurrent Engineering on the design, allowing designers dedicated in certain areas to design preferred items. In order to achieve product Top-Down Design, the procedure before building a model is to build the Master Part skeleton model. The Skeleton Design of the system is shown as Figure 4; it divides Master Part into first Master-Part, second Master-Part, and third Master-Part.

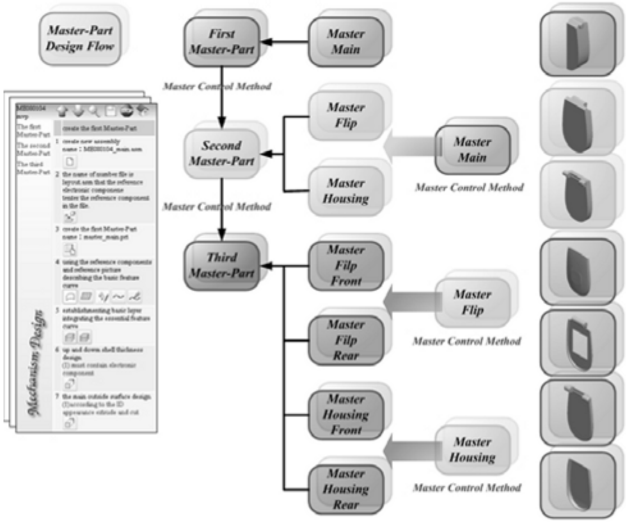


Figure 4. Top-Down Design

2.2 Master Control Method

Master Control Method is not a command or function of Pro/Engineer; it skillfully uses Pro/Engineer to share the information in merge function to achieve the feature relationships between slave components and master component. Furthermore, this method uses one referencing component to control many components, and this referencing component is referred to as Master Part. By using Master Part to control several components, all the components or sub-assemblies can be controlled by Master Part. Thus, if a component is changed, the whole product automatically changes subsequently.

3 Technical Document Database

The technical document database in this study serves as a knowledge management center. The main goal is to provide referencing advices and create a new knowledge-feedback way for enterprises and individuals. This system divides the technical document database into two segments, which are technical document and discussion of problems; it integrates and links with design process and feature database. There are five stages in the function plan, which are the creation, classification, sharing, acquisition, and application of design knowledge.

4 Mechanism Features Database

Mechanism Feature Databases use various parameterized designs to control customized features; the so-called "parameterized" means putting restrictions to each feature. By changing the dimensions of each similar and repetitive feature, it can restrict and automatize the feature design process. In application, one can construct specific frequently used features of dissimilar sizes to comply to different business fields, and even save it in its database. After that, selecting the parameters in the UDF (User Defined Feature) from the database can save the time in establishing features, achieving the goal of design standardization. As shown in Figure 5, putting the parameterized design into feature database saves the time of feature element. Furthermore, it use Browser interface to find associated parameters, making the feature design precise and fast, as Figure 6 shows.

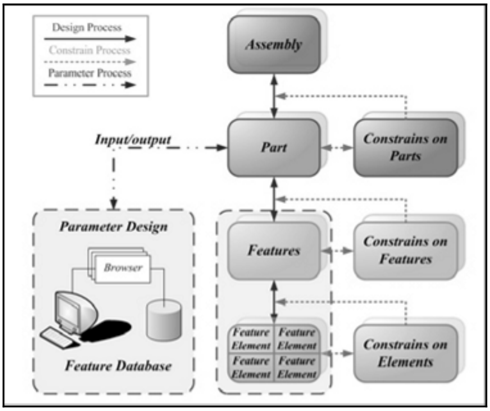


Figure 5. Feature Database

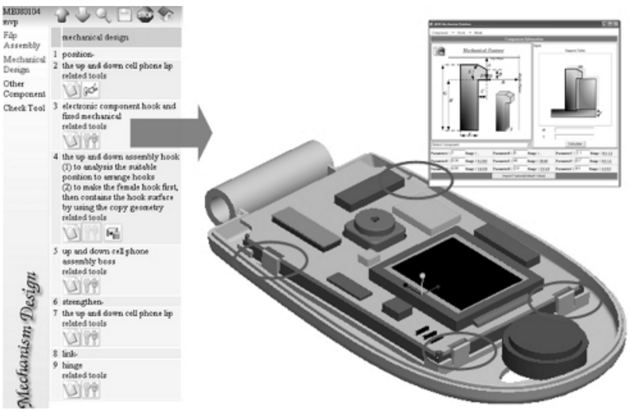


Figure 6. Multi-Classification

5 Results and Case Studies

There will show and introduce Collaborative Mechanism Design Navigating System, and cell phones will be used as the case studies. In the end, comparisons of using Collaborative Mechanism Design Navigating System to complete the design of cell phones are made. Figure 7 is a procedure of collaborative mechanism design system; users are required to proceed the design job step by step.

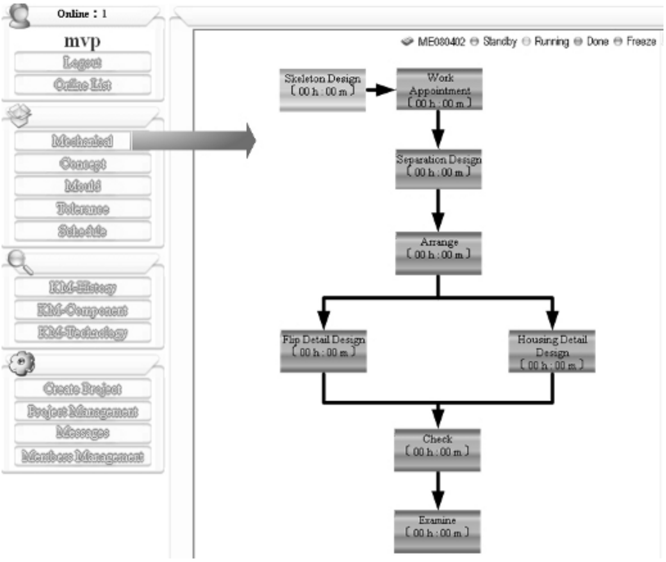


Figure 7. Procedure of Collaborative Mechanism Design System

Skeleton Design

In order to achieve product Top-Down Design, the procedure before building a model is to build the Master Part skeleton model. Firstly, when using Skeleton Design procedure to ask designers to build first Master-Part, the main job is to draw the outline of product from ID image, and then produce a basic 3D model. Also, we take the assembling relations of the parting surface space between flip and housing and the position of parting line into consideration; for example, the clearance between two parting surfaces or the position of parting line. Then, the second Master-Part merge first Master-Part by Master Control Method so that parameter relations can be established between second Master-Part and first Master-Part.

Layer Classification

As Figure 8 shows, this system bases on the design requirements of cell phone products and uses Layer Classification to implement a concise and convenient procedure. Moreover, this system carries out component layers information in individual components and uses Layer Classification to display or hide classified geometry. This type of second development methods speeds up the development time and provides a clear and simple design interface.

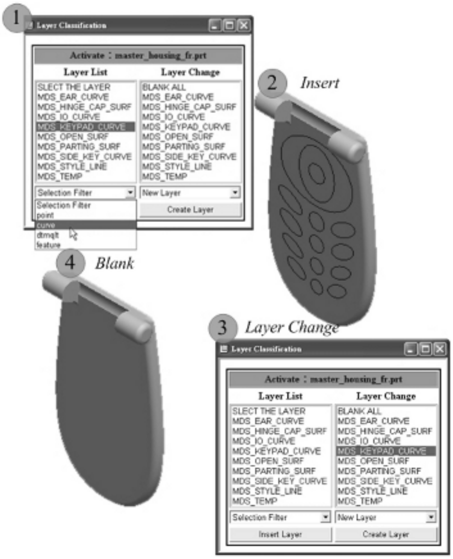


Figure 8. Layer Classification

<u>Design Separation</u>

Collaborative Design is an efficient methodology in speeding up product development and design. This system manages the authorized right of designers, and easily proceeds the designing jobs on the sub-components with the assistance of Master Control Method.

After establishing a Master Part on the procedure, fast merging and relations establishment is carried out on the other sub-components of Master Part. Then, the designing authorized right are appointed to individual designer, as Figure 9 shows; this way, appointed workers are able to clearly know the distribution of components and the situation of designers.

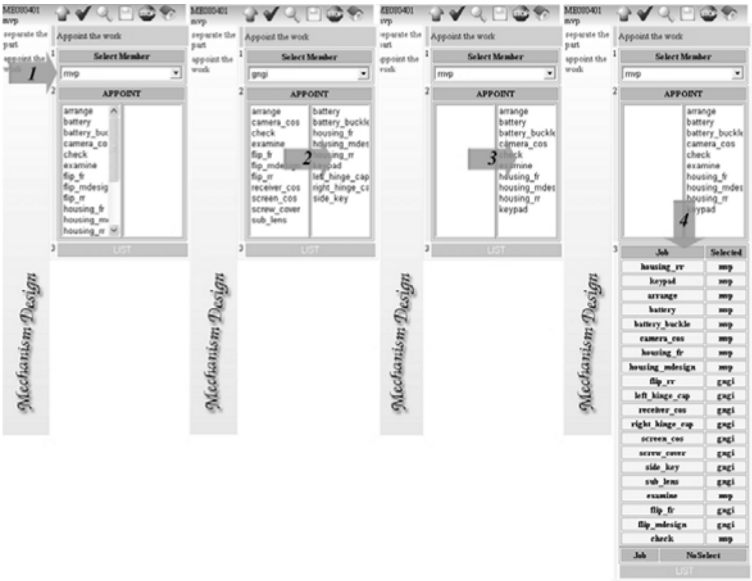


Figure 9. Work Appointment

Detail Design

After Design Separation, actual mechanism designing can be started, and this is where the assembling, positioning, fixing, and strengthening problems of component mechanism between components are taken into account.

Comparison of the System

Figure 10 is the upper part of the flip cell phone. Designer A is in charge of the collaborative design after Skeleton Design by assigning work to two designers to perform Detail Design. With the help of the time comparison chart, efficiency conditions can be known when carrying out the mechanism design. Generally speaking, usually a considerably amount of time is consumed when designing the complex skeleton model, especially if it didn't go through the systematic plan, which often results to the design modification problems caused by artificial negligence. From the comparisons of each procedure in the statistical time chart, it shows that under the conditions of using or not using the navigating system, systematic design can save about 30% of time, as shown in Table 1 and Table 2. In addition, by adding collaborative design in the Skeleton Design procedure saves 60% and more time in cell phone designing, as Table 3 shows.

Figure 10. collaborative design of the cell phone component

Table 1. Comparison of the Design Time

	Skeleton Design	Separation Design	Detail Design	Total Time
without using navigating system	21hr	2.5hr	7.5hr	31.0 hr
Mechanism-design navigating system	16.5hr	1.5hr	4.8hr	22.8 hr
collaborative navigating system	16.5hr	0.8hr	4.8hr	**1.1** hr

Table 2. Comparison of the Collaborative Design Time

	Total Time(Hour)	Total Time (Working Day)	Saving Time %
without using navigating system	31.0 hrs	1.30 days	--
mechanism-design navigating system	22.8 hrs	0.95 days	26%
collaborative navigating system	22.1 hrs	0.92 days	29%

Table 3. Comparison of the Total Design Time

	Item	Separation Design	Saving Time %
without using navigating system	1	2.5hr	--
mechanism-design navigating system	2	1.5hr	40%
collaborative navigating system	--	0.8hr	68%

6 Conclusions

Entering products into the market in the shortest amount of time possible is something every enterprise is pursuing. With the help of navigating system with knowledge management, the efficiency of the whole design is enhanced and the chances of artificial errors are reduced. Moreover, feature databases are able to carry out modeling jobs rapidly. By integrating technical databases, learning design experiences from other designers becomes very easy; no longer will intangible assets be lost when enterprise staff quit their jobs. Also, the technical databases that are established will help newcomers quickly begin product designing and manufacturing in order to maintain the advantages of marketing competition.

From the comparisons of case study, it shows that this flip cell phone requires approximately 31 hours of mechanism designing. The development of collaborative mechanism design navigating system assisting group of designers with collaborative design can shorten about 30% time in designing.

7 Acknowledgement

This work is supported by Taiwan Green Point Enterprise Co., Ltd., R&D Center of Mold and Molding Technology, Ministry of Education, Taiwan, and National Science Council in Taiwan (Grant NSC 96-2221-E-033-016).

8 References

[1] R. I. Winner, J. P. Pennell, H. E. Bertrent, M. M. G. Slusarczuk, “ The Role of Concurrent Engineering in Weapons System Acquisition”, IDA Report R338, Alexandria VA: Institute for Defense Analyses, 1988.

[2] W. Huifen, Z. Youliang,C. Jian, S. F. Lee, W. C. Kwong, “Feature-based collaborative design”, Journal of Materials Processing Technology,Vol.139, pp.613–618 ,2003.

[3] Y. Yang, S. Zhang, Z. Li, S. Gao, “Task Planning for Top-down Collaborative Assembly Design”, CSCWD Technical Program, 2006.

[4] N. Aleixos, P. Company, M. Contero, “Integrated modeling with top-down approach in subsidiary industries”, Computers in Industry ,Vol.53,PP.97–116,2004.

[5] M. Tang, Y. Wen, X. Mi, J. Dong, "Parametric Modeling with User-Defined Features", In Proceedings of the Sixth International Conference on Computer Supported Cooperative Work in Design, July 2001.

[6] W. R. Jong, C. H. Wu, H. H. Liu, M. Y. Li, "A Collaborative Navigation System for Concurrent Mold Design", International Journal of Advanced Manufacturing Technology, Accept.

WIW - A Web-Based Information System for Profile of Wind

Wu Xiao Bing[1], Adans Iraheta Marroquín, and Moacyr Fauth da Silva Jr.

Faculty of Engineering and Architecture, University of Passo Fundo (UPF), Brazil

Abstract: A Web-based Information System (WIS) is developed in an environment of Internet, utilizing the World-Wide-Web to get the public universal or private restricted accesses. Web-based Information system for profile of Wind (WIW) is a virtual information system developed to process the years collected data in the Laboratory of Anemometric of University of Passo Fundo, in Brazil. The data are measured by sensors at the heights of 20, 30 and 40 meters on an anemometric tower with 40 meters. The collected data are saved in a Datalogger, then they are transferred to a local computer by a Ramcard, and with WIW, they are uploaded to a special database and processed. Some analyses of the profile of wind during 2000-2006 are shown as an example. The system was written in PHP, with the template engine Smarty, and database MySQL.

Keywords: Web-based system, WIW, Profile of Wind, PHP/MySQL, Sustainable Development, Alternative Energies.

1 Introduction

Web technology is very popular today. It is not just meaning many web sites are developed, or a lot of information is published in Internet, above all many applications have been implemented in the environment of Internet, too, with its universal accessibility.

Xu, et al [10] present Co-OPERATE, a web-based system for better co-ordination of manufacturing planning and control activities across the supply network. Peter Anderson [13] presents a virtual emergency management information system prototype, which is developed in Canada to deal with the obstacles for integrating and coordinating emergency operations posed by disasters, especially when essential emergency management staffs are prevented from reaching emergency operations centers. Kneale, et al [4] show VELNET (Virtual Environment for Learning Networking), a virtual learning environment developed and used successfully to teach

[1] Corresponding Author E-mail: wu@upf.br

computer networking. Burn, et al [2] define the virtual organizations as the electronically networked organizations that transcend conventional organizational boundaries, with linkages which may exist both within and between organizations; they also present different models of virtuality which can exist within the electronic market. On exploitation of the energy of wind, Ale [1] shows some important studies have been done in Rio Grande do Sul (RS) State of Brazil on the potential of aeolian and solar energy [8], and systems for the agricultural electrification. In the works, the clear notion in implanting the aeolian systems is identified to collaborate with the energy supply of the State. The studies are based on the data of 42 stations distributed throughout the State of the RS in an area of 282,480 km^2. Ale [1] demonstrated the viability in implanting the alternative systems in the cities of Mostardas and Tavares in the RS, as the average speed of wind over 5 m/s.

2 Web-based Information Systems

A web-based system is a computer software system, which is not at the same local of an end user, but on a long physical distance. It is virtual, and of the universal accessibility. It can be hosted in a professional data center with special infrastructure to guarantee its availability for 24 hours a day, 7 days a week, with a high security.

Web-based Information System (WIS) is a virtual information system, which does the same papel as a traditional Information System (IS), the main difference is the way to access the system by the end users and the way to build it. The Figure 1 shows the idea of WIS.

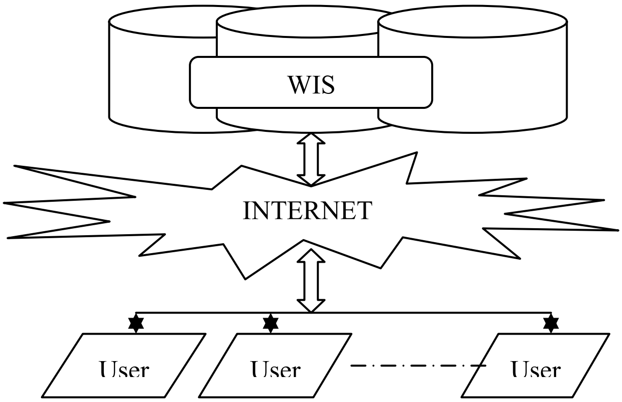


Figure 1. Web-Based Information Systems

3 System Structure

Based on the concept of WIS, we develop a web-based information system for profile of wind (WIW). The system consists of 6 modules: Database, Data Transfer, Data Edition, Data Process, Web Search and Users Control. The Figure 2 shows the structure of the system.

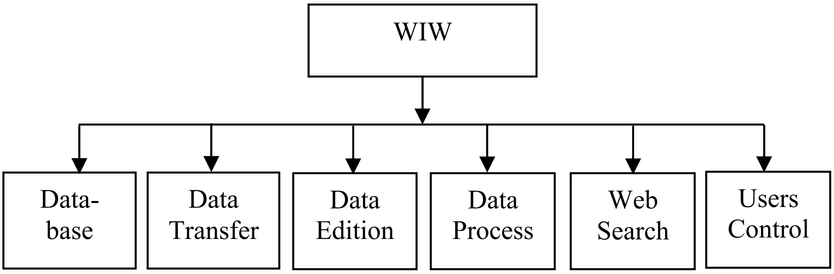


Figure 2. Structure of WIW

- **Database** is developed to organize and manage all the original measured data.
- **Data Transfer** is responsible to create a data table, upload a data file and transfer the measured data into the database.
- **Data Edition** is developed to search, edit and pre-process some special cases manually and internally by a system administrator.
- **Data Process** is utilized to process all the data in the database according to the necessity of research interesting.
- **Web Search** is used to search and show the wind information from the database by the Internet.
- **Users Control** is used to manage and control the access of the authorized users for the system WIW.

4 Module System

The Web-based information system for profile of wind (WIW) is developed in a module way. It consists in four parts:

- Template Engine
- Declaration

- Main code
- Module code

Template Engine is used to develop the presentation layer of the web application, defines the layout of the end-user's viewing pages. In the system, we applied the Smarty Template Engine to realize it. With the tag replacing, we get the dynamic contents in the web page; and with its quick and painless development and deployment of application, we get high-performance, scalability, security and future growth. It is of the following features: caching, security, easy to use and maintain, variable modifiers, filters, plugging, add-ons, debugging, compiling, and performance [14].

Declaration consists of two parts: Declaration for code environment and for html presentation style. To set up the working environment of script code, we need to define the globe variables, data base accessing information, application structure information, module accessing information, template engine information, etc. To control the presentation of contents to the end-user, we used CSS files to define the style of the web pages.

PHP codes controls the behaving of the system, it is separated into two parts: main code and module code. It works in a module way. The following figure shows how it works.

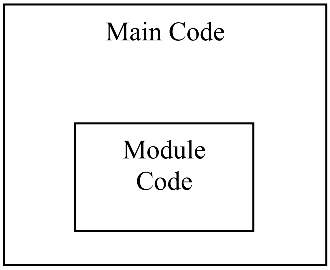


Figure 3. Structure of WIW

Each module code is organized in a file, and is inserted in the main code to define the specified action. Each activity of the system is dealt with by a module code, such as upload a set of data, process a set of data, and so on. The main code is a container and controller of module code file. It calls the defined module code dynamically according to the necessity of the system, to realize the specific action and producing the desired result (content) dynamically on the web page.

With the separation of PHP code, environment definition, style definition, and the html presentation, the system has a great flexibility to deal with the activities for the end users. For example, any alteration of html presentation layout (front end) does not need any changes of PHP code (back end).

The module structure of the system makes the development and maintenance work easier. A large activity can be broken into a set of small activities, and delivered them to different programmers.

Comparing with the traditional information system, the virtual way may publish the information by web, and can be accessed by everyone who is connected with the Internet. With the low cost of professional hosting and availability of free and complete development environment, it can reduce significantly the requirements and costs of infrastructure of information technology, especially helpful for a small business.

5 Implementation

The WIW is coded in PHP [15], with database MySQL [16]. To separate the front-end presentation and back-end PHP code, we utilize Smarty engine to realize it. All the files are hosted in a server of the University of Passo Fundo. The measured data and processed information maybe available for every one on web in the future.

5.1 Collection Of Data

The data shown in this paper are collected at the Laboratory of Anemometric of the Faculty of Engineering and Architecture, at the University of Passo Fundo. The measurements are done by sensors at the heights of 20, 30 and 40 meters on an anemometric tower with 40 meters. The Figure 4 shows the tower built especially for the wind profile study and the process of collection of data.

The main instrument to measure the wind profile is the anemometer, from which we get the measured data, such as the wind speed, direction, etc., produced from the movement of the shovels of the anemometers installed at 20, 30, 40 meters height in the anemometric tower.

The measured signals are sent to a Dataloger by a Router (switch). They are stored in Dataloger for a period until five days. With a Ramcard, the signals were transferred to a local computer with a special reader, and then upload to the database with the web-based information system (WIW).

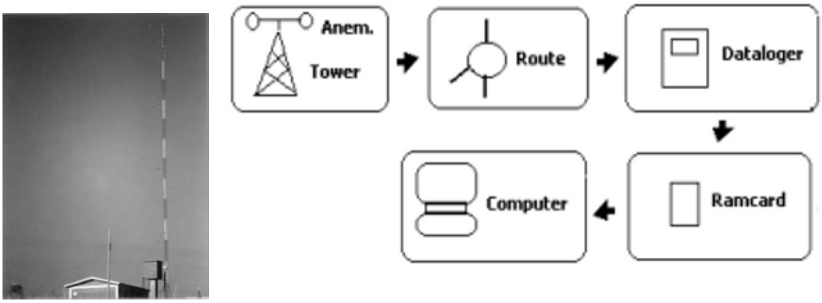


Figure 4. Anemometric tower with 40 meters and its process of data collection

5.2 Profiles Of Wind

Figures 5-11 present the different profiles of wind at 20, 30 and 40 meters, during 2000 – 2006, as an example. The principal object of research at the beginning is to obtain information about the behavior of wind and its possibility to generate energy. But the results can be used in the other fields, such as the influence of winds in buildings, thermal comfort, boundary limit atmospheric studies, weather, etc.

In the Figure 5, it shows the wind profile of the month of December of the year 2000. It's evident that the intensity of most wind data is over 4.0 m/s. This is a good signal to approve determined energetic potential. The maximum values are over 5.0 m/s and the minimum values are about 3.0 m/s.

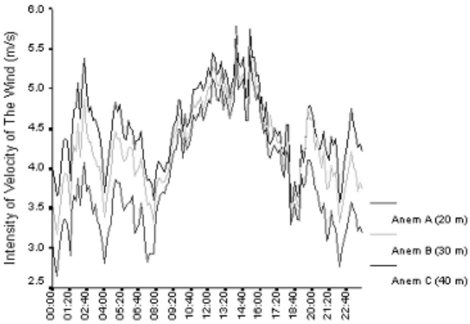


Figure 5. December – 2000

Figure 6. December – 2001

In the Figure 6, it shows the month of December of the year 2001, the situation of the winds in general, it's different than 2000, however, it also shows almost all values that are over 3.0 m/s, and most of them are over 5.0 m/s. It's interesting to observe in the figure, the periods where the wind increases. In the case of the Figure 7, evidently such increasing is earlier, between 07:00 to 10:00. After that, the wind comes down, and increases again at 12:00. All data shows that in December of 2002, the intensity is much stronger than the ones of 2000 and 2001.

Figure 7. December – 2002

Figure 8. December – 2003

In the year 2003 (Figure 8) the period of intense wind is much long, during 07:00 and 16:00. It is observed that the values of the velocity of wind get more than 5.5 m/s at altitudes of 20 and 30 meters, and over of 6.5 m/s at the altitude of 40 meters.

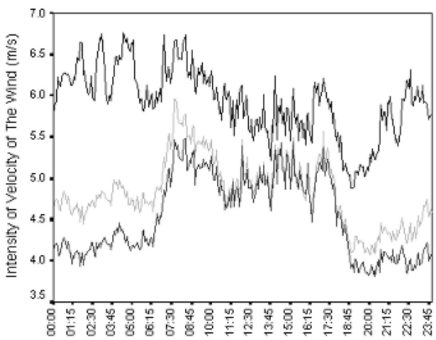


Figure 9. December – 2004

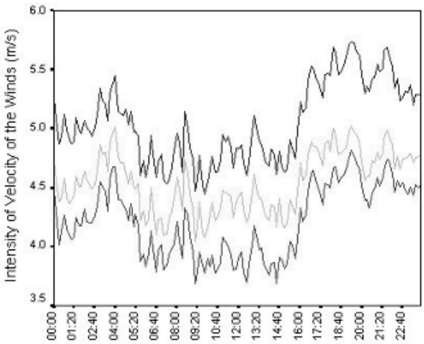


Figure 10. December – 2005

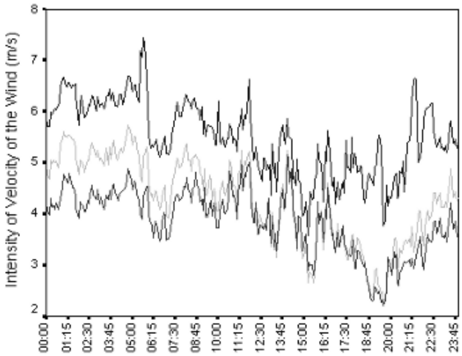


Figure 11. December – 2006

A similarly analysis can be sight in the Figure 9 (2004), Figure 10 (2006) and Figure 11 (2006). From the profiles of wind, we can see some good intensity of the wind in some periods. When the intensity of the wind increases, the energetic potential increases, too. It is also clear that the information of wind profiles can be useful in other fields, such as thermal comfort, buildings, and weather.

6 References

[1] Alé, J. V.; 1998 Wind and Solar Systems for Rural Electrification in State of Rio Grande do Sul, Brazil, ASME 1988.
[2] Burn, J., Marshall, P., & Wild, M. (1999). Managing Knowledge for Strategic Advantage in the Virtual Organization. Proceedings of the 1999 ACM SIGCPR conference on Computer personnel research, New Orleans, Louisiana, United States, April 08 - 10, 19 – 26.
[3] Gamma, E. etc.; 1994; Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wealey, Reading, MA, 1994.,
[4] Kneale, B., De Horta, A. Y., & Box, I. (2004); "VELNET (Virtual Environment for Learning Networking)". Proceedings of the sixth conference on Australian computing education - Volume 30; Dunedin, New Zealand, 161 – 168.
[5] Kumar, V., Maheshwari, B., & Kumar U. (2002); "Enterprise resource planning systems adoption process: a survey of Canadian organizations". International Journal of Production Research, Vol. 40, No. 3, 509-523.
[6] Park, K., & Kusiak, A. (2005); "Enterprise resource planning (ERP) operations support system for maintaining process integration". International Journal of Production Research, Vol. 43, No. 19, October, 3959-3982.
[7] Powell A., Piccoli, G., & Ives, B. (2004); "Virtual teams: A review of current literature and directions for future research". The DATA BASE Advances in Information Systems, Vol. 35, No. 1, Winter.

[8] Sadhu, D. P.; 1998; Wind Power Potential of Rio Grande do Sul State of Brazil, Solar & Wind Technology Vol. 5 Nº 4, pp. 405-410, 1988.
[9] Song L. & Nagiy, R. (1997); "Design and Implementation of a Virtual Information System for Agile Manufacturing". IIE Transactions, Vol. 29, No. 10, October, 839 – 857.
[10] Xu, H. Q., Ristic M., Besant, C., & Pradoux, C. (2005); "A Web-based system for manufacturing co-ordination in complex supply networks". International Journal of Production Research, Vol. 43, No. 10, May, 2049-2070.
[11] Yacktman, D. and Nghiem, A. D.; 1996; "Patterns in Web-Based Systems: Using design patterns with NeXT´s WebObjects framework". Object Magazine, July, 1996, Vol. 6, No. 4, pp 56-58.
[12] Wang, C., Xu, L., Liu, X., & Qin, X. (2005). ERP research, development and implementation in China: an overview. International Journal of Production Research, Vol. 43, No. 18, September, 3915-3932.
[13] Anderson P.; http://epix.hazard.net/vemis/
[14] http://www.smarty.net/docs.php.
[15] http://www.php.net/docs.php.
[16] http://dev.mysql.com/doc/.

Author Index