IntechOpen

Supply Chain the Way to Flat Organisation

Edited by Yanfang Huo and Fu Jia

Supply Chain,
The Way to Flat Organisation

Edited by

Yanfang Huo
and
Fu Jia

Supply Chain the Way to Flat Organisation
http://dx.doi.org/10.5772/105
Edited by Yanfang Huo and Fu Jia

First published in Croatia, 2009 by INTECH d.o.o.
eBook (PDF) Published by IN TECH d.o.o.
Place and year of publication of eBook (PDF): Rijeka, 2019.
IntechOpen is the global imprint of IN TECH d.o.o.
Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

Supply Chain the Way to Flat Organisation
Edited by Yanfang Huo and Fu Jia

p. cm.

ISBN 978-953-7619-35-0

eBook (PDF) ISBN 978-953-51-6406-7

Preface

With the ever-increasing levels of volatility in demand and more and more turbulent market competition, there is a growing recognition that individual business no longer compete as stand-alone entities but rather as supply chains. And supply chain management (SCM) has been both an emergent field of practice and an emerging academic domain to help to firms to satisfy the customer needs more responsively, with improved quality, reduction cost and higher flexibility.

According to one American professional association, SCM can be defined as a field which encompasses the planning and management of all activities involved in sourcing, procurement, conversion, and logistics management activities. Importantly, it also includes coordination and collaboration with channel partners, which can be suppliers, intermediaries, third-party service providers, and customers. In essence, SCM integrates supply and demand management within and across companies. Facing this challenge, the companies should make some fundamental changes, which involves not only the total transparency through Information sharing, but also process integration, organizational structures reengineering, and performance measures change as well. Only the organizations will win who can better structure, co-ordinate and manage the relationships with their partners in a network committed to better, closer and more agile relationship with the final customers.

Although lots of researches and practices have been devoted in this field, neither perspective is fully mature but each has considerable promise. Mainly concerned on the operation and control of supply, the book collected some latest development and findings. It consists of 20 chapters, each addressing a certain aspect of supply chain management, including the application and development ICT and the RFID technique in SCM, the new mathematical tools and techniques for SCM modeling and control, and some emerging issues in the academic research and practices of supply chain management. Each chapter gives the reader background information on a subject and proposes an original solution. This should serve as a valuable tool for professionals in this interdisciplinary field. Hopefully, readers will contribute their own discoveries and improvements, innovative ideas and concepts, as well as novel applications and business models related to the field of supply chain management. A brief introduction to each chapter is summarized in the following.

Chapter 1 is about the optimal inventory control strategy of a serial supply chain. A two-level model was suggested, one level to determine the optimal control strategy using a nonlinear integer-programming model solved by intelligent algorithms of GA, random-PSO and PEA, and the other to obtain the performance measurements of the optimized supply chain by simulation of the general push/pull model.

Chapter 2 explored a synergistic approach towards autonomic event management in supply chains aiming at improving the qualities of supply chain event management (SCEM), especially with regard to approaching self-X properties and automation. The holistic approach leveraged various computing paradigms of granular, semantic web, service-oriented, space-based, etc.

Chapter 3 proposed the use of enterprise input-output (EIO) models to represent and analyze physical and monetary flows between production processes, including logistics ones. Based on the use of EIO models, a set of complete and complementary tools able to analyze the problem according to different perspectives and point of views are presented.

Chapter 4 focused on the operational activities of the supply chain dynamics, and proposed dynamic models for multipurpose systems considering production ratios. By introducing new parameters, the models go from a simple linear supply chain based on material flows to a nonlinear one considering orders handling and based on traffic flow theory, both are applied to two simulation case studies.

Taking optimization based e-sourcing models as objective, Chapter 5 reviewed three popular e-sourcing techniques with their underlying mathematical programming models that are used to solve the winner determination problems, and presented two future directions also: global sourcing and robust sourcing.

RFID has thought to be "the first important technology of the twenty-first century", and lots of researches have been done in this field. Chapter 6 suggested a Domain Engineering Process for RFID Systems Development in Supply Chain, which defines a systematic process to perform domain engineering which includes the steps of domain analysis, domain design, and domain implementation.

In Chapter 7, some issues of the return policies and collaboration in supply chain are reviewed, involving an overview of the benefits and costs of returns policies, the different kinds of returns policies that are required to coordinate the supply chain for the different types of products and the impact of demand uncertainty and retailing competition on returns policies.

Chapter 8 focuses on the problem of managing at the operational level supply chains, and described it by a modular model based on the first order hybrid Petri net formalism, which can effectively describe the operational management policies and the inventory control rules, and enables the designer to impose an optimal SC dynamics according to given objective functions.

Chapter 9 described a statistical physics approach to understanding the supply chain oscillations, models of both normal modes and External interventions are demonstrate by inventory oscillations model and a fluid-flow model separately. It is the first time that the general approach together with its applications has been assembled in one place, along with a number of possible extensions.

Information Technique is one of the most enablers of SCM development. In Chapter 10, a framework that enhances the agility of SCM with IT is presented.

A simulation known as The Trading Agent Competition: Supply Chain Management Game (TAC SCM) was sponsored by a group of universities and research centers to compete against each other to prove mechanisms for supply chain situations since 2003. Chapter 11 presented the deepest analysis about the construction of the Tiancalli agents since 2005, intending to describe the effort and experience during the three years participating on TAC SCM.

Chapter 12, investigating the methods of supply chain integration for manufacturing industry in the background of China, proposed a three-echelon theoretical framework for supply chain integration based on Thorn's model and presented the relative key techniques of each level.

Referring to develop integrated supply chains significantly more flexible, responsive and agile than traditional supply chains, Chapter 13 discussed two new approaches-Dynamic Agility Index and Fuzzy Association Rule Mining - for modeling and evaluating agility in dynamic integrated supply chains.

Chapter 14 described the design and implementation of the MIDAS supply chain system by Using Web Services and the Service Oriented Architecture.

In Chapter 15, a distributed supply chain planning system for multiple companies with Limited Local Information using an augmented Lagrangian relaxation method has been proposed.

While by introducing Fuzzy mixed integer Linear Programming to tactical Supply Chain Planning, a multi-echelon, multi-product, multi-level and multi-period SC planning model was established in Chapter 16, given lack of knowledge (demand, process and supply uncertainties).

Chapter 17 explored the research issues on collaborative product design and development based on CM principles, which then be introduced in four areas separately-configuration identification, configuration change control, configuration status accounting, and configuration audits.

Chapter 18 illustrated the RFID and EPC potential for business processes and presented RFID@B2B, a new method to improve the supply chain performances using RFID technology.

Chapter 19, focusing on the issues and potential solutions for a range of security vulnerabilities of RFID systems, analyzed the underlying vulnerabilities that exist in RFID systems, illustrated the threats of possible attacks, and provided corresponding countermeasures.

In Chapter 20, aiming at the problems in traditional knowledge retrieval, an approach is put forward to supply chain knowledge management construction by introducing ontology, which consists of construction of domain ontology, formalization of ontology model, and development of supply chain knowledge management system based on ontology.

We would like to thank all the authors for their excellent contributions in the different areas of supply chain management. It is their knowledge and enthusiastic collaboration that lead to the creation of this book, which we are sure that will be very valuable to the readers.

December 2008

Editors

Yanfang Huo
School of Management, Tianjin University
P.R. China

Fu Jia
Cranfield School of Management, Cranfield University
UK

Contents

1

Optimal Control Strategy for Serial Supply Chain

Min Huang[1,2], W.H.IP[3], Xingwei Wang[1] and Jianqing Ding[1]
[1] College of Information Science and Engineering, Northeastern University, Shenyang, Liaoning
[2]Key Laboratory of Integrated Automation of Process Industry, Northeastern University, Shenyang, Liaoning,
[3]Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon,
[1,2]P.R.China
[3]Hong Kong

1. Introduction

With the emerging of global economy and the development of the technology in computer and communications, the enterprises are facing to new opportunities but also more challenges, which led to the concept of SC (Supply Chain). By controlling and collaborating each part of the supply chain, SCM (Supply Chain Management) reaches the aim of reducing the cost, improving the quality as well as service level, and so on, further enhances the integrated competition ability of the whole supply chain. SCM includes many aspects of the management activity, and the inventory management is one of the key aspects among them.

In this chapter, the research focuses on inventory control strategy optimization related to activities, such as the purchase, production, storage and transportation of material, work in process and finished goods inventory within up-stream and down-stream enterprises along the serial supply chain. The main research aspects are as follows:

1. To review the state of art of inventory control strategies of supply chain.
2. Supply chain inventory control strategy optimization is based on the simulation of supply chain inventory system. So, in this chapter, the general model of serial supply chain inventory control strategy is addressed.
3. Model for single objective control strategy optimization is established, which describes the optimization problem of serial supply chain inventory control strategy with the objective of minimum cost and the constraint of the customer service level and average input standard deviation.
4. The algorithm of GA (Genetic Algorithm), Random-PSO (Particle Swarm Optimization) and PEA (Pheromone Evolutionary Algorithm) are designed for the model. Simulation studies suggested that each of the algorithms can solve this problem efficient, and Random –PSO algorithm is most efficient one.

2. The general control model for serial supply chain

2.1 The different control strategies for serial supply chain

Under globalization and the rapid development of computers and information technology, all enterprises face new chances as well as more challenges. This breeds the concept of a serial supply chain, a value-added chain that is composed of a series of enterprises: raw material suppliers, parts suppliers, producers, distributors, retailers, and transportation enterprisers. Clients finally get their products, which are manufactured and handled systematically by the enterprises of the chain, started from either the raw material suppliers or the parts suppliers. This series of activities are the total activities of a complete serial supply chain, that is, from the supplier's suppliers to the clients' clients [1, 2].

SCM aims at decreasing the system cost, increasing the product quality, and improving service level by collaborating and controlling the conduct of each entities of the supply chain. The goal is to upgrade the overall competitive ability of the whole system. Hence, the inventory management of the serial supply chain is important. The inventory control strategy of an enterprise affects the cost and the revenue indirectly. Therefore, the target of an optimal inventory is both to maximize the degree of clients' satisfaction and to minimize the overall cost [3].

Inventory decouples the supplying, producing, and selling processes of an enterprise. Each process operates independently. This helps to reduce the effect that comes from the variation of demand forecast, and makes good use of resources when variation happened due to demand changes and market changes. On the other hand, capital is needed for setting up an inventory. The cost includes the capital used for inventory and products in process, the space used for inventory, the expenses on management, maintenance, and discarding of defected products. Inappropriate inventory management even affects the operation efficacy of the enterprise.

The characteristic of uncertainty of a supply chain increases the overall inventory of the whole chain system. It also brings unnecessary cost to the node enterprises of the supply chain. In order to avoid the "bullwhip effect" caused by the uncertainty of demand and supply, the traditional inventory strategy has to be revised. Inventory control of a supply chain can be improved by strategies like shared technology, contract system, and integrated enterprises. Thus, the competitive ability of a supply chain is enhanced.

Generally speaking, there are two kinds of production inventory systems: the push and the pull system. The current worldly popular production inventory control systems of Manufacturing Resource Planning (MRPII) and Just-in-time (JIT) belong to the system, respectively. The push production control system adopts a central control method and organizes production by forecasting the future demand. Therefore, production lead time is estimated in advance. The pull production control system adopts a distributed control method and production is organized according to the real demand [4]. Each method has its own advantages [5, 6]. Peoples try to combine the two methods to attain better performances [7-9], CONWIP (CONstant Work in Process), proposed by Spearman et al. in 1990, is an example of combined push/pull control method [10]. In 2001, Gaury et al. proposed a methodology to customize pull control systems[11]. In 2003, Ovalle and Marquez suggested the model of CONWIP control system for a serial supply chain and also shows the corresponding simulation analysis [12]. But, up till this moment, all researches on control system of supply chain only deal with single specific control strategy like the push system, the pull system, the classic combined push/pull CONWIP control system[10,13-14], or the simple combination of push and pull system. A generally used model of push/pull control strategy and its research is still absent. The aim of the research is to establish a generally used method of the inventory system of a serial supply chain to replace those traditional classic control systems.

In the systems of Kanban and CONWIP, system performances rely on the card quantities. Similarly, in the serial push/pull system, distribution of circulating cards (the number of circulating cards at different stages) determines the control model, guides the production time and production quantity of the generally used system. In the push system, the card numbers between each pair of nodes on the SC is infinite. Therefore, determination of circulating cards becomes a key factor affecting the operating efficacy of a generally used inventory system of a serial supply chain.

This chapter proposes an optimal control model that tackles a series of multi-stages of inventory control system of a supply chain. The model is based on the combination of nonlinear integer programming and the generally used push/pull system of the inventory of a serial supply chain. It determines the distribution of circulating cards by integrating the intelligent algorithms and simulation analysis. Both the case studies have proved that the results from intelligent algorithms are reasonable and effective.

2.2 The general control model for serial supply chain

Figure 1 shows that there are n nodes on the whole serial supply chains. Each node represents upstream/downstream node enterprises like raw material suppliers, manufacturers, distributors, retailers, and clients. Since the final target of a supply chain is to satisfy the clients' demands, each enterprise operates its production and sales under the generally used push/pull inventory system control. That is, production of each node enterprise is affected by the raw material supply of the upstream enterprise and the demand of the downstream enterprise. One important goal of the supply chain is to reach a win-win status, to maximize the profits of the whole chain instead of any individual enterprise. In order to control the quantity of products in process of the chain, there is a fixed product standard on the feedback for the demand on upstream enterprise i from downstream enterprise j and marks it as card number K_{ij}^{u}. Only when the real quantity of products in

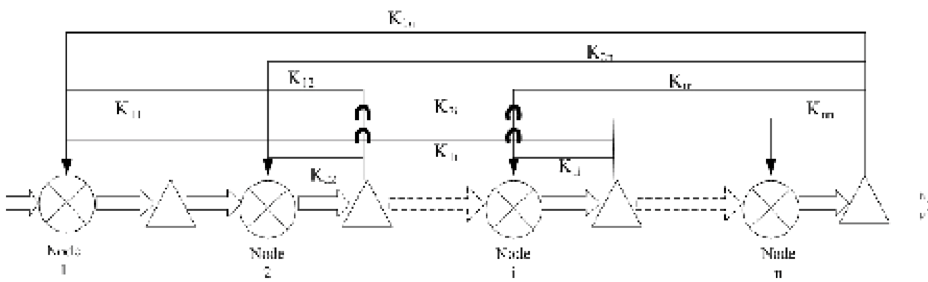


Fig. 1. The general control model for serial supply chain

process of node enterprise i is less than the forecast product quantity of each downstream enterprise, then it is allowed to proceed with the manufacture. Once the product of a unit is allowed to be manufactured by node enterprise i, the node's free card is attached to its manufactured container, then the product has finished processing it is sent to the next node enterprise i+1. The attached circulating card is detached and returns to node enterprise i as a free card and authorizes further manufactures of other products. When the value of K_{ij}^{u} is ∞, it means that there is no feedback control on upstream node *i* from downstream node *j*. If there is no feedback control on node *i* from all downstream nodes, then node *i* is under the push control.

The commonly used push/pull system has the following properties and assumptions:

1. Clients' demands satisfy the normal distribution with upper and lower bound.
2. The supply chain produces only one kind of product.
3. All upstream enterprises can obtain the return cards from any downstream manufacturing enterprise node without delay (no return delay).
4. There are sufficient materials for the initial node enterprise of the supply chain.

Description of the Control Strategy

In order to describe the control strategy, two definitions are needed.

Definition 1:

The card number matrix K

$$K=\begin{bmatrix} K_{11}K_{12}.........K_{1n} \\ K_{22}......K_{2n} \\ . \\ . \quad .. \quad ... \quad K_{nn} \end{bmatrix} \tag{1}$$

The element K_{ij}^{u} of matrix K represents the card number in the control cycle sent by downstream node *j* to upstream node *i*. $K_{ij}^{u} = \infty$ implies that there is no pull control from node j to node *i*. Since all nodes are under the control of their downstream nodes, so the lower triangular matrix of K is meaningless. Thus, values of these elements are fixed as –1. So, we have equation (2):

$$K_{ij} = \begin{cases} -1 & (i > j) \\ 1 \sim K_{ij}^{u} + 1 & (i \le j) \end{cases} \tag{2}$$

where K_{ij}^{u} represents the upper limit of the card number, K_{ij}^{u} + 1 represents +∞.

Definition 2:

Control matrix M: matrix that describes the control model.

$$M=\begin{bmatrix} M_{11}M_{12}......M_{1n} \\ M_{21}M_{22}......M_{2n} \\ . \\ . \\ M_{n1}M_{n2}......M_{nn} \end{bmatrix} \tag{3}$$

Matrix M in (3) shows what kind of control strategy each node of the supply chain has. If the element M_{ij} of matrix M equals to 1, it means there is pull control on upstream node *i* from downstream node *j*. If the element M_{ij} of matrix M equals to 0, it means that there is no pull control on upstream *i* from downstream node *j*. So we have equation (4):

$$M_{ij} = \begin{cases} 1 & (K_{ij} \neq \infty) \\ 0 & (K_{ij} = \infty) \end{cases} \quad i \leq j \tag{4}$$

Property 1: When the sum of all elements of row *i* of matrix M equals to 0, there is no pull control on node *i* from all its downstream nodes, which means that node *i* is under the push control from its upstream node.

2.3 The method for determine the optimal control strategy

The main problem of determining the optimal control strategy is how to choose an appropriate control strategy so that some goals of the supply chain are reached while some constraints are satisfied. A two-level model are presented to cope with this problem. The first-level model is a mathematic programming model, which optimizes the distribution of circulating cards by guaranteeing the goals of the supply chain and satisfying the certain constraints.

The second-level model is the general control model for serial supply chain which is used for the analysis of the inventory system of the whole serial supply chain. Definitions of the variables for analysis is given here and the relationship among the variables is illustrated in figure 2.

(1) Logistics variables

P_t^i quantity of products in process of node *i* of period t

Y_t^i product inventory of node *i* of period t

S_t^i transportation quantity from node *i* to *i* +1 of period t

O_t^i exported quantity of product of node *i* of period t

I_t^i imported quantity of raw material of node *i* of period t

X_t^i the real usable quantity of raw material of node *i* of period t

(2) Technology flow variables

OP_t^{cust} clients' demands of period t

D_t^i quantity of inflow orders of node *i* of period t

APC_t^i number of available cards of node *i* of period t

OP_t^i quantity of processed orders (manufacturing quantity) of node *i* of period t

B_t^n quantity of accumulated orders of node n of period t

DS_t^i the expected quantity of transportation of node i to *i* +1 of period t

TY_t^i quantity of available overall product of node *i* of period t

L^i period of manufacturing products of node *i*

Ls^i period of transporting products of node *i*

MLP^i product-load ability of node *i*

UC^i vessel capacity of node *i*

K matrix of circulating cards, element K_{ij} represents the number of cards in the control cycle of node *i* sent by node *j*

M the control matrix mark *i*, it is the last node which exerts the pull control on node *i*

(3) Capital flow variables

CR_t^i quantity of cash demand of node *i* of period t

ICR_t^i inventory value of node *i* of period t

R_t^i compulsory cash-in of node *i* of period t

Py_t^i compulsory payment of node *i* of period t

Pm_t^i product price of transportation unit of node *i* of period t

$Pwip_t^i$ price of products in process unit of node *i* of period t

mr_t^i margin benefits of node *i* of period t

$CumP_t^i$ accumulative benefits of node *i* of period t

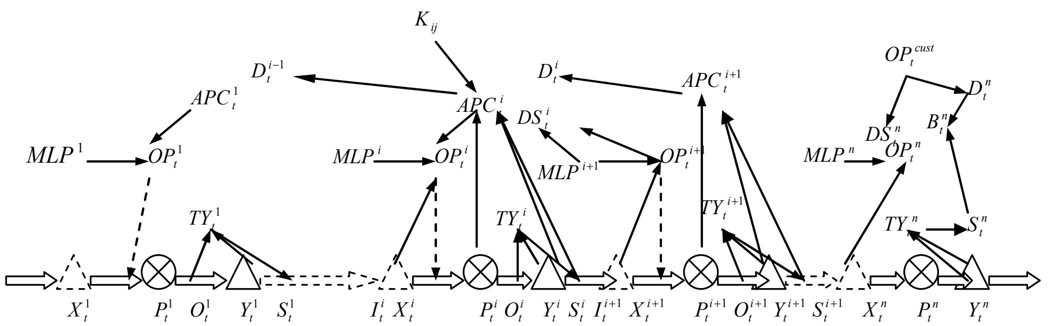


Fig. 2. Illustration of the relationship among the variables

In order to analysis the supply chain, the following performance measurements are defined according to the above relationship among the variables:

1. Service level (%): (the customer satisfy percentage of the supply chain)
   The satisfy percentage of the last node of the supply chain is considered.

$$S_l = \frac{100 * \left(\sum_{t=1}^{T} D_t^{\mathrm{n}} - B_T^n \right)}{\sum_{t=1}^{T} D_t^n} \tag{5}$$

2. Standard deviation of inputs:

$$SDO = \sqrt{\frac{1}{T-1} \sum_{t=1}^{T} \left(OP_t^1 - \sum_{t=1}^{T} OP_t^1 / T \right)^2} \tag{6}$$

3. Overall cost of the supply chain:

$$C = \frac{\sum_{t=1}^{T} \sum_{i=1}^{n} CR_t^i}{T} \tag{7}$$

The parameters delivered from the first-level model to the second-level model are circulating card distributions of each node. The parameters delivered from the second-level model to the first-level model are the service level, standard deviation of the input, and average overall cost of the supply chain.

3. The model and algorithm for optimal control strategy problem

3.1 The model for optimal control strategy problem

Usually, the goal of the supply chain operation is to minimize the total cost with the constraint of service level and input standard deviation. Under this circumstance, the first-level model is a nonlinear integer programming model. It optimizes the distribution of circulating cards by guaranteeing that the average overall cost of the supply chain of the system is minimized, under the constraints of service level S_{l_0} and the input standard deviation SDO_0. In another word, solve upper triangular elements K_{ij} of matrix K. So, the first-level model is given as follows:

$$\min C(K) \tag{8}$$

$$st: \quad \begin{cases} S_l(K) \geq S_{l_0} \\ SDO(K) \leq SDO_0 \end{cases} \tag{9}$$

$$K_{ij} \text{ is the integer between 1 and } K_{ij}^u + 1 \text{ where } i = 1, \cdots, n, j \geq i \tag{11}$$

where K_{ij} represents the upper bound of the card number and K_{ij} +1 represents +∞.

The first-level model is a mixture of combination problem and integer programming. If there are n node enterprises, the card number that needed to be determined will be $n(n+1)/2$. When node enterprise *i* is under the control of node enterprise *j* in the supply chain and the upper bound of card number is K_{ij}^u, there are $\prod_{i=1}\prod_{j\geq i}\left(K_{ij}^u + 1\right)$ states of searching spaces when the constraints are not considered. To process each state is possible only if both the scale of the supply chain and the cards number are small. Heuristic algorithms are needed to solve practical problems. So intelligent algorithm is applied to solve the first-level model to determine k. The GA, Random-PSO and PEA are used in this research.

A simulation is used in the second-level according to the general control model for serial supply chain to determine the performance measurements of the supply chain system under a given card distribution *K*, because the performance measurements of the supply chain, the cost, the service level and standard deviation of input are implicit functions of card distribution $k_{ij}\left(\forall i = 1, \cdots, n, j \geq i\right)$.

3.2 The intelligent algorithms for optimal control strategy problem

In this section, three intelligent algorithms, GA, Random-PSO and PEA are used designed for the model (8)-(11). The performance of them and the comparison among them are given.

3.2.1 Genetic algorithm

This section give the design and analysis of GA for the above model (8)-(11) [15,16].

3.2.1.1 Coding

Integer coding is adopted considering the characteristic of the problem. Each bite represents the element of the upper triangular part of matrix K accordingly, and there are $n(n+1)/2$ bites in total. Figure 3 is the illustration of coding.

K_{11}	K_{12}	K_{13}	K_{14}	…	……	……..	K_{nn}

Fig. 3. The illustration of coding

The range of each bite is from 1 to $K_{ij}^{u}+1$ where the value of K_{ij}^{u} is determined by the ability of each production node. For instance, if the production ability of the first node of a supply chain is 30, $K_{ij}^{u} \geq 30$. If the production ability of the second node is 20, K_{ij}^{u} will contain the control cycles of the two nodes, that is, $k_{12}^{u} \geq$ (30 + 20). Let $K_{ij}^{u}+1$ of a certain node as an infinite integer means that its downstream nodes have no limit of card number on it.

Property 2: According to this coding rule, there are different upper and lower bound for each bit of individual as each bit may include different number of nodes.

3.2.1.2 Fitness Function

Due to the minimizing property of the objective function, fitness functions. It is obtained by equation (12) as follows:

$$F(K) = f_{\max} - f(K) \tag{12}$$

$$f(K) = C(K) + \alpha_1 * [SDO(K) - SDO_0]^{+} + \alpha_2 * [S_{l_0} - S_l(K)]^{+} \tag{13}$$

where $\alpha_1 * [SDO(K) - SDO_0] + \alpha_2 * [S_{l0} - S_l(K)]$ are the penalties for not satisfying constraints of (9) and (10), α_1 and α_2 are penalty coefficients, and $f_{\max}$ is a given large value to guarantee that the overall fitness value is non-negative, and

$$[y]^{+} = \begin{cases} y & y > 0 \\ 0 & otherwise \end{cases}$$

3.2.1.3 Operators design

Each initial solution is obtained by creating an integer within the range of 1 to $K_{ij}^{u}+1$ randomly for each bit.

Two points crossover is adopted here. Two intersecting points are chosen from the chromosomes. Crossover is taken in the space between the two intersecting points and the rest is still inheriting the parent genes.

Mutation is also applied. First, create a number between 0-1 randomly. When the number created is smaller than the mutation probability, mutation will happen by creating an integer that lies within the limits of the circulating cards randomly.

The commonly used roulette mechanism is used as the choice strategy and the biggest iteration number is chosen as the criterion for algorithm termination.

3.2.1.4 The Elitist Mechanism

In order to maintain the best chromosome of each generation, an elitist mechanism is used in the choice process. If the best chromosome of the last generation is not duplicated into the next generation, the next generation will randomly delete a chromosome so that the best chromosome of the last generation will be duplicated directly.

3.2.1.5 Numerical Analysis

In order to analyze the performance of the algorithm, the problems with 4-nodes (10^{18}), 6 nodes (10^{40}) are analyzed here..

Example 1 is a supply chain with 4 nodes, the ability of each node enterprise is same, the custom demand is a normal distribution which mean of 4.0 and variance of 1.0, upper and lower bound are 8.0 and 0.0, the parameters are shown in table 1. The supply chain should ensure that the customer service levele is not less than 90% and the input standard deviation is less than 2.5.

Node enterprise	1	2	3	4
Quantity of products in process P_0^i	8	8	8	8
Product inventory Y_0^i	12	12	12	12
Transportation quantity S_0^i	4	4	4	4
Exported quantity of product O_0^i	8	8	8	8
Imported quantity of raw material I_0^i	8	8	8	8
The real usable quantity of raw material X_0^i	9	9	9	9
Quantity of processed orders OP_0^i	3	3	3	3
Period of manufacturing products L^i	1	1	1	1
Period of transporting products Ls^i	2	2	2	2
Product-load ability MLP^i	30	30	30	30
Container capacity UC^i	1	1	1	1
Margin benefits mr_t^i	1	1	1	1
Product price of transportation unit Pm_t^i	2	2	2	2
Price of products in process $Pwip_0^i$	1	1	1	1

Table 1. The initial value of parameters for each node enterprise in 4 nodes problem

According to the parameter set in table 1, the upper-lower bound of card numbers for different control segments are shown in table 2. So the size of this problem is $41^4 * 71^3 * 106^2 * 141 = 1.602 \times 10^{18}$, while the constraint is not considered.

The number of nodes	1	2	3	4
Lower and upper bound	1~41	1~71	1~106	1~141

Table 2. The lower and upper bound of card for 4 nodes problem

To analyze the performance of the algorithm, the algorithm is run for 100 times. The best solution is the best one within 100 runs. The best rate is the rate to reach the best value within 100 runs.
Taking reasonable parameter of GA, the best solution K=[∞, ∞, ∞, ∞, 16, 29, 43, 14, 31, 16] for 4 nodes problem is obtained.
For the supply chain with 6 nodes enterprises, the parameters are shown in table 3, the upper-lower bound of card number for different problem are shown in table 4.

Node enterprise	1	2	3	4	5	6
Quantity of products in process P_0^i	8	8	8	8	8	8
Product inventory Y_0^i	12	12	12	12	12	12
Transportation quantity S_0^i	4	4	4	4	4	4
Exported quantity of product O_0^i	8	8	8	8	8	8
Imported quantity of raw material I_0^i	8	8	8	8	8	8
The real usable quantity of raw material X_0^i	9	9	9	9	9	9
Quantity of processed orders OP_0^i	3	3	3	3	3	3
Period of manufacturing products L^i	1	1	1	1	1	1
Period of transporting products Ls^i	2	2	2	2	2	2
Product-load ability MLP^i	30	30	30	30	30	30
Container capacity UC^i	1	1	1	1	1	1
Margin benefits mr_t^i	1	1	1	1	1	1
Product price of transportation unit Pm_t^i	2	2	2	2	2	2
Price of products in process $Pwip_0^i$	1	1	1	1	1	1

Table 3. The initial value of parameters for each node enterprise in 6 nodes problem

The number of nodes	1	2	3	4	5	6
Lower and upper bound	1~41	1~71	1~106	1~141	1~175	1~210

Table 4. The lower and upper bound of card for 6 nodes problem

The custom demand is a normal distribution with mean of 4.0, variance of 1.0, upper-lower bound of 8.0 and 0.0, constraints condition are service level no less than 90% and the input standard deviation less than 2.5.

The scale of the problem is $41^6 * 71^5 * 106^4 * 141^3 * 175^2 * 210 = 1.951 \times 10^{40}$ while the constraint is not considered. Taking reasonable population size and iterative number, the best solution K=[∞, ∞, ∞, ∞, ∞, ∞, ∞, 62, 93, 104, 170, 21, 63, 95, ∞, 23, 65, ∞, 25, ∞, ∞] for 6 nodes problem is obtained.

The comparison of these two scales of problems is shown in table 5 taking reasonable parameters of NP(Number of Populations), NG(Number of Generations), PC(Probability of Crossover), PM(Probability of Mutation) obtained by simulation.

Scale of the problem	NP	NG	PC	PM	Best fitness	Best rate	T(s)
10^{18} (4- node)	200	150	1.0	0.3	892611.49	0.92	21
10^{40} (6-ode)	200	150	1.0	0.3	890379.75	0.86	39

Table 5. Comparison of the results of different scale of problems for GA

Table 5 showed that, after the scale of the 6-node problem has expanded by 10^{22} times as compared with a 4-node problem, the CPU time is increased by 18S and the best rate is decreased by 6%. Though there are expanded complexities of the problem, the algorithm still possesses high best rate. The increase in the CPU time is within an acceptable range. Moreover, time increase is mainly caused by the influences of the expansion of scale of the problem on simulation. Thus, the performance of the algorithm is not greatly affected by the scale of the problem and it is still a fairly stable algorithm.

3.2.2 Random-PSO algorithm

This section shows the solutions of the above model for single objective optimal control strategy problem by Random-PSO algorithm.

PSO algorithm is bring forward by Eberhart and Kennedy in 1995[17, 18]. Originally, PSO algorithm was proposed to simulate the movement of bird swarm. People observed animal society behavior, found that in a group information share was propitious to evolvement [19], that is the basic of PSO algorithm.

PSO is based on group intelligence, its unit is the swarm, then establish simple rules for each unit, so that the whole swarm could have complex characters for solving complex optimal problems. Because of the simple concept and easily to realize, PSO develop quickly in short time, soon recognized by international evolvement calculation field, and applied in many field like electric power optimization, TSP optimization, neural networks training, digital circuit optimization, function optimization, traffic accident exploration, parameter identification.

Classical PSO optimal algorithm described as follow:
Suppose the search space is D dimension, the position of the ith particle in particle swarm is expressed as $X_i = (X_{i1,} X_{i2,....} X_{iD})$, the speed of the ith particle is expressed as $V_i = (V_{i1,} V_{i2,......} V_{iD})$, the best position of ith particle searched so far is denoted as $P_i = (P_{i1,} P_{i2,......} P_{iD})$, the best position of the whole swarm have searched so far is denoted as $P_g = (P_{g1,} P_{g2,......} P_{gD})$. For every particle, the d dimension $(l \le d \le D)$ changes according to the equation as follow [20]:

$$V_{id} = wV_{id} + c_1 r_1 (P_{id} - X_{id}) + c_2 r_2 (P_{gd} - X_{id}) \tag{14}$$

In equation (14), V_{id} denotes the speed of the ith particle at d dimension, here: w is inertia weight, c1 and c2 are acceleration constants, r1 and r2 are random number in [0, 1] used to adjust the relative importance of P_{id} and P_{gd}, so that could obtain the next movement position of the particle:

$$X_{id} = X_{id} + V_{id} \tag{15}$$

The first part of the equation (14) is the former speed of the particle; the second part is "cognition", express the think of the particle; the third part is "social", express the information share and the cooperation between the particle [21]. "Cognition" part is explained by "law of effect" of Thorndike[22]. It is a fortified random action will possibly appear in future. The action here is "cognition", and we suppose that getting correct knowledge is enhanced, this model supposes that the particle is inspirited to reduce deviation. "Social" part is explained by vicarious fortified of Bandura[23]. According to the anticipation of this theory, when the observer observe a model intensifying an action that will increase the probability of this action coming, that means the particle 's cognition will be imitated by other particle. According to equation (14) and (15) to iterate, finally obtain the optimum solution of the problem.

3.2.2.1 Coding

Integer coding is adopted according to the characteristic of the problem as show in section 3.2.1.1.
For describing the problem easily, we change the coding into string, the unit of solution is denoted as: $X = (x_1, x_2, ..., x_m)$, here m is the length of the string, m=n(n+1)/2; every element x_i in vector X correspond to the element of the upper triangular matrix of K, $x_1 = K_{11}$, $x_2 = K_{12}$, $x_3 = K_{22}$, $\cdots x_i = K_{hl}$, $x_m = K_{nn} (h \le l)$ ·, $x_i \in V_i$, $V_i = (x_{i1}, x_{i2}, \cdots x_{ik}) = \left[1, K_{hl}^u + 1\right]$. V_i is the space of the i th bit of gene, k_i denote the size of this space.

3.2.2.2 Fitness Function

Due to the minimizing property of the objective function, fitness function is defined as the one in section 3.2.1.2.

3.2.1.3 Random-PSO algorithm design

Classical PSO algorithm is an effective method for searching continuous function extreme, but the research in discrete field is few. In 1997, Kennedy, Eberhart proposed "a discrete binary version of the particle swarm algorithm", namely PSO-SV algorithm, it used to solve binary space optimal problems, that first start to utilize PSO to solve the discrete problems [21,24]. This method can only solve the binary space optimal problems, though the performance of the algorithm is excellent, its application area is restricted, for many-dimensions discrete space optimal problems, it is nail-biting. Take the problem's particularity of account, here adopt a Random-PSO algorithm[25], and use it to solve the combinatorial optimization of actual circulating cards, which is the fixing of the circulating card number in every node enterprise of supply chain inventory control strategy. The standby card number's range of every unit of the solution constitute the local search space, the global search space is consisted by K_{ij}^{u} +1 local search space (K_{ij}^{u} is the digit capacity of the solution), that is all the standby card number's range constitute global search space. The structure of solution of the problem is shown in figure 4[25].

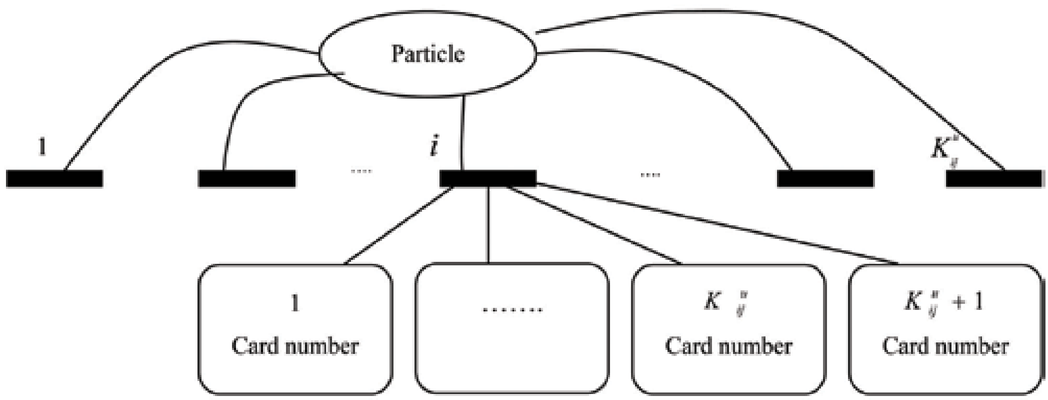


Fig. 4. Structural diagram of solution space

Every particle denotes the whole solution of the problem. The solution is consisted of three levels, first level is particle level, second level is every unit of solution, third level is card number; the card number of every unit constitute a local search space, the particle firstly search in the local space, choose a card number for every unit, then the card number at all unit constitute a solution. It is easy to see that the card number at all units constitute the global search space.

The speed and position of the particle update as follow[25]:

$$V_{id_j} = r\overset{\bullet}{*}V_{id_j} + r_1\overset{\bullet}{*}(P_{id} - X_{id_j}) + r_2\overset{\bullet}{*}(P_{gd} - X_{id_j})$$

$$= \begin{cases} r\overset{\bullet}{*}V_{id_j} + (p_{id} - X_{id_j}) + (p_{gd} - X_{id_j}) & r_1 > 0.5, r_2 > 0.5 \\ r\overset{\bullet}{*}V_{id_j} + (p_{id} - X_{id_j}) & r_1 > 0.5, r_2 \le 0.5 \\ r\overset{\bullet}{*}V_{id_j} + (p_{gd} - X_{id_j}) & r_1 \le 0.5, r_2 > 0.5 \\ r\overset{\bullet}{*}V_{id_j} & r_1 \le 0.5, r_2 \le 0.5 \end{cases} \quad (16)$$

$$X_{id_j} = \begin{cases} 0 & X_{id_j} < 0 \\ J_d & X_{id_j} > J_d \\ X_{id_j} + V_{id_j} & 0 \le X_{id_j} \le J_d \end{cases} \tag{17}$$

$$num_{id_j} = \begin{cases} num_{id_j} + 1 & X_{id_j} = j \\ num_{id_j} & X_{id_j} \ne j \end{cases} \tag{18}$$

$$F_{id_j} = \frac{num_{id_j}}{J_d + 1} \tag{19}$$

The normalization of F_{id_j} :

$$sum_{id} = \sum_{j=0}^{J_d} F_{id_j} \tag{20}$$

$$P_{id_j} = \frac{F_{id_j}}{sum_{id}} \tag{21}$$

To generate a random number in 0-1 for every unit of every particle, denote as rand,

$$X_{id} = \arg_j (P_{id_j} < rand \le P_{id_{j+1}}) \tag{22}$$

X_{id} is the code of the circulating card chosen by the *d*th unit of the *i*th particle. Here "$\dot{*}$" is different from the normal product, it is a binocular operator, two parts of their operands can not reverse, the former part is a random number control the effect of the other one which is a integer: $V_{id} \in [-2J_d, 2J_d]$, P_{id} P_{gd} $X_{id_j} \in \{0,...J_d\}$; $r \in (0,1)$ is inertia factor, used to adjust the speed, $r_1, r_2 \in (0,1)$ are random numbers, used to adjust the extreme of particle and the global extreme.; num_{id_j} note the times of card number which is j at the *d*th unit of the *i*th particle, P_{id_j} is the frequency of card number which is j at the *d*th unit of the *i*th particle, the probability is bigger as this value for the card number being j.

3.2.1.4 The Procedure for Random-PSO

The main procedure for Random-PSO is as follows[25]:

Step1: NC ← 0 (NC is iterative number)

To produce a random number j from every unit of particle i, $j \in \{0, \ldots J_i\}$, X_{id_j}=j constitute the initial position of the particle, equation (19)-(20) produce the initial solution, assign this value to particle extreme, take the better one as the global extreme; Let the initial speed v=0;

Step 2: if get the maximal iterative number, go to step 7, else go to step 3;

Step 3: use the control matrix K as the parameter to call the simulation, obtain three economic indexes, and calculate the fitness value.

Step 4: compare the currently particle fitness value and the particle extreme for every particle i, if the currently particle fitness value is better, then update P_{id};

Step 5: compare the currently particle fitness value and the global extreme, if the currently particle fitness value is better, then update P_{gd};

Step 6: update the V_{id_j} and X_{id_j} follow equation (16) and (17), and produce a new solution from equation (18)-(22), go to step 2;

Step 7: output the optimal objective function value and the card combination.

3.2.1.5 Numerical Analysis

In order to test the efficiency of the random-PSO algorithm, two problems in section 3.2.1.5 is used here. The comparison of the two problems is shown in table 6 taking reasonable parameters of NP and NG obtained by simulation.

Problem scale	NP	NC	Best fitness	Best rate	T(s)
10^{18} (4 nodes)	150	100	892611.49	0.94	13
10^{40} (6 nodes)	150	150	890379.75	0.90	54

Table 6. Comparison of the results of different scale of problems for random_PSO

Table 6 showed that, after the scale of the 6-node problem has expanded by 10^{22} times as compared with a 4-node problem, the CPU time is increased by 41s and the optimization percentage is decreased by 4%. Though there is the expanded complexity of the problem, the algorithm still possesses a better optimization percentage. The increase in the CPU time is within an acceptable range. Moreover, time increase is mainly caused by the influences of the expansion of scale of the problem on simulation. Thus, conclusively speaking, the optimal performance of the algorithm is not greatly affected by the scale increase of the problem.

3.3.3 Pheromone evolutionary algorithm

This section shows the solutions of the above model for single objective optimal control strategy problem by PEA.

Huang et al.[26] propose the evolutionary algorithm based on pheromone. It is a process of probability choices by making use of pheromone, which is one of the important concepts in the algorithm of ant system[27]. In the ant system, pheromone is evenly distributed in the solution space, and through the positive feedback of pheromone, the algorithm is converged to the optimal point of the whole set. In the discrete problem of evolutionary algorithm, the mutation operators every time undergo only one or several emergence in gene positions.

These mutation operators are lack of direction and the mutation is not even. Therefore the ability for searching optimization is weak. If pheromone is induced into the mutation operators in every mutation, the mutation operators will be more directed by making use of the changes in pheromone, which is strongly related to the fitness value. These changing pheromone guides the mutation of each gene position. Thus the mutation operators raise greatly their abilities for searching optimization. With the help of the mutation operators, the algorithm can steadily converged to the optimal of the whole search space.

The evolutionary algorithm combines the concept of pheromone and the mutation operators of genetic algorithm. The probability field represents pheromone. A series of code that represents the solution of the problem is used to represent the gene series. Same probability is given to each bit of the gene in the series for initiation. Then another series is created randomly from these probabilities and their fitness values are calculated. These fitness values help to adjust the distribution of chosen probabilities. The above process is repeated until the probability distribution is steady. The series that combines the possible values of the greatest chosen probabilities is the final solution. Here, through adjustment of the probability field, each gene position of the gene series undergoes mutation and each gene can inherit their fathers' characteristics. Under the operation of directional mutations, the algorithm is converged finally.

3.3.3.1 Coding

Integer coding is adopted according to the characteristic of the problem as show in section 3.2.1.1.

Following this coding rule, each solution unit has different upper bound of card number according to the numbers of nodes it has. For the sake of convenience, the above coding is changed into serial coding. Each solution unit can be described as: $X = (x_1, x_2, ..., x_m)$ where m is the length of the series and m = n (n + 1)/2. Each element x_i of vector X corresponds to the element of the upper triangular part of matrix K, that is,

$$x_1 = K_{nn}, x_2 = K_{12}, x_3 = K_{22}, \cdots, x_i = K_{hl}, x_m = K_{nn} (h \le l)$$

$x_i \in V_i, V_i = (x_{i1}, x_{i2}, ... x_{ik_i}) = \left[1, K_{hl}^u + 1\right]$. V_i is the space of the *i*th bit of gene, k_i denote the size of this space.

V_i corresponds to a probability distribution. $P_i = (p_{i1}, p_{i2}, ..., p_{ik_i})$ where p_{ij} corresponds to x_{ij} and $\sum_{j=1}^{k_i} p_{ij} = 1$. P_i is the probability field of V_i and $P = \{P_i\}$.

3.3.2.2 Fitness function

Due to the minimizing property of the objective function, fitness function is defined as the one in section 3.2.1.2.

3.3.3.3 Generating new generation

The value of X is formed according to the probability field and x_i is chosen by roulette method[26].

$$\text{If } \sum_{j=1}^{l-1} P_{ij} < rand[0,1] \le \sum_{j=1}^{l} P_{ij}\text{, then } x_i = x_{il} \tag{23}$$

Where $P_{ij} = \sum_{h=1}^{j} p_{ih}$, $P_{i0} = 0$. The values of $x_i (i = 1, 2, 3, \cdots, m)$ are obtained by the above method. Thus *X* is formed.

3.3.3.4 Adjustment of the probability field

In order to let the probability field converge, continuous adjustment is needed. Assume that the fitness value of a certain solution unit X_{old} of the last generation is $fitness_{old} = F(X_{old})$ where $F(X)$ is a function of fitness value that is determined by the practical problem.

When the fitness value of the unit X_{new} of the preceding generation $fitness_{new} = F(X_{new})$, the probability field P of x_{ij} is adjusted as below:

The probabilities p_{ij} that correspond to x_{ij} are changed into $p_{ij} \cdot (1 + \Delta p_{ij})$. Then p_{i1} , p_{i2} … p_{ik_i} . undergo the unified process and $i = 1, 2, 3, \cdots, m$ .Here,

$$\Delta p_{ij} = \alpha \cdot \arctan\left(\frac{fitness_{new}}{fitness_{old}} - 1 \right) \tag{24}$$

Δp_{ij} is the amount for probability adjustment. α is the adjustment coefficient. The choices of adjustment coefficient of the probability field and the stoppage parameters are important and delicate. If the adjustment coefficient of the field is too big, then the speed of convergence is too fast, it will easily give the local optimization. If the coefficient is too small, then the speed is too slow, and the efficiency of the algorithm becomes low. That is why the substantial adjustment has to be done according to the practical problem of the instance simulation. The continuous adjustment of P increases the probability of getting solution units with high fitness values. At the same time, through unified process, the probability of having solution units with low fitness values decreases. This situation gives a very high probability of getting the optimal solution. The function in probability adjustment is a tangent function, $\arctan(x)$. It prevents the value of Δp_{ij} being too large to find the global optimal solution.

Actually, in the adjustment process, the fitness values are the directed values that operate mutation in the probability field. Mutations happen at each gene position of every generation simultaneously. Therefore, the genes have learned their fathers' characteristics thoroughly and this hastens the algorithm's speed of convergence.

3.3.3.5 Initialization

In order to guarantee the algorithm searches within global space, at the initiation stage, each solution should have the same probability of being chosen. That is p_{ij} which should have even distribution.

$$p_{ij} = 1/k_i \, , I = 1, 2, 3, \cdots, m; j = 1, 2, 3, \cdots k_i \tag{25}$$

3.3.3.6 The termination criterion

The largest generation is chosen as the termination criterion.

3.3.3.7 The Procedure for PEA

The main steps in the flow of the solution are:

Step 1: NC ← 0 (NC are the iterative steps)
The algorithm based on pheromone evolution produces the population X of the initial solutions and their corresponding initial probability field P. The actual procedure is to create a random number q for each x_{ij} of every solution unit x_i. $q \in V_i$ and $x_{ij} = q$forms the initial solution X which becomes the input for the simulation. The probability field is initialized according to formula (25);
Step 2: The fitness values $fitness_{old}=F(X_{old})$ of each solution of the last generation X_{old} are calculated; then turn to step 3;
Step 3: If the termination criterion is satisfied, turn to step 6; or else, create a solution of the new generation X_{new} and process it in the simulation. Then the fitness values $fitness_{new}=F(X_{new})$ of each solution of the present generation are calculated;
Step 4: The new probability field that corresponds to X_{new} is obtained by adjusting the probability field according to formula (24);
Step 5: Let $X_{old} \leftarrow X_{new}$, then turn to step 2;
Step 6: The optimal value of the objective function and the corresponding circulating card number are output.

3.3.3.8 Numerical Analysis

In order to test the efficiency of the PEA algorithm, two problems in section 3.2.1.5 is used here. The comparison of the two problems is shown in table 7 taking reasonable parameters of NP, NG, PC, PM obtained by simulation.

Problem scale	NP	NG	Best fitness	Best rate	T(s)
10^{18} (4 nodes)	200	150	892611.49	0.92	19
10^{40} (6 nodes)	200	150	890379.75	0.88	41

Table 7. Comparison of the results of different scale of problems for PEA

Table 7 has shown that, after the scale of the 6-node problem has been expanded by 10^{22} times as compared with a 4-node problem, the CPU time is increased by 22s and the optimization rate is decreased by 4%. Though there is the expanded complexity of the problem, the algorithm still possesses a better optimization rate and the increase in the CPU time is within an acceptable range. Moreover, time increase is mainly caused by the influences of the expansion of scale of the problem on simulation. Thus, conclusively speaking, the optimal performance of the algorithm is not greatly affected and it is still a fairly stable algorithm.

3.3.4 The comparison among the different algorithms

Finally the GA, Random-PSO and PEA is compared. The results are shown in Table 8.

Algorithm (problem scale)	NP	NG	Best rate	T（s）
GA (10^{18})	200	150	0.92	21
Random-PSO (10^{18})	150	100	0.94	13
PEA (10^{18})	200	150	0.92	19
GA (10^{40})	200	150	0.86	39
Random-PSO (10^{40})	150	150	0.90	54
PEA (10^{40})	200	150	0.88	41

Table 8. Comparison among GA, Random-PSO and PEA for different scale of problems

Table 8 has shown that, the Random_PSO has highest optimization rate when the scale of problem increase, and the CPU time of the three algorithms is similar. Therefore, the Random-PSO is more effective than GA and PEA for this kind of problem.

3.4 Conclusions

Determination of the optimal control strategy is a key factor for a successful supply chain. This chapter has made researches on the optimal inventory control strategy of a serial supply chain and presents the description of a two-level model. The first-level model is a nonlinear integer-programming model. Its main purpose is to determine the optimal control strategy which gives the minimal overall cost of a supply chain under some constraints. These constraints include the customer service no less than the given value and the standard deviation of input less than a given value. When the inventory control strategy is given, the second-level model is used to obtain the performance measurements of the supply chain. The first-level model reaches optimization through the algorithm based on intelligent algorithm. The intelligent algorithms of GA, random-PSO and PEA is considered in this study. The second-level model implements the general push/pull model of inventory of a serial supply chain by simulation. The main characteristic of the second-level model is that, the choice of control, push or pull, of a node is determined by whether it is under the feedback control of its downstream nodes. This has the potential to be an efficient quantitative tool for more complex SC analysis in the global business environment.

Instances of different scales of problems are analyzed. The results shown the effectiveness and the efficiency of the method. The cost of the whole supply chain is minimized while satisfying the customer demands and limiting the "bullwhip effect". It balances production rhythm and shared benefits of each node enterprise of a supply chain, which gives a quantitative support of rational organization of purchase, production, transportation, and sales.

Finally, comparison study of different scales of problem for three intelligent algorithms is given. Results shows that the random_PSO has highest optimization rate when the scale of problem increase, and the CPU time of the three algorithms is similar. It suggested that the random-PSO is more effective than the other two algorithm for this kind of problem.

This work was supported by the National Natural Science Foundation of China (Grant no. 70671020, 70721001, 70431003, 60673159), the Program for New Century Excellent Talents in University (Grant no. NCET-05-0295, NCET-05-0289), Specialized Research Fund for the Doctoral Program of Higher Education (Grant no. 20070145017, 20060145012), the National High-Tech Research and Development Plan of China (Grant No. 2006AA01Z214), the Key Project of Chinese Ministry of Education (Grant No. 108040), the Natural Science Foundation of Liaoning Province (Grant No. 20062022).

4. Reference

Chopra S , Meindl P, Supply Chain Management (3rd Edition) , Prentice Hall, 2006

Hugos M, Essentials of Supply Chain Management (2nd Edition), Wiley , 2006.

Dolgui A, Soldek J, Zaikin O, Supply chain optimization: product/process design, facility location and flow control, New York: Springer, 2005.

Gstettner S, Kuhn H, Analysis of Production Control Systems Kanban and CONWIP. International Journal of Production Research, 1996, 34(11): 3253-3273.

Deleersnyder J. L., Hodgson T. J., Mueller H, et al. Kanban Control Led Pull System: Analysis Approach. Management Science, 1989, 35(9): 1079-1091.

Sarker B. R., Fitzsimmons J. A., The Performance of Push and Pull Systems: A Simulation and Comparative Study. International Journal of Production Research, 1989, 27(11): 1715-1732.

Flapper S. D. P., Miltenburg G. J.,Wijugard J., Embedding JIT into MRP. International Journal of Production Research, 1991, 29(2): 329-341

Larsen N. E. Alting L. Criteria for selecting production control philosophy. Production Planning & Control, 1993, 4(1): 54-68

Villa A., Watanabe T. Production management: Beyond the Dichotomy Between “Push” and “Pull”. Computer Integrated Manufacturing Systems, 1993, 6(1): 53-63

Spearman, M. L., Woodruff, D. L., and Hopp, W. J. CONWIP: a Pull Alternative to Kanban. International Journal of Production Research, 1990, 28(5): 879-894.

Gaury E, Kleijnen JPC and Pierreval H, A methodology to customize pull control systems, Journal of the Operational Research Society, 2001, 52, 789–799.

Ovalle O R., Marquez A C., Exploring the Utilization of a CONWIP System for Supply Chain Management. International Journal of Production Economics,2003,83:195-215

Herer Y. T., Masin M. Mathematical Programming Formulation of CONW IP Based Production Lines and Relationships to MRP. International Journal of Production Research, 1997, 35 (4): 1067-1076.

Bose, G.J., Rao, A. Implementing JIT with MRP Creates Hybrid Manufacturing Environment. Industrial Engineering. 1988, 20(9): 49-53.

Michalewicz, M, Genetic algorithms + data structure = evolution programs. (3rd rev. and extended ed). Berlin; Hong Kong: Springer, 1996

Mitchell, M. An introduction to genetic algorithms. The MIT Press, Cambridge, Massachusetts, London, England, 1996.

Kennedy J , Eberhart R. Particle swarm optimization. Proc IEEE International Conference on Neural Networks, Perth, Australia 1995,1942-1948.

Eberhart R C, Kennedy J A. A new optimizer using particle swarm theory. Proc 6th Int. Symposium on Micro Machine and Human Science. Nagoya, 1995. 39-43.

Wilson E O. Sociobiology : The New Synthesis. .MA: Belknap Press, 1975.

Shi Y., Eberhart R. A modified particle swarm optimizer. Proc IEEE Int. Conf Evolutionary Computation. Anchorage, 1998. 69-73.

Kennedy J. The particle swarm: Social adaptation of knowledge. P roc IEEE Int. Conf on Evolutionary Computation. Indiam apolis, 1997. 303-308.

Thorndike E L. Animal Intelligence: Empirical Studies. New York: MacMillan, 1911.

Bandura A .Social Foundations of Thought and Action :A Social Cognitive Theory. New Jersey: Prentice 2 Hall, 1986.

Kennedy J. Eberhart R.A Discrete binary version of the particle swarm algorithm. In: Proceedings of the IEEE International Conference on Systems. Man and Cybernatics, 1997, 5:4104-4108.

Huang M, Wu X J, Wang X W, Ip W X, Yung K L, The risk programming of virtual enterprises based on algorithms of Min / Max / Mean / Random-PSO, 1st IEEE Conference on Industrial Electronics and Applications (ICIEA 2006), 2006.05, Singapore, 997-100

Huang Z. C., Wu F. C., Hu X. L., An evolutionary algorithm to integer programming problem based on pheromone, Reaserch on Computer Applying, 2001, 7(2), 27-29.

Dorigo M, Maniezzo V. A colony ant system: optimization by a colony of cooperating agent, IEEE Tran. On System, Man and Cybemetics Part B,1996, 26(1),34-37.

2

A Synergistic Approach towards Autonomic Event Management in Supply Chains

Roy Oberhauser
Aalen University
Germany

1. Introduction

Supply Chains (SCs), due to their very nature and intent (e.g., embracing change in markets, products, manufacturing, partners, globalization) in conjunction with market pressures, will face ongoing challenges that are necessarily reflected on the Information Technology (IT) infrastructure used to manage and optimise their operations.

Supply Chain Management Software (SCMS) typically covers the various functional aspects of SCs, including integration technology. The result of the IT integrations is a form of an information supply chain, including computational representations of physical SC entities. For purposes of this chapter SCMS will be considered to incorporate any ERP solutions and/or IT infrastructure utilized to enable the information integration required to support the SCs. Current SC IT challenges include decision making, collaboration, and attaining qualities such as scalability, performance, integratability, correctness, and reliability in the face of the perpetual dynamics and increasing complexity of SCs.

To avoid disruptions to SCs, Supply Chain Event Management (SCEM) considers the set of possible event scenarios and plans solutions. Events can be either representations of real-world events or can be introduced as a side-effect of the Information Systems (IS) supporting the SC (IT events). SCs can achieve their goals for optimal management of operations only to the extent and degree that they manage and automate the necessary information flow, especially with regard to managing unexpected events. The effective handling of potentially disruptive events is vital to achieving the aforementioned qualities, yet the ongoing change (mirrored in the IT systems) in entities and the properties and relations thereof, necessarily limits the sufficiency and totality of predefined solutions. A synergistic approach that leverages various computing paradigms can provide improved SCEM solutions.

In the face of potentially disruptive SC and IT events (referred to as SCEs in this chapter), autonomic computing (AC), inspired by the human autonomic nervous system, with its stated goals of self-configuration, self-optimization, self-healing, and self-protection (also known as self-X), would appear to be a synergistic candidate for improving SCEM. While some properties defined for autonomic systems[1] may not be applicable to SCEM, others will be beneficial. A partial application of AC techniques to achieve improved reactive event

---

[1] http://www.research.ibm.com/autonomic/overview/elements.html

management might be both practical and beneficial to SCEM. However, the changeability, heterogeneity, distribution, internationalization/localization, support, governance issues, and partner interdependencies in SCMS (both from an IT and linguistic/cultural viewpoint) makes SCEM and self-X attainment in SC and SCMS far more challenging compared to that of a self-contained rigid system.

Granular Computing (GC) is a paradigm that concerns itself with the processing of complex information entities called information granules, recognizing that at different abstraction levels of data, different relationships can be inferred (Pedrycz, 2001), (Bargiela et al., 2003), (Pedrycz et al., 2008). The meaning and impact of an SCE is also dependent on the granularity at which it is viewed, and other implications and trends may be detected at various abstraction levels.

To enable internationalization and decoupled SC partner agents to autonomically collaborate to address SCEs, it is imperative that the meaning for shared concepts be defined. The Semantic Web (SemWeb) adds machine-processable semantics to data (Berners-Lee et al., 2001). SemWeb computing (SWC) allows for greater and improved automation and integration of information in large information SCs due to its formal structuring of information, clearly defined meanings and properties of domain concepts, and standardized exchange formats. One of the issues facing SemWeb is the creation and adoption of standardized ontologies in OWL (Web Ontology Language) (McGuinness et al., 2004) for the various industry domains to precisely define the semantic meaning of the domain data – standardization is laborious and adoption is slow. However, to address both the challenge of SCEM to avoid disruptive impacts and the challenge of SCMS to achieve self-X and other qualities in a heterogeneous, changing, loosely-coupled and global environment, a transitional hybrid stage is proposed. A high-value event-specific subset tailored to the SCMS is tackled first that enables the collaborative involvement of partner agents (computing or human). In other words, if the partners have no agreement on a common meaning of an event, the concepts necessary to diagnose the indicative problem, and the meanings of the actions required in a solution, then the required collaborative and (partially to completely) automatable solutions for interdependent and non-trivial situations will continue to be elusive.

Additionally, to enable collaboration, partner exchangeability, and sharing across heterogeneous IT partner services and data, standardized access protocols for SCMS and SCEM is desirable if not essential. Service-oriented Computing (SOC), with its reliance on Web Services (WS), provides platform-neutral integration for arbitrary applications (Alonso et al., 2003).

Furthermore, Space-Based Computing (SBC) is a powerful paradigm for coordinating autonomous processes by accessing a distributed shared memory (called a tuple space) via messaging, thereby exhibiting linear scalability properties by minimizing shared resources. Tuple spaces implement a shared data repository of tuples (an ordered set of typed fields) that can be accessed concurrently in a loosely-coupled way based on the associative memory paradigm for parallel and distributed computing first presented by (Gelernter, 1985).

This chapter explores the potential for SCs that a synergistic approach to SCEM (SASCEM) that leverages various computing paradigms provides for improving the qualities of SCEM, especially with regard to approaching self-X properties and automation.

The rest of the chapter is organized as follows: Section 2 presents a review of the literature. In Section 3 the solution approach is presented. Section 4 presents initial implementation work based on the solution approach. In Section 5 preliminary results which evaluated certain performance and scalability characteristics are discussed, followed by a conclusion.

2. Literature review

(Mischra et al., 2003) describes an agent-based decision support system for a refinery SC, where agents collaborate to create a holistic strategy using heuristic rules. (Bansal et al., 2005) present a model-based framework for disruption management in SCs, generalizing the approach of (Mischra et al., 2003).

Related to SCs, Value-Added Networks (VANs) are hosted service offerings that add value to common networks by acting as an intermediary between business partners for sharing proprietary or standards-based data via shared business processes. As such they can be viewed as supporting informational SCs. Work on modelling collaborative decision making in VANs includes MOFIS (Naciri et al., 2008) and could be applied to improving SCEM, e.g., via integration of the concepts in a SASCEM.

Complex Event Processing (CEP) (Luckham, 2002) is a concept to deal with meaningful event detection and processing using pattern detection, event correlation, and other techniques to detect complex events from simpler events. Besides the research work that considers various aspects of CEP (e.g., high volume, continuous queries), commercial products include the TIBCO Complex Event Processing Suite.

The Resource Event Agent (REA) model aims at providing a basic generic shared data model that can describe economic phenomena of several different systems, both within and between enterprises of many different types (McCarthy, 1982). Work includes (Haugen et al., 2000) who present a semantic model for SC collaboration, (Hessellund, 2006) discusses SC modelling extensions to REA, while (Jaquet et al., 2007) presents a semantic framework for an event-driven operationalization and extension of the REA model that preserves flexibility and heterogeneity. An extended REA approach and hybrid/partial semantic formalization of events are congruent with a SASCEM.

Multi-Agent Systems (MAS) have been researched extensively, as has MAS in combination with SCs. Agent-based event management approaches includes Sense, Think & Act (ST&A), which exhibits function-driven, goal-driven (local goals), and collaborative goal-driven (global goals) behaviours (Forget et al, 2006). Agent-oriented supply-chain management is explored in (Fox et al., 2000) among others. (Adla, 2008) proposes an integrated deliberative and reactive architecture for SCM for supporting group decision making. Although this work has typically not utilized SOC and SWC, enabling and leveraging the integration of such problem-solving approaches is one goal of a SASCEM.

Work on semantic enhancement of tuple spaces includes sTuples (Khushraj et al., 2004), which extends the object-oriented JavaSpace implementation (Freeman et al., 1999) with an object field of type DAML-OIL Individual. (Tolksdorf et al., 2005) and (Tolksdorf et al., 2005a) describe work on Semantic Tuple Spaces. The Triple Space Computing (TSC) project[2] aims to develop a communication and coordination framework for the WSMX Semantic Web Service platform (Bussler et al., 2005) (Simperl, 2007). However, there has been insufficient exploration of the application of semantically-enhanced tuple spaces for collaborative event-based problem solving in general, and for SCEM in particular.

---

[2] http://tsc.deri.at

With regard to partner communication interoperability, the issue of scalable server-side push notification protocol over HTTP for Space-based Computing (SBC) is explored in (Kahn et al., 2007) but lacks standardization. Agent-interoperability via Web Services has been explored, e.g., JADE WSIG (Greenwood, 2005), but its application to SCs is still hampered due to a lack of standardization, e.g., by FIFA (Greenwood et al., 2007).

3. Solution

To achieve improved and more holistic solutions for SCEs while exhibiting AC and other expected qualities, the SASCEM is a synthesis of various areas of computing, specifically granular (GC), semantic web (SWC), service-oriented (SOC), space-based (SBC), event-based (EBC), context-aware (CAC), multi-agent (MAC), and autonomic computing (AC) as shown in (Fig. 1).

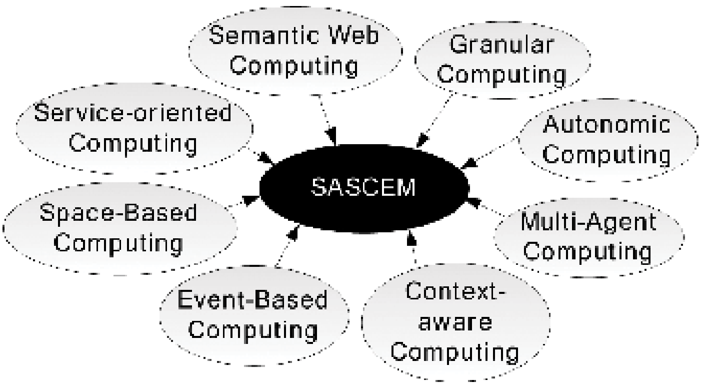


Fig. 1. Synergistic Solution Approach to SCEM

Solution constraints include heterogeneity, e.g., in partner agent implementations, rule-based techniques, platform software, and the adaptive and dynamic specialization of problem-solving for SCs. Additionally, it is assumed that for non-trivial SCs, no complete autonomic problem-solving for SCEM is as yet practical, thus the involvement of humans to the necessary degree is subsumed.

Principles that guided the solution approach include shared-nothing, decentralization, loose-coupling, standards-based communication, exchangeability (e.g, of collaborative decision making agent techniques), and enabling hybrid subsets for practical collaborative problem solving in SCs.

A simplified distributed SC solution infrastructure is shown in (Fig. 2). Using the SBC paradigm, tuple spaces are used to store event and event-relevant data, without deciding on meaning. Separate Semantic Web-aware tuple spaces are then used for collaboration on event diagnosis, problem prescription, and prognosis. Proactions or reactions are then initiated by partner agents and may involve the invocation of Partner or Infrastructure Services. Infrastructure Services and Partner Services provide the integration and access to SC (partner) functionality in accordance with the SOC paradigm. Heterogeneous interoperability and accessibility is supported via standards-based Web Services protocols, such as SOAP and REST (zur Muehlen et al., 2005). While an Enterprise Service Bus (ESB) is possible, its use depends on the SCMS and SCEM needs. In place of WS, Semantic Web Services (SWS), which envisions enabling automatic and dynamic interaction between software systems (Studer et al., 2007), might be a consideration; however, since the data repository can be readily accessed using simpler WS interfaces, a pragmatic approach utilizing the minimal amount of SemWeb to the extent needed to enable partner collaboration is currently preferable until SWS maturity and adoption has progressed.

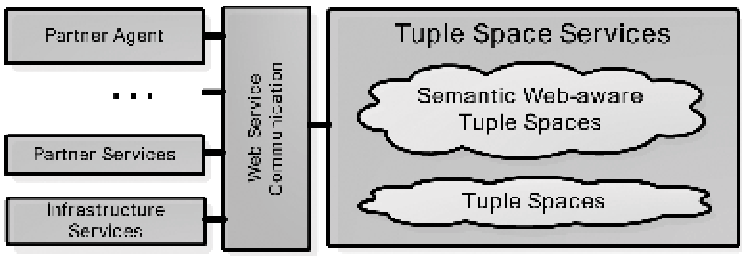


Fig. 2. Solution Infrastructure of the SASCEM (simplified)

The details of the solution approach will follow the event process steps shown in (Fig. 3).

Fig. 3. Event Process Steps in the SASCEM

3.1 Event acquisition

The acquisition of SCEs can come from sensors, partner machines and IT systems or services, and other event producers. In accord with EBC, the functionality of SCEM is triggered and invoked in response to the generation of SCEs. The events can be simple events to complex events inferred from simpler events, as considered in CEP. To enable the advantages of GC, these SCEs should be retained in their original state and supplementary complex events generated when these are detected via pattern matching or other CEP techniques by partner agents or other components. CEP and GC can be incorporated in (Partner or Infrastructure) Services or Agents.

3.2 Event storage

The event data is stored as a tuple in a tuple space following the SBC paradigm. This allows decoupled partner agents to flexibly subscribe to and be notified of relevant events. The tuple can be retrieved over time by various partner agents.
The data model is a hybrid that keeps data-only SCE tuples separate from the SemWeb tuple space. The SASCEM uses a hybrid transitional approach of communication between agents, supporting a blend of SemWeb and other data exchange in the tuple spaces. This allows the original event data to be viewed at different times, at different granularity levels, and to have multiple and even contradictory interpretations by diverse partner agents.
Registration for notifications by partner agents can be based on event data arrival, event data changes, etc., independent of semantic events. Thus partner agents without semantic

awareness but, e.g., with viable event handling rules and heuristics, can participate and support SCEM. Those partner agents with SemWeb capabilities can collaborate in the SemWeb tuple space and create and adjust the semantic meaning of the event data, type, attributes, and relations at different levels of abstraction and perhaps in different ontologies. This includes analysis and processing with regard to the event's relation to a problem (if any), diagnosis, prognosis, prescription, actions, etc. necessary to resolve it.

3.3 Contextual annotation

Contextual annotation of the event supports the retrieval of relevant data close to the occurrence of the event, and helps to determine its meaning and implications as well as infer complex events. As events are diagnosed over time, it may be determined by partner agents that certain information which is applicable and relevant should be gathered and other information may be determined to be irrelevant. CAC is thus utilized to annotate contextual and environmental information with the event, and those services registered for the event are notified. If no RDF(S)[3] (Brickley et al., 2004) information is provided with the event, then this too could be annotated to provide a uniform way of describing information resources associated with the event.

3.4 Event diagnosis

The correct diagnosis of SCEs is dependent on appropriate knowledge and rules, and due to the partner interdependency of SCs, collaborative effort to achieve AC is necessary. Diagnostic MAC enables the various partner agents to specialize in their particular knowledge without the limitations that a centralized single agent would incur. In order for heterogeneous partner agents to collaborate to achieve (semi-)autonomic behaviour, SWC is utilized to allow for a standardized and extensible approach for giving meaning to the events. A SemWeb-enabled tuple space (SWETS) provides a shared data storage where the meaning of the data types is defined and collaborative event analysis and interpretation is thus enabled. SemWeb-aware agents using inference engines can collaborate at various abstraction levels using GC paradigms. Complex events can be inferred from simple events, e.g., regarding their timing, sequence, patterns, or trends, and CEP could be utilized. If the collaborative diagnosis relates the event to a(n) (unknown) problem, processing continues, otherwise it is completed. Multiple and even contradicting diagnoses are allowed and may occur. Note that this situation may in turn create a new event which in turn goes through the processing steps.

Ontologies are minimally necessary for the intersection set of concepts necessary for SCEM between partner agents. In this regard, full ontologies that cover all possible concepts in the SC can - but must not necessarily, be avoided. A partial application of SemWeb appears practical and reasonable at this time, given some current practical limitations with regard to payoff vs. effort, standardization, maturity, industrial usage, training, tooling, etc. Yet the intersection of concepts between partners requires a formal definition and agreement in order for collaborative and automated SCEN to be enabled.

[3] Resource Description Framework (Schema)

While agents are often considered to be artificial computational entities that perform tasks with a degree of autonomy, in the SASCEM agents include the set of human agents as well for problem solving, supporting a hybrid spectrum from completely manual to automatable diagnosis and solutions, since each SC is unique and for non-trivial dynamic SCs new events and problems may occur that require human intervention before they become automatable.

3.5 Problem prescription

Using the SWETS, the agents, based on the possible diagnoses, collaboratively decide on a prescription consisting of a set of actions, e.g. using (Adla, 2008) or other decision techniques, and incorporating AC techniques where applicable.

3.6 Problem prognosis

Separately from the prescription, the forecasted impact, side-effects, and success chances of the diagnosis and/or the prescription in the form of a prognosis could optionally be (collaboratively) determined and placed in the SWETS, perhaps triggering new events.

3.7 Proactions and reactions

Based on the prescription and/or prognosis, the reactions are executed by the appropriate agent(s), using partner or infrastructure services as needed, and preventative proactions can be executed to limit the impact of side effects, repeated problems, etc.

4. Solution implementation

The prototype implementation of the SASCEM currently includes an adaptation of an open source tuple space implementation (XSpace[4]). Hybrid support for SWETS is currently dependent on the outcome of a tsc++[5] evaluation and integration. Apache Axis2[6], which supports asynchronous WS, was used for WS communication.

To illustrate the SASCEM implementation and for prototype testing purposes, an ontology (Fig. 4) for a software SC was created using Protege 3.3.1. First it will be described in prose, followed by OWL abstract syntax. Work on SCM ontologies includes (Haller et al., 2008).

BusinessObjects can depend on other BusinessObjects and have Suppliers, Consumers, and Producers. A Service is a BusinessObject with a Protocol, including human and organizational services, and can be specialized as a WS or a SWS.

Products and Information are Artifacts, which are BusinessObjects. Systems, Hardware, and Software are Products and Products may have a Configuration. A Patch is Software. A Document is Information.

Events may refer to one or more BusinessObjects and be associated with one or more Problems. Problems refer to a Quality that is affected, may include a Diagnosis and a Prognosis. A Diagnosis may include a Prescription that may refer to a set of Actions and may refer to a Patch and/or Configuration.

[4] http://xspacedb.sourceforge.net/

[5] http://tsc.sti2.at/

[6] http://ws.apache.org/axis2/

Fig. 4. Partial Software Supply Chain Event Management Domain Ontology

An alphabetical listing in OWL abstract syntax follows (Listing 1):

```
Class(Action partial owl:Thing)
Class(Artifact partial BusinessObject)
Class(BusinessObject partial restriction(hasEvent minCardinality(0))
         owl:Thing
         restriction(version cardinality(1))
         restriction(name cardinality(1))
         restriction(depends minCardinality(0))
         restriction(hasConsumer minCardinality(0))
         restriction(hasProducer minCardinality(1))
         restriction(hasSupplier minCardinality(0)))
Class(Configuration partial owl:Thing)
Class(Consumer partial restriction(hasBusinessObject minCardinality(1))
       owl:Thing
       restriction(name cardinality(1))
       restriction(homepage cardinality(1)))
```

```
Class(Diagnosis partial restriction(hasPrescription minCardinality(0))
        owl:Thing)
Class(Document partial Information)
Class(Event partial restriction(hasProblem minCardinality(0))
        restriction(hasBusinessObject minCardinality(0))
        owl:Thing)
Class(Format partial owl:Thing)
Class(Hardware partial Product)
Class(Information partial Artifact
        restriction(hasFormat cardinality(1)))
Class(Patch partial Software)
Class(Prescription partial restriction(hasPatch maxCardinality(1))
        restriction(hasAction minCardinality(0))
        restriction(description cardinality(1))
        owl:Thing)
Class(Problem partial restriction(hasEvent minCardinality(0))
        owl:Thing
        restriction(description cardinality(1))
        restriction(hasImpact minCardinality(0))
        restriction(hasSolution minCardinality(0))
        restriction(hasQuality minCardinality(0)))
Class(Product partial restriction(hasConfiguration minCardinality(0))
        Artifact)
Class(Prognosis partial owl:Thing
        restriction(description cardinality(1)))
Class(Protocol partial owl:Thing)
Class(Quality complete oneOf(Functionality
        Reliability
        Usability
        Efficiency
        Maintainability
        Portability))
        SubClassOf(Quality owl:Thing)
Class(SemanticWebService partial WebService)
Class(Service partial restriction(hasProtocol cardinality(1))
        BusinessObject)
Class(Service partial restriction(hasProtocol cardinality(1))
        BusinessObject)
Class(Software partial Product)
Class(Supplier partial owl:Thing
        restriction(name cardinality(1))
        restriction(homepage cardinality(1)))
Class(System partial Product)
Class(WebService partial Service)
```

Listing 1. Partial Software Supply Chain Event Management Domain Ontology

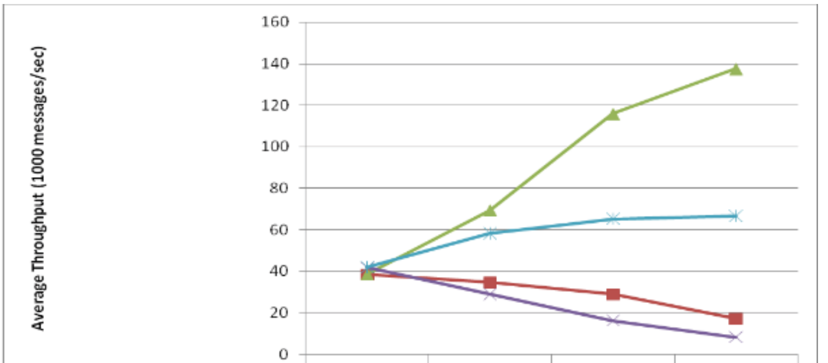
5. Results

Preliminary results considered the viability of the solution architecture and prototype implementation used for this peer-based middleware combination of a tuple space, relational database, message broker, and asynchronous Web Services infrastructure for addressing the SC qualities in scalability on a per-agent and a system level before integrating true SemWeb-aware problem-solving agents. For this, two key throughput scenarios were measured consisting of the event message into the tuple space (put scenario) and the notify scenario to other agents (notify scenario).

The test configuration consisted of 2,4 GHz Dual Core Opteron 180 PCs running Windows XP Pro SP2, 3.3GB RAM, 100 Mbit LAN, JRE 1.6.0_07, and Apache Axis2 0.93. One server PC ran Xspace 1.1, Jboss 4.0.3, and HSQLDB 1.8.0. The averages over three runs were used for all results (Fig. 5).

For the notify scenario, 1000 SOAP messages containing an event to put into the tuple space were sent from a single producer PC to the server, with a Message-Driven Bean, upon receiving the put, notifying agents (via asynchronous SOAP messages) on either 1, 2, 4, or 8 consumer PCs. Note that all the throughput results exclude and ignore the server and the producer PCs, but only consider the notification throughput on the consumers. The results show that asynchronous notifications by the tuple service to 1 to 2 and 4 peers regarding the put allowed an almost linear scalability, with a reduction at 8 peers due to full CPU utilization on the server.

For the put scenario, 1000 SOAP messages containing an event to put into the tuple space were sent from either 1, 2, 4, or 8 producer PCs to the same tuple space on the server PC. The results show a significant reduction in cumulative throughput with each added peer, which can be explained by the transactional bottleneck of the puts to the relational database on the server. These results and storage options, including persistence requirements on a per tuple basis, will be taken into account and optimization opportunities considered in future work.



	1	2	4	8
Notify Throughput per Peer	38452	34678	28950	17190
Cumulative Notify Throughput	38452	69356	115799	137521
Put Throughput per Peer	41875	29185	16336	8346
Cumulative Put Throughput	41875	58370	65345	66769

Fig. 5. Average throughput vs. number of peers for web service notifications

Since for SASCEM the number of notifications is expected to be much higher than the number of generated events, the nearly linear scalability for notifications show that the SBC and EBC foundation for SASCEM is viable for SCEM.

6. Conclusion

The increasing reliance on SCs, coupled with increasing complexity, dynamism and heightened quality expectations, are necessarily reflected in the SCMS and implicitly in the need for improved SCEM to limit disruptions and achieve self-X qualitites. A novel synergistic approach to SCEM, as presented in this chapter (SASCEM), leverages the computing paradigms of granular, semantic web, service-oriented, space-based, event-based, context-

aware, multi-agent, and autonomic computing to create a holistic solution approach that can change how SCEM is approached. Within the SASCEM, the hybrid approach to SWC makes adoption practical and viable in the near term. Preliminary results show sufficient performance and scalability qualities for such an SBC infrastructure to address SCEM.
The scope of applicability for this approach goes beyond SCEM, and could be applied to event management in general outside of SCs. Moreover, SCMS might be architected differently where a SASCEM adopted.
Future work includes integrating SemWeb-based problem-solving agents with Semantic Web-aware tuple spaces and evaluating the solution with regard to real-world problem-solving scenarios.

7. Acknowledgements

Thanks to Tobias Gaisbauer for his assistance with the experiments and implementation.

8. References

Adla, A. (2008). Agent-Based Group Decision Making, *Proceedings of The Third International Conference on Software Engineering Advances (ICSEA 2008)*, pp.146-151, ISBN 978-0-7695-3372-8, Sliema, Malta, October 2008, IEEE Computer Society

Alonso, G.; Casati, F.; Kuno, H. & Machiraju, V. (2003). *Web Services – Concepts, Architectures and Applications*, Springer Verlag, Berlin

Bansal, M.; Adhitya, A.; Srinivasan, R. & Karimi, I.A. (2005). An Online Decision Support Framework for Managing Abnormal Supply Chain Events, 38th European Symposium of the Working Party on Computer Aided Process Engineering: ESCAPE-15, 29 May - 1 June 2005, Barcelona, Spain, Elsevier Science B.V., ISBN 0444519912

Bargiela, A. & Pedrycz, W. (2003). *Granular Computing: An Introduction,* Dordercht: Kluwer Academic Publishers

Berners-Lee, T.; Hendler, J. & Lassila, O. (2001). The Semantic Web, *The Scientific American*, May 2001.

Brickley, D. & Guha, R.V. (2004). RDF Vocabulary Description Language 1.0, RDF Schema, W3C Recommendation, February 2004, http://www.w3c.org/TR/rdf-schema/

Bussler, C.; Kilgarriff, E.; Krummenacher, R.; Martin-Recuerda, F.; Toma, I. & and Sapkota, B. (2005) WSMX Triple-Space Computing, http://wsmo.org/TR/d21/v0.1, June 2005, D21 v0.1

Forget, P. ; D'Amours, S. & Frayret, J.-M. (2006). Collaborative event mangement in supply chains: an agent-based approach, *Information Technology for Balanced Manufacturing Systems*, volume 220, *IFIP International Federation for Information Processing*, 2006, Springer, Boston, USA, pp. 89–98

Fox, M. S.; Barbuceanu M. & Teigen, R. (2000). Agent-oriented supply-chain management. *International Journal of Flexible Manufacturing Systems*, 12(2-3), pp. 165–188

Freeman, E.; Hupfer, S. & Arnold, K. (1999). *JavaSpaces Principles, Patterns and Practice*, The Jini Technology Series, Addison-Wesley, Reading MA, USA.

Gelernter, D. (1985). Generative communication in Linda, *ACM Transactions on Programming Languages and Systems*, volume 7, number 1, January 1985, pp. 80-112

Greenwood, D. (2005). JADE Web Service Integration Gateway (WSIG), JADE Tutorial, AAMAS 2005

Greenwood, D.; Lyell, M.; Mallya, A. & Suguri, H. (2007). The IEEE FIPA approach to integrating software agents and web services, *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, Honolulu, Hawaii, 2007, Article No. 276, ISBN:978-81-904262-7-5

Haller A.; Gontarczyk, J. & Kotinurmi P. (2008). Towards a complete SCM ontology: the case of ontologising RosettaNet, *Proceedings of the 23rd Annual ACM Symposium on Applied Computing*, ACM, Fortaleza, Ceará, Brazil, 2008, pp. 1467-1473

Hessellund, A. (2006). Supply Chain Modeling with REA. *TR 2006-80*, 2006, The IT University of Copenhagen, Denmark

Jaquet, M. & Østerbye, K. (2007). A Property Driven Approach Towards Creating an Adaptive REA Framework, *Proceedings of the REA-25 Conference*, June 2007, Delaware, USA

Kahn, E.; Riemer, J. & Lechner, L. (2007). XVSMP/Bayeux: A Protocol for Scalable Space Based Computing in the Web, *Proceedings of the 16th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises* (WETICE 2007), pp. 68-73

Khushraj, D.; Lassila, O. & Finin, T. (2004). "sTuples: Semantic Tuple Spaces," First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'04), pp.268-277

Luckham, D. (2002). *The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems*, Addison-Wesley Professional, ISBN 0-201-72789-7

McCarthy, E. W. (1982). The REA Accounting Model: A Generalized Framework for Accounting Systems in a Shared Data Environment. *The Accounting Review*, July 1982, pp. 554-78

McGuinness, D. L. & van Harmelen, F. (2004). OWL Web Ontology Language: Overview, W3C Recommendation, February 2004, http://www.w3c.org/TR/owl-features/

Naciri, S.; Yoo, M. & Glardon, R. (2008). MOFIS: new conceptual modeling framework for handling value adding networks complexity, *Proceedings of The Third International Conference on Software Engineering Advances (ICSEA 2008)*, pp.516-522, ISBN 978-0-7695-3372-8, Sliema, Malta, October 2008, IEEE Computer Society

Pedrycz, W. (2001). *Granular Computing: An Emerging Paradigm*, Heidelberg: Physica-Verlag

Pedrycz, W.; Skowron, A. & Kreinovich, V. (2008). *Handbook of Granular Computing*, Wiley-Interscience, ISBN: 978-0-470-03554-2

Simperl, E.; Krummenacher, R. & Nixon, L. (2007). A coordination model for triplespace computing, *Proceedings of the 9th International Conference on Coordination Models and Languages (Coordination)*, Springer Verlag, June 2007

Studer, R.; Grimm, S. & Abecker, A. (2007) *Semantic Web Services: Concepts, Technologies, and Applications*, Springer, ISBN-10: 3540708936

Tolksdorf, R.; Bontas, E. P. & Nixon, L. J. (2005). Towards a Tuplespace-Based Middleware for the Semantic Web, *Proceedings of the 2005 IEEE/WIC/ACM international Conference on Web intelligence*, (September 19 - 22, 2005). Web Intelligence. IEEE Computer Society, Washington, DC, pp. 338-344

Tolksdorf, R.; Nixon, L.; Bontas, E. P.; Nguyen, D. M. & Liebsch, F. (2005a). Enabling real world Semantic Web applications through a coordination middleware, *Proceedings of the 2nd European Semantic Web Conf.* ESWC'05, 2005

zur Muehlen, M.; Nickerson, J. V. & Swenson, K. D. (2005). Developing Web Services Choreography Standards - The Case of REST vs. SOAP, *Decision Support Systems*, Volume 40 , Issue 1, July 2005, Elsevier Science Publishers B. V. Amsterdam, The Netherlands, ISSN:0167-9236

3

Managing Logistics Flows Through Enterprise Input-Output Models

V. Albino[1], A. Messeni Petruzzelli[1] and O. G. Okogbaa[2]
[1]DIMeG, Politecnico di Bari, Bari,
[2]DIMSE, University of South Florida,
[1]Italy
[2]USA

1. Introduction

Nowadays, the management of logistics flows is becoming a crucial activity for competitiveness. In fact, globalization is changing the way in which companies organise their production and distribution activities, considerably increasing the spatial complexity of supply chains (see also Choi & Hong, 2002; Stephen, 2004). Therefore, firms have to redesign their supply chains, both global (Meixell & Gargeya, 2005) and local (Carbonara et al., 2001), in order to sustain competitiveness and to deal with the new geography of customers and suppliers (see also Hulsmann et al., 2008; Keane & Feinberg, 2008).

In this economic scenario, logistics activities cannot be more considered as a derived demand, but as a key factor for achieving competitive advantage (Hesse & Rodrigue, 2004; Gunasekaran & Cheng, 2008). In fact, the reduction of transportation time and costs can lead supply chains to improve their effectiveness and efficiency. With this regard, in the literature several studies have focusing their attention on the analysis of logistics performance, providing measures and indicators, supporting managers and policy makers in the identification of logistics strategies and policies (see also Lai & Cheng, 2003; Lai et al., 2004).

Furthermore, globalization has moved competition from single companies to whole supply chains, thus requiring a joint design and management of logistics flows (Xu & Beamon, 2006; Yi & Ozdamar, 2007). Therefore, in order to guarantee the integrated and effective organization of logistics services, their management and coordination is generally assigned to specific actors, namely third-party logistics (3PL) provider or logistic service provider (LSP) (e.g. Hertz & Alfredsson, 2003; Carbone & Stone, 2005; Kim et al., 2008), which constitute the interconnectedness among the different actors of the supply chain. This new generation of actors is called into being to provide a total logistics service enabling faster movement of goods, shorter turnaround time, more reliable delivery, and reducing the number of transfers.

Moreover, the growing attention towards the environmental sustainability has forced organizations to manage their logistics activities evaluating the environmental effects (e.g. Jayaraman & Ross, 2003; Wang & Chandra, 2007). In fact, international trades, global activities of multinationals, and the division of labour/production are strongly increasing these negative effects, which are also accentuated by the growing market share of the most energy intensive modes of transportation (truck and air[1]) and the relative decline of other modes (ship and rail[2]) (EEA, 2004). The EU White Paper on Transport Policy (CEC, 2001) recognises that transport energy consumption is increasing and that 28% of CO_2 emissions are now transport-related. Carbon dioxide emissions continue to rise, as transport demand outstrips improvements in energy-related emissions. The sector with the largest projected increase in EU-15 emissions is transport.

In this scenario, consumers and governments are pressing companies to re-design and carefully manage their logistics networks, in order to reduce the environmental impact of their products and processes (Thierry et al., 1995; Quariguasi Frota Neto et al., 2008).

In the present paper, we propose the use of enterprise input-output (EIO) models to represent and analyse physical and monetary flows between production processes, including logistics ones. In particular, we consider networks of processes transforming inputs into outputs and located in specific geographical areas.

The paper is structured as follows. In the following section, a brief review of EIO models is presented. Then, in Section 3 some possible application fields of EIO models are identified. Sections 4 and 5 describe the basic equations of EIO models and their use. In Section 6 and 7 EIO models are applied to represent and analyse transportation processes, both at an aggregate and disaggregate level, and logistics services markets, respectively. Finally, the main findings and results are summarized into discussion and conclusions (Section 8).

2. Enterprise input-output models

The input-output (IO) approach has been typically applied to analyse the structure of economic systems, in terms of flows between sectors and firms (Leontief, 1941). So doing, analysing the interdependencies among entities, economists and managers can evaluate the effect of technological and economic change at regional, national, and international level.

According to the different level of analysis, IO models can be highly aggregated or disaggregated. Miller and Blair (1985) use a disaggregated level and consider the pattern of materials and energy flows amongst industry sectors, and between sectors and the final customer. A higher level of disaggregation is useful to define a model better fitting real material and energy flows. However, the drawback of working on a high level of disaggregation is represented by the lack of consistency in the input coefficients. In fact, it is sufficient that technological changes happen in a process to modify the input coefficients. On the other hand, because of the small scale, it is easy to know which technological changes are employed in one or more processes and the modifications to apply to the technical coefficients.

EIO models constitute a particular set of IO models, useful to complement the managerial and financial accounting systems currently used extensively by firms (Grubbstrom & Tang, 2000; Marangoni & Fezzi, 2002; Marangoni et al., 2004). In particular, Lin and Polenske (1998) proposed a specific IO model for a steel plant, based on production processes rather than on products or branches. Similarly, Albino et al. (2002, 2003) have developed IO models for analyzing in terms of material, energy, and pollution flows the complex dynamics of

[1] Air transport is growing by 6–9 % per year in both the old and new EU Member States.

[2] The market shares of modes such as rail are increasing only marginally, if at all.

global and local supply chains, and of industrial districts, respectively. Moreover, EIO models based on processes have been adopted to evaluate the effect of different coordination policies of freight flows on the logistics and environmental performance of an industrial district (Albino et al., 2008).

At the single firm's level the EIO model can be useful to coordinate and manage internal and external logistics flows. At the level of the whole industrial cluster the enterprise input-output model can be effective to analyse logistics flows and to support coordination policies among firms and their production processes.

As in the case of industrial districts, EIO models can be applied to contexts highly characterized by the geographical dimension, such as the local and global supply chains. For better addressing the spatial dimension the EIO approach can be integrated with GIS technology, geographically referring all the inputs and outputs accounted in the models (e.g. Van der Veen & Logtmeijer, 2003; Zhan et al., 2005; Albino et al., 2007).

This paper aims at investigating logistics related issues adopting EIO models. To cope with this aim, transportation is modelled as a process (or input) both at an aggregate and disaggregate level, providing the other processes with the logistics services necessary to convey products from origins to destinations. In the former, transportation is modelled as a single process (or input) that supplies all the other production processes involved in the chain. Alternatively, it can be modelled considering all the tracks representing the transportation network through which products flow to and from production processes using the disaggregate approach.

These two approaches are used to pursue different system goals. In particular, the aggregate model is used to analyse the logistics flows from a managerial perspective. In fact, economic and operational performance can be evaluated. Whereas, the adoption of a disaggregated approach permits a more space-oriented analysis. Specifically by modelling all the tracks it is possible to examine issues related to traffic congestion, transportation infrastructure availability, and pollutant emissions in specific geographical areas.

3. EIO models for logistics: a framework of analysis

As stated in the previous section, EIO models are accounting and planning tools aimed at describing production process and analyzing their reciprocal interdependences. Here, we intend to shed further light on the adoption of EIO models to manage logistics flows, providing a framework that identifies their main application fields and explains their usefulness.

In particular, we can consider two main perspectives under which the production processes and related logistics flows can be investigated: i) a spatial and ii) an operational perspective. In the former, the processes are described referring to their location into a specific geographical area. This approach can be effective to examine space-related issues, such as traffic congestion, pollutant emissions, transportation infrastructure, and work force availability. In this case, the analysis is applied to the set Π_G, constituted by all the processes π_i (i=1,…,n) located in the area G.

Adopting an operational perspective, goals oriented to maximize the efficiency and effectiveness of the processes belonging to a specific supply chain can be pursed. Therefore, the application field is related to the set Π_{SC}, constituted by all the processes π_i (i=1,…,n) belonging to the supply chain SC. Moreover, considering the logistic flows associated to the production processes, a further application can be represented by the analysis of all the flows between processes π_i (i=1,…,n) managed by a specific logistic provider. Thus, the set Π_{LP}, constituted by all the flows ω_{ij} (i=1,…,n and j=1,…,m) managed by a specific logistic provider LP, can be studied.

These application fields are not mutually exclusive. In fact, they can be combined in order to provide more specific and complex analysis. For instance, we can consider the set $\Pi_G \cap \Pi_{sc}$, represented by all the processes located in the area G and involved in the supply chain SC. Then, we can describe the generic process π_i belonging to this set adopting both an operational and geographical perspective. In particular, all its inputs and outputs are described taking into account the nature and their origins and destinations.

In Figure 1, the process π_i is represented, identifying its main output (x_i), the inputs supplied by other processes belonging to $\Pi_G \cap \Pi_{sc}$ (z_{1i}, z_{2i},…, z_{ni}), the wastes and by products produced by π_i (w_1, w_2,…, w_n), and the other primary inputs required by π_i and supplied by processes that are not included into the set to $\Pi_G \cap \Pi_{sc}$ (r_1, r_2,…, r_s).

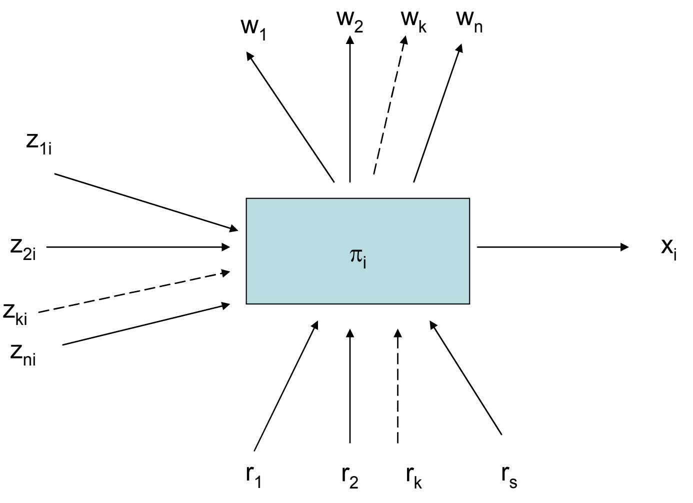


Fig. 1. Inputs and outputs of the process π_i.

This representation can be useful for accounting purposes, since it permits to identify the outputs produced by the process and all the required inputs. However, in order to take into account the spatial characteristics of inputs and outputs they have to be geographically referred, considering their origins and destinations. All the processes belonging to $\Pi_G \cap \Pi_{sc}$ can be geo-referred as well as the flows between them.

In fact, the primary input r_k can be supplied by distinct origins. Thus, we can distinguish the input on the basis of its origins, being r_{kA} and r_{kB}, where A and B represent two distinct locations. Moreover, also the main output can be delivered to different destinations. In particular, these destinations can belong or not to the considered set of processes. In the latter, we indicate as f_i the output produced by p_i and destined outside the boundary of the system. Therefore, the main output can be distinguished on the basis of the destinations. For instance, we can have f_{iC} and f_{iD}. The same consideration can be applied to wastes and by products (w_{kG}, w_{kF}).

In Figure 2 the process π_i is represented considering the geographical locations of inputs and outputs.

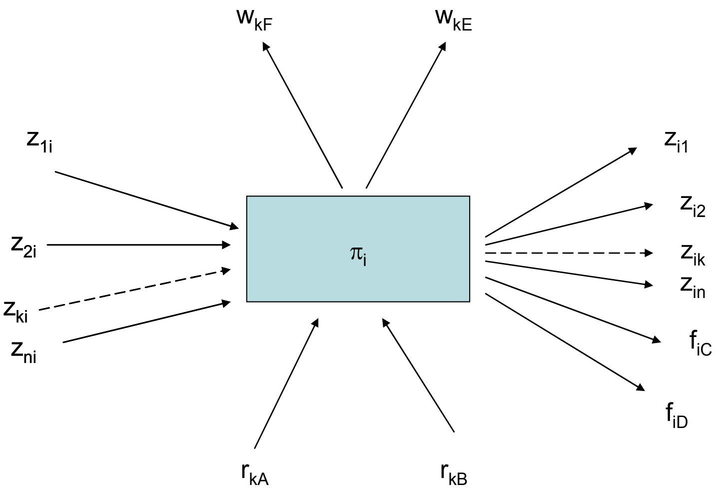


Fig. 2. Inputs and outputs of the process p_i, distinguished by geographical locations.

Therefore, these two distinct representations permit to move from a physical and monetary description of the processes (Figure 1) to a spatial one (Figure 2).

4. EIO models and production processes: basic equations

Let us consider a set of production processes. This set can be fully described if all the interrelated processes as well as input and output flows are identified and modelled.
Let Z_0 be the matrix of domestic (i.e. to and from production processes within the set) intermediate deliveries, f_0 is the vector of final demands (i.e. demands leaving the set), and x_0 the vector of gross outputs. If n processes are distinguished, the matrix Z_0 is of size n x n, and the vectors f_0, and x_0 are n x 1. It is assumed that each process has a single main product as its output. Each of these processes may require intermediate inputs from the other processes, but not from itself so that the entries on the main diagonal of the matrix Z_0 are zero.
Of course, also other inputs are required for the production. These are s primary inputs (i.e. products not produced by one of the n production processes). Next to the output of the main product, the processes also produce m by-products and waste. r_0 and w_0 are the primary input vector, and the by-product and waste vector of size s x 1 and m x 1, respectively.
Define the intermediate coefficient matrix A as follows:

$$A \equiv Z_0 \hat{x}_0^{-1}$$

where a "hat" is used to denote a diagonal matrix. We now have:

$$x_0 = Ax_0 + f_0 = (I - A)^{-1} f_0$$

It is possible to estimate R, the s x n matrix of primary input coefficients with element r_{kj} denoting the use of primary input k (1,…, s) per unit of output of product j, and W, the m x n matrix of its output coefficients with element w_{kj} denoting the output of by-product or waste type k (1,…, m) per unit of output of product j. It results:

$$r_0 = R\, x_0$$

$$w_0 = W x_0$$

Note that the coefficient matrices A, R, and W are numerically obtained from observed data. A change in the final demand vector induces a change in the gross outputs and subsequently changes in the input of transportation, primary products, and changes in the output of by-products and waste.

Suppose that the final demand changes into $\overline{f}$, and that the intermediate coefficients matrix A, the primary input coefficients matrix R, and the output coefficients matrix W, are constant (which seems a reasonable assumption in the short-run), then the output changes into:

$$\overline{x} = (I - A)^{-1}\overline{f}$$

Given this new output vector, the requirements of primary products and the outputs of by-product and waste are:

$$\overline{r} = R\overline{x}$$

$$\overline{w} = W\overline{x}$$

where $\overline{r}$ gives the new s x 1 vector of primary inputs, and $\overline{w}$ the new m x 1 vector of by-products and waste types.

The enterprise I-O model can be also adopted to account the monetary value associated with each production process. In particular, let p_0 be the vector of the prices with element p_i denoting the unitary price of the main product at the end of the process i. Thus, considering the vector of the gross outputs x_0, we can compute the vector y_0, representing the total revenues associated with each gross output as follows:

$$y_0 = \hat{x}_0 p_0$$

Moreover, we can define the matrix B, where the generic element b_{ij} is expressed as:

$$b_{ij} = a_{ij}\frac{p_i}{p_j}$$

Then, we have:

$$y_0 = By_0 + \hat{f}_0 p_0 = \left(I - B\right)^{-1}\hat{f}_0 p_0$$

If n production processes are considered, the matrix B is of size n x n, and the vectors $\hat{f}_0 p_0$ and y_0 are n x 1. Moreover, we can define the vector of the prices p_0^w, where p_i^w represents the unitary price associated to the wastes and by-products of each process. In particular, waste and by-product will have non-positive and non-negative price respectively. Hence, considering the vector w_0, we can identify the vector y_0^w representing the total revenues associated with each waste and by-product as follows:

$$y_0^w = \hat{w}_0 p_0^w$$

Of course, costs are sustained by the production processes. Let in_0 be the vector of the costs associated to the primary inputs, including wages and salaries, and an_0 the vector of investments amortization. Then, the profit (pt) for all the production processes can be computed as:

$$pt = \sum_{i=1}^{n}\left(y_i + y_i^w - \sum_j p_j z_{ji} - in_i - am_i\right)$$

5. EIO models for a supply chain stage

In the present section, we propose a theoretical example, aimed at describing the physical and monetary flows associated with a network of production processes, not including transportation, taking into account the geographical location of inputs and outputs. For the sake of simplicity, a supply chain stage is considered.
Let us consider three production processes, π_1, π_2, and π_3, belonging to Π_{sc} and exchanging products as shown in Figure 3.

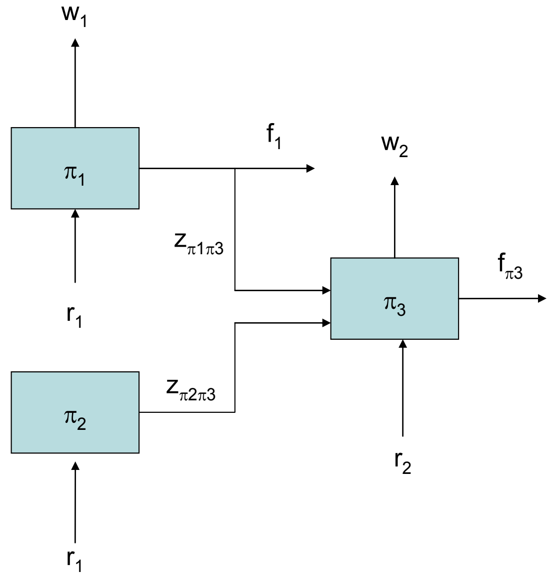


Fig. 3. Inputs and outputs of production processes in a supply chain stage.

Adopting the EIO models, the balance table accounting the materials flows of the supply chain stage is reported in Table 1.

Processes	π_1	π_2	π_3	f_0	x_0
π_1			$a_{\pi_1\pi_3}x_{\pi_3}$	f_{π_1}	x_{π_1}
π_2			$a_{\pi_2\pi_3}x_{\pi_3}$		x_{π_2}
π_3				f_{π_3}	x_{π_3}
Primary inputs					
r_1	$r_{1\pi_1}x_{\pi_1}$	$r_{1\pi_2}x_{\pi_2}$			
r_2			$r_{2\pi_3}x_{\pi_3}$		
Wastes and by-products					
w_1	$w_{1\pi_1}x_{\pi_1}$				
w_2			$w_{2\pi_3}x_{\pi_3}$		

Table 1. Balance table for the supply chain stage in Figure 3.

As previously explained, the same type of input and output can be characterised by different origins and destinations. Let us assume that the final demand f_3 is delivered to the geographical destinations A and B, the primary input r_2 comes from the geographical origins C and D, and the waste w_1 is destined to the geographical destinations E and F (Figure 4).

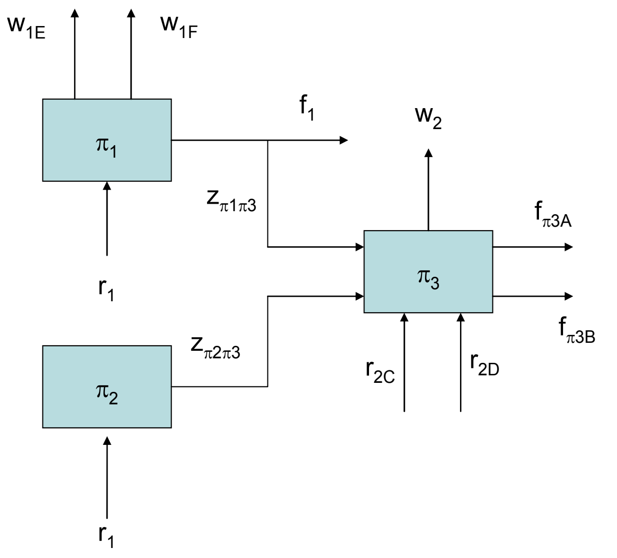


Fig. 4. Inputs and outputs of production processes distinguished by geographical origins and destinations.

On the basis of this representation, it is possible to define the related balance table, reported in Table 2.

Processes	π_1	π_2	π_3	f_{0A}	f_{0B}	x_0
π_1			$a_{\pi_1\pi_3}x_{\pi_3}$	f_{π_1}		x_{π_1}
π_2			$a_{\pi_2\pi_3}x_{\pi_3}$			x_{π_2}
π_3				$f_{\pi_3 A}$	$f_{\pi_3 B}$	x_{π_3}
Primary inputs						
r_1	$r_{1\pi_1}x_{\pi_1}$	$r_{1\pi_2}x_{\pi_2}$				
r_{2C}			$r_{2C,\pi_3}x_{\pi_3}$			
r_{2D}			$r_{2D,\pi_3}x_{\pi_3}$			
Wastes and by-products						
w_{1E}	$w_{1E,\pi_1}x_{\pi_1}$					
w_{1F}	$w_{1F,\pi_1}x_{\pi_1}$					
w_2			$w_{2\pi_3}x_{\pi_3}$			

Table 2. Balance table for the supply chain stage in Figure 4.

Balance tables referring to the monetary flows among processes can be similarly computed.

6. EIO models for logistics flows in a supply chain stage

The flows of materials among processes and their final outputs require to be conveyed from origins to destinations. Therefore, in order to effectively describe and analyse the network of production processes, transportation has to be considered. For the sake of simplicity, a supply chain stage is analyzed.

In EIO models, transportation can be modelled as: i) a production process or ii) a primary input, which provides other processes with inputs consisting of logistics service, in terms of the distance covered to convey all main products to their destinations.

In particular, the transportation system can be modelled as a single production process (T) that supplies all the other production processes involved in the supply chain stage and requires inputs such as workforce, fuel, and energy, as shown in Figure 5.

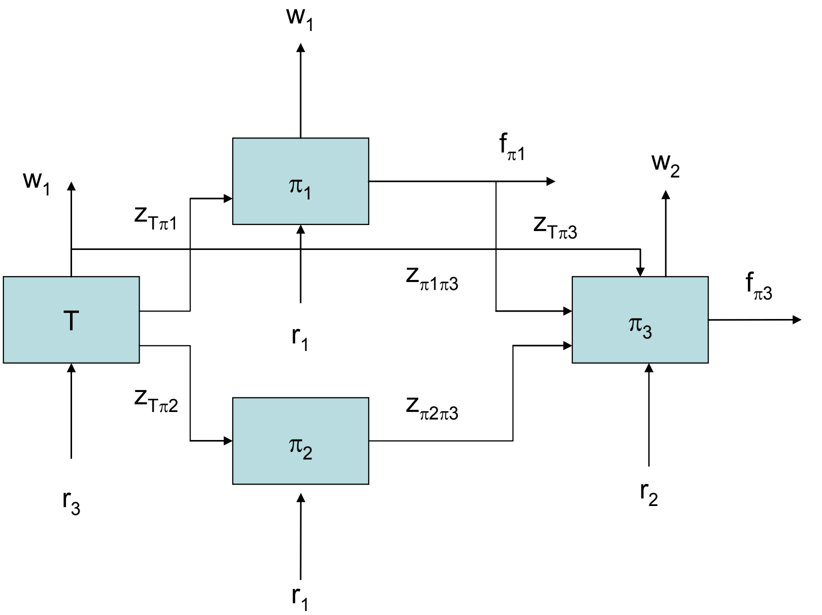


Fig. 5. Inputs and outputs of production processes, including transportation.

Following this approach, the balance table can be represented as shown in Table 3.

Processes	π1	π2	π3	T	f0	x0
π1			$a_{\pi_1\pi_3}x_{\pi_3}$		f_{π_1}	x_{π_1}
π2			$a_{\pi_2\pi_3}x_{\pi_3}$			x_{π_2}
π3					f_{π_3}	x_{π_3}
T	$a_{T\pi_1}x_{\pi_1}$	$a_{T\pi_2}x_{\pi_2}$	$a_{T\pi_3}x_{\pi_3}$			x_T
Primary inputs						
r1	$r_{1\pi_1}x_{\pi_1}$	$r_{1\pi_2}x_{\pi_2}$				
r2			$r_{1\pi_3}x_{\pi_3}$			
r3				$r_{3T}x_T$		
Wastes and by-products						
w1	$w_{1\pi_1}x_{\pi_1}$			$w_{1T}x_1$		
w2			$w_{1\pi_3}x_{\pi_3}$			

Table 3. Balance table for the supply chain stage in Figure 5.

Logistics flows can be also modelled adopting a disaggregate approach, i.e. a single transportation process can be associated to each origin and destination materials flow (Figure 6). Moreover, transportation processes can also be distinguished on the basis of the logistics flow, if materials and the trucks load capacity are different.

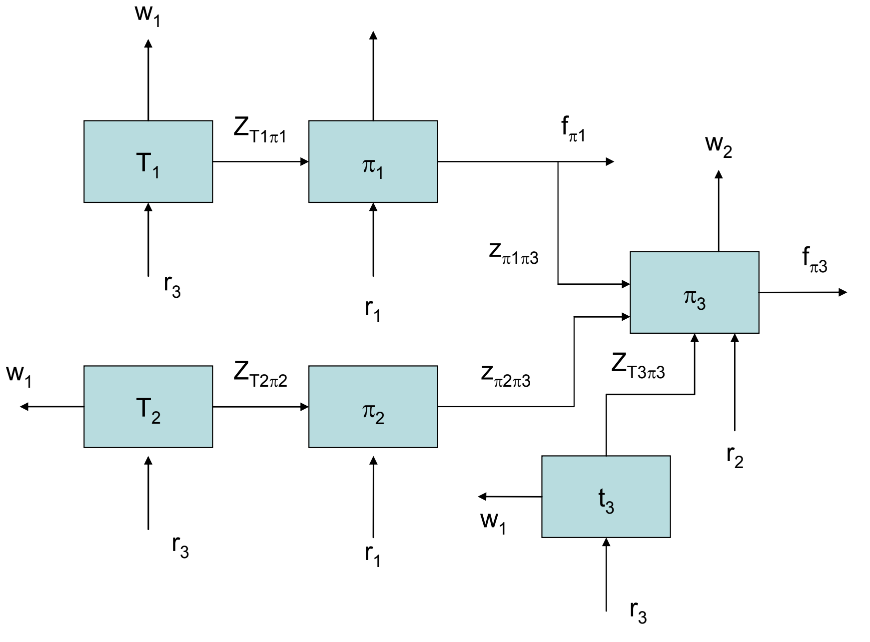


Fig. 6. Inputs and outputs of production processes, including transportation for each origin-destination flow.

In this case, the balance table is reported in Table 4.

Processes	π_1	π_2	π_3	T_1	T_2	T_3	f_0	x_0
π_1			$a_{\pi_1\pi_3}x_{\pi_3}$				f_{π_1}	x_{π_1}
π_2			$a_{\pi_2\pi_3}x_{\pi_3}$					x_{π_2}
π_3							f_{π_3}	x_{π_3}
T_1	$a_{T_1\pi_1}x_{\pi_1}$							x_{T_1}
T_2		$a_{T_2\pi_2}x_{\pi_2}$						x_{T_2}
T_3			$a_{T_3\pi_3}x_{\pi_3}$					x_{T_3}
Primary inputs								
r_1	$r_{1\pi_1}x_{\pi_1}$	$r_{1\pi_2}x_{\pi_2}$						
r_2			$r_{2\pi_3}x_{\pi_3}$					
r_3				$r_{3T_1}x_{T_1}$	$r_{3T_2}x_{T_2}$	$r_{3T_3}x_{T_3}$		
Wastes and by-products								
w_1	$w_{1\pi_1}x_{\pi_1}$			$w_{1T_1}x_{T_1}$	$w_{1T_2}x_{T_2}$	$w_{1T_3}x_{T_3}$		
w_2			$w_{2\pi_3}x_{\pi_3}$					

Table 4. Balance table for the supply chain stage in Figure 5.

As stated at the beginning of the section, transportation can be alternatively modelled as a primary input. Therefore, no inputs, wastes, and by-products related to transportation are considered. In Figure 7 and Table 5, the supply chain stage and the balance table referred to this case are represented.

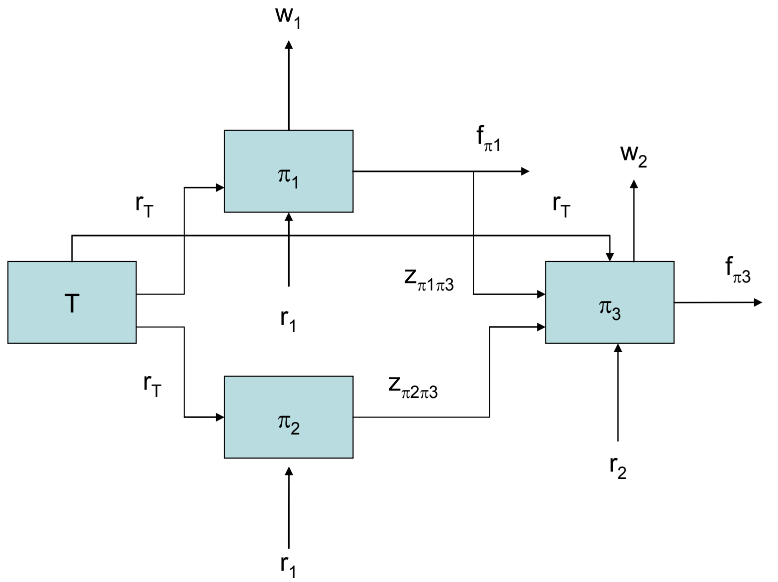


Fig. 7. Inputs and outputs of production processes, including transportation as a primary input.

Processes	π_1	π_2	π_3	f_0	x_0
π_1			$a_{\pi_1\pi_3} x_{\pi_3}$	f_{π_1}	x_{π_1}
π_2			$a_{\pi_2\pi_3} x_{\pi_3}$		x_{π_2}
π_3				f_{π_3}	x_{π_3}
Primary inputs					
r_1	$r_{1\pi_1} x_{\pi_1}$	$r_{1\pi_2} x_{\pi_2}$			
r_2			$r_{2\pi_3} x_{\pi_3}$		
T	$r_{T\pi_1} x_{\pi_1}$	$r_{T\pi_2} x_{\pi_2}$	$r_{T\pi_3} x_{\pi_3}$		
Wastes and by-products					
w_1	$w_{1\pi_1} x_{\pi_1}$				
w_2			$w_{2\pi_3} x_{\pi_3}$		

Table 5. Balance table for the supply chain stage in Figure 7.

Also in this case, logistics flows can be modelled using a disaggregate approach, distinguishing different transportation inputs, according to the origin-destination materials flow.

The proposed EIO models can be adopted to analyse the logistics flows of a supply chain stage located in a specific geographical area. Therefore, we can consider a set of production processes belonging to $\Pi_G \cap \Pi_{sc}$.

However, these models are not able to make distinction about primary inputs, wastes, by-products, and outputs transportation. To make distinction, we add virtual processes located within the considered geographical area G or on its boundaries, depending on where the primary input is available (within or outside the area). Each virtual process, corresponding to a specific primary input, is characterised by geographical information about its location and it has an output that can be transported to all the production processes requiring that input. For each virtual process no inputs are allowed from the production processes.

Let us consider *h* virtual processes corresponding to *s* primary inputs from outside the geographical system. Then, we introduce Z_0^* and x_0^* as the matrix of domestic intermediate deliveries and the vector of gross outputs including the *h* virtual processes, respectively. If *n* processes are distinguished, including transportation processes, the matrix Z_0^* is of size (*n*+*h*) x (*n*+*h*) and the vector x_0^* is (*n*+*h*) x 1.

Define the intermediate coefficient matrix *A** as follows:

$$A^* \equiv Z_0^* \hat{x}_0^{-1*}$$

The apex * can be extended with similar meaning to all variables as needed.

The same approach can be used to model wastes and by-products transportation.

Let us consider two production processes, π_j and π_k, two virtual processes, v_1 and v_2, corresponding to two primary inputs, r_1 and r_2, respectively, and the process T having, for the sake of simplicity, no intermediate deliveries from processes π_j and π_k, and no primary inputs. Moreover, each process, primary input, waste, and by-product is characterised by a single location, and no imports are considered from outside G, unless the two primary inputs. Finally, let us assume that the final demand $f_{\pi k}$ is delivered to the geographical destination A and the waste w_1 is destined to the geographical destination B (Figure 8).

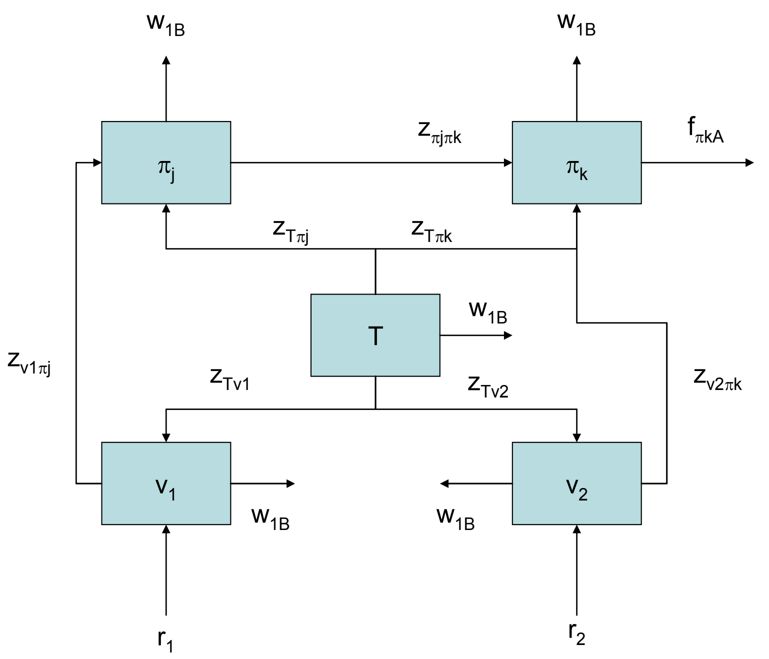


Fig. 8. Inputs and outputs of production processes, including transportation and virtual processes.

In Table 6 the balance table referred to the supply chain stage depicted in Figure 8 is reported.

Process	π_j	π_k	T	v_1	v_2	f_{0A}	x_0
π_j		$a_{\pi_j\pi_k} x_{\pi_k}$					x_{π_j}
π_k						$f_{\pi_k A}$	x_{π_k}
T	$a_{T\pi_j} x_{\pi_j}$	$a_{T\pi_k} x_{\pi_k}$					x_T
v_1	$a_{v_1\pi_j} x_{\pi_j}$						x_{v_1}
v_2		$a_{v_2\pi_k} x_{\pi_k}$					x_{v_2}
Primary inputs							
r_1				$r_{1v_1} x_{v_1}$			
r_2					$r_{2v_2} x_{v_2}$		
Wastes and by-products							
w_{1B}	$w_{1B,\pi_j} x_{\pi_j}$	$w_{1B,\pi_k} x_{\pi_k}$	$w_{1B,T} x_T$	$w_{1B,v_1} x_{v_1}$	$w_{1B,v_2} x_{v_2}$		

Table 6. Balance table for the supply chain stage in Figure 8.

As previously explained, the main output of process T is represented by the total distance covered by transportation means to deliver products from origins to destinations. Thus, considering the distance between the processes, as provided in Table 7, we can compute, for instance, $z_{T\pi_j}$ as: $z_{T\pi_j} = d_1 \cdot \frac{z_{\pi_j\pi_k}}{C}$

where C represents the transportation means' load capacity.

From/to	π_j	π_k	v_1	v_2
π_j		d_1	d_2	d_3
π_k	d_1		d_4	d_5
v_1	d_2	d_4		d_6
v_2	d_2	d_4	d_6	

Table 7. Distance between processes.

Moreover, the distances between the processes can be distinguished into the different paths covered to convey products, which are constituted by the track connecting the processes, as shown in Table 8.

From/to	π_j	π_k	v_1	v_2
π_j		θ_1-θ_2	θ_1-θ_3	θ_1-θ_4
π_k	θ_2-θ_1		θ_2-θ_3	θ_2-θ_4
v_1	θ_3-θ_1	θ_3-θ_2		θ_3-θ_4
v_2	θ_4-θ_1	θ_4-θ_2	θ_4-θ_3	

Table 8. Paths covered by transportation means.

Therefore, $z_{T\pi_j}$ results:

$$z_{T\pi_j} = \left(d_{\theta_1} + d_{\theta_2}\right) \cdot \frac{z_{\pi_j\pi_k}}{C} = d_{\theta_1} \cdot \frac{z_{\pi_j\pi_k}}{C} + d_{\theta_2} \cdot \frac{z_{\pi_j\pi_k}}{C}$$

where d_{θ_1} and d_{θ_2} represent the lengths of the track $\theta_1$and θ_2, respectively.

On the basis of these considerations, the transportation process can be modelled adopting a disaggregate approach into *l* processes (θ_k , k=1,…, l), corresponding to the *l* tracks covered by transportation means to deliver products. In this case, the balance table reported in Table 4 can be described as in Table 8.

Process	π_j	π_k	θ_1	θ_2	θ_3	θ_4	v_1	v_2	f_{0A}	x_0
π_j		$a_{\pi_j\pi_k}x_{\pi_k}$								x_{π_j}
π_k									$f_{\pi_k A}$	x_{π_k}
θ_1	$a_{\theta_1\pi_j}x_{\pi_j}$						$a_{\theta_1 v_1}x_{v_1}$			x_{θ_1}
θ_2	$a_{\theta_2\pi_j}x_{\pi_j}$	$a_{\theta_2\pi_k}x_{\pi_k}$						$a_{\theta_2 v_2}x_{v_2}$		x_{θ_2}
θ_3							$a_{\theta_3 v_1}x_{v_1}$			x_{θ_3}
θ_4								$a_{\theta_4 v_2}x_{v_2}$		x_{θ_4}
v_1	$a_{v_1\pi_j}x_{\pi_j}$									x_{v_1}
v_2		$a_{v_2\pi_k}x_{\pi_k}$								x_{v_2}
Primary inputs										
r_1							$r_{1v_1}x_{v_1}$			
r_2								$r_{2v_2}x_{v_2}$		
Wastes and by-products										
w_{1B}	$w_{1B,\pi_j}x_{\pi_j}$	$w_{1B,\pi_k}x_{\pi_k}$	$w_{1B,\theta_1}x_{\theta_1}$	$w_{1B,\theta_2}x_{\theta_2}$	$w_{1B,\theta_3}x_{\theta_3}$	$w_{1B,\theta_4}x_{\theta_4}$	$w_{1B,v_1}x_{v_1}$	$w_{1B,v_2}x_{v_2}$		

Table 8. Balance table in the disaggregated representation of transportation processes.

Adopting the disaggregate approach, each transportation input is related to a given track θ_k (k=1,…, h). Then, each process delivering products to two or more final destinations can be distinguished according to their final destinations, maintaining the same geographical location. In fact, let us assume to have three production processes (π_1, π_2, and π_3) and two tracks (θ_1 and θ_2), with the balance table reported in Table 9, where θ_1 and θ_2 connect π_1 with π_2, and π_1 with π_3, respectively.

	π_1	π_2	π_3	F
π_1	0	z_{12}	z_{13}	0
π_2	0	0	0	f_2
π_3	0	0	0	f_3
θ_1	t_{11}	0	0	
θ_2	t_{21}	0	0	

Table 9. Balance table.

If f_2 increases, then z_{12} must increase and, consequently, x_1 increases. However, only t_{11} must increase to permit to serve more output of π_1 to π_2. Then, the process p_1 has to be distinguished into p_{12} and p_{13}, according to its final destinations, as shown in Table 10.

	π_{12}	π_{13}	π_2	π_3	F
π_{12}	0	0	z_{12}	0	0
π_{13}	0	0	0	z_{13}	0
π_2	0	0	0	0	f_2
π_3	0	0	0	0	f_3
θ_1	$t_{1,12}$	0	0	0	
θ_2	0	$t_{2,13}$	0	0	

Table 10. Balance table in the case of process π_1 distinguished according to its final destinations.

Now, $t_{1,12}$ and $t_{2,13}$ represent the distance covered by the transportation mean to convey the output of process π_1 to the process π_2 through the track θ_1, and the output of process π_1 to the process π_3 through the track θ_2, respectively.

Finally, it is important to compute and geo-refer the pollution caused by the transportation means along the tracks. Then, let us define the waste vector w_0^θ of size *m* x 1 and the *m* x *h* matrix W^θ of output coefficients with element w_{kj}^θ denoting the output of waste type k (1,…,m) per unit of transportation input j.

It results:

$$w_0^\theta = W^\theta \theta_0$$

Then, pollution caused by transportation can be easily computed and geo-referred.

7. EIO models and logistics services markets

As stated in Section 1, logistics services are generally managed by logistics providers (3PL), which own the key competencies and capabilities necessary to assure their effectiveness and efficiency.

Referring to transportation services, three distinct actors can be identified: i) suppliers, which have to deliver products to one or more customers; ii) customers, which require products from one or more suppliers; iii) 3PL providers, which provide the transportation service and coordinates logistics flows between suppliers and customers.

The interaction among these different actors represents what is generally defined as a logistics services market. In particular, it can be composed by actors belonging both to the same companies, or different and independent ones.

Adopting EIO models it is possible to describe these markets modelling each actor as a different production process.

Let us consider three suppliers (S_1, S_2, and S_3), two customers (C_1, C_2, and C_3), and two logistics providers, which are represented by two distinct transportation processes (T_1, and

T_2), forming the logistics services market depicted in Figure 9. For the sake of simplicity, no primary inputs, wastes, and by-products are considered.

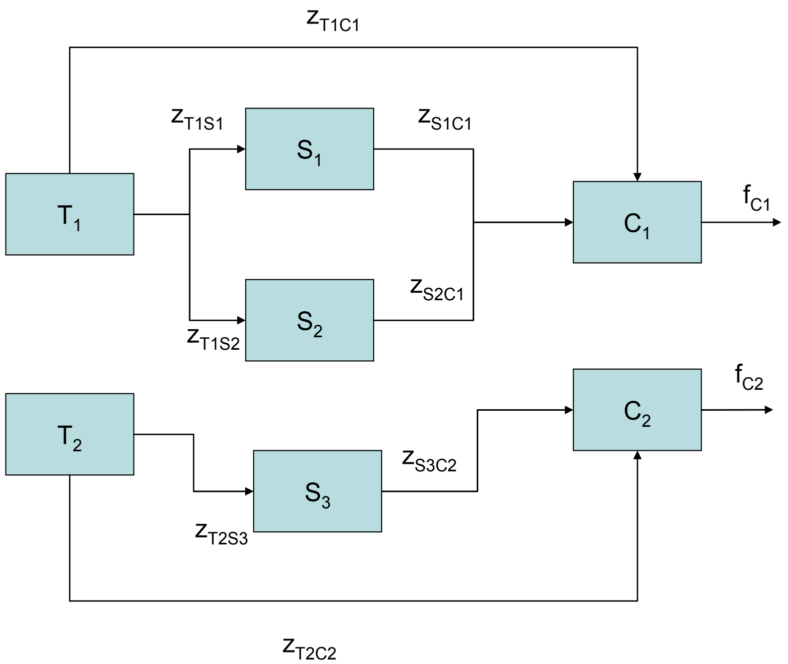


Fig. 9. A logistics services market.

The EIO model permits to account the demand and supply of logistics services, as shown in the balance Table 11.

Process	S_1	S_2	S_3	C_1	C_2	T_1	T_2	f_0	x_0
S_1				$z_{S_1C_1}$					x_{S_1}
S_2				$z_{S_2C_1}$					x_{S_2}
S_3					$z_{S_3C_2}$				x_{S_3}
C_1								f_{C_1}	x_{C_1}
C_2								f_{C_2}	x_{C_2}
T_1	$z_{T_1S_1}$	$z_{T_1S_2}$		$z_{T_1C_1}$					x_{T_1}
T_2			$z_{T_1S_3}$		$z_{T_1C_2}$				x_{T_2}

Table 11. Balance table for the logistics services market in Figure 9.

This representation can be useful also as a planning tool, to analyse the economic and environmental performance of logistics services markets, investigating how it is affected by the different degree of cooperation among the actors. In fact, different market organizations can be proposed and investigated, on the basis of the collaboration degree of the actors, and following approaches aimed at minimizing the number of trips, such in as in the case of consolidation strategies, and at creating logistics networks specialized, for instance, by geographical areas, types of product, and services.

8. Discussion and conclusions

The present paper has proposed the use of EIO models to describe and analyse logistics flows to support managers and policy makers in the definition of policies for their management and coordination. In particular, different approaches and models have been proposed.

Two main perspectives have been used to analyse logistics flows, such as a spatial and operational one, pursuing different goals. In fact, geo-referring the production processes belonging to a supply chain, and their inputs and outputs, spatial-oriented analyses can be performed in order to deal with issues related to traffic congestion, pollutant emissions, transportation infrastructures, and work force availability. Therefore, on the basis of these analyses policies aimed at improving the transportation sustainability and reducing its negative impact on the environment can be identified.

Differently, the adoption of an operational framework of analysis can permit to describe logistics flows involved in specific supply chains, or, more in detail, managed by specific actors, in order to analyse and improve logistics economic performance. For instance, solutions aimed at consolidating the flows and reducing the number of trips and the transportation costs can be achieved.

Moreover, transportation has been modelled both as a process and as a primary input. The difference between these two approaches depends on the inclusion of inputs required and wastes produced by transportation. In fact, in the former all the transportation inputs, such as workforce and fuel, and its pollutant emissions are considered. Therefore, this approach can be useful for 3PL to evaluate the economic and environmental performance of their activities. On the contrary, the modelling of transportation as a primary input can be used to investigate its impact on the supply chains. In fact, in this case the model can be a suitable accounting and planning tool for actors representing the demand of logistics services, in order to analyse, for instance, how logistics affects their profit and operations.

Transportation has been also modelled at an aggregate and disaggregate level. The aggregate level of analysis permits to investigate managerial-oriented issues, giving a holistic view of the transportation to evaluate its global role on the supply chains' efficiency and effectiveness. Differently, the disaggregate approach permits a more in-depth analysis, since it describes all the tracks covered by transportation means, thus allowing policy makers to propose actions for the logistics management according to a social and environmental perspective.

Finally, EIO models have been used to represent logistics service markets, in order to account and analyse the demand and supply of transportation. This can be useful for both logistics providers and customers to organize markets, on the basis of different criteria (e.g. consolidation, specialization, geographical areas) and to improve their performance.

The present study contributes to extend the existing framework on logistics management, providing a set of complete and complementary tools, based on the use of EIO models, able

to analyse the problem according to different perspectives and point of views. The power of the described methodology is strictly related to the possibility to investigate the impact of logistics on the whole performance systems, thanks to the direct and indirect relationships and interdependences among production processes.
Further researches should be devoted to apply the models to actual cases, in order to show their effectiveness as both accounting and planning tools and to strength the relevance of the proposed approach.

9. References

Albino, V., De Nicolò, M., Garavelli, A.C., Messeni Petruzzelli, A., Yazan, D.M. (2007) Rural development and agro-energy supply chain. An application of enterprise input-output modelling supported by GIS, *16th International Input-Output Conference*, July 2-6, Istanbul, Turkey.

Albino, V., Dietzenbacher, E., Kühtz, S. (2003) Analyzing Material and Energy Flows in an Industrial District using an Enterprise Input-Output Model, *Economic Systems Research*, Vol. 15, pp. 457-480.

Albino, V., Izzo, C., Kühtz, S. (2002) Input-Output Models for the Analysis of a Local/Global Supply Chain, *International Journal of Production Economics*, Vol. 78, pp. 119-131.

Albino, V., Kuhtz, S., Messeni Petruzzelli, A., (2008) Analysing logistics flows in industrial clusters using an enterprise input-output model, *Interdisciplinary Information Sciences*, Vol. 14, pp. 1-17.

Carbonara, N., Giannoccaro, I., Pontrandolfo, P. (2001) Supply chains within industrial districts: theoretical framework, *International Journal of Production Economics*, Vol. 76, pp. 159-176.

Choi, T.Y Hong, Y. (2002) Unveiling the structure of supply networks: case studies in Honda, Acura, DaimlerChrysler, *Journal of Operations Management*, Vol. 20, pp. 469–493.

Commission of the European Communities (2001) *European Transport Policy for 2020: Time to Decide* (Luxembourg: Office for Official Publications of the European Communities).

European Energy Agency (2004) *EEA signals 2004, A European Environment Agency Update on Selected Issues* (Copenhagen: European Energy Agency).

Grubbström, R.W., Tang, O. (2000) An Overview of Input-Output Analysis Applied to Production-Inventory Systems, *Economic Systems Research,* Vol. 12, pp. 3-25.

Gunasekaran, A., Cheng, T.C.E. (2008) Special issue on logistics: new perspectives and challenges, *Omega*, Vol., 36, pp. 505-508.

Hertz S., Alfredsson M. (2003) Strategic development of third party logistics provider, *Industrial Marketing Management*, Vol. 30, pp. 139-149.

Jayaraman, V., Ross, A. (2003) A simulated annealing to distribution network design and management, *European Journal of Operational Research*, Vol. 144, pp. 629-645.

Kim, C., Yang, K.H., Kim, J. (2008) A strategy for third-party logistics systems: a case analysis using the blue ocean strategy, *Omega*, Vol. 36, pp. 522-534.

Lai, K.H., Cheng, T.C.E. (2003) A Study of the Freight Forwarding Industry in Hong Kong, *International Journal of Logistics: Research and Applications*, Vol. 7, pp. 72-84.

Lai, K.H., Ngai, E.W.T., Cheng, T.C.E. (2004) An empirical study of supply chain performance in transport logistics, *International Journal of Production Economics*, Vol. 87, pp. 321-331.

Leonief, W.W. (1941) *The Structure of the American Economy*. Oxford University Press, New York.

Lin, X., Polenske, K.R. (1998) Input-Output Modelling of Production Processes for Business Management, *Structural Change and Economic Dynamics,* Vol. 9, pp. 205-226.

Marangoni, G., Colombo, G., Fezzi, G. (2004) Modelling Intra-Group Relationships, *Economic Systems Research*, Vol. 16, pp. 85-106.

Marangoni, G., Fezzi, G. (2002) I-O for Management Control: The Case of GlaxoSmithKline, *Economic Systems Research*, Vol. 14, pp. 245-256.

Meixell, M.J., Gargeya, V.B. (2005) Global supply chain design: a literature review and critique, *Transportation Research*, Vol. 41, pp. 531-550.

Miller, R.E., Blair, P.D. (1985) *Input Output Analysis: Foundations and Extensions.* Prentice Hall, Englewood Cliffs.

Quariguasi Frota Neto, J., Bloemhof-Ruwaard, J.M., van Nunen, J.A.E.E., van Heck, E. (2008) Designing and evaluating sustainable logistics networks, *International Journal of Production Economics*, Vol. 111, pp. 195-208.

Stephen, T. (2004) Further reflections on the supply chain complexity gap, *Supply Chain Europe*, Vol. 13, pp. p. 38.

Thierry, M., Salomon, M., van Nunen, J.A.E.E., van Wassenhove, L. (1995) Strategic issues in product recovery management, *California Management Review*, Vol. 37, pp. 114-135.

Van der Veen, A., Logtmeijer, C. (2003) *How vulnerable are we for flooding? A GIS approach*, (No. EUR Report 20997 EN). Office for Official Publications of the European Communities, European Commission, Brussels.

Wang, Y., Chandra, L.S. (2007) Using e-business to enable customised logistics sustainability, *International Journal of Logistics Management*, Vol. 18, pp. 402-419.

Xu, L, Beamon, B.M. (2006) Supply chain coordination and cooperation mechanisms: an attribute-based approach, *The Journal of Supply Chain Management*, Winter 2006, pp. 4-12.

Yi, W., Ozdamar, L. (2007) A dynamics logistics coordination model for evacuation and support in disaster response activities, *European Journal of Operational Research*, Vol. 179, pp. 1177-1193.

Zhan, G., Polenske, K.R., Ning, A. (2005) Evaluation of yellow-dust storms and countermeasure policies in North China using regional input-output models, *15th International Input-Output Conference*, 27 June – 1 July, Beijing, P.R. of China.

4

Dynamic Analysis and Control of Supply Chain Systems

Alejandro Rodríguez-Angeles, América Morales Díaz and Arturo Sánchez
Centro de Investigación y de Estudios Avanzados
Depto. Ing. Electrica, Unidad Saltillo & Unidad Guadalajara
México

1. Introduction

Globalization has changed in the last decade the shape of world trade, from independent and local markets to highly interrelated value chains demanding a large variety of goods with high quality standards. A supply chain is thus understood as a highly interconnected demand network composed of different stages or nodes that include raw materials, distributors of raw materials, manufacturers or producers, distributors of manufactured products, retailers and customers. Between these stages or nodes there is an interchange of information (orders) and flows of materials. The above mentioned elements of a supply chain and their relations give rise to complex structures, whose interactions affect the performance of the entire system.

Nowadays the markets require flexibility, speed and productivity in order to satisfy an environmentally-conscious consumer demanding a larger variety of manufactured goods. These trade conditions impose low costs and effectiveness in production. Therefore, most supply chains strive for minimizing raw material and finished products inventories in fast distribution networks (Fujimoto, 2002), (Steidtmann, 2004) while fulfilling astringent manufacturing rules. This tendency has resulted on what is called synchronization of production and distribution at the supply chain. Supply chain synchronization occurs when the consumer business world is linked together by technology, making each of the constitutive parts, consumers, suppliers, producers, associates and distributors synchronize with the whole. Thus, when the consumer places a request for an end-product, there is a synchronized retailer or distributor there to deliver it. For example (Koudal, 2003) studied the demand and supply dynamics in the automotive value chain, yielding flexibility and fast consumer respond. In textile industry, TAL Appareal Group applied electronic and communications platforms to evolve into a flexible manufacturer, growing from a single local textile mill to a global multinational company (Koudal & Wei-teh, 2005). From an economic point of view the potential impact of performance improvements on production systems is tremendous.

The arising dynamics complexity of highly synchronized supply chains poses demanding control challenges for their operation and management, for which representative models are required. Two decision levels can be considered in the supply chain operation: the first is the tactical and is referred to the decision making process that optimizes the supply chain performance, and the second involves the operational activities. A well-constructed supply chain model, either for the tactical or operative levels, must preserve relevant information associated with the flow and transformation of raw materials to finished products and their distribution to the consumers.

Several approaches in supply chain modelling have been considered, from quantitative models (Tayur et al., 1998) to fuzzy models (Petrovic et al., 1999). Extensive overviews can be found in (Daganzo, 2002) and (Shapiro, 2007). Most of the proposed supply chain models are deterministic, nonetheless there are a few that takes into account demand uncertainty in an stochastic framework (e.g. Lababidi et al., 2004) .

Supply chain management has been extensively studied. Most works focused on the tactical area and mainly deal with production planning and scheduling. In planning and scheduling of supply chains one of the goals is to determine the inventory levels to satisfy the market demands in a timely and cost effective manner. Inventory levels planning and orders handling are addressed in (Perea et al., 2001) and (Cetinkaya & Lee, 2000). In (Perea et al., 2001) a dynamic model based on balances of inventories and orders is presented, together with a decentralized decision making process that considers production policies. The same model is used in (Perea et al., 2003) to develop a model predictive control strategy, improving the planning forecast in a time horizon.

Some others works are focused on the supply chain dynamics and its operational level. In (Lin et al. 2004) an analysis and control for bullwhip effect is presented based on z-transform. By using traffic dynamics (Nagatani & Helbing, 2003) proposed a continuous time dynamic model to represent the behaviour of the inventories and the production rates. Moreover a control strategy and production policies to bring an unstable system into the stable region are proposed. The aim is to control the production rates, while inventories are, to some extend, free to evolve in a bounded region, which may result in unsuitable stock levels regarding a reliable and competitive performance of the system. Analytical conditions for absolute and convective instabilities were presented in (Helbing et al., 2004). Finally, for the dynamic management of supply chain networks, (Dunbar & Desa, 2005) demonstrated the application of nonlinear distributed model predictive control strategies. (Lefeber, 2004) also considers traffic flow theory to propose a supply chain model based on differential equations, taking into account delays on handling orders and inventory level dynamics.

This chapter is focused on the operational activities of the supply chain dynamics. Similarly to those works developed in (Perea et al., 2001), (Helbing et al., 2004) and (Lefeber, 2004), dynamic models for multipurpose systems are proposed considering production ratios. First a dynamic model dealing only with material flows is introduced, and then it is extended to consider order flows and delays on handling materials and orders. In agreement with physical limitations, bounds in the inventory levels and flow rates are taken into account. Once the models are introduced, a bounded and smooth controller is proposed. This controller regulates the inventory levels while synchronizing the flows at the whole supply chain system. Closed loop convergence to desired inventory levels and production rates stability is proved. The proposed modelling and control techniques are applied to two simulation case studies: a petrochemical multiproduct plant and a crude oil blending and distribution system. The chapter closes with some general conclusions and trends in supply chain systems.

2. Supply chain dynamic models

A supply chain model must capture relevant activities associated with the inventory levels, and the flow and transformation of good from the raw material stage to the final costumer.

The entities that may form a supply chain, i.e. suppliers, distributors, producers, etc., they can be classified as producer or non producer nodes. The main difference among them is that producer nodes may vary their production ratio, thus varying its inventory level. Non producer nodes may vary their inventory levels only by ordering from upstream entities.

Although inventory and material flow dynamics may be enough for modelling some supply chain systems, there are situations in which information flows, such as orders, are needed to properly represent the system dynamics. This section presents both types of models, one entirely based on inventory levels and material flows, and one that considers orders handling.

2.1 Linear supply chain systems: material flow rate model

The models are based on macroscopic mass balance equations. Therefore the mass variation is equal to the difference between the material inflow and outflow rate (see Figure 1) and is represented by equation (1). These models represent systems in which the material delivery flow rate corresponds exactly and without delay to the demanded one. Notice that the above considerations render linear dynamic models based on differential equations.

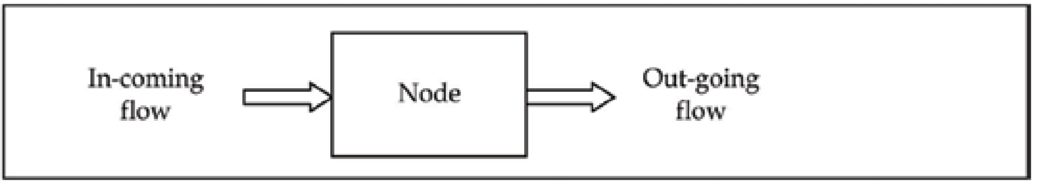


Fig. 1. Mass balance representation in a supply chain node.

By considering the dynamics on the properties of interests, i.e. inventory levels and material flows, a supply chain consists from n-nodes, with inventory levels N_i, for $i = 1,\ldots,n$, and depending if the node is a producer or not, a production or incoming flow rate λ_i; as it is depicted in Figure 2.

The change in the inventory level is represented by the difference between the production or incoming rate λ_i, depending whether it is a producer or non producer node respectively, and the delivery flow rate $\lambda_{d\_p,i}$.

$$\frac{dN_i}{dt} = \lambda_i - \lambda_{d\_p,i} \tag{1}$$

Where $\lambda_{d\_p,i}$ represents the total demand of products to node i, and considers the demand of all the r nodes requiring products or material from node i, with individual demanding rates λ_i and with the product ratio $F_{i,j}$, this is

$$\lambda_{d\_p,i} = \sum_{j=1}^{r} F_{i,j}\lambda_j \tag{2}$$

The product ratio $F_{i,j}$ allows modelling multipurpose or multiproduct systems. The product ratios represent the amount of product i that is required by the j-node to produce a unit of product. The product ratio allows changing recipes or product proportion, such that different products can be easily considered. Furthermore, these product ratios define the synchronized behaviour among the material flows at the supply chain.

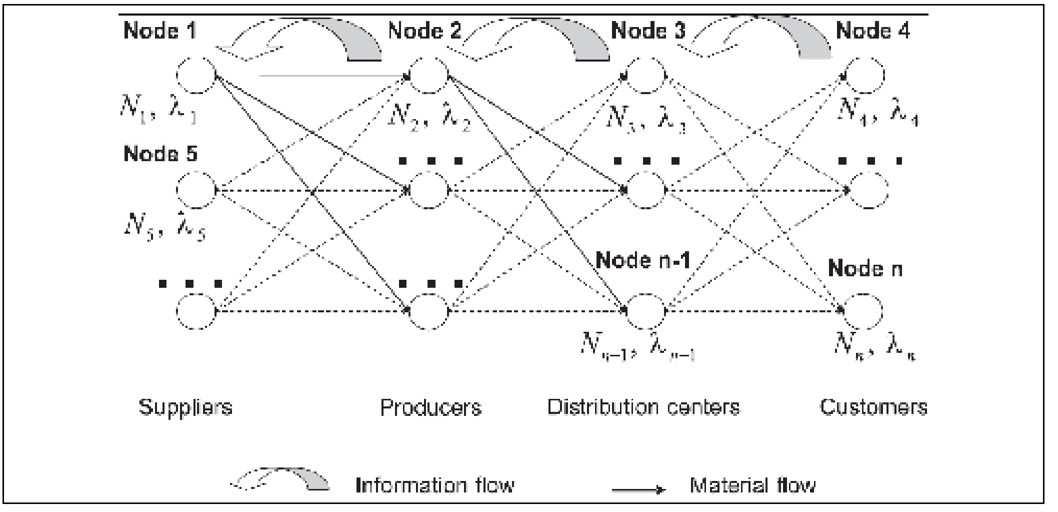


Fig. 2. Inventories, information and material flow in a supply chain.

The production or incoming rate λ_i constitutes the control action to vary the inventory dynamics in a node-i, whether it is a producer or non producer node respectively.
For a non producer node its incoming rate corresponds only to the product or materials that are received from its suppliers. Contrary, in a producer node the production rate varies accordingly to production policies and posses dynamical behaviour. In a producer node a change in the production rate involves several activities that require an adaptation time T_i. Moreover, if W_i denotes the control action that varies the production rate λ_i in a producer node, then its dynamics can be represented by

$$\frac{d\lambda_i}{dt} = \frac{1}{T_i}(W_i - \lambda_i) \tag{3}$$

Notice that because of physical or operational boundedness of the system the inventories, production and incoming rates, must be bounded accordingly, this is

$$\begin{aligned} N_{i,\min} \le N_i \le N_{i,\max} \\ 0 \le \lambda_i \le \lambda_{i,\max} \end{aligned} \tag{4}$$

where $N_{i,\min}, N_{i,\max}$ are the minimum and maximum allowed inventory level respectively, and $\lambda_{i,\max}$ is the maximum production or incoming rate.
In case of producer nodes, their dynamics includes inventory levels and production rate, thus, they are modelled by equations (1), (2) and (3). Meanwhile for non producers only inventory levels dynamics (1) and (2) are considered, which vary thru ordering products or material from upstream nodes.
The model represented by equations (1)-(4) disregard delivering delays, nevertheless, the adaptation time T_i plays a role that induces a kind of delay. The adaptation time parameter T_i allows reproducing bullwhip phenomena among other phenomena in this type of model.

At model (1), (2), (3) and (4) raw materials storage at a producer node is consider as an independent entity, which renders extra degrees of freedom to the system regarding control policies. This consideration establishes an advantage and at the same time a challenge that allows regulating inventory levels and simultaneously supply chain synchronization.
The consumer demand is the variable that imposes the policy operation and is traduced to flow rates for each node, taking into account the production ratios $F_{i,j}$. For example, in the automotive supply chain several types of cars have to be built, the net flow of finished cars per time unit is the costumer demand. But at upstream entities of the supply chain this is traduced in the corresponding part rate, e.g. 4 wheels and one motor per car, and so on.

2.2 Nonlinear supply chain systems: orders flow model

A supply chain as that depicted in Figure 2 includes information flows, particularly orders among downstream and upstream entities. In some situations such information flows have to be taken into account, either to strive for performance improvement or because not all orders are fulfil immediately, then an unattended orders balance appears.
Accumulation of material N_i and unattended orders O_i can be modelled similarly as balances between incoming and outgoing flows as follows

$$\frac{dN_i}{dt} = \lambda_i - y_i \tag{5}$$

$$\frac{dO_i}{dt} = \delta_i - y_i \tag{6}$$

Where y_i represents the shipped product (attended orders) and δ_i represents the total product demanding rate, which correspond to all orders requiring product or material from the ith-node. Similarly to the total product demand (2) product ratios $F_{i,j}$ may be considered, this is

$$\delta_i = \sum_{j=1}^{r} F_{i,j}\lambda_j$$

In the above model representation (5), (6), when an order is put in node i, it will take an interval of time until it is attended, this is represented by a time delay parameter denoted by τ_i. Furthermore, following traffic flow theory (Lefeber, 2005) a shipping function that depends on the inventory level and the accumulated orders is introduced, this is

$$y_i = \frac{N_i}{(N_i + 1)} \frac{O_i}{\tau_i} \tag{7}$$

Equation (7) reflects the outgoing product rate in node i and it is bounded from below by $y_i \geq 0$, thus only positive flows are allowed. The outgoing product policy depends on inventory level N_i such that the product will be shipped only if there is in stock. Since (7) is a nonlinear function on N_i, then the resulting supply chain model is of nonlinear type.
The production rate λ_i in a producer node, as in the case of model (1), (2), (3) and (4), varies accordingly to production policies. Therefore the dynamics in the flow rate is the same as in

equation (3). For simplicity in case of producer nodes it is assumed that upstream nodes deliver the demanded product or material as soon as it is required for production; this is a producer has storage facilities which guarantee the required materials. Then these storage facilities are the ones that can have unattended orders from their suppliers. Furthermore, it is considered that producer nodes modify its production rate to deliver the material amount that is required such that orders accumulation is not an issue in a production facility. Then a producer node is modeled by equation (5), (3), and (7) with $y_i = \delta_i$.

For non producer nodes, their dynamics is given only by inventory levels and balance of unattended order, equations (6) and (7). Therefore accumulation of unattended orders appears only in non producer nodes.

3. Supply chain regulation, control strategies

When regulating the supply chain dynamics, the goal is to establish flow rates that guarantee that the inventory levels reach and keep a desired level, while a final costumer demand is satisfied. On the other hand, as it was mentioned the supply chain must satisfy physical constraints, such as bounded inventory levels, and maximum production or incoming flow rates, see equation (4). Therefore the control actions to regulate the supply chain must be designed taking into account such constraints.

This section presents two bounded control strategies for inventory regulation and flow rate synchronization. For these purposes PI techniques are considered to vary the production or incoming flow rates. Stability is proved by standard analysis based on linearization theory and pole placement techniques. First the controller for linear supply chain systems is presented and then it is extended for nonlinear supply chains.

3.1 Control strategy by flows variation

For producer nodes, variation on the production rate λ_i, whose dynamics is given by equation (3), is achieved by the control action or production policy W_i. Meanwhile for non producer nodes, inventory levels regulation is achieved by direct variation of their incoming flow rates λ_i. For each of these two classes of systems similar control actions are proposed as follows.

Since the production rate λ_i in (2, 3) is bounded by (4), then the control action must be bounded accordingly at a producer node, this is

$$0 \leq W_i \leq \lambda_{i,\max} \tag{8}$$

To satisfy the above constraint the control action W_i for a producer node is proposed as

$$W_i = \lambda_{i,\max}\left(2 - \frac{1}{1+e^{-\alpha_i(N_i - N_{c,i})}} - \frac{1}{1+e^{-\alpha_i N_{c,i}}}\right) \tag{9}$$

Meanwhile for a non producer node its inventory dynamics (1) is modified by its incoming rate λ_i, such that it is proposed to vary as

$$\lambda_i = \lambda_{i,\max}\left(2 - \frac{1}{1+e^{-\alpha_i(N_i - N_{c,i})}} - \frac{1}{1+e^{-\alpha_i N_{c,i}}}\right) \tag{10}$$

where $\lambda_{i,\max}$ is the maximum production or incoming rate in the node, see (4), α_i is a parameter that regulates the convergence rate of N_i. The larger α_i, the faster the convergence rate of N_i to its desired value $N_{d,i}$. However, too large values of α_i can induce instability. $N_{c,i}$ is proposed as a modified PI control, it acts as a nominal reference and forces N_i to a desired constant value $N_{d,i}$. $N_{c,i}$ allows introduction of several production policies as the ones presented in (Helbing, 2003).

$$N_{c,i} = N_{d,i} - K_{P,i}(N_i - N_{d,i}) - K_{I,i}\int (N_i - N_{d,i})dt \tag{11}$$

Where $K_{P,i}$ and $K_{I,i}$ are the positive proportional and integral control gains, respectively. The integral action renders a steady error equal to zero around the equilibrium point, while the proportional action regulates the convergence rate.

The controllers (9) and (10) use exponential functions to render a bounded control action, while allowing a smooth and fast convergence. The term in between parenthesis is bounded in [0,1], after multiplying this term by $\lambda_{i,\max}$, the physical and operational limitations on the production or incoming rate, see (4), are recovered.

Stability analysis

The goal of this section is to establish closed loop stability and synchronization conditions for the supply chain at an equilibrium point.

The controllers (9) and (10) induce nonlinearities in the closed loop system, so that, linearization techniques are considered for the stability analysis.

Theorem 1

The equilibrium point (λ_i^*, N_i^*) of the closed loop formed by a producer node (1 - 3) and (9) is given by $\lambda_i^* = \lambda_{d\_p,i}$ for the production rate, and $N_i^* = N_{d,i}$ for the inventory level. Meanwhile, the closed loop of a non producer node (1) and (10) implies the same equilibrium point as for a producer node, it is $\lambda_i^* = \lambda_{d\_p,i}$ and $N_i^* = N_{d,i}$.

Proof:

First for a producer node, from (1) and (3) it follows that the equilibrium conditions are

$$0 = \lambda_i^* - \lambda_{d-p,i} \qquad 0 = \frac{1}{T_i}\left(W_i - \lambda_i^*\right) \tag{12}$$

thus $\lambda_i^* = \lambda_{d\_p,i}$ and simultaneously $W_i = W_i(N_i^*) = \lambda_i^*$ such that $W_i(N_i^*)$ must be constant. By substitution of (9) and (11), it follows that because $\lambda_{i,\max}$ and α_i are constants, then $W_i(N_i^*)$ is constant if and only if $N_i^* = N_{d,i}$.

A similar analysis allows concluding the closed loop equilibrium point for non producer nodes. ■

Note that the equilibrium point of both producer and non producer nodes implies that $\lambda_i^* = \lambda_{d\_p,i}$, therefore λ_i synchronizes with the total demanding rate $\lambda_{d\_p,i}$, given by equation (2), and such that instantaneous consumption and supply chain synchronization are achieved. Simultaneously the inventory levels fulfills $N_i^* = N_{d,i}$, thus inventory regulation is obtained.

Theorem 2

The closed loop system formed by a producer node (1 - 3) with the controller (9) and (11) is locally asymptotically stable and converge to the equilibrium point $\lambda_i^* = \lambda_{d\_p,i}$ and $N_i^* = N_{d,i}$, if the gains α_i, $K_{P,i}$, and $K_{I,i}$ satisfy

$$\alpha_i \geq 1 \tag{13}$$

$$\zeta_1 \leq K_{P,i} < \zeta_2 \tag{14}$$

$$0 < K_{I,i} \leq \frac{1}{4}\frac{K_{P,i}}{T_i} \tag{15}$$

$$\zeta_1 = \left| \frac{\left(1+e^{-(\alpha_i N_{d,i})}\right)^2 \left(\alpha_i \lambda_{i,\max} T_i - 1\right)}{\alpha_i \lambda_{i,\max} T_i \left(e^{-(\alpha_i N_{d,i})} - 1\right)^2} \right|$$

$$\zeta_2 = \left| \frac{\left(1+e^{-(\alpha_i N_{d,i})}\right)^2}{\left(e^{-(\alpha_i N_{d,i})} - 1\right)^2} \right|$$

Furthermore, the above conditions ensure an overdamped closed loop system, (i.e. the eigenvalues are negative and purely real), avoiding large overshoots and keeping physical and operations constraints given by (4).

Proof:

In the equilibrium point $\lambda_i^* = \lambda_{d\_p,i}$, $N_i^* = N_{d,i}$, the closed loop is linearize around small deviations $\delta\lambda_i$ and δN_i as

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ \frac{1}{T_i} W_i'(N_i^*) & -\frac{1}{T_i} \end{bmatrix} x \tag{16}$$

where $x = \begin{bmatrix} \delta N_i & \delta\lambda_i \end{bmatrix}^T$ and whose eigenvalues $s_{1,2}$ are given by

$$s_{1,2} = \frac{-\frac{1}{T_i} \pm \sqrt{\frac{1}{T_i^2} + \frac{4}{T_i} W_i'(N_i^*)}}{2} \tag{17}$$

Then the closed loop is asymptotically stable and overdamped if the derivative of the control function $W_i'(N_i^*)$ fulfills

$$-\frac{1}{4T_i} \leq W_i'(N_i^*) < 0 \tag{18}$$

which imposes conditions on α_i, $K_{P,i}$, and $K_{I,i}$. Replacing (11) in (9) yields

$$W_i'(N_i^*) = \alpha_i \lambda_{i,\max} \left(-\frac{1}{4}\left(1 + K_{P,i} + tK_{I,i}\right) + \frac{\left(K_{P,i} + tK_{I,i}\right) e^{-(\alpha_i N_{d,i})}}{\left(1 + e^{-(\alpha_i N_{d,i})}\right)^2} \right) \tag{19}$$

where t represents the integration time. When $N_i^* = N_{d,i}$ there is not integral action, thus it can be taken $t = 0$, and from (19) and by algebraic manipulation it follows that sufficient conditions on $K_{P,i}$, for (18) being satisfied, are given by (14).
By defining the regulation error $e_i = N_i - N_{d,i}$, replacing it in (11) and taking first derivative with respect to time it is obtained

$$\frac{dN_{c,i}}{dt} = -\alpha_i K_{P,i} \frac{de_i}{dt} - K_{I,i} e_i \tag{20}$$

In the equilibrium point N_i becomes constant, i.e. $N_i^* \to N_{d,i}$, thus (20) equals to zero and by Laplace transform it is obtained the pole

$$\frac{dN_{c,i}}{dt} = -\alpha_i K_{P,i} \frac{de_i}{dt} - K_{I,i} e_i \tag{21}$$

Considering that the pole in (21) must verify condition (18) to limit the dynamics of the closed loop, then $K_{I,i}$ must satisfy condition (15).
Finally, since only $K_{P,i}$ through the condition (14) depends on α_i, it can be to some extend freely chosen. Thus for convenience and to obtain fast convergence it is taken that $\alpha_i > 1$. ■
Theorem 3
The closed loop system formed by a non producer node (1) with the controller (10) is locally asymptotically stable and converge to the equilibrium point $\lambda_i^* = \lambda_{d\_p,i}$ and $N_i^* = N_{d,i}$, if the gains α_i, $K_{P,i}$, and $K_{I,i}$ satisfy

$$\alpha_i \geq 1 \tag{22}$$

$$K_{P,i} < \left| \frac{\left(1 + e^{-(\alpha_i N_{d,i})}\right)^2}{\left(e^{-(\alpha_i N_{d,i})} - 1\right)^2} \right| \tag{23}$$

$$K_{I,i} > 0 \tag{24}$$

Furthermore, the above conditions ensure an overdamped closed loop system.
Proof:
It follows as for Theorem 2 ■

3.2 Control strategy by orders and flows handling

Due to the similarities between the models for linear supply chain systems (1 – 4) and the model that considers orders handling (5, 6, 3 and 7), the controllers (9, 10 and 11) still apply. This section shows that the unattended orders modify the closed loop equilibrium point and the value conditions on the control gains, but the control architecture (9, 10 and 11) is robust enough to deal with order handling and the implicit delays modelled by the shipping function y_i given by (7) and the production adaptation parameter T_i.
Stability analysis
The stability analysis follows as the one presented in Section 3.1, first the equilibrium point is determined and then stability and convergence conditions are provided.

Theorem 4

The equilibrium point (λ_i^*, N_i^*) of the closed loop formed by a producer node (5), (3), and (7) with $y_i = \delta_i$, and the controller (9) is given by $\lambda_i^* = \delta_i$ for the production rate, and $N_i^* = N_{d,i}$ for the inventory level. Meanwhile, the closed loop of a non producer node (5), (6), (7), and (10) posses the equilibrium point $\lambda_i^* = \delta_i$, $N_i^* = N_{d,i}$, $O_i^* = \left[\delta_i \tau_i \left(N_{d,i}+1\right)\right] / N_{d,i}$.

Proof:

Follows as the proof of Theorem 1, with the supply chain dynamics equal to zero. ■

Theorem 5

The closed loop system formed by a producer node (5), (3), and (7) with $y_i = \delta_i$, and the controller (9) is locally asymptotically stable and converge to the equilibrium point $\lambda_i^* = \delta_i$, $N_i^* = N_{d,i}$ if the gains α_i, $K_{P,i}$, and $K_{I,i}$ satisfy

$$\alpha_i > 0 \tag{25}$$

$$\zeta_3 \le K_{P,i} \tag{26}$$

$$0 < K_{I,i} < \frac{K_{P,i} W_i(N_{d,i})}{N_{d,i}(1+N_{d,i})} \tag{27}$$

$$\zeta_3 = \left| \frac{\left(1+e^{-(\alpha_i N_{d,i})}\right)^2 \left[\left(\alpha_i \lambda_{i,\max} N_{d,i}\left(N_{d,i}+1\right)\right) - 4W_i'(N_{d,i})\right]}{\alpha_i \lambda_{i,\max} N_{d,i}\left(N_{d,i}+1\right)_i \left(e^{-(\alpha_i N_{d,i})}-1\right)^2} \right|$$

With $W_i'(N_{d,i})$ as in (19). Furthermore, the above conditions ensure an overdamped closed loop system, avoiding large overshoots and keeping physical and operations constraints given by (4).

Proof:

Follows as for Theorem 2 with

$$\dot{x} = \begin{bmatrix} -\dfrac{W_i(N_{d,i})}{N_{d,i}(N_{d,i}+1)} & 1 \\ \dfrac{1}{T_i} W_i'(N_i^*) & -\dfrac{1}{T_i} \end{bmatrix} x \tag{28}$$

Then by evaluating the system eigenvalues it follows that the closed loop is asymptotically stable and overdamped , if the following condition fulfills

$$\frac{W_i(N_{d,i})}{N_{d,i}(N_{d,i}+1)} - W_i'(N_{d,i}) \le 0 \tag{29}$$

The rest of the proof follows as for Theorem 2, yielding the stability and convergence conditions (25), (26) and (27). ■

Theorem 6

The closed loop system formed by a non producer node (5), (6) and (7) with the controller (10) is locally asymptotically stable and converge to the equilibrium point $\lambda_i^* = \delta_i$, $N_i^* = N_{d,i}$ and $O_i^* = \left[\delta_i \tau_i \left(N_{d,i}+1\right)\right]/N_{d,i}$, if the gains α_i, $K_{P,i}$, and $K_{I,i}$ satisfy

$$\alpha_i > 0 \tag{30}$$

$$K_{P,i} \leq \left| \frac{\left(1+e^{-(\alpha_i N_{d,i})}\right)^2 \left[\left(\alpha_i \lambda_{i,\max} N_{d,i}\left(N_{d,i}+1\right)\right) - 4\lambda_i'(N_{d,i})\right]}{\alpha_i \lambda_{i,\max} N_{d,i}\left(N_{d,i}+1\right)_i \left(e^{-(\alpha_i N_{d,i})}-1\right)^2} \right| \tag{31}$$

$$K_{I,i} \leq \frac{K_{P,i}\phi}{N_{d,i}(N_{d,i}+1)} \tag{32}$$

$$\phi = 2\sqrt{\tau_i N_{d,i}\left[N_{d,i}^{\;2}(1-\tau_i) + \tau_i \lambda_i(N_{d,i})\right]} + N_{d,i}^{\;2}(1-2\tau_i) + \lambda_i(N_{d,i})$$

Furthermore, the above conditions ensure an overdamped closed loop system.
Proof:
It follows as for Theorem 5. ■

4. Study case application

At this section the modelling and control techniques proposed in sections 2 and 3 are applied to a simulation case study. For the linear supply chain a multiproduct petrochemical plant is considered. Meanwhile, for the nonlinear supply chain with orders handling an extraction, blending and distribution system for crude oil is studied.

4.1 Linear supply chain: multiproduct petrochemical plant

The proposed modelling and controller techniques for linear supply chains, Sections 2.1 and 3.1, are tested by simulations on a multi-product petrochemical company, which produces different grades of polyethylene products, see Figure 3. The numbers on the left upper side of the nodes identifies the numbering used through the simulations and figures presenting plots of the results.

Hexane and catalyst are imported, whereas ethylene is obtained from a local refinery. The production of ethylene and butane is carried out by independent production plants. There exist intermediate storages for the hexane, ethylene, butane and catalyst feedstocks. Only five demand sources are taken into consideration, from D1 to D5. The reactors R1 and R2 produce different polymeric products depending on the fed material (production ratio) and operation conditions. Each reactor produces two polymers: R1 produces A1 and A2, and R2 produces B1 and B2 in a cyclic way, according to a given schedule; R1, R2 and their storages are of multi-product kind, thus have different stocks per product. For supply chain synchronization, it is considered that the demanded product is supplied to the costumer only during the production time of the corresponding product.

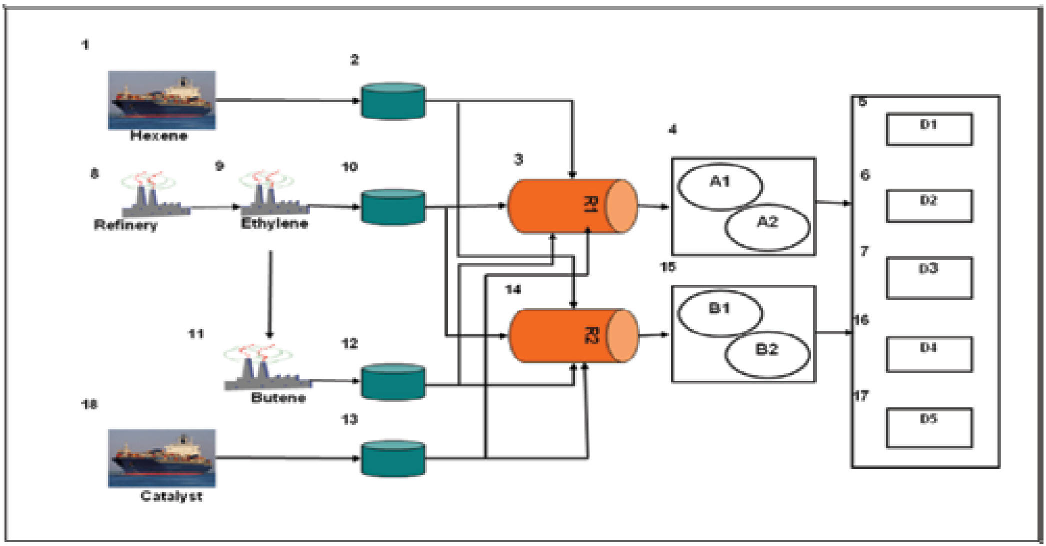


Fig. 3. Supply chain for a multi-product polyethylene petrochemical plant

The simulated period is 10 hrs. Reactors R1 and R2 change from producing product A1 to A2 and B2 to B1 at t=5 hrs respectively. The production ratios and demands for the four products are listed in Table 1. The rest of the production ratios are independent of the kind of product, such that $F_{8,9} = 1.5$ and $F_{9,11} = 0.8$, and those for storage purposes are all equal to 1. The plant capacity per reactor is 34.24 [MT/hr].

For reactor R1 and product A1 a desired inventory level of $N_{d,3} = 400$ is considered, while for product A2, $N_{d,3} = 395$. For the distributor of products A1 and A2, $N_{d,4} = 440$ for A1 and $N_{d,4} = 420$ for A2. Similarly for reactor R2 and its distributor $N_{d,14} = 420$ and $N_{d,15} = 1000$ for B1, $N_{d,14} = 420$ and $N_{d,15} = 1300$ for B2.

	A1	A2		B1	B2
$F_{2,3}$	0.25	0.4	$F_{2,14}$	0.6	0.4
$F_{10,3}$	0.15	0.2	$F_{10,14}$	0.15	0.1
$F_{12,3}$	0.5	0.3	$F_{12,14}$	0.1	0.3
$F_{13,3}$	0.2	0.1	$F_{13,14}$	0.15	0.2
D_1	5	8	D_4	9	14
D_2	3	6	D_5	13	8
D_3	12	11			

Table 1. Demands in [MT/hr] and production ratios for products A1, A2, B1 and B2.

The storage capacity of the plant is of 2000 [MT] for nodes 2, 10, and 12; 10000 [MT] for nodes 4, 13 and 15; and for nodes 3, 8, 9, 11 and 14 of 500 [MT]. According to a monthly schedule with daily resolution, inventory levels on nodes 1 and 13 must be of 3000 [MT] and 2500 [MT], respectively. The desired inventory levels for the rest of the nodes are listed in Table 2. Note that the planning or scheduling instance must take into account the physical and operational limitations listed in Table 2 to provide feasible and attainable inventories.

The initial values at t=0 hrs and the operational limitations are listed in Table 2. For the multi-product reactor R1 its initial values are for product A1 $N_3(0) = 405$, while for product A2, $N_3(0) = 390$. For the distributor the initial values are $N_4(0) = 435$ for A1 and $N_4(0) = 425$ for A2. Similar for R2 and its distributor the initial values are $N_{14}(0) = 385$ and $N_{15}(0) = 990$ for B1, $N_{14}(0) = 415$ and $N_{15}(0) = 1310$ for B2.

Node	1	2	3	4	8	9	10	11	12	13	14	15
$N_i(0)$	3000	992			377	377	360	356	360	2500		
$\lambda_i(0)$			24		36	39		33			22.8	
T_i			0.5		0.1	0.1		0.2			0.1	
$\lambda_{i,\max}$		120	34.24	40	50	60	50	55	120	60	34.24	60
$N_{d,i}$		1000			370	380	365	360	365	2500		

Table 2. Initial values $N_i(0)$, $\lambda_i(0)$, maximum production and incoming rates $\lambda_{i,\max}$ [MT/hr], desired inventories $N_{d,i}$ [MT], and adaptation time T_i [hr].

Note that the initial values for the inventories are near to the desired ones as to generate illustrative curves with small oscillations and fast convergence. Nevertheless the controller can deal with large differences on the initial inventories and productions rates with respect to the desired ones.

The bounds for the control gains were calculated according to Theorems 2 and 3, such that the gain values (Table 3) were chosen inside the corresponding bounds.

Node	2	3	4	8	9	10	11	12	14	15
α_i	1	2	10	1	2	1	1	1	5	1
$K_{P,i}$	0.1	0.9	1	0.5	0.3	0.1	0.2	0.1	0.6	0.1
$K_{I,i}$	0.07	0.08	0.07	0.1	0.04	0.07	0.04	0.07	0.04	0.07

Table 3. Control gain values

Figure 4 presents the inventory levels for producer N_3, N_9, and non producer N_4, N_{10} nodes. Note that all inventories converge to their desired values with smooth response. The inventory N_9 shows higher oscillations than the others during transient (t < 1) because it is the node most to the left of the shown ones, such that it is affected by the dynamic changes of all the related downstream nodes. This is the phenomena that origins the bullwhip effect in large supply chains. Notice that the inventories N_3, N_4 implies individual stock levels for the products A1 and A2.

From Figures 5 and 6 notice that all rates touch their boundaries at transient and when changes in production from A1 to A2 are required (t=5 hrs), which shows that the physical and operational bounds, equation (4), are held. As a result of changing the production from A1 to A2, λ_3 changes its value, while the inventories of the products A1 and A2 converge to their desired values, see Figure 4. The producer λ_9, λ_{11}, and incoming λ_{10} rates are shown in Figure 6, where it is shown that λ_9 synchronizes to its demanding rates, with $F_{9,10} = 1$ and $F_{9,11} = 0.8$. Also notice that λ_3, and λ_4 synchronizes between them accordingly to the production ratio $F_{3,4} = 1$, see Figure 5. Meanwhile the incoming rate λ_4 synchronizes to the total demand of product A1 of 20 [MT/hr] and A2 of 25 [MT/hr], such that instantaneous consumption synchronization is achieved.

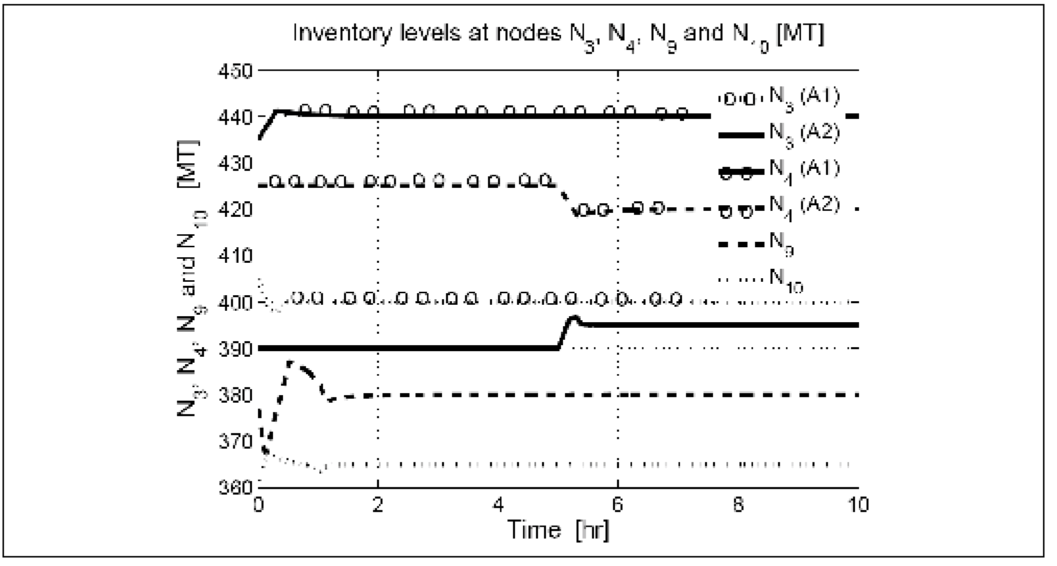


Fig. 4. Inventory levels N_3, N_4, N_9, and N_{10}.

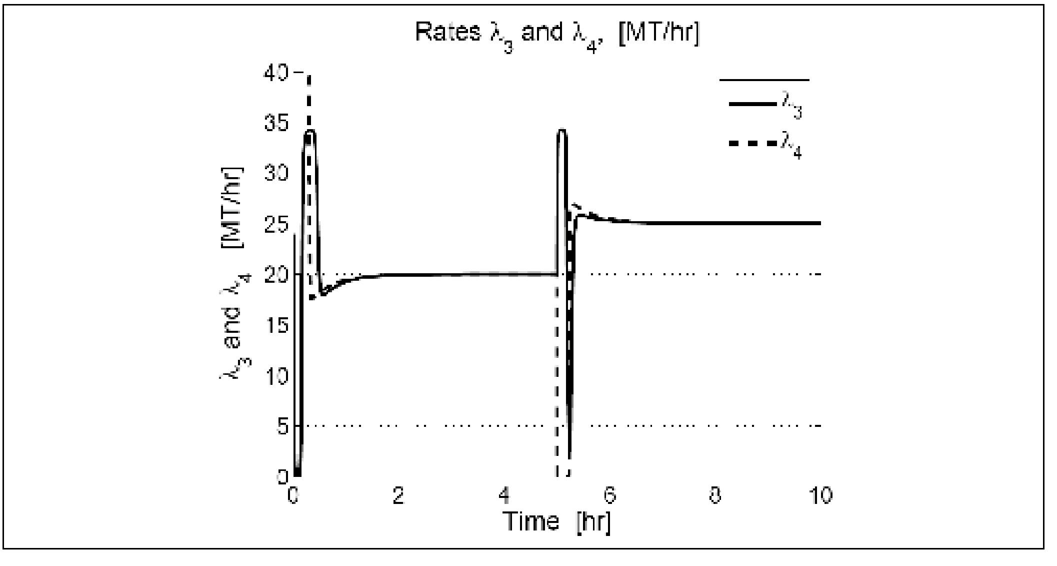


Fig. 5. Production λ_3 and incoming λ_4 rates.

Figure 7 shows a comparison study for different control gains. Note that although the PI control action is filtered by the bounded function, see equations (9), (10) and (11), the behaviour of the PI actions is preserved. A bigger overshoot but faster convergence is obtained when increasing the proportional action and a smaller stationary error is achieved by increasing the integral action.

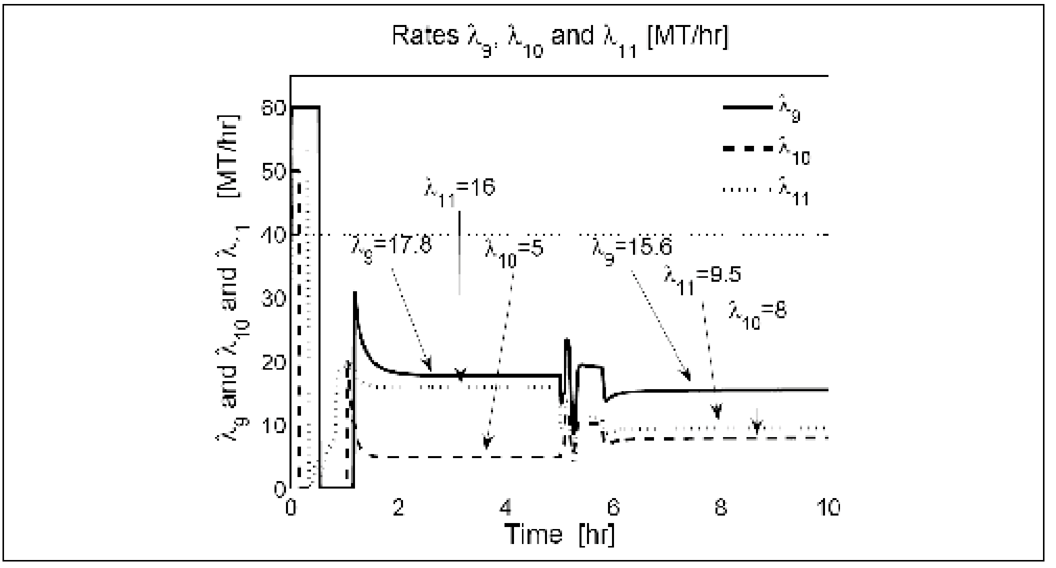


Fig. 6. Production λ_9, λ_{11}, and incoming λ_{10} rates.

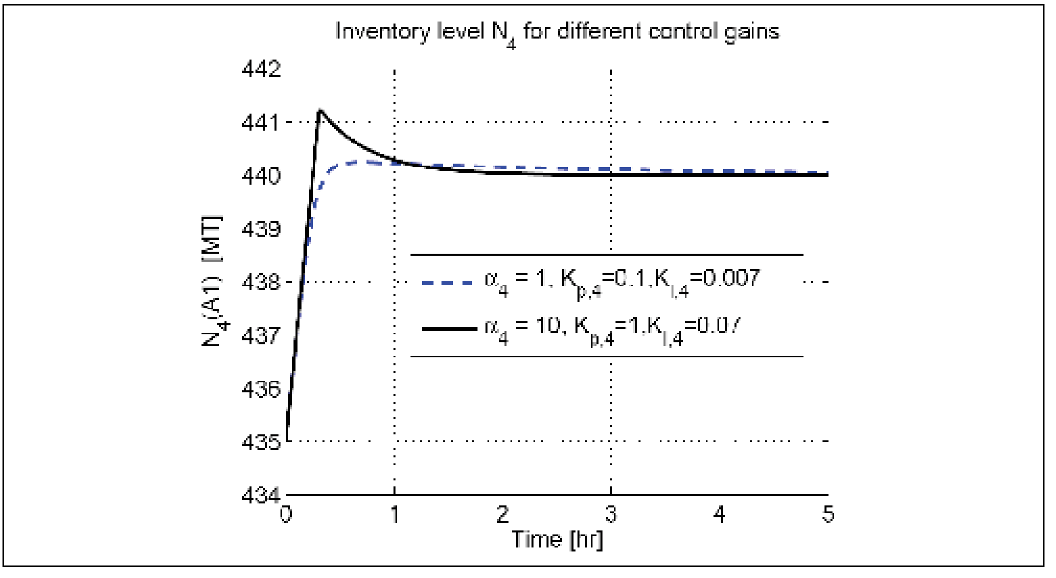


Fig. 7. Comparison study for different gain values.

4.2 Nonlinear supply chain: crude oil blending and distribution system

The modelling techniques and controller proposed in Section 2.2 and 3.2 for nonlinear supply chains that consider orders handling are tested by simulations on a blending and distribution crude oil system. The goal is to keep a desired inventory level, while the demanding rate at the distribution side must be satisfied. Figure 8 shows the crude oil supply chain system, where the numbers on the left side identifies the nodes. The initial and desired inventory values in [MT] are listed at Table 4 as well as the adaptation time T_i for producers and delay time for processing orders τ_i in [hr].

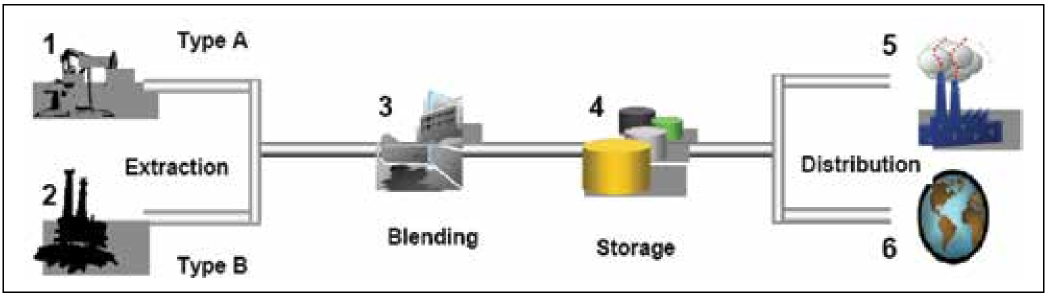


Fig. 8. Extraction, blending and distribution system for crude oil.

Node	$N_{d,i}$	$N_i(0)$	T_i	τ_i	δ_i	$\lambda_{i,\max}$
1	8000	9000				400
2	6000	7000				400
3	2000	4000	2			400
4	3000	4000		0.25		500
5	1500	2000		0.25	150	200
6	1000	2000		0.25	200	250

Table 4. Initial values and desired inventories.

The bounds for the control gains were calculated accordingly to Theorems 5 and 6 such that the gain values (Table 5) were chosen inside the corresponding bounds.

Node	1	2	3	4	5	6
α_i	1	0.1	1	1	2	1
$K_{P,i}$	0.97	0.6	2.02	0.11	0.14	0.11
$K_{I,i}$	$3\text{x}10^{-6}$	$3\text{x}10^{-6}$	0.25	0.22	0.28	0.22

Table 5. Control gain values

Figures 9 and 10 show the results obtained at the closed loop, notice that the inventory levels converge to their desired values, while the control actions follow the demands, thus demands are satisfied. The demand and control actions show the delay introduced by considering traffic flow theory at the dynamic models, nonetheless at steady state they synchronizes each other. Several conditions, such as noise at demands and sudden changes on desired inventories, have been tested by simulations, showing robustness of the closed loop system; however for the brevity of space the plots are skipped. The oscillations during transients at demands and control actions are greatly affected by the delays modelled by T_i and τ_i, such that bullwhip phenomena can arise.

5. Conclusions and trends

Supply chains are challenging systems due to their dynamical behavior, ranging from fast changing demand, transport and delivery delays, inventory management, and production

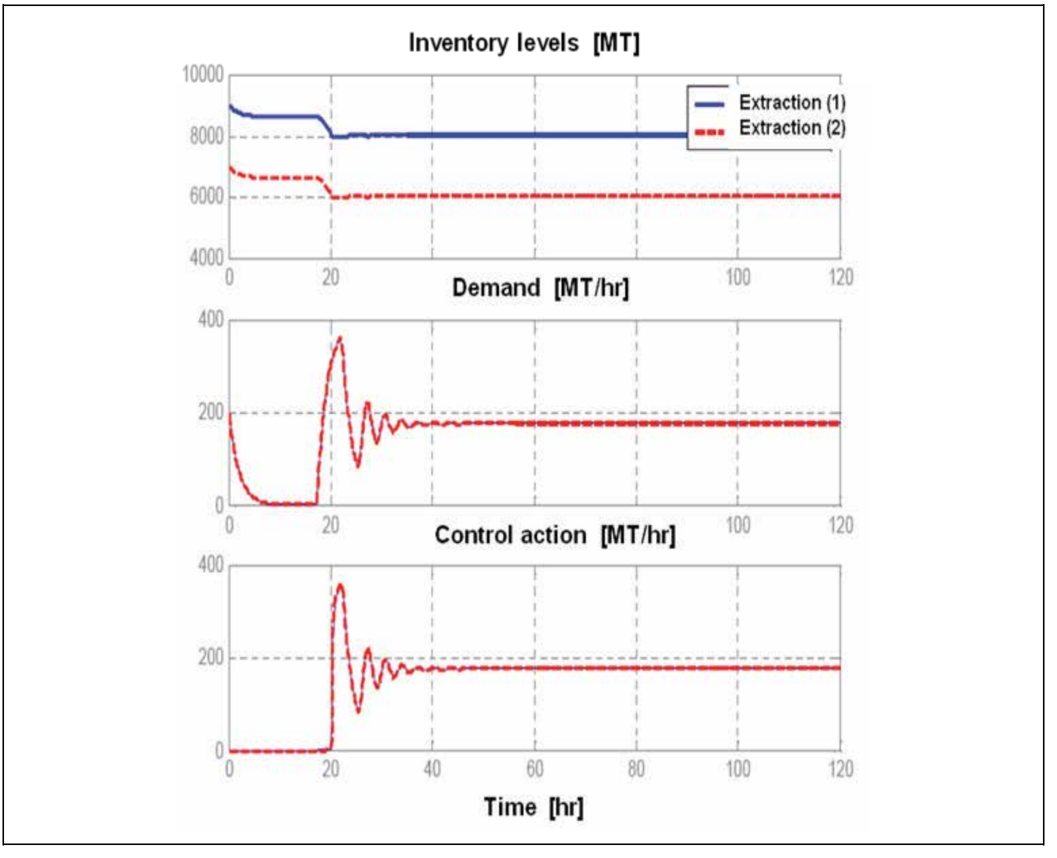


Fig. 9. Supplier nodes 1 and 2.

phenomena. Appropriate control and operation of such systems may contribute to achieve the enterprise objectives.

In this work, the proposed supply chain models go from a simple linear supply chain based on material flows to a nonlinear one considering orders handling and based on traffic flow theory. Nonetheless the simplicity of both proposed models, they reproduced production and transport delays by introduction of time parameters. However, there are several supply chain phenomena that are still to be modeled, such as recycling material loops, hybrid systems with continuous and discrete time dynamics, stochastic behavior, among many others. Differential equation models, as the ones considered here, may be useful on modeling such phenomena, but there are other approaches such as fuzzy sets, neuronal networks, to mention a few, that could be considered.

Furthermore, both the model and control of supply chains must consider the physical and operative constraints. For this purpose, bounded control actions have to be proposed. Several trends on control of supply chains go from classic control theory to time discrete systems, fuzzy strategies and even evolutive algorithms. Due to the type of presented supply chain models it is straightforward to consider classic control techniques such as the PI controllers introduced at this work. An important point when proposing control strategies is the capability of determining stability and operative conditions for the system.

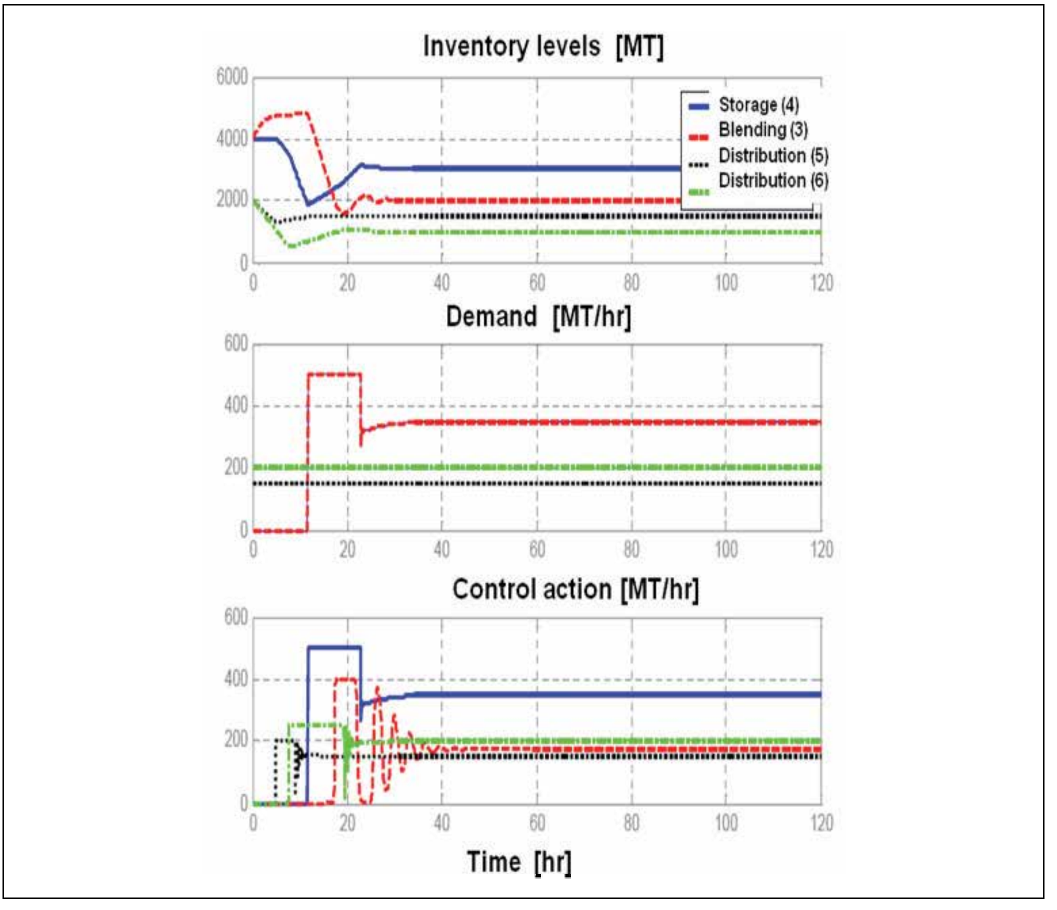


Fig. 10. Nodes 3, 4, 5, and 6.

This is a key issue to proceed for implementation of the controller to a real supply chain system, otherwise, instability or critical operative conditions may arise. Classic control techniques have the advantage of a well developed variety of stability analysis tools, like pole placement and linearization techniques in which this work is based on to obtain tuning guidelines and stability conditions for the closed loop system.

Simulation study cases have been considered for the purpose of demonstration of the closed loop performance achieved by the proposed models and controllers. In general the simulations allow concluding stability of the supply chain system, convergence of the inventory levels to the desired values and synchronization of the material flows. Thus everything what was predicted by the stability analysis has been confirmed by the simulation case studies. Although comparison and analysis of simulation results are valid to draw some conclusions, development of scale models or prototypes has to be considered for a better understanding of supply chains and more realistic implementation of the controllers and management strategies of such systems.

6. References

Cetinkaya, S., & Lee, C., (2000). Stock replenishment and shipment scheduling for vendor managed inventory systems, *Management Science*, 46, 2, electronic ISSN 217-232, 217-232.

Daganzo, C.F. (2002). *A Theory of Supply Chains*, Springer-Verlag, ISBN: 3-540-00288, Berlin, Germany.

Dunbar, W.B. & Desa, S. (2005). Distributed Model Predictive Control for Dynamic Supply Chain Management, Int. Workshop on Assessment and Future Directions of NMPC, Freudenstadt-Lauterbad, Germany, August, 26-30.

Fujimoto, H. (2002). Supply chain for the synchronization of production and distribution. *Osaka Keidai Ronshu*, 53, 3, 307-315.

Helbing, D., (2003) Modeling supply networks and business cycles as unstable transport phenomena. *New Journal of Physics,* 5, 90, ISSN 1367-2630, 90.1-90.28.

Helbing, D., Lämmer, S., Seidel, T., Seba, P. & Platkowski, T. (2004). Physics, stability, and dynamics of supply chain networks. *Physical Review* E, 70, 6, 066116(1-6), ISSN: 1550-2376.

Koudal, P., (2003). Integrating demand and supply chains in the global automotive industry: building a digital loyalty network at General Motors. Deloitte Research and Stanford Global Supply Chain Management Forum, Research article, www.dc.com/research

Koudal, P., & Wei-the ,V. L., (2005). The power of synchronization: the case of TAL Apparel Group, Deloitte Research, Research article, www.dc.com/research

Lababidi, H., Ahmed, M., Alatiqi, I., & Al-Enzi, A., (2004). Optimizing the supply chain of a petrochemical company under uncertain operating and economic conditions, *Industrial & Engineering Chemistry Research,* 43, 1, ISSN: 0888-5885, 63-73.

Lefeber, E., (2004). Nonlinear models for control of manufacturing systems, *Nonlinear dynamics of production of production systems,* Ed. G. Radons and R. Neugebauer, ISBN: 3-527-40430-9, 71-83.

Lin, P., Wong, D., Jang, S., Shieh, S., & Chu, J., (2004). Controller design and reduction of bullwhip effect for a model supply chain system using z-transform analysis, *Journal of Process Control,* 14, 5, ISSN: 0959-1524, 487-499.

Nagatani T. & Helbing, D. (2004). Stability analysis and stabilization strategies for linear supply chains. *Physica A*, 335, 3-4, 644-660, ISSN: 0378-4371.

Perea-Lopez, E.; Grossman, I.E., Ydstie, B.E. & Tahmassebi, T. (2001). Dynamic Modeling and Decentralizes Control of Supply Chain. *Ind. Eng. Chems. Res.*, 40, 15, 3369-3383, ISSN: 1520-5045.

Perea-Lopez, E., Ydstie, B.E. & Grossman, I.E.; (2003). A model predictive control strategy for supply chain optimization. *Comp. & Chem. Eng.*, 27, 15, 1201-1218, ISSN: 0098-1354.

Petrovic, D., Roy, R., & Petrovic, R., (1999). Supply chain modeling using fuzzy sets, *International Journal of Production Economics,* 59, 1-3, ISSN: 0925-5273, 443-453.

Shapiro, J. F. (2007). *Modeling the Supply Chain*. Duxbury Thomson Learning Onc., ISBN: 0-495-12611-x, h., Pacific Belmont CA.

Steidmann, C. (2004). *Synchronicity: an emerging vision of the retail future*. Deloitte Research, Research Brief Report, May, www.dc.com/research.

Tayur, S., Ganeshan, R., & Magazine, M., (1998). *Quantitative models for supply chain management,* Boston, Kluwer Academic Publishers, ISBN: 0-7923-8344-3.

5

Optimization Based e-Sourcing

Kameshwaran Sampath[1] and Lyès Benyoucef[2]
[1]*Indian School of Business*
[2]*INRIA Nancy-Grand Est*
[1]*India*
[2]*France*

1. Introduction

Sourcing or procurement is the process by which a company obtains goods and services for its manufacturing and operations. The materials procured could range from raw materials, components, and sub-assemblies, to office supplies and furniture. The services procured can be as vital as design and R&D to daily operations like IT and logistics. e-Sourcing or e-procurement refers to *online* procurement of the above direct and indirect inputs by an industrial buyer. The other predominantly used terms for this process are *procurement auctions*, *reverse auctions*, and *e-auctions*. We use the above terms in this chapter with the following definition of Minahan (2001): *the process of utilising Web-based technologies to support the identification, evaluation, negotiation, and configuration of optimal groupings of trading partners into a supply chain network, which can then respond to changing market demands with greater efficiency*.

e-Sourcing of production and non-production goods and services has been in practice since early 1990s, especially among the *Fortune 2000* companies. It was widely accepted then that web based sourcing can provide following advantages (Minahan 2001):

- Identify and negotiate with a broad range of qualified suppliers;
- Reduce process costs for sourcing engagements;
- Shorten sourcing cycles by 25% to 30%;
- Reduce time-to-market cycles by 10% to 15%;
- Negotiate an average of 5% to 20% unit price reductions;
- Extend strategic sourcing to a wider range of products and services; and
- Enhance collaboration and knowledge sharing.

The market analysts' predictions about the worth of online business transactions were in trillions of USD by 2003/4. However, the e-bust that happened in 2000, followed by the market studies of the real world implementations showed that these figures are indeed exaggerated and overstated, if not false (Emiliani & Stec, 2004; 2005). Irrespective of the figures projected and achieved, the use of e-sourcing is growing with steady incremental gains rather than abrupt exponential profits.

- Ariba[1], a leading provider of online spend management solutions, has enabled sourcing of USD 250 Billion worth of goods and services till date, using its *Ariba Sourcing Solution*. The total annual savings generated is over USD 15 Billion (Ariba, 2007).

[1] http://www.freemarkets.com

- GlaxoSmithKline achieved a 5,452% annualized return-on-investment using Emptoris[2] (another leader in providing sourcing solutions and a pioneer in the use of optimization for strategic sourcing in industry).
- Motorola received the prestigious Franz Edelman Award for Achievement in Operations Research and the Management Sciences in 2004 for the application of optimization bid analysis with Emptoris to save USD 600 million.
- More than 60 Fortune 500 companies use CombineNet[3] for their most advanced strategic sourcing activities, with greater than 45x average return on investment.
- Volume-discount and combinatorial auctions benefited Mars Inc. and its suppliers (Hohner et al., 2003).
- Chilean government used combinatorial auctions to assign catering contracts for the supply of school meals to children and resulted in a 22% cost savings (Epstein et al., 2002).

In this chapter, we present different optimization based e-sourcing auctions from the literature and industry best practices. We also extend how these mechanisms could be used in global sourcing and the future research to include risk mitigation.

Auction is a market mechanism with well-defined set of rules for determining the terms of an exchange of something for money (McAfee & McMillan, 1987). Procurement auctions are *reverse* auctions in the sense that the buyer is the auctioneer and the sellers (suppliers) are the bidders. Traditionally, auctions for procurement at the industrial scale were mainly used by government for purchasing goods and services. The main reasons are openness and fairness of auctions, and till today, government purchasing happens through auctions. However, for many years, auctions played a relatively minor role in industrial procurement (Rothkopf & Whinston, 2007).

The industrial approach to procurement was to develop long-term cooperative relationships with few suppliers. The advent of Internet and the advancement of e-commerce changed this approach radically. The long term strategic partnerships with suppliers are still prevalent for sourcing of custom designed goods and services. On the other hand, for sourcing of commoditized goods and services, e-sourcing through auctions are being increasingly used by industries (Kouvelis et al., 2006). Also among the academics, there is a recent growing interest in this Internet based business process that led to new generation of procurement techniques like *combinatorial* (Cramton et al., 2006), *volume discount* (Eso et al., 2005) and *multi-attribute* (Bichler et al., 1999; Kameshwaran et al., 2007).

e-Sourcing and in general, e-auctions, are being studied by scholars from different disciplines such as economics, operations research, management science, information systems, and computer science (Rothkopf & Whinston 2007). Many works from operations research and management science community approach e-sourcing from the perspective of solving a *supply chain optimization* problem. This chapter also adopts the same perspective and other economic issues like *information asymmetry*, *adverse selection*, and *moral hazards* (McAfee & McMillan, 1987) are not addressed here.

The chapter is organized as follows. Section 2 describes the dynamics of the sourcing process, briefly outlines the design issues, and introduces the three different sourcing formats considered in this chapter. The three sourcing techniques are presented in detail in

[2] http://www.emptoris.com
[3] http://www.combinenet.com

sections 3 to 5. For each of the techniques, the mathematical programming formulation is presented. Sourcing based on volume discounts, one of earliest formats in business tradition, is presented in Section 3. The more recent and popular combinatorial sourcing is described in Section 4. Multi-attribute sourcing and its extension configurable bids are discussed in Section 5. Section 6 is devoted to global sourcing, where the traditional industrial procurement format of a single buyer with multiple suppliers is extended to multiple factories of the same organization procuring from multiple suppliers. Section 7 discusses the robustness approach to e-sourcing to design a risk-tolerant sourcing network that will operate at acceptable levels under a wide range of pre-identified random scenarios. Final notes are given in Section 8 and references are listed in Section 9.

2. Dynamics, design issues, and taxonomy

2.1 Dynamics

The overall industrial sourcing dynamics can be described as a standard three-step process: (1) Pre-auction stage, (2) auction stage, and (3) post-auction stage. This is adapted from the three-step process described in Caplice & Sheffi (2007) for procurement of transportation services.

Pre-auction stage

The buyer forecasts the demand for the planning horizon and determines which suppliers to invite for the sourcing auction. Common practice is to retain most incumbents (to maintain long term buyer-supplier relationship) and invite some new suppliers. A set of mandatory supplier selection criteria (Weber et al. 1991) is used to identify these potential suppliers. The buyer also decides the format of the auction and bid structure (to be described in detail).

Auction stage

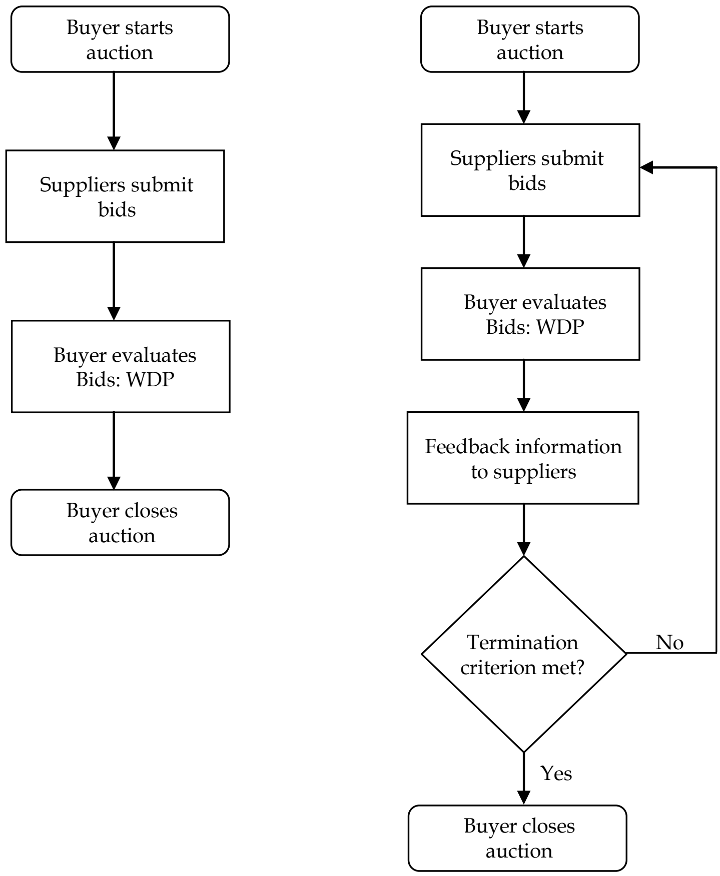
The demand, auction format, and the bid structure are communicated to the pre-selected suppliers in the form of RFQ or RFP through the use of faxed lists, spreadsheets, online web pages, email, or direct EDI connections. The suppliers conduct their own analysis and prepare the bid according the required format. The bidding phase in the auction could be *single round* or *multiple rounds* (see Figure 1). Once the bids are received, the buyer solves the *winner determination problem* (WDP), where the bids are evaluated to determine the winning the suppliers (also known as *bid evaluation problem*). In the single round auction, the auction closes after this stage and the suppliers are intimated of their status. In multi-round auctions, the winners determined are provisional winners and the suppliers are given feedback information, using which they can resubmit bids. Once a termination criterion is met the auction is closed. The criterion could be that no new bids from suppliers or the upper bound on the number of rounds reached.

Post-auction stage

Once the auction is closed, the results of the WDP are uploaded to the downstream planning, execution, auditing, and payment systems.

2.2 Design issues

The sourcing process with RFX generation and bidding by suppliers is inherently based on auctions and hence the design principles for sourcing generally follow auction design. As mentioned above, auctions can be categorized based on the dynamics as: (1) *single-round or one-shot* auctions and (2) *multiple-round* or *progressive* or *iterative* auctions. Single-round auctions are sealed bid auctions. Multi-round auctions can be sealed bid or open bid, but has



(a) Single-round auction (b) Multi-round auction

Fig. 1. Single-round and multi-round auction dynamics

multiple rounds of bidding phases. At the end of each bidding phase, there will be flow of information from the auctioneer to the bidders. This will help the bidders to prepare their bids for the next bidding phase. The design parameters of single-round auctions are *bid structure*, *winner determination policy*, and *pricing policy*. The bid structure specifies the format of bids, the winner determination policy describes the technique to determine the winners, and the pricing policy determines the price(s) of the winning good(s).

On the other hand, design of multi-round auctions is relatively non-trivial, which includes the specification of bid structure, winner determination technique at each bidding round, information exchange at the end of each round, termination condition, and the pricing policy. Multi-round auction has many advantages over its one-shot counterpart (Cramton 1998), especially in sourcing (Parkes & Kalagnanam 2005). There are many design methodologies for multi-round auctions for sourcing (Bikhchandani & Ostroy 2006, Kameshwaran et al. 2005, Parkes & Kalagnanam 2005).

The sourcing process with the RFX and the bidding, only borders on auctions and are indeed less formally structured than auctions. The auction design is generally based on the

principles of *mechanism design*. Mechanism design (Mas-Colell et al., 1995) is the sub-field of microeconomics and game theory that considers how to implement good system-wide solutions to problems that involve multiple self interested agents, each with private information about their preferences. The mechanism design methodology has also been found useful in designing e-markets (Varian, 1995). One of the main assumptions in mechanism design is that the rules of the auction are a common knowledge to all the participating agents. In sourcing, though the rules of bid submission are common knowledge, rules of winner determination may not be revealed to the suppliers. The purchasing manager may take into account several business rules and purchasing logic in winner determination, which are not generally revealed to the suppliers. Moreover, the criteria and the constraints can be modified by the auctioneer (buyer), based on the received bids. Here, we do not follow the mechanism design approach.
Pricing policy is another design issue, which dictates the price to be paid to the winning suppliers for the supply of the winning goods. The commonly used pricing policy in current e-sourcing systems is the *pay-as-bid* or *first price* policy where the suppliers are paid the cost quoted in their respective bids. There are non-trivial pricing policies such as VCG (Ausubel & Milgrom, 2006), where the price is function of the price quoted by the other suppliers. Though this pricing policy has certain desirable economic features, it is not widely used in practice.

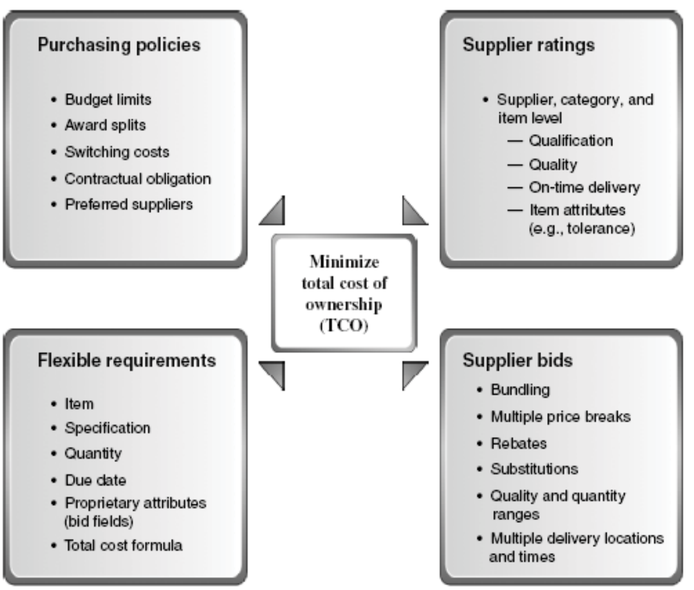


Fig. 2. Factors considered by Motorola in awarding business to suppliers (Source: Metty et al. (2005))

2.2 Bid structure and winner determination

We consider only the design issues related to bid structure and winner determination technique, from the perspective of the buying organization (auctioneer). The bid structure dictates how a bid is defined. For example, it could be as simple as a unit price for an item or set of attributes like unit price, lead time, quantity, etc. A *rich* bid structure is advantageous to both the buyer and suppliers. The buyer has more negotiating parameters rather than just unit price and hence can optimize the total cost or procurement. Suppliers, on the other hand, can differentiate themselves from their competitors with value added services rather than competing on just cost. The earlier e-sourcing techniques achieved cost reduction to buyers by squeezing the profit margins of the competing suppliers. Many historical and incumbent suppliers did not prefer the online sourcing as they felt that the buyers used it to wring price concessions from them in the presence of new suppliers (Jap, 2002). However, e-sourcing evolved with a rich set of bid structures, providing a win-win situation to both the buyers and sellers, and thereby achieving overall supply chain efficiency. Figure 2 shows the factors considered by Motorola to minimize the total cost of ownership while awarding business to the suppliers (Metty et al., 2005).

The winner determination problem faced by the buyer at the end of the bidding phase (or at the end of every bidding phase in progressive auctions) is an optimization problem. The problem is to determine the set of winning suppliers and their respective winning items, such that the total procurement cost is minimized subject to various business constraints and purchasing policies. Indeed, one of the earliest applications of linear programming is winner determination (also referred as bid evaluation) (Gainen et al., 1954). Many commercial bid analysis products from companies like Emptoris, and CombineNet use optimization techniques like linear programming, combinatorial optimization, and constraint programming. Optimization also allows for addition of business constraints and purchasing logic as side constraints in winner determination, which is a new development in sourcing auctions (Rothkopf & Whinston, 2007). Some of the commonly used business rules are:

- Limiting the number of winning suppliers in a given range;
- Limiting the business awarded (in terms of quantity or worth) to a winning supplier in a given range;
- Guaranteeing a minimum amount of business to incumbent suppliers;

Note that the above business rules need not be disclosed to the suppliers and often many of them are experimented with WDP like analyzing *what-if* scenarios.

2.3 Taxonomy

We categorize the e-sourcing techniques based on the bid structure and the winner determination policy, which also implicitly depends on the number of goods purchased and their respective quantities. We broadly classify e-sourcing under three categories: (a) Volume discount sourcing, (b) Combinatorial sourcing, and (c) Multi-attribute and multi-criteria sourcing. We describe each of the above in detail in the following sections.

3. Volume discount sourcing

Volume or quantity discounts in sourcing is a long established business tradition. Buyers expect discounts for buying large quantities and the suppliers provide discounts to price discriminate from the competing suppliers. Studies by Lippman (1969), Prikul & Aras (1985), Jucker & Rosenblatt (1985), and Dolan (1987) focus on how buyers determine the economic order quantities with quantity discounts. On the other hand, Crowther (1964), Monahan (1984), Lee & Rosenblatt (1986), and Kim & Hwang (1988) focus on the supplier's perspective of formulating the form of quantity discount pricing schedule.

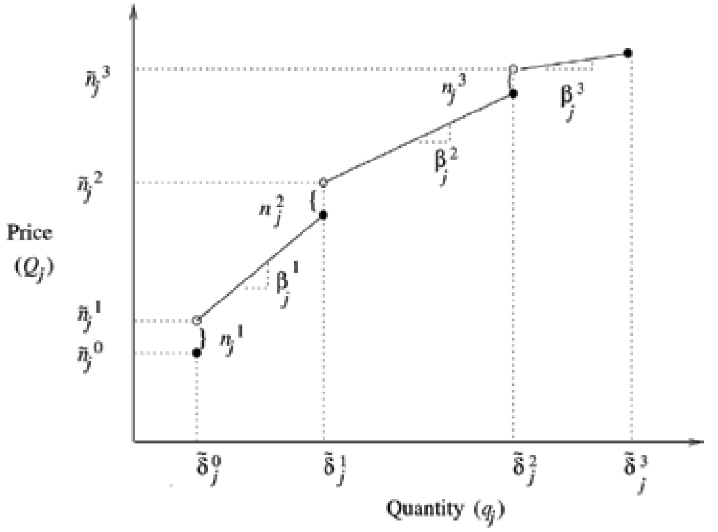


Fig. 3. Piecewise linear cost function as bid

According to Sadrian & Yoon (1994), the rationale behind quantity discount models is derived from the numerous economic advantages gained from buyers ordering larger quantities of products. With larger orders, both the supplier and the buyer should be able to reduce their order processing costs. These might include packaging and handling costs, administrative costs, and shipping costs. However, the major benefits come from the savings in the supplier's manufacturing costs. A supplier that produces the item itself for these larger orders will require fewer manufacturing setups and have larger production runs. These savings are especially significant if the supplier experiences high product-specific setup costs. On the other hand, increased order size results in a higher inventory holding cost for the buyer. Therefore, the supplier should compensate for this extra cost with an attractive quantity discount pricing schedule to induce buyers to increase their order size.

Munson and Rosenblatt (1998) provided a classification of different quantity discounts. The form of discount could be *all-units* or *incremental*. In all-units form, the discount is applicable to all the units purchased, whereas in incremental, the discount is applicable only to the additional units that has crossed the break-up point. The *item aggregation* describes whether the discount applies to one or multiple products (bundling). A *business volume discount* represents item aggregation where the price breakpoints are based on the total dollar volume of business across all products purchased from the supplier. The *time aggregation* describes whether the discounts apply to individual purchases or multiple purchases over a given time frame. Finally, the number of price breakpoints may be one, multiple, or infinite (as represented by a continuous price schedule). In this section, we present incremental discount form applied to sourcing of multiple units of a single item (Kameshwaran & Narahari, 2009b). Our focus here is on solving the WDP faced by the buyer, when the suppliers submit non-convex piecewise linear quantity discount price functions.

3.1 Sourcing of multiple units of a single item

Consider an industrial procurement of multiple units of a single item. The demanded item can be a raw material or a sub-component and let the demanded number of units be *B*. Let *N* be the number of potential suppliers, out of whom the winning suppliers have to be selected. Usually, the buyer prefers multi-sourcing as the supply failure in a single supplier scenario will be disruptive. Further, the demand B could be formidably large for a single supplier. With multi-sourcing, the buyer needs to determine the winning suppliers and also their winning quantity. This provides an additional negotiable parameter to the suppliers in addition to unit price. The bid thus is a price function defined over quantity. The price function that is commonly used in industry for long-term strategic sourcing is *piecewise linear* (Davenport & Kalagnanam, 2002).

3.2 Piecewise linear bid

The bid submitted by supplier j is a cost function Q_j defined over the supply quantity $[a_j, z_j]$. Q_j is piecewise linear and nonconvex, as shown in Figure 3. It can be represented compactly by tuples of break points, slopes, and costs at break points: $Q_j \equiv ((\tilde{\delta}_j^0 = a_j, ..., \tilde{\delta}_j^{l_j} = z_j), (\beta_j^1, ..., \beta_j^{l_j}), (\tilde{n}_j^0, ..., \tilde{n}_j^{l_j}))$. For notational convenience, define $\delta_j^s \equiv \tilde{\delta}_j^s - \tilde{\delta}_j^{s-1}$ and n_j^s as the jump cost associated with linear segment s. Note that, by this definition, $n_j^0 = \tilde{n}_j^0$.

Notations

$[a_j, z_j]$	Supply range
Q_j	Cost function defined over $[a_j, z_j]$
l_j	Number of piecewise linear segments in Q_j
β_j^s	Slope of Q_j on $(\tilde{\delta}_j^{s-1}, \tilde{\delta}_j^s)$
δ_j^s	$\equiv \tilde{\delta}_j^s - \tilde{\delta}_j^{s-1}$
n_j^s	Fixed cost associated with segment s
$\tilde{n}_j^s$	$Q_j(\tilde{\delta}_j^s) + n_j^s$
Q_j	$\equiv ((\tilde{\delta}_j^0 = a_j, ..., \tilde{\delta}_j^{l_j} = z_j), (\beta_j^1, ..., \beta_j^{l_j}), (\tilde{n}_j^0, ..., \tilde{n}_j^{l_j}))$

The function is assumed to be strictly increasing, but not necessarily marginally decreasing as shown in the figure. The assumed cost structure is generic enough to include various special cases: *concave, convex, continuous,* and $a_j = 0$. The cost structure enables the suppliers to express their *volume discount* or *economies of scales* and/or the production and logistics constraints. The volume discount strategy, which is *buy more and pay less* can be expressed with marginally decreasing cost functions. The discontinuities with jump costs in the cost structure can capture the production and transportation constraints.

3.3 Winner determination problem

Let J be the set of N suppliers. The index j denotes a supplier from set J. As each supplier can only submit one bid, we use the index j to denote both the supplier and his bid. The winner determination problem (WDP) faced by the buyer is to minimize the total cost of procurement with the following decisions: (1) select a set of winning bidders $J' \subseteq J$ and (2) determine the trading quantity q_j for each winning bid $j \in J'$. The above decisions are to be made subject to the following constraints:

- *Supply Constraint:* For every winning bid $j \in J'$, $q_j \in [a_j, z_j]$, and for losing bids, $q_j = 0$.
- *Demand Constraint:* The total quantity procured should satisfy the demand of the buyer: $\sum_{j \in J'} q_j \geq B$.

The WDP is a nonconvex piecewise linear knapsack problem (Kameshwaran & Narahari, 2009a), which is *NP*-hard. It is a minimization version of a nonlinear knapsack problem with a demand of B units. Each bid corresponds to an *item* in the knapsack. Unlike traditional knapsack problems, each item j can be included in the knapsack in a pre-specified range $[a_j, z_j]$ and the cost Q_j is a function of quantity included.

The cost function Q_j of Figure 3 is nonlinear but due to the piecewise linear nature, the WDP can be modelled as the following MILP.

$$\min \sum_{j \in J} \left(n_j^0 d_j^0 + \sum_{s=1}^{l_j} \left(n_j^s d_j^s + \beta_j^s \delta_j^s x_j^s \right) \right) \tag{1}$$

subject to

$$d_j^1 \leq d_j^0 \quad \forall j \in J \tag{2}$$

$$x_j^s \leq d_j^s \quad \forall j \in J; 1 \leq s \leq l_j \tag{3}$$

$$x_j^s \geq d_j^{s+1} \quad \forall j \in J; 1 \leq s < l_j \tag{4}$$

$$\sum_{j \in J} \left(a_j d_j^0 + \sum_{s=1}^{l_j} \delta_j^s x_j^s \right) \geq B \tag{5}$$

$$d_j^s \in \{0,1\} \quad \forall j \in J; 0 \leq s \leq l_j \tag{6}$$

$$x_j^s \in [0,1] \quad \forall j \in J; 1 \leq s \leq l_j \tag{7}$$

The decision variable x_j^s denotes the fraction of goods chosen from the linear segment s of bid j. For this setup to make sense, whenever $x_j^s > 0$ then $x_j^{s-1} = 0$, for all s. To enable this, binary decision variable d_j^s is used for each segment to denote the selection or rejection of segment s of bid j. The winning quantity for bid j is $a_j d_j^0 + \sum_{s=1}^{l_j} \delta_j^s x_j^s$ with cost $n_j^0 d_j^0 + \sum_{s=1}^{l_j} \left(n_j^s d_j^s + \beta_j^s \delta_j^s x_j^s \right)$. The binary decision variable d_j^0 is also used as an indicator variable for selecting or rejecting bid j, as $d_j^0 = 0$ implies that no quantity is selected for trading from bid j.

3.4 Business constraints

The business rules and purchasing logic can be added as side constraints to the WDP. For the above procurement scenario, the relevant business constraints are restricting the number of winning suppliers in a given range [*LB, UB*] and guaranteeing a minimum volume (or monetary business worth) *MIN_QTY* (*MIN*_VAL) for a set of incumbent suppliers $J' \subset J$.

$$LB \leq \sum_{j \in J} d_j^0 \leq UB \tag{8}$$

$$\sum_{j \in J'} \left(a_j d_j^0 + \sum_{s=1}^{l_j} \delta_j^s x_j^s \right) \geq MIN\_QTY \tag{9}$$

$$\sum_{j \in J'} \left(n_j^0 d_j^0 + \sum_{s=1}^{l_j} \left(n_j^s d_j^s + \beta_j^s \delta_j^s x_j^s \right) \right) \geq MIN\_VAL \tag{10}$$

The above constraints can be added as side constraints to the WDP. Usually one of the (9) or (10) is used. Business rule that limits the winning quantity or business value for a winning supplier can be implicitly included by suitably modifying the supply range $[a_j, z_j]$.

3.5 Algorithms

Dynamic programmic based exact and approximation algorithms were proposed in (Kameshwaran & Narahari, 2009a) and a Benders' decomposition based exact algorithm was proposed in (Kameshwaran & Narahari, 2009b) to solve the WDP formulated as (1)-(7). Similar procurement scenarios have been considered in the literature with various assumptions. Kothari et al. (2003) expressed the cost function using fixed unit prices over intervals of quantities (piecewise linear but continuous with no jump costs) and approximation algorithms based on dynamic programming were developed for solving the WDP. Procurement with nonconvex piecewise linear cost functions was considered by Kameshwaran & Narahari (2005) with the additional business constraint of restricting the number of winning suppliers. A Lagrangian based heuristic was proposed to solve the WDP. Eso et al. (2005) considered the quantity discount procurement of heterogeneous goods and column-generation based heuristic was proposed to solve the WDP.

3.6 Other discount based sourcing techniques

In the above, we briefly discussed about volume discounts offered while procuring multiple units of a single item. Eso et al. (2005) considered buying multiple items with volume discounts for each item. There are two kinds of discounts for procuring multiple units of multiple items: *Business volume discounts (*Sadrain & Yoon ,1994) and *total quantity discounts* (Goossens et al. 2007). In the business volume discounts, the discounts are based on the total monetary worth of the purchase rather than on the quantity. This discount structure is applicable in telecommunication sourcing. In total quantity discounts, discount is based on the total quantity of all items purchased. This discount is used in chemical and also in telecom capacity sourcing. Exact algorithms based on brand and bound were proposed in (Goossens et al., 2007) to solve this problem. For a special case with single unit demand for multiple items, a suite of branch-and-cut algorithms was proposed in (Kameshwaran et al., 2007).

4. Combinatorial sourcing

Consider a sourcing scenario where the buyer wants to buy a set of heterogeneous items. Two immediate approaches to procure them are in *sequence* (sequential procurement with one after another) and in *parallel* (all items are procured simultaneously by conducting a sourcing auction for each item separately). The third option is to conduct a *combinatorial auction* where the supplier can bid on a combination of items by providing a single bid price (Cramton, 2006). Thus the bid price is conditional on winning the entire combination of items. These auctions are ideal for scenarios in which synergies exist between the items. Suppose a supplier obtains more profit by selling a set of items together, then he can submit this *all-or-nothing* combinatorial bid by providing a discounted price on that entire package. The supplier can submit more than one bid and the items in different bids can be overlapping.

Combinatorial auctions were initially used in selling scenarios like airport slot allocation (Rassenti et al., 1982) and radio spectrum auctions (Rothkopf et al., 1998). The sourcing applications mainly include procurement of transportation services (Caplice & Sheffi, 2006), in addition to direct sourcing of industrial inputs (Hohner et al., 2003). In this following, we present various combinatorial bids and the respective WDP formulations.

4.1 Static package bids

Let the items to be procured be indexed by i, each with demand d_i. A bidder j bids on a *package* or *bundle* of items, providing a single bid price for that bundle. Let the package be indexed by k. As mentioned above, the bidder can submit different packages as bids with possibly overlapping items. The winner determination problem can be formulated as the following 0-1 integer program.

$$\min \sum_{j} \sum_{k} C_j^k y_j^k \tag{11}$$

subject to

$$\sum_{j} \sum_{k:i\in k} \delta_{ij}^k y_j^k = d_i \quad \forall i \tag{12}$$

$$y_j^k \in \{0,1\} \quad \forall j,k \tag{13}$$

where the notations are:

Indices

i Item identification

j Supplier identification

k Package identification

Decision variables

y_j^k = 1 if supplier j is assigned package k = 0, otherwise

Data

C_j^k Bid price for package k of supplier j

δ_{ij}^k Volume of item i as a part of package k for supplier j

The objective function (11) minimizes the total procurement cost. The constraint (12) enforces the demand requirements of the buyer. The above formulation allows for each supplier to win more than one package bids. This is *OR* bidding language (implying logical OR). Another popular bidding language used in practice is *XOR*, which allows at most one winning package bid for each supplier. For a more detailed discussion about the bidding languages, see Nisan (2000). The XOR constraint can be easily included as follows:

$$\sum_k y_j^k \le 1 \quad \forall j \tag{14}$$

The above formulation is more appropriate for unit demand d_i=1 for each item i (hence $\delta_{ji}^k = 1$). For multi-unit demands, flexible package bids are beneficial, as the buyer can choose the winning quantity for each supplier.

4.2 Flexible package bids

With flexible package bids, supplier j can provide supply range $[LB_{ji}^k, UB_{ji}^k]$ for item i as a part of package k. The formulation for the WDP is as follows:

$$\min \sum_j \sum_k \sum_i C_{ij}^k x_{ij}^k \tag{15}$$

subject to

$$\sum_j \sum_{k:i\in k} x_{ij}^k = d_i \quad \forall i \tag{16}$$

$$LB_{ji}^k y_j^k \le x_{ij}^k \le UB_{ji}^k y_j^k \quad \forall i,k,j \tag{17}$$

$$y_j^k \in \{0,1\} \quad \forall j,k \tag{18}$$

$$x_{ji}^k \ge 0 \quad \forall i,k,j \tag{19}$$

where the additional decision variable and data are:

x_{ij}^k	Decision variable that denotes the winning quantity for item i from package k of supplier j
C_{ij}^k	Unit bid price for item i from package k of supplier j

4.3 Business constraints

Several business rules are used in combinatorial sourcing. We will need additional decision variables and data to add the business rules as side constraints to the WDP.

Additional decision variables

w_j^i	= 1 if supplier j supplies item i, = 0 otherwise
z_j	= 1 if supplier j is a winning supplier, = 0 otherwise

Additional data

L_i	Item limit of suppliers who can supply item i
$[S', S'']$	Range of number of overall winning suppliers

[*Min_Vol, Max_Vol*]	Minimum and maximum volume guarantee
[*Min_Val, Max_Val*]	Minimum and maximum business guarantee
M	A large constant
F_j	Fixed cost of developing supplier j
F_j^i	Fixed cost of developing supplier j for item i

To limit the number of suppliers at the item level and at the whole sourcing level, following side constraints can be added:

$$x_{ij}^k \le M \ w_j^i \quad \forall i,k,j \tag{20}$$

$$y_j^k \le M \ z_j \quad \forall k,j \tag{21}$$

$$\sum_j w_j^i \le L_i \quad \forall i \tag{22}$$

$$S' \le \sum_j z_j \le S'' \tag{23}$$

$$w_j^i \in \{0,1\} \quad \forall j,i \tag{24}$$

$$z_j \in \{0,1\} \quad \forall j \tag{25}$$

Minimum and maximum volume (business) guarantees can be enforced with the following constraints:

$$Min\_Vol \ z_j \le \sum_k \sum_i x_{ij}^k \le Max\_Vol \ z_j \quad \forall j \tag{26}$$

$$Min\_Val \ z_j \le \sum_k \sum_i C_{ij}^k x_{ij}^k \le Max\_Val \ z_j \quad \forall j \tag{27}$$

Including new suppliers into the sourcing network may incur extra fixed costs. This cost is associated with developing and maintaining a long-term relationship with a new supplier. This is due to the joint technology transfer, engineering, and quality programs with the supplier to enable him to meet the buyer's business and product and requirements. Sometimes the fixed cost could at product level. The fixed cost business constraints, however, need to be added at the objective function.

$$\min \sum_j \sum_k C_{ij}^k x_{ij}^k + \sum_j \sum_i F_{ij} w_j^i + \sum_j F_j z_j \tag{28}$$

4.4 Algorithms

Winner determination problems for combinatorial bids are well studied among the current bid structures. As noted in (Sandholm et al., 2005), three different approaches have been pursued in literature: (1) algorithms that find a provable optimal solution but the computational time dependent on problem instances (Sandholm, 2006), (2) algorithms that are fast with guaranteed computational time but can only find a feasible, not necessarily an optimal solution (Lehmann et al., 2002), and (3) restricting the bundles on which bids can be submitted so that the problem can be solved optimally and provably fast (Rothkopf et al., 1998; Muller, 2006). Combinatorial sourcing are supported and conducted by many commercial providers like CombineNet, Manhattan Associates, JDA, NetExchange, and Trade Extensions.

5. Multi-attribute and multi-criteria sourcing

In industrial procurement, several aspects of the supplier performance, such as quality, lead time, delivery probability, etc have to be addressed, in addition to the qualitative attributes of the procured item. A multi-attribute bid has several dimensions and this also allows the suppliers to differentiate themselves, instead of competing only on cost. Multi-attribute auctions deal with trading of items which are defined by multiple attributes. They are considered to play significant role in the commerce conducted over the WWW (Teich et al., 1999; Bichler, 2001). A multi-attribute auction as a model for procurement within the supply chain was studied in (Che, 1993). It is a one-shot auction in which the suppliers respond to the scoring function provided by the buyer. Multi-attribute auction for procurement proposed in (Branco, 1997) has two stages: A supplier is chosen in the first stage and the buyer bargains with the chosen supplier in the second stage to adjust the level of quality. The other approach in designing multi-attribute auctions is combining multi-criteria decision analysis and single-sided auction mechanisms.

5.1 Scoring function

Evaluating the bids by taking into account different factors is a multi-criteria decision making (MCDM) problem. MCDM has two parts: *multi-attribute decision analysis* and *multiple criteria optimization*. Multi-attribute decision analysis techniques are often applicable to problems with a small number of alternatives that are to be ordered according to different attributes. Two commonly used multi-attribute decision techniques (Belton 1986) are *multi-attribute utility/value theory* (MAUT) (Keeney & Raiffa, 1976) and the *analytic hierarchy process* (AHP) (Saaty, 1980). They use different techniques to elicit the scores or weights, which denote the relative importance among the attributes. MAUT allows one to directly state the scores or estimate as a utility function identified through risk lotteries. AHP uses paired comparisons of hierarchical attributes to derive weights as ratio-scale measures. An insightful comparison of both techniques is presented in (Belton 1986). For a comprehensive study of different multi-attribute decision analysis techniques the reader is referred to (Olson 1996).

Multi-attribute decision analysis has been used in traditional supplier/vendor selection problems (Ghodsypour & O'Brien, 1998; Benyoucef et al., 2003). Multi-attribute auction based on MAUT for e-procurement was proposed in (Bichler et al., 1999). The bids submitted by the suppliers are in the form of (attribute, value) pairs. Each attribute has a set of possible values. Thus a bid is an ordered tuple of attribute values.

Indices

i Attribute identification

K_i Set of possible values for attribute i
j Supplier identification
Multi-attribute bid from j
V_j $(v_{1j}, \ldots, v_{ij}, \ldots)$ where $v_{ij} \in K_i$

The buyer assigns weights to the attributes indicating their relative importance and has a scoring function for each attribute. The scoring functions essentially convert each attribute value to a virtual currency, so that all attribute values can be combined into a single numerical value that quantifies the bid. The combination rule generally used is the weighted additive combination.
Scores and weights
S_i Scores for values of attribute i: $S_i(v_{ij}) \in \mathrm{R}$
w_i Weight for attribute i
Additive scoring function for bid V_j
$\sum_i w_i S_i(v_{ij})$

The above weighted scoring function implicitly assumes *preferential independence of all attributes* (Olson 1996). In other words, the preference for any value of an attribute is independent of any value of any other attribute. However, in many real world applications, interactions exist between attribute values. Such preferential dependencies require non-linear scoring functions, which are seldom used in practice. For a more comprehensive study on the design of multi-attribute auctions see (Bichler, 2001). IBM Research's ABSolute decision engine (Lee et al., 2001) provides buyers, in addition to standard scoring mechanisms, an interactive visual analysis capability that enables buyers to view, explore, search, compare, and classify submitted bids.
An iterative auction mechanism to support multi-attribute procurement was proposed in (Beil & Wein, 2003). The buyer uses an additive scoring function for non-price attributes and announces a scoring rule at the beginning of each round. Through inverse optimization techniques, the buyer learns his optimal scoring rule from the bids of the suppliers. The mechanism is designed to procure a single indivisible item. An English auction protocol for multi-attribute items was proposed in (David et al., 2002), which again uses weighted additive scoring function to rank the bids. All the above mechanisms solve the incomparability between the bids, due to multiple attributes, by assigning a single numerical value to each bid and then ranking the bids by these values. Multi-criteria auction proposed in (Smet, 2003) is an iterative auction which allows incomparability between bids and the sellers increment their bid value by bidding more in at least one attribute. Iterative multi-attribute auctions for procurement were proposed in (Parkes & Kalagnanam, 2005) for procuring a single item. The bid consists of a price for each attribute and the iterative format provides feedback to the suppliers to update their bid prices.

5.2 Multi-criteria optimization for bid evaluation

In multiple criteria decision making situations with large or infinite number of decision alternatives, where the practical possibility of obtaining a reliable representation of decision maker's utility function is very limited, multiple criteria optimization techniques are useful approaches. Multiple attributes can be used both in bid definition and bid evaluation (winner determination). In the following, we describe the use of multiple criteria in bid evaluation using *goal programming* (adapted from Kameshwaran et al. (2007)). In (Beil &

Weun, 2003), the attributes are distinguished as *endogenous* (bidder controllable) and *exogenous* from the bidders' perspective. Attributes in bid definition (or RFQ) provide a means to specify a complex product or service, whereas in bid evaluation, the buyer can use multiple attributes to select the winning bidders. Therefore in bid definition, all attributes should be endogenous for the bidders, whereas in bid evaluation, the buyer can use some exogenous attributes to select the winners. In the MCDM literature, the words criteria and attribute are used interchangeably, and are defined as descriptors of objective reality which represent values of the decision makers (Zeleny, 1982).

We associate the word attribute with the RFQ and bids i.e. the buyer declares in the RFQ various attributes of the goods. We use the word criteria to indicate the objectives defined by the buyer for evaluating the bids. For example, if the attributes defined in the RFQ are cost, delivery lead time, and delivery probability, and then the criteria used by the buyer for evaluating the bids can be total cost, delivery lead time, and supplier credibility. With the above norm established, a criterion for evaluating the bids may consist of zero, one, or many attributes defined in the RFQ. For example, the criterion that the winning supplier should have high credibility, is not an attribute defined in the RFQ but private information known to the buyer. On the other hand, minimizing cost of procurement is a function of many attributes defined in the RFQ. Thus criterion is used here in the sense of an objective.

Multiple criteria optimization problems can be solved using various techniques like goal programming, vector maximization, and compromise programming (Steuer, 1986; Romero, 1991). We describe here the use of (goal programming) GP to solve the bid evaluation problem. Unlike many multiple criteria optimization techniques which require special software tools, GP can be handled by commercial linear and nonlinear optimization software packages with minimal modifications. In GP, the criteria are given as goals and the technique attempts to simultaneously achieve all the goals as closely as possible. For example, the cost minimization criterion can be converted to the goal: *Cost* ≤ \$20, 000, where \$20, 000 is the target or aspiration level. When the target levels are set for all criteria, GP finds a solution that simultaneously satisfies all the goals as closely as possible: It is more of a *satisficing* technique than an optimizing technique. The goal g can be any of the following types:

- greater than or equal to ($\geq t_g$)
- less than or equal to ($\leq t_g$)
- equality ($=t_g$)
- range ($\in [t_g^{'}, t_g^{''}]$)

The t_g's are the target or aspiration levels. Without loss of generality let us assume the following goal structure for the procurement problem:

$$
\begin{aligned}
&\text{goal}\,\{\mathbf{c_1X} = f_1\} && (f_1 \geq t_1)\\
&\text{goal}\,\{\mathbf{c}_2\mathbf{X} = f_2\} && (f_2 \leq t_2)\\
&\text{goal}\,\{\mathbf{c}_3\mathbf{X} = f_3\} && (f_3 = t_3)\\
&\quad\vdots\\
&\text{goal}\,\{\mathbf{c}_G\mathbf{X} = f_G\} && (f_G \in [t_G^{'}, t_G^{''}])
\end{aligned}
\tag{29}
$$

$$
\text{subject to} \quad \mathbf{X} \in F \tag{30}
$$

The $\mathbf{X}$ is the vector of decision variables belonging to the feasible set F. The constraint set $\mathbf{X} \in F$ can be explicitly defined by linear inequalities. For brevity, we will use the above implicit representation. To convert the above GP to a single objective mathematical program, a deviational variable is defined for each goal. It essentially measures the deviation of the respective goal from its target value. Following goal constraints are added to the constraint set (30):

$$\begin{gathered}
\mathbf{c_1 X} + \gamma_1^{+} \geq t_1 \\
\mathbf{c}_2\mathbf{X} - \gamma_2^{-} \leq t_2 \\
\mathbf{c}_3\mathbf{X} + \gamma_3^{+} - \gamma_3^{-} = t_3 \\
\vdots \\
\mathbf{c}_G\mathbf{X} + \gamma_G^{+} \geq t_G^{'} \\
\mathbf{c}_G\mathbf{X} - \gamma_G^{-} \leq t_G^{''} \\
\text{all } \gamma \geq 0
\end{gathered} \tag{31}$$

The range goal gives rise to two constraints but the other goals lead to only one each. The γ_g^{+} measures the deviation away from the goal in the positive direction and γ_g^{-} is for the negative direction. The above goal constraints do not restrict the original feasible region *F*. In effect, they augment the feasible region by casting *F* into a higher dimensional space (Steuer, 1986). The GP techniques vary by the way the deviational variables are used to find the final solution. We present here the weighted GP technique for solving the bid evaluation problem.

Weighted GP (WGP) or *Archimedian* GP uses weights, given by the buyer, to penalize the undesirable deviational variables. The buyer specifies the weight $\kappa_g^{+/-}$ for goal g. The weights measure the relative importance of satisfying the goals. The GP (29) will then be the following single objective programming problem:

$$\min \sum_{g} \kappa_g^{+/-} \gamma_g^{+/-} \tag{32}$$

subject to

$$(31) \text{ and } \mathbf{X} \in F \tag{33}$$

The goals are generally incommensurable (for example, cost minimization is measured in currency whereas minimizing lead time is measured in days) and the above objective function is meaningless as the weighted summation includes different units. The most intuitive and simplest way would be to express g as percentage rather than as absolute value (Romero, 1991). For e-sourcing, the buyer can specify maximum deviation allowed for a goal and then use the percentage of deviation in the objective function.

The multi-attribute sourcing techniques described in this section are extremely useful for sourcing complex goods and services, but they are not wide spread in practice as one would expect. The main hurdle is the lack of exposition of the purchase managers and vendors to these techniques. It is only a matter of time till they are convinced of the profitability of these techniques at the cost of the high complexity, like in the case of combinatorial and volume discount auctions.

5.3 Configurable bids

Configurable bids are used for trading complex configurable products and services like computer systems, automobiles, insurances, transportation, and construction (Bichler et al., 2002). Configurable bids are an extension of multi-attribute bids. A multi-attribute bid is a set of attribute-value pairs, where each pair denotes the value specified by the bidder for the corresponding attribute. In a configurable bid, the bidder can specify multiple values for an attribute. The buyer can configure the bid optimally by choosing an appropriate value for each of the attributes.

Indices

i Attribute identification

k Value identification

j Supplier identification

Configurable bid from j

$\{c_{ij}^{k}\}_{i,k}$ where c_{ij}^{k} is the cost of value k for attribute i

Decision variables

x_{ij}^{k} = 1 if value k is chosen for attribute i for supplier j

The above bid structure implicitly assumes that the total cost is the sum of the individual costs incurred for each attribute. This may not be realistic but on the other hand, defining a cost function over a space of attribute-value pair is pragmatically impossible for the buyer. For example, a bid for 10 attributes with 5 values for each should consider a space of 9.7 million possible configurations. The additive cost structure generally works fine, except for certain constraints. For example, while configuring a computer system, a particular operating system may require a minimum amount of memory but not vice versa. Such logical constraints are not uncommon. Also, such logical constraints can be used to model non-additive cost structures like discounts and extra costs. The logical constraints can be converted into linear inequalities (probably with additional binary variables) and hence can be added to the winner determination problem. Buyer's constraints like homogeneity of values for a particular attribute in multi-sourcing can also be added as constraints to the optimization problem.

The configurable bids and in general, multi-attribute sourcing is not widely used in practice despite the theoretical popularity. Even the laboratory experiments showed encouraging results. Multi-attribute auctions with three different settings were experimented in laboratories: (1) with buyer's scoring function fully revealed for two attributes (Bichler, 2000), (2) with buyer's scoring function not revealed for three attributes (Strecker, 2003), and (3) with partial revelation of the scoring function for three attributes (Chen-Ritzo et al., 2005). All the three showed that multi-attribute auction formats outperform single attribute auctions. Though rarely used in practice currently, one can expect to see its wide spread usage in near future.

6. Global sourcing

Advent of global markets enhanced the emergence of global firms which have factories in different countries. Manufacturers typically set up foreign factories to benefit from tariff and trade concessions, low cost direct labor, capital subsidies, and reduced logistics costs in foreign markets (Ferdows, 1997). Global sourcing is used as a competitive strategy by firms to face the international competition, where suppliers located worldwide are selected to meet the demands of the factories, which are also located internationally (Gutierrez & Kouvelis, 1995; Velarde & Laguna, 2004). The main reasons are lower costs, improved quality, operational flexibility, and access to new technology.

Global sourcing is also used synonymously with outsourcing by some authors. In this chapter, global sourcing is used to denote *international sourcing* or *international purchasing*. In particular, we define global sourcing as procuring from a set of suppliers located worldwide to meet the demands of a set of factories, which are also located worldwide. Thus, there is no single buyer, but a set of buyers (factories belonging to the same company). Consider a company with many factories located domestically in a region. The purchasing department usually aggregates the demands of all the factories (to gain volume discount) and conducts e-sourcing auction for procurement. There is no distinction between the different factories from the suppliers' perspective, as usually they belong to the same region. Consider a multinational company with a set of factories located worldwide. The classical way of managing a multinational is to operate each firm as a domestic firm in its respective country. In the last two decades, global firms started adopting integrated management strategies, which blurs the national borders and treat the set of factories from different countries as a part of the same supply chain network. Global sourcing is one such integrated strategy, where suppliers located worldwide are selected to meet the demands of the factories, which are also located internationally. In this section, we present the design of *global sourcing network*, which is the equivalent to the winner determination problem in the global sourcing scenario.

Global sourcing network (GSN) is a set of suppliers in various countries to support the demands of the firm's international factory network. There are two kinds of decisions that are made in the design of GSN:

- *Supplier selection*: The subset of suppliers to be included in the sourcing network. This is a strategic investment decision that is made at the beginning of planning horizon, which incurs the one-time supplier development costs to the firm.
- *Order allocation*: The allocation of orders from the selected suppliers to the factories to meet the demand at the factories. This is a tactical decision, influenced by the procurement costs.

The first decision is implemented before the planning horizon and the second is implemented during it. This is a single-period problem as there is only one order allocation. The supplier selection decision is assumed fixed and irreversible during the planning horizon i.e. no new suppliers can be added once the decision is made. Each supplier has a fixed development cost, which is the cost of including the supplier in the network. The objective is to minimize the total procurement cost that includes both the supplier development costs and the order allocation costs. Hence, both the decisions are contingent on each other and are made in tandem. In addition to the suppliers, we consider two other sources of supply: *Redundant inventory* and *spot purchase*. Redundant inventory is a part of strategic decision, which once invested incurs a fixed cost irrespective of whether it is used or not. Thus it has a fixed cost and a maximum capacity associated with it. Spot purchase is another option that has no strategic component. If all other sources are unavailable, the organization can always go for this sure but costlier option. We assume that the capacity is infinite. The cost incurred due to lost in sales or unmet demand can also be modelled using this option. It essentially has the same characteristics: *No fixed cost*; *no upper limit*; *sure but costlier option*. All the above can be summarized as follows.

Parameters

- *International factory network*: The number of factories and their locations are assumed to be known and fixed. Index i is used as the factory identifier.
- *Potential suppliers*: The potential global suppliers are assumed to be known and their locations are fixed. Suppliers are identified by index j.
- *Demand*: The demand for the item to be sourced at factory i is d_i.
- *Supply*: The available supply quantity from supplier j is given as range $[a_j, z_j]$, which denotes the minimum and maximum quantity that can be procured from the supplier.
- *Supplier development costs*: The fixed cost of developing supplier j is Fc_j if he is accepted in the sourcing network.
- *Procurement costs*: Unit cost of procurement from supplier j for factory i is c_{ij}.
- *Redundant inventory*: A possible investment in redundant inventory for each factory i with capacity r_i and total cost Ic_i. It is more realistic to assume different levels of investments with varying capacity and cost $\{(r_i^l, Ic_i^l)\}$. For the sake of brevity, we assume only one level of investment for each factory. The proposed model can be easily extended to include various levels.
- *Spot purchase*: For each factory i, there is a sure source of supply with unit cost Sc_i and infinite capacity. Penalty incurred due to lost sales of unmet demand can also be modelled similarly. We have just restricted to one option of this kind per factory for the sake of brevity.

The design of GSN involves identifying an optimal set of suppliers, order allocation from the winning suppliers, investments in the redundant inventories, and the quantity to be spot purchased for the factory network, such that the total cost of procurement is minimized.

Decision variables

x_i = 1 if supplier j is included in the network, = 0 otherwise

y_{ij} Quantity supplied from supplier j to factory i

u_i = 1 if investment is made for redundant inventory at factory i

w_i Spot purchase quantity at factory i

MILP formulation

$$\min \sum_j Fc_j x_j + \sum_i \sum_j c_{ij} y_{ij} + \sum_i Ic_i u_i + \sum_i Sc_i w_i \tag{34}$$

subject to

$$\sum_j y_{ij} + w_i + r_i u_i \geq d_i \quad \forall i \tag{35}$$

$$a_j x_j \leq \sum_i y_{ij} \leq z_j x_j \quad \forall j \tag{36}$$

$$x_j \in \{0,1\} \quad \forall j \tag{37}$$

$$y_{ij} \geq 0 \quad \forall i, j \tag{38}$$

$$u_i \in \{0,1\}, w_i \geq 0 \quad \forall i \tag{39}$$

The above problem is the same as the capacitated version of the well studied *facility location* problem (Drezner & Hamacher, 2002) with the suppliers as the facilities and the factories as the markets with demands. The developing cost of a supplier is the fixed cost associated with opening of a new facility. Many of the algorithms for facility location problem can be adapted for solving the design of GSN problem.

7. Robust e-sourcing

Current supply chains are characterized by leanness and JIT principles for maximum efficiency, along with a global reach. This makes the supply chain highly vulnerable to exogenous random events that create deviations, disruptions, and disasters.

- A strike at two GM parts plants in 1998 led to the shutdowns of 26 assembly plants, which ultimately resulted in a production loss of over 500,000 vehicles and an $809 million quarterly loss for the company.
- An eight-minute fire at a Philips semiconductor plant in 2001 brought its customer Ericsson to a virtual standstill.
- Hurricanes Katrina and Rita in 2005 on the U.S. Gulf Coast forced the rerouting of bananas and other fresh produce.
- In December 2001, UPF-Thompson, the sole supplier of chassis frames for Land Rover's Discovery vehicles became bankrupt and suddenly stooped supplying the product.

Much writings in the recent past as white papers, thought leadership papers, and case studies on supply chain risk management have emphasized that redundancy and flexibility are pre-emptive strategies that can mitigate loses under random events. But this is against the leanness principles and increases the cost. It is required to trade-off between the leanness under normal environment and robustness under uncertain environments. It is in this context; this section briefly introduces *robustness*, a characteristic of winner determination that is almost neglected in current e-sourcing. Caplice & Sheffi (2006), who were directly involved in managing more than hundred sourcing auctions for procuring transportation services, emphasize on the significance of robustness in bid evaluation. The supplier bankruptcy, transportation link failure, change in demand are common sources of uncertainties that are need to be taken into account during bid evaluation.

7.1 Deviations and disruptions

The uncertainties in supply chains might manifest in the form of *deviations*, *disruptions*, or *disasters* (Gaonkar & Viswanadham, 2004). The deviations refer to the change in the certain parameters of the sourcing network like the demand, supply, procurement cost, and transportation cost. The deviations may occur due to macroeconomic factors and the default sourcing strategies may become inefficient and expensive under deviations. Disruptions change the structure of the supply network due to the non-availability of certain production, warehousing and distribution facilities or transportation options due to unexpected events caused by human or natural factors. For example, Taiwan earthquake resulted in disruption of IC chip production and the foot-and-mouth disease in England disrupted the meat supply. Under such structural changes, the normal functioning of supply chain will be momentarily disrupted and can result in huge losses. The third kind of risk is a disaster, which is a temporary irrecoverable shut-down of the supply chain network due to unforeseen catastrophic system-wide disruptions. The entire US economy was temporarily shutdown due to the downturn in consumer spending, closure of international borders and shut-down of production facilities in the aftermath of the 9/11 terrorist attacks. In general, it is possible to design supply chains that are robust enough to profitably continue operations in the face of expected deviations and disruptions. However, it is impossible to design a supply chain network that is robust enough to react to disasters. This arises from the constraints of any system design, which is limited by its operational specification.

First, we characterize the deviations and disruptions that can happen in a sourcing network. The three parameters that influence the sourcing decision are: *Demand*, *supply*, and *procurement cost*. The demand is the buyer's parameter, whereas the supply and the cost are given in the bids by the suppliers. In terms of bid evaluation as a mathematical program, the objective coefficients are the costs and the demand-supply parameters are the right hand side constants of the constraints. The optimal solution to the above mathematical program obviously depends on the three parameters. However, all the three are subject to deviations:

- *cost deviation* due to macroeconomic change or exchange rate fluctuations
- *supply disruption* due to supplier bankruptcy
- transportation link failure due to natural calamity or port strike, leading to *supply disruption*
- *supply deviation* due to upstream supply default
- *demand deviation* due to market fluctuation

The above deviations and disruptions are realized after the bid evaluation but before the physical procurement. Thus, these deviations can render the optimal solution provided by the bid evaluation costly and inefficient, and even sometimes infeasible and inoperable. To handle unforeseen events in sourcing network or in general, supply chain network, there are two obvious approaches: (1) to design networks with built in risk-tolerance and (2) to contain the damage once the undesirable event has occurred. Both of these approaches require a clear understanding of undesirable events that may take place in the network and also the associated consequences and impacts from these events. We show here how we can design a risk-tolerant sourcing network by taking into account the uncertainties in bid evaluation.

7.2 Bid evaluation under uncertainty

Bid evaluation problem is an optimization problem and hence we can draw upon the optimization techniques that can handle randomness in data. The decision-making environments can be divided into three categories (Rosenhead et al., 1972): certainty, risk, and uncertainty. In *certainty* situations, all parameters are deterministic and known, whereas risk and uncertainty situations involve randomness. In *risk* environments, there are random parameters whose values are governed by probability distributions that are known to the decision maker. In *uncertainty* environments, there are random parameters but their probability distributions are unknown to the decision maker.

The random parameters can be either continuous or discrete scenarios. Optimization problems for risk environments are usually handled using stochastic optimization and that for uncertain environments are solved using robust optimization. The goal of both the stochastic optimization and robust optimization is to find a solution that has acceptable level of performance under any possible realization of the random parameters. The acceptable level is dependent on the application and the performance measure, which is part of the modelling process.

Stochastic optimization problems (Birge & Louveaux, 1997; Kall & Wallace, 1994) generally optimize the expectation of the objective function like minimizing cost or maximizing profit. As probability distributions are known and expectation is used as the performance measure, the solution provided is ex-ante and the decision maker is risk neutral. Robust optimization (Kouvelis & Yu, 1994) is used for environments in which the probability information about the random events is unknown. The performance measure is hence not expectation and various robustness measures have been proposed. The two commonly used measures are minimax cost and minimax regret. The minimax cost solution is the solution that minimizes the maximum cost across all scenarios, where a scenario is a particular realization of the random parameters. The minimax regret solution minimizes the maximum regret across all scenarios. The regret of a solution is the difference (absolute or percentage) between the cost of that solution in a given scenario and the cost of the optimal solution for that scenario. Both the approaches have been used to solve the sourcing problem with randomness.

A robust optimization based approach for uncapacitated version of the sourcing problem with exchange rate uncertainty (cost deviation) was considered in (Gutierrez & Kouvelis, 1995; Kouvelis & Yu, 1997). The uncertainties were modelled using discrete scenarios and minimax regret criterion was used to determine the robust solution. In (Velarde & Laguna, 2004), the deviations in both demand and exchange rates were considered. The randomness was modelled using discrete scenarios with probabilities. The objective function had two components: expected cost and variability (that measures the risk). In the following we outline a robust optimization based approach to solve the bid evaluation problem.

7.3 Robustness approach to bid evaluation

The objective here is to propose a robust optimization methodology to design a sourcing network that is risk-tolerant. The choice of robust optimization is due to the fact that managers are more concerned about the outcome of a random event than its probability of occurrence (March & Shapira, 1987). Hence, the optimization of expected cost approach, which implicitly assumes the decision maker to be risk neutral, is not directly applicable. It was also noted by Gutierrez & Kouvelis (1995) that decisions of the managers are not evaluated by their long term expected outcome but by their annual or half-yearly performance. Hence, robust optimization that directly works with the outcome of the random events, rather than probability and long-run expected outcomes, is more appropriate for e-sourcing.

In the proposed methodology, the randomness is modelled via discrete scenarios. The advantage with discrete scenarios is that one need not concern about the source of the scenario, but rather work with the scenario directly. For example, a supplier might get disrupted due to several reasons: Bankruptcy, transportation link failure, upstream supply failure, etc. The buyer needs to only concern about the scenario of a particular supplier failing rather than the sources that would cause it. Working at the level of scenarios is complicated for probability models, as one has to derive the probability of a scenario from the probabilities of the random events that are responsible for that scenario. With no probability information required for robust information, discrete scenario modelling is more appropriate for sourcing. In the following, we abstract the bid evaluation problem to be an optimization problem without specifying the bid structure and the business constraints.

Indices

s Scenario identifier

Data

D^s Demand vector in scenario s

A^s Supply vector in scenario s

C^s Cost vector in scenario s

Notation

X A solution vector to the bid evaluation problem

X^s Optimal solution vector to bid evaluation in scenario s

X^R Robust solution vector

$Z^s(\cdot)$ Cost of solution $(\cdot)$

A scenario s is characterized by a 3-tuple of vectors: $\{D^s, A^s, C^s\}$. Thus, any change in demand or supply or cost or their combinations represents a scenario. By definition, any two scenarios will differ in at least one of the D, A, C vectors. Let s=0 denote the default or unperturbed scenario. The lowest cost L^s for scenario s is its optimal cost:

$$L^s = Z^s(X^s) \tag{40}$$

From the optimality of X^s, it follows that for any solution X, $L^s \leq Z^s(X)$. The relative regret of solution X for scenario s is:

$$r^s(X) = \frac{Z^s(X) - L^s}{L^s} \tag{41}$$

Let U^s denote the maximum cost that will be incurred for scenario s. If $U^s - L^s$ is negligible, then the scenario is not sensitive to the solution. On the other hand if $U^s >> L^s$, the scenario s has to be judiciously handled, even if it is a low probable event, as it might end up with huge increase in the cost. The objective function is robust optimization is a robustness measure. The two commonly used measures are *minimax cost* and *minimax regret*. The minimax cost solution is the solution that minimizes the maximum cost across all scenarios, where a scenario is a particular realization of the random parameters. The minimax regret solution minimizes the maximum regret across all scenarios. The minimax regret objective is given by:

$$\min_X \max_s r^s(X) \tag{42}$$

In general, the minimax versions are overly conservative as the emphasis is on the worst possible scenario, which may occur very rarely in practice. Hence, a solution that is good with respect to the worst-case scenario may perform poorly on the other commonly realizable scenarios. Another measure of robustness is to constrain the regret within pre-specified value $p^s : r^s(X) \leq p^s$ (Snyder & Daskin, 2006). Small values of p^s make the solution X to perform close to that of the optimal solution X^s for scenario s. Thus, by judiciously selecting p^s, the buyer can characterize the importance of scenario s. To implement the above, the following constraints are included in the formulation for robust design.

$$Z^s(X) \leq (1 + p^s)L^s \quad \forall s \tag{43}$$

Note that L^s is the optimal cost for scenario s, and hence for robust design, one needs to solve the bid evaluation problem for each of the scenarios. For any solution X, $L^s \leq Z^s(X) \leq U^s$. Combining with constraint (43), one can derive the maximum value for p^s:

$$p^s \leq \frac{U^s}{L^s} - 1 \qquad \forall s \tag{44}$$

The $\{p^s\}$ are the input parameters to be provided by the buyer to define the *acceptable levels* of operation for different scenarios. Determining U^s will aide the buyer in choosing an appropriate value for p^s. Let X^R be a robust solution that satisfies the constraints (xx). Then,

$$r^s(X^R) \leq p^s \qquad \forall s \tag{45}$$

Given a set of robust solutions $\{X^R\}$, the buyer can choose the best one based on different business criteria. The robust solution ensures that the sourcing network operates at the predetermined operating levels under a wide range of pre-identified scenarios. Thus, the usefulness of the approach clearly depends on the number and the nature of the representative scenarios identified. However, constraint set (43) requires that the winner determination problem needs to be solved for every scenario. As noted in the previous sections, winner determination problems are computationally challenging for complex bid structures and in the presence of business constraints. Thus, the need for solving it for each of the scenarios limits the number of scenarios that can be considered in the robust design.

8. Final notes

This chapter was devoted to optimization based e-sourcing models. It reviewed three popular e-sourcing techniques with their underlying mathematical programming models that are used to solve the winner determination problems. The volume discount and combinatorial sourcing are actively used in business-to-business commerce saving billions of dollars annually. The multi-attribute sourcing technique is yet to catch up, but one can expect increased use in near future given its popularity in literature and encouraging results in laboratory experiments.

All the three techniques reviewed in the chapter are provided commercially as e-sourcing tools by many vendors. We also presented in this chapter two future directions, which are inevitable in the evolution of e-sourcing techniques and tools. One is global sourcing, which connects multiple suppliers with multiple factories, all located internationally. The second is the robust sourcing that takes into account deviations and disruptions, which can render the solution provide by traditional e-sourcing tools inefficient and costly. We presented both the above in the framework of optimization based e-sourcing and hence the currently available methodologies and tools can be adapted to include them.

9. References

Ariba. (2007). *Ariba Corporate fact sheet 2007*. http://www.freemarkets.com

Ausubel & Milgrom. (2006). The lovely but lonely Vicrey auction. In: *Combinatorial Auctions*, Cramton, P., Shoham, Y., & Steinberg, R. (Eds.), pp. 17-40, MIT Press, Cambridge.

Beil, D. R. & Wein, L. M. (2003). An inverse-optimization-based auction mechanism to support a multiattribute RFQ process. *Management Science*, Vol. 49, No. 11, pp. 1529–1545.

Belton, V. (1986). A comparison of the analytic hierarchy process and a simple multi-attribute value function. *European Journal of Operational Research*, Vol. 26, pp. 7–21.

Benyoucef, L., Ding, H. & Xie, X. (2003). Supplier selection problem: selection criteria and methods. *Rapport de recherche 4726*, INRIA, Lorraine, France.

Bichler, M., Kaukal, M. & Segev, A. (1999). Multi-attribute auctions for electronic procurement. *Proceedings of the First IBM IAC Workshop on Internet Based Negotiation Technologies*, pp. 18–19, Yorktown Heights, NY.

Bichler, M. (2000). An experimental analysis of multi-attribute auctions. *Decision Support Systems*, Vol. 29, pp. 249–268.

Bichler, M. (2001). *The Future of e-Markets: Multidimensional Market Mechanisms*. Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.

Bichler, M., Kalagnanam, J. & Lee, H. S. (2002). RECO: Representation and evaluation of configurable offers. In: *Computational Modeling and Problem Solving in the Networked World: Interfaces in Computing and Optimization*, Bhargava, H. K. & Ye, N. (Eds.), Kluwer Academic Publishers.

Bichler, M. & Kalagnanam, J (2005). Configurable offers and winner determination in multi-attribute auctions. *European Journal of Operational Research*, Vol. 160, No. 2, pp. 380–394.

Bikhchandani, S. & Ostroy, J. (2006). Ascending price Vickrey auctions. *Games and Economic Behavior*, Vol. 55, Iss. 2, Pp. 215-241.

Birge, J. R. & Louveaux, F. V. (1997). *Introduction to Stochastic Programming*. Springer-Verlag, New York.

Branco, F. (1997). The design of multidimensional auctions. *Rand Journal of Economics*, Vol. 28, No. 1, pp. 63–81.

Caplice, C. & Sheffi, Y. (2006). Combinatorial auctions for truckload transportation. In: *Combinatorial Auctions*, Cramton, P., Shoham, Y., & Steinberg, R. (Eds.), pp. 539-571, MIT Press, Cambridge.

Che, Y. K. (1993). Design competition through multidimensional auctions. *RAND Jounal of Economics*, Vol. 24, pp. 668–680.

Chen-Ritzo, C., Harrison, T. P., Kwasnica, A. M., & Thomas, D. J. (2005). Better, Faster, Cheaper: An Experimental Analysis of a Multiattribute Reverse Auction Mechanism with Restricted Information Feedback. *Management Science*, Vol. 51, No. 12, pp. 1753–1762.

Cramton, P. (1998). Ascending auctions. *European Economic Review*, Vol 42, pp. 745–756.

Cramton, P., Shoham, Y. & Steinberg, R. (2006). *Combinatorial Auctions*. MIT Press, ISBN. 0262033429, Cambridge.

Crowther, J. F. (1964). Rationale for quantity discounts. *Harvard Business Review*, March-April, pp. 121-127.

Davenport, A.J. & Kalagnanam, J.R. (2002). Price negotiations for procurement of direct inputs. In: *Mathematics of the Internet: E-auction and Markets*, Dietrich, B. & Vohra, R.V. (Eds.), pp. 27–43, Springer, New York.

David, E., Azoulay-Schwartz, R., & Kraus, S. (2002). An English auction protocol for multi-attribute items. In: *Agent-Mediated Electronic Commerce*, Padget, J. A. & Shehory, O. (Eds.), Vol. 2531, Lecture Notes in Computer Science, pp. 52–68, Springer.

Dolan, R. J. (1987). Quantity discounts: Managerial issues and research opportunities. *Marketing Science*, Vol. 6, No. 1, pp. 1-23.

Drezner, Z. & Hamacher, H. W. (2002). *Facility location: Applications and theory*. Springer-Verlag, Berlin.

Emiliani, M. and Stec, D. (2004), Aerospace parts suppliers' reaction to online reverse auctions. *Supply Chain Management: An International Journal*, Vol. 9, No. 2, pp. 139-153.

Emiliani, M. & Stec, D. (2005). Wood pallet suppliers' reaction to online reverse auctions. *Supply Chain Management: An International Journal*, Vol. 10, No. 4, pp. 278–288.

Epstein, R., Henrıquez, L., Catalan, J., Weintraub, G. Y., & Cristian Martınez. (2002). A combinational auction improves school meals in Chile. *Interfaces*, Vol. 32, No. 6, pp. 1–14.

Eso, M., Ghosh, S., Kalagnanam, J. & Ladanyi, L. (2005). Bid evaluation in procurement auctions with piecewise linear supply curves. *Journal of Heuristics*, Vol. 11, No. 2, pp. 147-173.

Ferdows, K. (1997). Making the most of foreign factories. *Harvard Business Review*, Vol. 75, pp. 73–88.

Gainen, L., Honig, D. & Stanley, E. D. (1954). Linear programming in bid evaluation. *Naval Research Logistics Quarterly*, Vol. 1, pp. 48–52.

Gaonkar, R. & Viswanadham, N. (2004). A conceptual and analytical framework for the management of risk in supply chains. *Proceedings of IEEE International Conference on Robotics and Automation*, Vol. 3, No. 26, pp. 2699 - 2704, IEEE ICRA'04, New Orleans, April 26 - May 1, 2004.

Ghodsypour, S. H. & O'Brien, C. (1998). A decision support system for supplier selection using an integrated analytic hierarchy process and linear programming. *International Journal of Production Economics*, Vol. 56–57, pp. 199–212.

Goossens, D. R., Maas, A.J.T., Spieksma, F.C.R. & van de Klundert, J.J. (2007). Exact algorithms for procurement problems under a total quantity discount structure. *European Journal of Operational Research*, Vol. 178, No. 2, pp. 603-626.

Gutierrez, G. J. & Kouvelis, P. (1995). A robustness approach to international sourcing. *Annals of Operations Research*, Vol. 59, pp. 165–193.

Hohner, G., Rich, J., Ng, E., Reid, G., Davenport, A. J., Kalagnanam, J. R., Lee, H. S. & An, C., (2003). Combinatorial and quantity-discount procurement auctions benefit Mars, incorporated and its suppliers. *Interfaces*, Vol. 33, No. 1, pp. 23–35.

Jap, S. D. (2002). Online, reverse auctions: Issues, themes, and prospects for the future. *Journal of the Academy of Marketing Science*, Vol. 30, No. 4, pp. 506–525.

Jucker, J. V. & Rosenblatt, M. J. (1985). Single-period inventory models with demand uncertainty and quantity discounts: Behavioral implications and a new solution procedure. *Naval Research Logistics Quarterly*, Vol. 32, pp. 537-550.

Kall, P. & Wallace, S. W. (1994). *Stochastic Programming*. Wiley, Chichester.

Kameshwaran, S., Benyoucef, L. & Xie, X. (2005). Design of progressive auctions for procurement based on Lagrangian relaxation. *Proceedings of IEEE International Conference on E-Commerce Technology (CEC 2005)*, IEEE Computer Society, pp. 9-16.

Kameshwaran, S., Narahari, Y., Rosa, C. H., Kulkarani, D. M. & Tew, J. D. (2007). Multi-attribute e-Procurement using Goal Programming. *European Journal of Operational Research*, Vol. 179, Iss. 2, pp. 518–536.

Kameshwaran, S., Benyoucef, L. & Xie, X. (2007). Branch-and-cut algorithms for winner determination in discount auctions. *Algorithmic Operations Research*, Vol. 2, pp. 112–128.

Kameshwaran, S. & Narahari, Y. (2009a). Nonconvex piecewise linear knapsack problems. *European Journal of Operational Research*, Vol. 192, Iss. 1, pp. 56–68.

Kameshwaran, S. & Narahari, Y. (2009b). Benders-based winner determination algorithm for volume discount procurement auctions. *International Journal of Logistics Systems and Management*, forthcoming.

Kim, K. H. & Hwang, H. (1988). An incremental discount pricing schedule with multiple customers and single price break. *European Journal of Operational Research*, Vol. 35, pp. 71-79.

Kothari, A., Parkes, D. & Suri, S. (2003). Approximately-strategyproof and tractable multi-unit auctions. *Proceedings 4th ACM Conference on Electronic Commerce (EC-2003)*, pp. 166–175, San Diego, June 2003, ACM.

Kouvelis, P. & Yu, G. (1997). *Robust Discrete Optimization and its Applications*. Kluwer, Boston, MA.

Kouvelis, P., Chambers, C. & Wang, H. (2006). Supply chain management research and Production and Operations Management: Review, trends, and opportunities. *Production and Operations Management*, Vol. 15, No. 4, pp. 449–469.

Lee, H. L. & Rosenblatt, M. J. (1986). A generalized quantity discount pricing model to increase supplier's profits. *Management Science*, Vol. 32, pp. 1177-1185.

Lee, J., Bichler, M., Verma, M., & Lee, H. S. (2001). Design and implementation of an interactive visual analysis system for e-sourcing. *Research report*, IBM Research, Yorktown Heights, NJ, USA.

Lehmann, D., O'Callaghan, L. I., & Shoham, Y. (2002). Truth revelation in rapid, approximately efficient combinatorial auctions. *Journal of ACM*, Vol. 49, No. 5, pp. 577–602.

Lippman, S. A. (1969). Optimal inventory policy with multiple set-up costs. *Management Science*, Vol. 16, pp. 118-138.

March, J. G. & Shapira, Z. (1987). Managerial perspectives on risk and risk taking. *Mangement Science*, Vol. 33, No. 11, pp. 1404-1418.

Mas-Colell, A., Whinston, M. & Green, J. (1995). *Microeconomic Theory*. Oxford University Press, New York.

McAfee, R. P. & McMillan, J. (1987). Auctions and bidding. *Journal of Economic Literature*, Vol. 25, No. 2, pp. 699–738.

Metty, T., Harlan, R., Samelson, Q., Moore, T., Morris, T., Sorensen, R., Schneur, A., Raskina, O., Schneur R., Kanner, J., Potts, K., & Robbins, J. (2005). Reinventing the Supplier Negotiation Process at Motorola. *Interfaces*, Vol. 35, No. 1, pp. 7-23.

Minahan, T. (2001). *Strategic e-Sourcing: A Framework for Negotiating Competitive Advantage*, Aberdeen Group.

Monahan, J. P. (1984). A quantity discount pricing model to increase vendor profits. *Management Science*, Vol. 30, pp. 720-726.

Muller, R. (2006). Tractable Cases of the Winner Determination Problem. In: *Combinatorial Auctions*, Cramton, P., Shoham, Y., & Steinberg, R. (Eds.), pp. 319-336, MIT Press, Cambridge.

Munson, C. L. & Rosenblatt, M. J., (1998). Theories and realities of quantity discounts: An explanatory study. *Production and Operations Management*, Vol. 7, No. 4, pp. 352–369.

Nisan, N. (2000). Bidding and allocation in combinatorial auctions. *Proceedings of the ACM Conference on Electronic Commerce*, pp. 1–12, ACM.

Olson, D. L. (1996). *Decision aids for selection problems*. Springer, New York.

Parkes, D. & Kalagnanam, J. (2005). Models for iterative multi-attribute procurement auctions. *Management Science*, Vol. 51, No. 3, pp. 435–451.

Pirkul, H. & Aras, O.A. (1985). Capacitated multiple item ordering problem with quantity discounts. *IIE Transactions*, Vol. 17, pp. 206-211.

Raiffa, H. & Keeney, R. (1976). *Decisions with Multiple Objectives*. Wiley, New York.

Rassenti, S. J., Smith, V. L. & Bulfin, R. L. (1982). A combinatorial auction mechanism for airport time slot allocation. *Bell Journal of Economics*, Vol. 13, pp. 402–417.

Romero, C. (1991). *Handbook of Critical Issues in Goal Programming*. Pergamon Press, Oxford.

Rosenhead, J., Elton, M., & Gupta, S. K. (1972). Robustness and optimality as criteria for strategic decisions. *Operational Research Quarterly*, Vol. 23, No. 4, pp. 413–431.

Rothkopf, M. H., Pekec, A., & Harstad, R. (1998). Computationally manageable combinatorial auctions. *Management Science*, Vol. 44, No. 8, pp. 1131–1147.

Rothkopf, M. H. & Whinston, A. B. (2007). On E-Auctions for Procurement Operations. *Production and Operations Management*, Vol. 1, No. 4, pp. 404–408.

Sadrian, A. A. & Yoon, Y. S. (1994). A procurement decision support system in business volume discount environments. *Operations Research*, Vol. 42, No. 1, pp. 14-23.

Saaty, T. L. (1980). *The analytic hierarchy process*. McGraw-Hill, New York.

Sandholm, T., Suri, S., Gilpin, A., & Levine, D. (2005). CABOB: A Fast Optimal Algorithm for Winner Determination in Combinatorial Auctions, *Management Science*, Vol. 51, No. 3, pp. 374-390.

Sandholm, T. (2006). Optimal Winner Determination Algorithms. In: *Combinatorial Auctions*, Cramton, P., Shoham, Y., & Steinberg, R. (Eds.), pp. 337-368, MIT Press, Cambridge.

Sheffi, Y. (2004). Combinatorial Auctions in the Procurement of Transportation Services. *Interfaces*, Vol. 34, No. 4, pp. 245-252.

Smet, de Y. (2003). Multicriteria auctions: An introduction. *Proceedings of the 14th Mini-EURO Conference*, HCP 2003.

Snyder, L. V. & Daskin, M. S. (2006). Stochastic p-robust location problems. *IIE Transactions*, Vol. 38, pp. 971–985.

Strecker, S. 2003. Preference revelation in multi-attribute reverse English auctions: A laboratory study. *Proceedings of International Conference on Information Systems (ICIS)*, pp. 271–282, Association for Information Systems, Seattle, WA.

Steuer, R. E. (1986). *Multiple criteria optimization: Theory, computation and application*. John Wiley and Sons, New York.

Teich, J., Wallenius, H. & Wallenius, J. (1999). Multiple issue auction and market algorithms for the World Wide Web. *Decision Support Systems*, Vol. 26, pp. 49–66.

Varian, H. R. (1995). Economic mechanism design for computerized agents. *Proceedings of the 1st Conference on USENIX Workshop on Electronic Commerce - Volume 1*, pp. 2-2, New York, July 11 - 12, 1995, USENIX Association, Berkeley, CA.

Velarde, J. L. & Laguna, M. (2004). A benders-based heuristic for the robust capacitated international sourcing problem. *IIE Transactions*, Vol. 36, No. 11, pp. 1125–1133.

Weber, C. A., Current, J. R. & Benton, W. C. (1991). Supplier selection criteria and methods. *European Journal of Operational Research*, Vol. 50, pp. 2-18.

6

A Domain Engineering Process for RFID Systems Development in Supply Chain

Leonardo Barreto Campos[1], Eduardo Santana de Almeida[2], Sérgio Donizetti Zorzo[3] and Silvio Romero de Lemos Meira[4]
[1]*Federal University of Vale do São Francisco (UNIVASF)*
[2]*Recife Center for Advanced Studies and Systems (CESAR)*
[3]*Federal University of São Carlos (UFSCar)*
[4]*Federal University of Pernambuco (UFPE)*
Brazil

1. Introduction

According to the Supply-Chain Council (1997), the supply chain encompasses every effort involved in producing and delivering a final product or service, from the supplier's supplier to the customer's customer. Supply Chain Management (SCM) includes managing supply and demand, sourcing raw materials and parts, manufacturing and assembly, warehousing and inventory tracking, order entry and order management, distribution across all channels, and delivery to the customer. In this context of several sources of information exchanging data dynamically in supply chain, the Radio Frequency Identification (RFID) appears as a technology able to identify objects such as manufactured goods, animals, and people. Thus, the goal of the RFID technology in supply chain management is to guarantee interoperability providing, for example, accurate and real-time information on inventory of the organizations, product recalls and communications among supply chain participants.

On the other hand, the RFID-based systems used in supply chain management were not considered by a specific software development process. In this scenario, a process is important and necessary to define how an organization performs its activities, and how people work and interact in order to achieve their goals. In particular, processes must be defined in order to ensure efficiency, reproducibility and homogeneity (Almeida, 2007). There are several definitions on software process (Osterweil, 1987), (Pressman, 2005), and (Sommerville, 2006). According to Ezran et al. (2002) software processes refer to all the tasks necessary to produce and manage software, whereas reuse processes are the subset of tasks necessary to develop and reuse assets (Ezran et al., 2002). The adoption of either a new, well-defined, managed software process or a customized one is a possible facilitator for success in reuse programs (Morisio et al., 2002). In supply chain domain, many scenarios and processes are repeatable among supply chain participants (sub-domains), for example, inventory management, shipment and delivery of the goods, and localization of a product. In this sense, software reuse – the process of creating software systems from existing software rather than building them from scratch – is a key aspect for improving quality and productivity in the software development.

In the context of software reuse, important research *including company reports* (Bauer, 1993), (Endres, 1993), (Griss, 1994), (Joos, 1994), (Griss, 1995), *informal research* (Frakes & Isoda, 1995), (Frakes & Kang, 2005) and *empirical studies* (Rine, 1997), (Morisio et al., 2002), (Rothenberger et al., 2003) have highlighted the relevance of a reuse process, since the most common way of software reuse involves developing applications reusing pre-defined assets. The software reuse processes literature focuses on two directions: *Domain Engineering* and, currently, *Software Product Lines* (in section 3 a more detailed discussion about it is presented). Thus, motivated by increasing utilization of the RFID technology in supply chain and the lack of specific development processes for the RFID-based systems development in supply chain, this chapter aims at defining a systematic process to perform domain engineering which includes the steps of domain analysis, domain design, and domain implementation.

In the next section we present the parts of the EPCglobal Network. Eight software reuse processes distributed in domain engineering and software product lines are discussed in the followed section. This is followed by an overview of the proposed domain engineering process. The Sections 5, 6 and 7 describe the domain analysis, domain design, and domain implementation steps respectively. Finally, the conclusion summarizes the contributions this work and directions for future works.

2. The EPCglobal network

One critical issue of the new technologies is their standardization. In case of the RFID systems, both EPCglobal and International Standards Organization (ISO) have adopted RFID in their standards. According (Sabbaghi & Valdyanathan, 2008) the most prominent industry standards for RFID are the EPCglobal specifications and standards for supply chain. The EPCglobal Inc is a nonprofit organization that was initiated in 2003 by MIT Auto-ID Center in cooperation with other research universities to establish and support the EPC Network as the global standard for the automatic and accurate identification of any item in supply chain. The EPCglobal is establishing the standards on how information is passed from RFID readers to various applications, as well as from application to application, in the supply chain. These standards are specified in EPCglobal Architecture Framework, or simply EPCglobal Network. Its is a collection of interrelated hardware, software, and data standards, together with core services that are operated by EPCglobal and is delegates, all in service of a common goal of enhancing business flows and computer applications through the use of Electronic Product Codes (Armenio, 2007). It is composed of five components: (i) Electronic Product Code, (ii) Identification System, (iii) EPC Middleware, (iv) Discovery Services, and (v) EPC Information Service.

Firstly, the Electronic Product Code (EPC) is defined as "a naming and identification scheme designed to enable the unique identification of all physical and virtual objects, assemblies and grouping of objects, and non-objects such as service" (Engels, 2003). It is incorporated into a RFID chip and attached to a physical object. An Electronic Product Code is comprised of header and more three distinct numbers: domain manager number, object class number, and serial number. In this way is possible to provide information about product or object such as your category, data and time of manufacture, final destination, etc. The Identification System consists of RFID tags and RFID readers. RFID Tag is an electronic device composed of microchip and an antenna attached to a substrate, as shown in Figure 1. On the other hand, the RFID readers create a radio frequency field that detects radio waves. When a tag passes through a radio frequency field generated by a compatible reader, the tag reflects back to the reader the identifying information about the object to which it is attached, thus identifying that object.

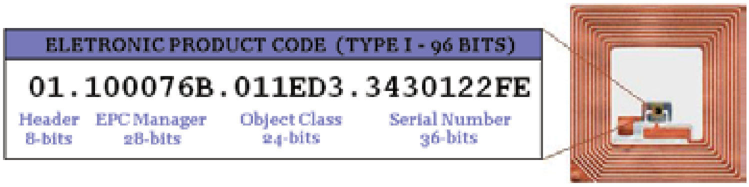


Fig. 1. Electronic Product Code and RFID Tag

Next, the EPC Middleware manages real-time read events and information, provides alerts, and manages the basic read information for communication to EPC Information Services and a company's other existing information systems. The Discovery Services returm locations that have some data related to an EPC (EPCglobal, 2005). In general, a Discovery Services may contain pointers to entities other than the entity that originally assigned the EPC code. Hence, Discovery Services are not universally authoritative for any data they may have about an EPC. The important service in Discovery Services is the Object Name Service (ONS) that, given an EPC, can return a list of network accessible service endpoints that pertain to the EPC in question. Finally, the EPC Information Service (EPCIS) provides an uniform programmatic interface to allow various clients to capture, secure, and access EPC-related data and the business transactions which that data is associated (Harrison, 2003). Companies that assign EPC numbers can maintain EPC Information Service servers with item information. Using EPC numbers does not require organizations to share EPC data or use other components of the system.

The EPCglobal Network presented previously contains several aspects that can be considered by Software Reuse. According to (Harrison, 2003), the hardware, software, and Interfaces defined in EPCglobal Network are management by applications with networked databases. In this sense, software reuse can be used in development of applications, reusing assets available in domains of the Supply Chain. For example, the localization of a product in supply chain is divided into six steps: (i) the RFID reader capture the EPC stored on the tag, (ii) EPC Middleware verify and validate the EPC, (iii) EPC Information Service search data related to EPC in local ONS and return the result, (iv) next, the supply chain participant authenticate it in the EPCglobal Network, (v) ONS search data related to EPC in external databases using EPCglobal Network infrastructure, and (vi) return the search results to application as shown in Figure 2.

This scenario and others situation are commons for some supply chain domains. Therefore, the goal of domain engineering process described in this chapter is to identify common and specific features, scenarios, domain-specific software architecture, and so on, for analysts and designers in a domain, as well as simplifying the identification and implementation of the software components.

The next section presents an analysis involving eleven software reuse processes discussing their, fundamentals, concepts, pros and cons that consists a base for the process defined in this chapter.

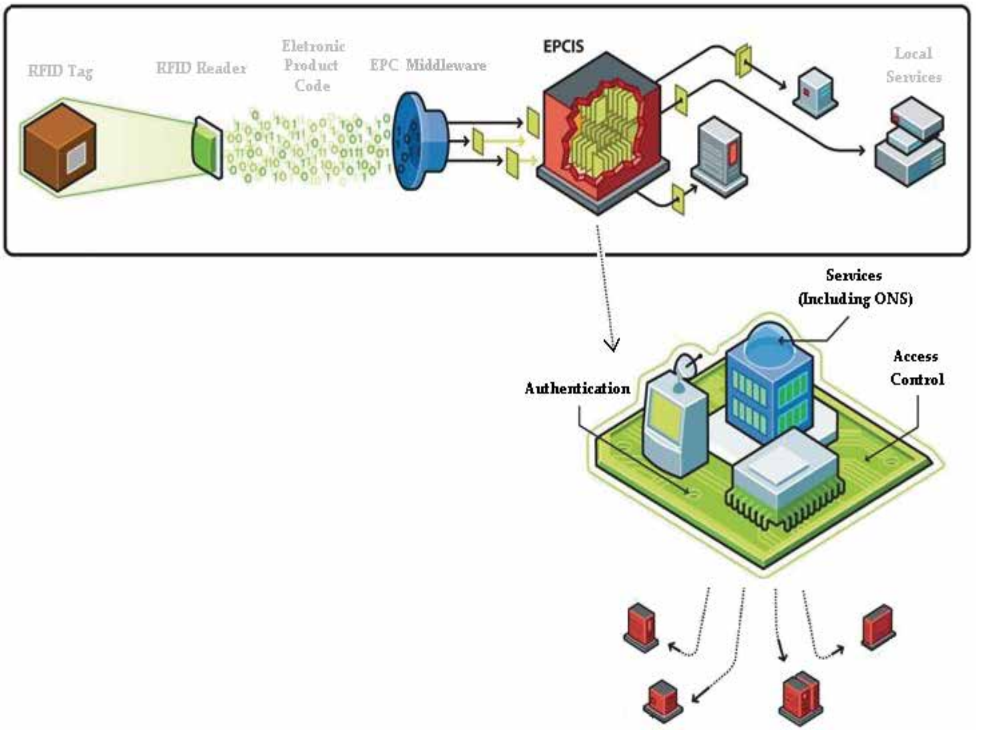


Fig. 2. Scenario of utilization of the EPCglobal Network

3. Software reuse processes

Since the time that software development started to be discussed within the industry, researchers and practitioners have been searching for methods, techniques, and tools that would allow for improvements in costs, time-to-market and quality (Almeida, 2007). The software reuse processes encompass concepts, strategies, techniques, and principles that enable developers to create new systems effectively using previously developed architectures and components. A software reuse process, besides presenting issues related to non-technical aspects (education, culture, organizational aspects, etc), must describe two essential activities: the development *for* reuse (Domain Engineering), which will be discussed next and the development *with* reuse (Application Engineering) which consists in building applications based on the assets produced in Domain Engineering.
In the state-of-art of the software reuse processes presented in (Almeida et al., 2005) and the discussions about it in (Frakes & Kang, 2005) is possible to note that several research studies aimed at finding efficient ways to develop reusable software. These work focus on two directions: Domain Engineering and, currently, Software Product Lines, as can be seen in the next sections.

3.1 Domain engineering

Domain Engineering is defined in (Czarnecki & Eisenecker, 2000) as "*the activity of collecting, organizing, and storing past experience in building systems or parts of systems in a particular domain in the form of reusable assets, as well as providing an adequate means for reusing these assets* *when building new systems*". In this context, this work analysed four domain engineering processes. The first domain engineering approach is called Draco. This approach is based on transformation technology and was developed by James Neighbors in his Ph.D. Work (Neighbors, 1989). The main idea introduced by Draco is to organize software construction knowledge into a number of related domains. Each Draco domain encapsulates the needs and requirements and different implementations of a collection of similar systems. In this work, is presented also, an initial direction to development software using Domain Engineering. However, his approach is very difficult to apply in the industrial environment due the complexity to perform activities such as writing transformations and using the Draco machine.

Described in the 90's, Feature-Oriented Domain Analysis (FODA) is a domain analysis method developed at the Software Engineering Institute (SEI). The method presented in (Kang et al., 1998) consists of two phases: Context Analysis and Domain Modelling. The major contribution of FODA method is the feature model. An important part of this model is the feature diagram that defines three types of features: mandatory, alternative, and optional features. Next, the Feature-Oriented Reuse Method (FORM) (Kang et al., 2002), seen as an elaboration of the FODA, to present four layers to classify features: capability, operating environment, domain technology, and implementations technique. Moreover, the processes do not include essential domain analysis techniques such as domain scoping.

Other important domain engineering method analysed was the Organization Domain Modeling (ODM) developed by Mark Simos. The ODM process described in (Simos et al., 1996) consists of three main phases: Plan Domain, Model Domain, and Engineer Asset Base. However, the phases do not present specific details on how to perform many of its activities. According to (Czarnecki & Eisenecker, 2000) the ODM provides a "*general high-level guidance in tailoring the method for application within a particular project or organization*".

3.2 Software product lines

A new trend started to be explored in software reuse process is the Software Product Line area. According to (Clements & Northrop, 2001), a software product line is "*a set of software-intensive systems sharing a common, managed set of features that satisfy the specific needs of a particular market segment or mission and that are developed from a common set of core assets in a prescribed way*". In this context, we analysed four software product lines processes. Firstly, we present the Family-Oriented Abstraction, Specification and Translation (FAST) process, described in (Weiss & Lai, 1999). The FAST process focuses on pattern for software production processes that strives to resolve the tension between rapid production and careful engineering. The FAST process consists of three well-defined sub-processes: domain qualification, domain engineering, and application engineering. On the other hand, some activities in the process such as in Domain Engineering are not as simple to perform, for example, the specification of an Application Modeling Language.

Presented in (Atkinson, 2000), the Komponentenbasierte Anwendungsentwicklung (KobrA) approach provides a generic assets that can accommodate variants of a product line through framework engineering. The gap in KobrA approach is does not present guidelines to perform systematic tasks such as domain analysis and domain design. An effort to apply the reuse concepts within the embedded systems domain is described in (Winter et al., 2002). The Pervasive Component Systems (PECOS) approach focuses on two issues: how to enable the development of families of PECOS devices? And how pre-fabricated components have

to be coupled? Some gaps was identified in PECOS, for example, in component development, there is not guidance on how the requirements elicitation is performed, and how the components are identified.
Finally, is presented in (Gomaa, 2005) the Product Line UML-Based Software Engineering (PLUS) approach considered the most current process related to product lines. PLUS extends UML by integrating several product line engineering techniques to support UML-based product line engineering. PLUS defines three general steps: requirements, analysis, and design. Modelling in order to provide various modelling techniques and notation for product line requirements engineering activity, use case modelling, and feature modelling. However, in requirements and analysis, activities related to scoping are not considered.

4. The domain engineering process

According to the previous Section, software reuse can be an important way to develop software, offering benefits related to quality, productivity and costs, mainly, using a well-defined reuse process. However, the current software reuse processes present gaps and lack of details in key activities such as, for example, domain engineering. Other requirements like application engineering, metrics, economic aspects, reengineering, adaptation, and quality are important as well, but focus on this chapter is on domain engineering. In this context, this section presents an overview of the proposed domain engineering process, its foundations, and steps.

4.1 Overview of the process

As defined in the previous chapter, domain engineering is the activity of collecting, organizing, and storing past experience in building systems or parts of systems in a particular domain in the form of reusable assets (i.e. reusable work products), as well as providing an adequate means for reusing these assets (i.e. retrieval, qualification, dissemination, adaptation, assembly, and so on) when building new systems (Czarnecki & Eisenecker, 2000, pp. 20). A domain engineering process should define three important steps: Domain Analysis (DA), Domain Design (DD), and Domain Implementation (DI). In general, the main goal of Domain Analysis is domain scoping and defining a set of reusable, configurable requirements for the systems in the domain. Next, Domain Design develops a common architecture for the system in the domain and devising a product plan. Finally, Domain Implementation implements the reusable assets, for example, reusable components, domain-specific languages, generators, and a production process (Czarnecki & Eisenecker, 2000).
Before presenting more details about the proposed domain engineering process, it is important to discuss two basic concepts which will be used next: (i) Domain: encompasses not only the real world knowledge in a given problem area, but also the knowledge about how to build software systems in that area, corresponding to the domain as a set of systems view, (ii) Feature: is widely used in domain analysis to capture the commonalities and variabilities of systems in a domain. In general, there are two definitions of features found in the reuse literature. The first says that an end-user-visible characteristic of a system, the FODA definition. The second definition about feature says that it is a distinguishable characteristic of a concept (e.g. system, component, etc) that is relevant to some stakeholder of the concept, the ODM definition (Simos et al., 1996). In this work, the first definition will be used, since it is the base for the domain analysis area. Other important concepts, more

specific to each step of the process, will be presented later, together with the process detailed description.

4.2 The foundations

A software development process can be understood as the set of activities needed to transform an user's requirements into a software system. The way it is done can change from process to process. For example, processes can be focused on domain engineering, services (Papazoglou & Georgakopoulos, 2003), or Model-Driven Development (MDD) (Schmidt, 2006), or use different process models, however, all kind of process is based on some foundations. In this section, the foundations for the domain engineering process will be presented.

Process Model. A software process model is an abstract representation of a software process (Sommerville, 2006). Each process model represents a process from a particular perspective, and thus provides only partial information about that process. There are different process models published in the software engineering literature, such as Waterfall model, Evolutionary development, Component-Based Software Engineering (CBSE), Incremental delivery, Spiral development, among others (Pressman, 2005). These process models are widely used in current software development practice. The domain engineering process defined in this thesis is based on the spiral process model (Boehm, 1988), however, it presents some characteristics of the CBSE process model, since reusable assets are used to develop applications. The spiral model proposed originally by Boehm (1988), rather than represent the software process as a sequence of activities with some backtracking from one activity to another, consists of a spiral, where each loop represents a phase of the software development process. The Figure 3 shows an overview of the process according to spiral process model. **Domain Driven.** Instead of traditional software development processes as, for example, the RUP, which is use-case driven, the domain engineering process is domain-driven, where the focus is on a set of applications for a particular domain. In this domain, the crucial points are: to identify common and specific features from existing, future, and potential applications; to organize this information in a domain model; next, to design the Domain-Specific Software Architecture (DSSA); and, finally, to implement reusable components for that domain. Even being domain driven, use cases are also used in the process as will be shown in section 6. **Software Architecture.** Software architecture involves the structure and organization by which modern system components and subsystems interact to form systems; and the properties of systems that can be better designed and analysed at system level (Kruchten et al., 2006).

C**omponent-Based Development (CBD).** Component-based Development techniques are important because in domain design, for example, it is interesting to modularize the architecture in well-defined components, which can be easily changed without affecting other parts of the architecture. Moreover, in domain implementation, whose goal is to develop reusable assets, an important way of doing it is through a set of domain-specific components, increasing the reuse potential. **Desig Pattern.** A design pattern is a larger-grained form of reuse than a class because it involves more than one class and the interconnection among objects from different classes. From the perspective of the domain engineering process, design patterns are important because they can be used to encapsulate the variability existing in domain analysis model and perform the mapping for design.

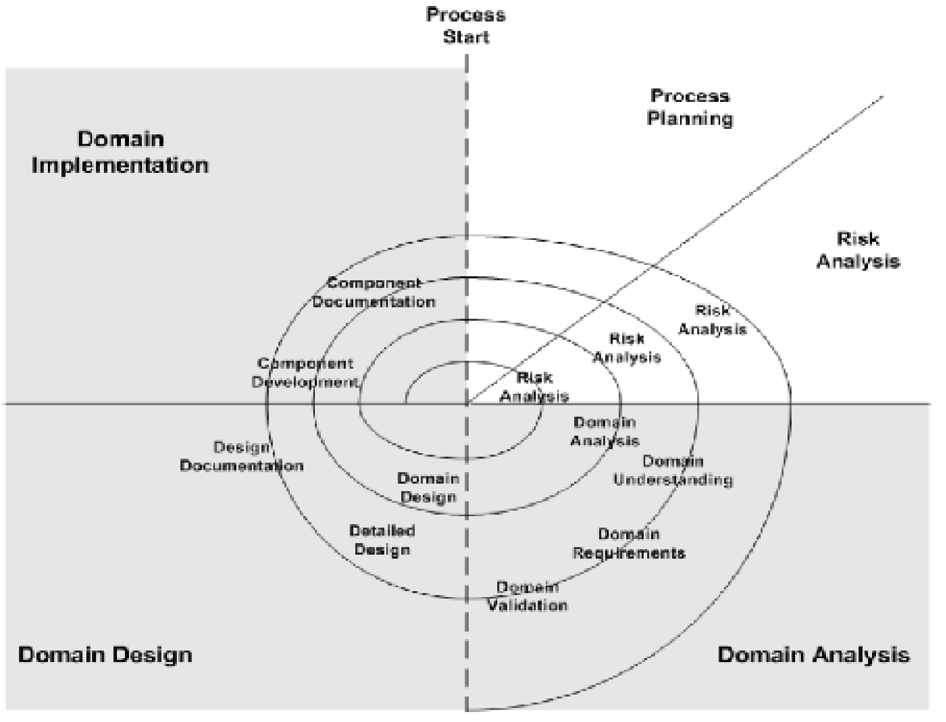


Fig. 3. Process model of the domain engineering process.

4.3 Steps of the domain engineering process

The process for Domain Engineering defined in this chapter is composed of three steps: Domain Analysis, Domain Design, and Domain Implementation. Due to the amplitude, each step is respectively divided in activities and sub-activities. The steps are defined to be used in sequence. However, even with less optimal results, they can be used separately. Therefore, in this chapter, each step will be treated as an approach for a different part of the domain engineering life cycle. In this sense, there is an approach for Domain Analysis, described in section 5, an approach for Domain Design, described in section 6, and, finally, an approach for Domain Implementation, described in section 7.

5. Domain analysis step

The Domain Analysis is considered an important phase in reuse-based software development because it is able to identify common and variable features, analyse the domain scope, define the people (stakeholders) who has a defined interest in the result of the project, etc. The first definition for domain analysis was presented by Neighbors as "*the activity of identifying the objects and operations of a class of similar systems in a particular problem domain*" (Neighbors, 1989). For Neighbors, the identification of objects and operations commons in domain minimizes effort in systems development. However, Neighbors not present steps of "*how to perform*" domain analysis.

Based in this gap, works as (Arango, 1989), (Prieto-Diaz, 1990) and (Almeida, 2006) focus on the outcome, not on the process. For Prieto-Diaz (1990), domain analysis is: "*to find ways to* *extract, organize, represent, manipulate and understand reusable information, to formalize the domain analysis process, and to develop technologies and tools to support it*". In this sense, this section presents a domain analysis approach for engineering RFID-based systems in supply chain. The goal this approach is identify and modelling commons and variables features presents in each domain supply chain (Campos & Zorzo, 2007). The domain analysis consists of four steps: Planning, Requirements, Domain modelling, and Documentation. The next sub-sections presents each step in details.

5.1 Planning

Firstly, ours goals whit domain analysis are: (i) description of the domain, (ii) identify the stakeholders, and (iii) domain scoping. Therefore, the planning step is based in three sub-activities (*P*):

P1. Domain: encompass to describe which supply chain will be applied the domain analysis (e.g., domain of the healthcare, automotive, food, etc). Next, is necessary to divide the supply chain in domains and describe it into four aspects (*A*): *A1. Activity:* that objective of sub-domain in supply chain? *A2. Input:* from who the sub-domain receives information in supply chain. *A3. Output:* to who the sub-domain send information in supply chain. *A4. Technology:* where and which objective to use RFID systems in sub-domain. Hence, the supply chain will be represented how domains that uses RFID systems in specifics activities, inputs, and outputs. *P2. Stakeholder analysis:* the stakeholders are "*people or someone who has a defined interest in the result of the project*". In this sense, many stakeholders can be identified in development and utilization of RFID-based systems in supply chain. For example, the RFID engineering, person that must be expert with EPCglobal Network, RFID readers, RFID tags and yours variabilities, installation, utilization, etc.

P3. Domain scope: this step consists in identify and discard domains in supply chain out of scope. Domains that do not send or receive data for other sub-domains are eliminated. Next, the domain scope analysis is made in terms of horizontal scope. This type of analysis has the goal answer the questions: How many different systems are in the domain? Finally, the last step, consists of analysing the domain scope in terms of vertical scope. Such, is important answer the questions: Which parts of these systems are in the domain? In this context, vertical domains contain complete systems. An organization which does not have any experience with Domain Engineering should choose a small but important domain. After succeeding with the first domain, the organization should consider adding more and more domains to cover its product lines.

5.2 Requirements

The second step in domain analysis is the requirements elicitation, or simply requirements. The goal is to describe the characteristics of the domain and to understand the users' needs. The requirements identification process includes stakeholders as manager, engineering, and end user identified in planning activity. The Requirements activity is not an easy task, mainly, because can exist some problems in potential since the domain contains several systems, (P*r*) as: *Pr1. Ambiguity*: stakeholders do not know what they really want. *Pr2. Redundancy:* requirements of stakeholders different interpreted of the same form. *Pr3.Conflicting Requirements:* Different stakeholders with conflicting requirements. *Pr4. External Factors:* domain requirements may be influenced by organizational and political factors. *Pr5. Stakeholders Evolution:* news stakeholders may emerge during the analysis process. *Pr6. Requirements Evolution:* change of the requirements during the analysis process.

Our effort to minimize errors is to make the requirements elicitation through features. The feature definition used in this chapter is in concordance with (Kang et al., 1998): "*an end-user-visible characteristic of a system*". After defining the form to extract the domain requirements, is necessary to define as to extract them. In this task, the approach uses the concept of the scenarios. The scenarios are descriptions of how a system is used in practice. Thus, the steps (*S*) for requirements elicitation from scenarios are: *S1. Initial stage*: systems stage at the beginning of the scenario. *S2: Events*: normal flow of events in the scenario. *S3. Alternative Events*: eventual events out the normal flow that can cause error. *S4. Finish stage*: systems stage on completion of the scenarios. *S5. Stakeholders*: to list the stakeholders that had participated in scenarios.

5.3 Domain modelling

The Domain Modelling is the third step in domain analysis. Your goal is identifying and modelling commons and variables requirements in domains. In RFID-based systems, the features will be based on the EPCglobal Network and the following sub-activities (*M*): *M1. Commonality analysis:* consist in identify which features are commons to all applications of the domain. There are different ways to identify common requirements. This approach uses a based-priority sub-domain-requirements matrix shown in Table 1. The idea is select requirements by priority for all stakeholders.

	Dom. 1	**Dom. 2**	**Dom. 3**	**. . .**	**Dom. n**
Req. 1	Pr2	Pr1	Pr2	-	-
Req. 2	X	Pr2	Pr3	-	-
Req. 3	Pr1	Pr1	Pr1	-	-
. . .	-	-	-	-	-
Req. n	-	-	-	-	-

Table 1. Structure of Based-Priority Domain-Requirements Matrix

The left column of the matrix lists the requirements of the considered sub-domains. The sub-domains themselves are listed in the top row. In the body of the matrix it is filled by priority of the requirements. The priorities (*Pr*) are classified as follows: *Pr1. High:* the requirement '*Pr1*' is mandatory for all sub-domains and is thus a candidate to be defined as a common domain requirement. *Pr2. Medium:* the requirements that assists high-priority requirements to keep the functionality of the systems. *Pr3. Low*: low-priority requirements to systems. After filling of the matrix, the domain analyst must define ideal priority for commons requirements.

The second sub-activity is *M2. Variability analysis:* this activity consists in identifying which features are variable to applications of the domain. According to (Svahnberg et al., 2001) in situations where a lot of effort has been made to preserve variability until very late in the development process, the systems provides greater reusability and flexibility. Finally, we have the sub-activity *M3. Domains modelling:* here, the commonalities and variabilities are modelled. The model may be applicable at a high level to a number of applications. In this approach the features may be *mandatory, optional, or features* or *alternative* as shown Figure 4 (Czarnecki & Eisenecker, 2000):

According to Czarnecki and Eisenecker (2000) a *mandatory feature* node is pointed to by a simple edge ending with a filled circle. An *optional feature* may be included in the description of a concept instance if and only its parent is included in the description. A concept may have one or more sets of direct *alternative features*. Finally, a concept may have one or more sets of direct *or-features*. However, if the parent of a set of or-feature is included in the description of a concept instance, then any non-empty subset from the set of or-features is included in the description; otherwise, none are included.

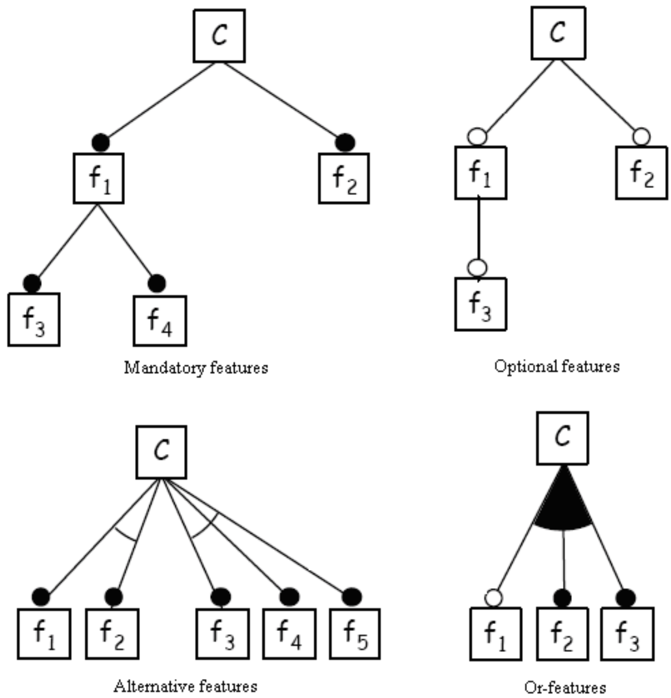


Fig. 4. Features types. Adapted (Czarnecki & Eisenecker, 2000)

Documentation. In this activity the requirements, identified in form of features, will be documented. According to (Czarnecki & Eisenecker, 2000) the template used for document features contain the fields:

6. The domain design step

The second phase of the domain engineering process defined in this chapter is the Domain Design. The key goal this phase is to produce the domain-specific or reference architecture, defining its main software elements and their interconnections in concordance with (Bosch, 2000). The concept of software architecture as a distinct discipline started to emerge in 1990, and in 1995 (Shaw & Garlan, 1996), the field had a strong grow with contributions from industry and academia, such as methods (Kazman et al., 2005) for software architecture. Our domain design approach use the following concept: "*A software architecture is a description of the subsystems and components of a software system and the relations between them. Subsystems and components are typically specified in different views to show the relevant functional and non-functional properties of a software system. The software architecture of a system is an artefact. It is the result of the software development activity*", presented by (Clements et al., 2004).

Feature Name:
Semantic Description
Each feature should have at least description describing its semantics
Rationale
A feature should have a note explaining why the feature is included in the model
Stakeholders and client programs
Each feature should be annotated with stakeholders (e.g., users, customers, developers, managers) who are interested in the feature and the client programs that need this feature
Exemplar applications
If possible, the documentation should describe features with known applications implementing them
Constraints
Constraints are had dependencies between variable feature. Two important kinds of constraints are mutual-exclusion constrains and required constrains
Open/closed attribute
Variation points should be market as open if new direct variable sub-feature (or features) are expected. On the other hand, marking a variation point as closed indicates that no other direct variable sub-feature (or feature) are expected
Priorities
Priorities may be assigned to features in order to record their relevance to the process

The main way of reusing a software architecture is to design a Domain-Specific Software Architecture[1] (DSSA) (Tracz, 1995) or Product-Line Architecture[2] (Dikel et al., 1997). The difference between software architecture in general and a DSSA is that a DSSA is used by all applications in the domain. In this sense, a DSSA for RFID-based Systems in Supply Chain is defined in the domain design phase and your goal is develop an assemblage of software components, specialized for a particular type of task (domain), generalized for effective use across that domain, composed in a standardized structure effective for building successful applications (Tracz, 2005). The next sections present the activities of the domain design: (i) Mapping, (ii) Components Design, (iii) Architecture Views, (iv) and, Architecture Documentation.

6.1 Mapping

The first activity in domain design is the mapping from requirements to reference architecture. An important issue considered in this activity is the variability. According to

1 Term used by the reuse community and adopted in this thesis.

2 Term used by the software product lines community. However, both present the same idea

(Svahnberg et al., 2001), variability is the ability to change or customize a system. Improving variability in a system implies making it easier to do certain kinds of changes. It is possible to anticipate some types of variability and construct a system in such a way that it facilitates this type of variability. In domains supply chains there are many variabilities, both in use of the RFID technology and in supply chain organization. Therefore, the requirements mapping must keep the variability in order to repeat the process for many different domains and offer a reference for it.

Other issue that the domain designer should consider is with components specifications. Some decisions (e.g. algorithms used in component development, objects and types of the component interfaces) can to restrict the component reuse. When these decisions conflict with specific requirements, the components reuse is limited or the system will be inefficient. An efficient way to minimize or eliminate these conflicts is using Design Pattern. According to Christopher Alexander (Alexander et al., 1977) "*each pattern describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million tomes over, without ever doing it the same way twice*". This way, the pattern is a description of the problem and the essence of its solution, so that the solution may be reused in different settings. In the software world, design pattern were popularized by the gang of four (Gamma et al., 1995). According to Gamma and his colleagues, design patterns describe a recurring design problem to be solved, a solution to the problem, and the context in which that solution works. From the perspective of the domain engineering process, design pattern are important because they can be used to encapsulate the variability existing in domain analysis model and perform the mapping for design. In general, a pattern has four essential elements: the pattern name, the problem, the solution, and the consequences.

In the approach presented in this chapter, Design Patterns are used, but together with useful guidelines that determine how and when patterns can be used to represent the different kinds of variability that can exist in a DSSA for RFID-based Systems in Supply Chain. In to order to design the variability of each module, we consider that it should be traceable from domain analysis assets (features) to architecture, according to alternative, or and optional features (Lee & Kang, 2004).

6.1.1 Alternative features

Alternative features indicate a set of features, from which only one must be present in an application. Thus, the following set of patterns can be used (Gamma et al., 1995): **Abstract Factory.** The abstract factory pattern provides an interface for creating families of related or dependent objects without specifying their concrete classes. Specifying a class name when the domain designer creates an object commits you to a particular implementation instead of a particular interface. In this way, this pattern can be used to create objects indirectly and assure that only one feature can be present in the application. In RFID-based systems in supply chain, there are several readers and simultaneous readings. Thus, the EPC Middleware must select only one EPC in case of various requisitions and to discard unnecessary information of the data bases as shows Figure 5.

Chain of Responsibility. This pattern avoids coupling the sender of a request to its receiver by giving more than one object a chance to handle the request. Objects in a chain of responsibility must have a common type, but usually they do not share a common implementation. In this sense, the same requisitions realized in distinct domains can be resolved by different objects. For example, the Discovery Service can want to be able to query the data in local ONS or in external databases. **Factory Method.** Defines an interface for creating an object, but let subclasses decide which class to instantiate. This pattern is similar to the abstract factory and can be used also for alternative features. Finally, **Observer** defines an one-to-many dependency between objects so that when one object changes state, all its dependents are notified and updated automatically. Using this pattern, features can be added to the application as a plug-in, after the deployment. In supply chains, the systems must be flexible the various patterns RFID tags used in identifying of the products or items.

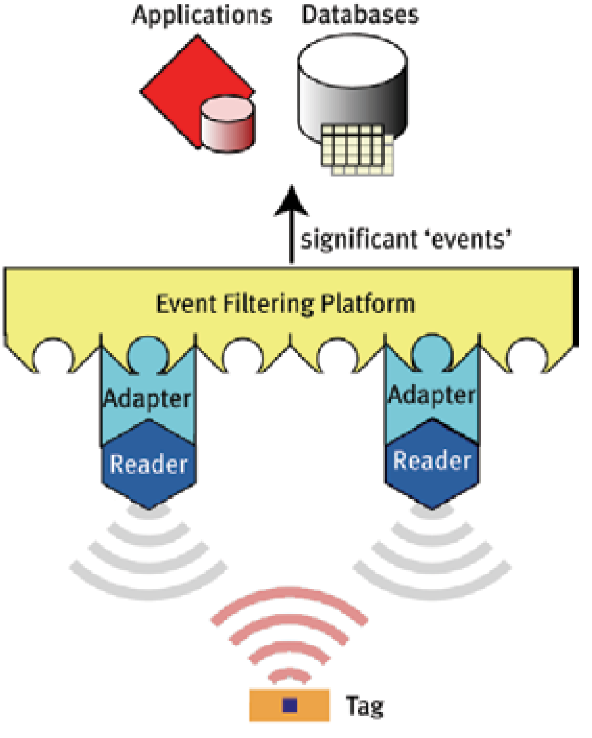


Fig. 5. Simultaneous readings. Adapted (Harrison, 2003)

6.1.2 Optional features

Optional features are features that may or may not be present in an application. For this type of feature, three patterns can be used (Gamma et al., 1995):

Decorator. Attaches additional responsibilities to an object dynamically. Decorators provide a flexible alternative to sub-classes for extending functionality. The decorator pattern can be used for optional features, mainly those that are additional features. Thus, if a feature is present, the ConcreteDecorator is responsible to manage and call the execution. **Prototype.** Specifies the kinds of object to create using a prototypical instance, and create new objects by copying this prototype. The prototype pattern specifies the kinds of objects to create using a prototypical instance, and creates new objects by copying this prototype. In this pattern, the prototype specifies how the interaction wit the feature should be, by defining a concrete prototype for each feature. When the EPC Information Service request data of any EPC to Object Naming Service and it does not provide information, data obtained of the external databases are copied in local server. **Observer.** This pattern can be used in the same way as in alternative features.

6.1.3 Or features

Or features represent a set of features, from which at least one must be present in an application. For this type of feature, three patterns can be used (Gamma et al., 1995):

Bridge. Decouples an abstraction from its implementation so that the two can vary independently. This pattern is appropriated where exist dependence on object representations or implementations, and dependence on hardware and software platform.

Builder. The pattern separates the construction of a complex object from its representation so that same construction process can create different representations. This pattern can be used to build composed features. Thus, for the remainder of the architecture, only the Director is available, being responsible to decide which features will be in the application and which will not.. For example, in the transport of pallets, the application decides what transport unit will be utilized (truck, ship, aeroplane, etc), and creates the object automatically considering its characteristics (size, weight, etc). **Singleton.** Ensures a class only has one instance, and provide a global point of access to it. This pattern also is strongly recommended to or features that interact with mandatory features.

6.2 Component design

In this activity, the goal is to create an initial set of interfaces and component specifications. This activity is composed of two steps: Identify Interfaces, and Component Specification. Firstly, is important understand the concept of interfaces. For (Szyperski et al., 2002) define interface as "*a set of operations, with each operation defining some services or function that the component will perform for the client*". In concordance with (Cheesman & Daniels, 2000) our work considers two types of interfaces: system and business. The business interfaces are abstractions of the information that must be managed by components. Our process for identifying them is the following: to analyse the feature model to identify classes (for each module and component); to represent the classes based on features with attributes and multiplicity; and refine the business rules using formal language. The system interfaces and their operations emerge from a consideration of the feature model and mainly of the use case model. This interface is focused on, and derived from, system interactions. Thus, in order to identify system interfaces for the components, the domain architect uses the following approach: for each use case, he considers whether or not there are system responsibilities that must be modelled. If so, they are represented as one or more operations of the interfaces (just signatures). This gives an initial set of interfaces and operations.

After identifying the interfaces, additional information for specifying components are necessary as, for example, interdependency of the components, and interfaces. The steps presented in this chapter to identifying the interfaces are in concordance with (Cheesman & Daniels, 2000). Firstly, for every component that is specified, the domain architect defines which interfaces its realizations must support (provided and required interfaces). Next, the restrictions of interaction between components must be specified. Unlike the traditional interactions in implementation level, interactions of components define restrictions on the specification level.

6.3 Architecture views

A good way of mapping requirements to implementation is across of the architecture views. The view must be defined as a representation of a set of system elements and the relations associated with them. A view constrains the types of elements, relations and properties that are represented in that view. In this work the four views considered are: (i) module view, (ii) process view, (iii) deployment view, and (iv) data view.

The module view shows structure of the system in terms of units of implementation (e.g. component, class, interfaces and their relations). It is essentials because represents the blueprints for software engineering. Despite of the EPCglobal Network to propose one architecture reference, it is can contain different modules in domains. The module view defines three types de relations in concordance with UML relations between modules (Jacobson et al., 1999): *is part of*, *depends*, *is a* as shown the Figure 6. The first relation "is part of" is used when a package contain sub-packages and class. The second relation "depends of" show the dependences between modules, for example, if the EPCIS need to update your data bases are necessary to authenticate in EPCglobal Network. Finally, the relation "is a" have as goal to represent specialization or generalization among modules, or interface realization. The notation more appropriate to represent module view is although UML diagrams as: package, components, class, and objects diagrams.

The second architecture view defined in this chapter is the runtime view. It shows the systems in execution, your properties, performance, and help to analyse some features in runtime. The best representation for this view is using the UML diagrams following: Interaction, Timing, State Machine, Activity, Communication, and Sequence diagrams. Together with the activity diagram, the state machine diagram to offer more features to describe the process exists in RFID-based systems of the supply chain. These diagrams depict behavioural features of the system or business process.

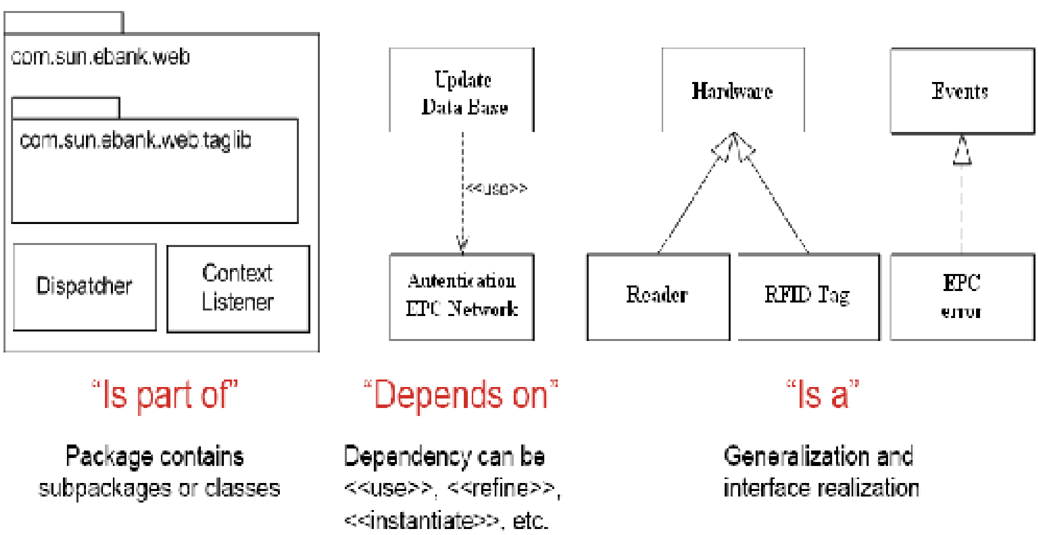


Fig. 6. UML Relations between modules

In deployment view our goal is to describe the hardware structure which the systems are running. Thus, is possible to verify the interconnection between EPC Information Services,

to analyse the performance of the EPCglobal Network, security, and access control to data bases. The UML 2.0 define the deployment diagram with goal of shows the physical deployment of the system, such as the computers, and devices (nodes) and how connect to each other.
Finally, the data view can be used to describe the data bases modelling and their relationship. The goal this view is to improve performance and adaptability of the systems, and to avoid redundancy and enforce consistency. In RFID-based supply chains context the data view is stronger used to represent the data bases that store information about each RFUD tag. The UML diagram that better show the data view representation is the class diagram. However, this view can also be represented entity-relationship diagram.

6.4 Architecture documentation

After defining the view, the domain designer will make the architecture documentation, especially, information that will be applied to more than a vision. In this sense, we define a template with goal of to assist architecture documentation.

Architecture Documentation
1. Guidelines
Describe the way that the architecture documentation is organized, including the use this document in Supply Chain.
2. Design Information
Show design information as, for example, EPC version, previous and auxiliary documents, design members, and goals in general lines.
3. Domain Information
Describe the domain that will project their quality attributes, functional and non-functional requirements with major relevance for supply chain designers.
4. Views Documentation
4.1 Name
4.2 Graphic Representation
4.3 Elements Description
4.4 Relationship of views
4.5 Others information
5. Relation between Analysis and Design
Show which requirements described in analysis phase are in architecture
6. Glossary
Glossary of the system and acronyms

7. The domain implementation step

The last phase of the domain engineering process for RFID-based systems development in supply chain is the Domain Implementation. In concordance with (Pohl et al., 2005), the goal of this step is to provide the implementation and documentation of the reusable assets described in previous step. The activities defined in this chapter for domain implementation step are in concordance with component-based development methods and software reuse processes, among this process are UML Components (Cheesman & Daniels, 2000) and Catalysis (D'Souza & Wills, 1998). The following sections show activities of the domain implementation.

7.1 Component implementation

In this activity, the software engineer, based on requirements, implements the software components through a set of well defined sub-activities. The approach is intended to be used in the scope of domain engineering, and therefore it depends on assets developed in domain analysis (feature model, requirements, domain use case model) and domain design (domain-specific software architecture, component specifications).

This activity is divided into two sets of sub-activities, each one with a different purpose. Sub-activities 1 to 4 deal with the *provided* services, i.e. when the software engineer wants to implement a component to be reused. Sub-activities 5 to 7 deal with *required* services, i.e. when the software engineer wants to reuse services from existent components. The first sub-activity is to describe the component, providing general-purpose information, such as the component vendor, version, package, among others. This information may be used to identify a component, an important issue when components are stored in a repository, for example.

In this second sub-activity, the software engineer should *specify the interfaces*. However, as mentioned before, the domain implementation method depends on artefacts developed in domain analysis and design, such as the domain-specific software architecture and component specifications. These artefacts already contain the interface specification, and so the software engineer only needs to review and refine them, if necessary. In the third sub-activity, the goal is to *implement the services* defined in the previous sub-activity, using any implementation technology, as well as the code to *register these services to be used by other components*, if a dynamic execution environment is used. In fourth sub-activity, which concludes the provided side of the component, the goal is to *build and install the component*. According to the implementation technology used, this involves compiling and packaging the component in a form that is suitable to be deployed in the production environment.

Sub-activities 1 to 4 deal with the *provided* side of a component. In order to implement the *required* side, three sub-activities should be performed: First, the software engineer needs to describe the component that will reuse other services. This is similar to first sub-activity, but with the focus on the services that are required. In this sub-activity, the code that accesses the required services is implemented. Here, different techniques can be employed, such as the use of adapters, wrappers, or other ways to implement this access. The main goal of this sub-activity is to provide low coupling between the required service and the rest of the code, so that it can be more easily modified or replaced. The last sub-activity corresponds to *building and installing the component* that reuses the services, which is similar to fourth sub-activity. Although these two sets of sub-activities (1-4 and 5-7) are focused on different aspects, in practice they will be present in most components, since normally each component has *both* provided and required interfaces.

7.2 Component documentation

When a component is designed and implemented, the developer has clearly in mind some test scenarios and specifics set of the use cases. Thus, case the client does not encompass the component goals, it will be used incorrect way. In this sense, the component documentation is presented by Sametinger (1997) as "*a direct path for information from author to customer, transfers knowledge efficiently. It is one of the most important ways to improve program comprehension [and reduce] software costs*"(Sametinger 1997). This way, Kotula (Kotula 1998) presents thirty nine interrelated patterns as solution for documentation of quality. It is grouped in six categories: (i) *Generative Patterns*: which describe high-level, pattern-creating pattern, (ii) *Content Pattern*: which describe the material that must be included in the documentation, (iii) *Structure Patterns*: which describe how the documentation must be organized, (iv) *Search Patterns*: which discuss the facilities needed to find specific information, (v) *Presentation Patterns*: describing how the documentation should be presented graphically; and (vi) *Embedding Patterns*: which provide guidelines for how to embed documentation content within source code.
Other the hand, (Taulavuori et al., 2004) says that "*definition of the documentation pattern is not sufficient for the adoption of a new documentation practice. An environment that supports the development of documentation is also required*". This way, Taulavuori et al., (2004) provide guidelines concerning how to document a software component. After to analyse patterns defined by Kotula (1998) and component documentation in the context of software product lines, described in (Taulavuori et al., 2004), this chapter defines the following template for component documentation.

Component Documentation
1.General Information
1.1 Name
Should be well defined and describe the component
1.2 Type
Expresses the way the component is intended to be used
1.3 Goal
Describe the relation with the RFID technology present in supply chain
1.4 History
Describes the life cycle of the component.
2. Interfaces
2.1 Required Interfaces
The interface information is here defined as including the interface name, type, description, behaviour and interface functions.

2.2 Provided Interfaces
The same that required interfaces
3. Standards
3.1 Protocol
Describe the interaction between two components needed to achieve a specific objective
3.2 Standards
What EPCglobal Network standards the component is using? What standards are necessaries? Standards can restrict the compatibility, structure and functionality of components.
4. Technical Details
The Technical Details describes details the of design and implementation component.
4.1 Development Environment
Defines the environment in which the component has been development.
4.2 Interdependencies
Describes component's dependency on the components
4.3 Prerequisites
Defines all the other requirements that component may have to operate.
4.4 Implementation
Implementation includes composition, context, configuration, and interface implementation. Composition information describes the internal structure of the component, which can be derived from the component's class diagram. The component's class diagram must be included, if possible, as well as the classes, operations and attributes.
4.5 Restrictions
Should describe all the items that will limit the provider's options for designing or implementing the components.
5. Information Non-functional
5.1 Modifiability
Define as the component can be adapted in new supply chains.
5.3 Expandability
Describes how new is often qualified only for OCM components.
5.3 Performance
Performance is a quality attribute that measures the component.. The measurements are

size of the component, prioritisations of events, and utilization in RFID systems.
5.4 Security
Define strategies to combat hackers. Including cryptography in RFID tags level.
6. General Information
6.1 Installation guide
Defines the operations that must be performed before the component can be used.
6.2 Support
Customer support includes the name of the contact person and an address or a phone number where the customer can get help.

8. Conclusion

As widely discussed in this chapter, the reuse process present gaps and lack of details in key activities such as, for example, domain engineering. In this sense, we believe that our approach can be useful to reduce the gaps and lack of details among the steps, and presenting a domain engineering process for RFID-based systems development presents in supply chain domain. This work can be seen as initial climbing towards the full vision for an efficient domain engineering process and interesting directions remain to improve what started and to explore new routes. Thus, the following issues should be investigated as future work: (i) Other RFID Standards, (ii) Other Directions in Software Architecture, for instance, Service-oriented or Model-Driven and (iii) Architecture Documentation.

9. References

Alexander, C. (1977). A Pattern Language: Towns, Buildings, Construction, Oxford University Press, 1977, pp. 1216

Almeida, E. S. et al., (2005). A Survey on Software Reuse Processes, *IEEE International Conference on Information Reuse and Integration,* pp. 66-71, Nevada, USA, August, 2005, Las Vegas

Almeida, E. S. et al., (2006). The Domain Analysis Concept Revisited: A Practical Approach, *9th International Conference on Software Reuse,* Lecture Notes in Computer Science, Torino, Italy, June, 2006, pp. 43-57

Almeida, E. S. (2007). RiDE – The RiSE Process for Domain Engineering, Ph. D. Thesis, Federal University of Pernambuco, Recife, Brazil, 2007

Arango, G. (1989). Domain Analysis: from art form to Engineering Discipline, *5th International Workshop on Software specification and design,* Pennsylvania, USA, May, 1989, Pittsburgh

Armenio, F et al., (2007) *The EPCglobal Architecture Framework,* EPCglobal Final Version, pp. 7, September 2007

Atkinson, C.; Bayer, J. & Muthig, D. (2000). Component-Based Product Line Development: The KobrA Approach, *First Software Product Line Conference,* Kluwer International Series in Software Engineering and Computer Science, pp. 19, Colocado, USA, August, 2000, Denver

Bauer, D. (1993). A Reusable Parts Center, *IBM Systems Journal,* Vol. 32, No. 04, pp. 620-624, September, 1993

Boehm, B. W. (1988). A Spiral Model of Software Development and Enhancement, *IEEE Computer,* Vol. 21, No. 05, pp. 61-72, May, 1988

Bosch, J. (2000). Design and Use of Software Architecture, Addison-Wesley, 2000, pp. 354

Campos, L. B. & Zorzo, S. D. (2007). A Domain Analysis Approach for Engineering RFID Systems in Supply Chain Management, *IEEE International Conference on System of Systems Engineering,* Texas, USA, pp. 165-171, April, 2007, San Antonio

Cheesman, J. & Daniels, J. (2000). UML Component A Simple Process for Specifying Component-Based Software, Addison-Wesley, 2000, pp. 208

Clements, P & Northrop, L. (2001). Software Product Lines: Practices and Patterns, Addison-Wesley, 2001, pp. 608

Clements, P. et al., (2004). Documenting Software Architectures: Views and Beyond, Addison-Wesley, 2004, pp. 512

Czarnecki, K & Eisenecker, U. W. (2000). Generative Programming: Methods, Tools, and Applications, pp. 832, Addison-Wesley, 2000

Dikel, D. (1997). Applying Software Product-Line Architecture, *IEEE Computer,* Vol. 30, No. 08, pp. 49-55, August, 1997

D'Souza, D. F. & Wills, A. C. (1998). Objects, Components and Frameworks with UML: The Catalysis Approach, Addison-Wesley, 1998, pp. 816

Endres, A. (1993). Lessons Learned in an Industrial Software Lab, *IEEE Software,* Vol. 10, No. 05, pp. 58-61, September, 1993

Engels, D (2003). The use of the Electronic Product Code, *White Paper of the Massachusetts Institute of Technology* (MIT), pp. 03, February 2003.

EPCglobal (2005). Object Naming Service (ONS) Version 1.0, EPCglobal Ratified Specifications, pp. 09, October, 2005

Ezran, M.; Morisio, M & Tully, C. (2002). Practical Software Reuse, pp. 374, Springer

Frakes, W. B. & Isoda, S. (1994). Success Factors of Systematic Software Reuse, *IEEE Software,* Vol. 12, No. 01, pp. 15-19, September, 1994

Frakes, W. B. & Kang, K. C. (2005). Software Reuse Research: Status and Future, *IEEE Transactions on Software Engineering,* Vol. 31, No. 07, pp. 529-536, July, 2005

Gamma, E. (1995). Design Patterns: Elements of Reusable Object-Oriented Software, Addison-Wesley, 1995, pp. 395

Gomaa, H. (2005). Designing Software Product Lines with UML: From Use Cases to Pattern-Based Software Architectures, Addison-Wesley, 2005, pp. 701

Griss, M. L. (1994). Software Reuse Experience at Hewlett-Packard, *16th IEEE International Conference on Software Engineering,* pp. 270, Italy, May, 1994, Sorrento

Griss, M. L. (1995). Making Software Reuse Work at Hewlett-Packard, *IEEE Software,* Vol. 12, No. 01, pp. 105-107, January, 1995

Harrison, M. (2003). EPC Information Service – Data Model and Queries, White paper of the Massachusetts Institute of Technology (MIT), pp. 03, October, 2003

Jacobson, I.; Booch, G. & Rumbaugh, J. (1999). The Unified Software Development Process, Addison-Wesley, 1999, pp. 463

Joos, R. (1994). Software Reuse at Motorola, *IEEE Software,* Vol. 11, No. 05, pp. 42-47, September, 1994

Kang. K. C. et al., (1998). FORM: A Feature-Oriented Reuse Method with domain-specific reference architectures, *Annals of Software Engineering Notes,* Vol. 05, No. 00, January, 1998, pp; 143-168

Kang, K. C.; Lee, J. & Donohoe, P. (2002). Feature-Oriented Product Line Engineering, *IEEE Software,* Vol. 19, No. 04, pp; 58-65, July/August, 2002

Kazman et al., (2005). A Basis for Analyzing Software Architecture Analysis Methods, *Software Quality Journal,* Vol. 13, No. 04, pp. 329-355, December, 2005

Kotula, J. (1998). Using Patterns To Create Component Documentation. *IEEE Software,* Vol. 15, No. 02, pp. 84-92, March/April, 1998

Kruchten, P.; Obbink, H. & Stafford, J. (2006). The Past, Present, and Future of Software Architecture, *IEEE Software*, Vol. 23, No. 02, pp. 22-30, March/April, 2006

Lee, K. & Kang, K. C. (2004). Feature Dependency Analysis for Product Line Component Design, *8th International Conference on Software Reuse,* Madri, Spain, July, 2004, pp. 69-85

Morisio, M; Ezran, M & Tully, C. (2002). Success and Failure Factors in Software Reuse, *IEEE Transactions on Software Engineering*, Vol. 28, No. 04, pp. 340-357, April, 2002

Neighbors, J. M. (1989). Draco: A Method for Engineering Reusable Software Systems, in Software Reusability Volume I: Concepts and Models, T. Biggerstaff, A. Perlis, 1989, pp. 425

Osterweil, L. (1987). Software Process are Software too, *9th International Conference on Software Engineering*, pp. 02-13, California, USA, March/April, 1987, Monterey

Papazoglou, M. P. & Georgakopoulos, D. (2003). Service-Oriented Computing, *Communications of the ACM,*Vol. 46, No. 10, pp. 25-28, October, 2003

Pohl, K.; Bockle, G. & van der Linden, F. (2005). Software Product Line Engineering: Foundations, Principles, and Techniques, Springer, 2005, pp. 467

Pressman, R. S. (2005). Software Engineering: A Practitioner's Approach , pp. 880, McGraw-Hill, 2005

Prieto-Diaz, R. (1990). Domain Analysis: An Introduction, *ACM SIGSOFT Software Engineering Notes,* Vol. 15, No. 02, pp. 47-54, April, 1990

Rine. D. C. (1997). Success Factors for Software Reuse that are Applicable Across Domains and Businesses, *ACM Symposium on Applied Computing*, pp. 182-186, California, USA, March, 1997, San Jose

Rothenberger, M. A.; Dooley, K. J.; Kulkarni, U. R. & Nada, N. (2003). Strategies for Software Reuse: A Principal Component Analysis of Reuse Practices, *IEEE Transactions on Software Engineering,* Vol. 29, No. 09, pp. 825-837, September, 2003

Sabbaghi, A. & Valdyanathan, G. (2008). Effectiveness and Efficiency of RFID Technology in Supply Chain Management: Strategic values and Challenges. *Journal of Theoretical and Applied Electronic Commerce Research,* Vol. 03, No. 02, August 2008, pp. 71-81, ISSN 0718-1876 Electronic Version

Sametinger, J. (1997). Software Engineering with Reusable Components, Springer-Verlag, 1997, pp. 275

Schmidt, D. C. (2006). Model-Driven Engineering, *IEEE Computer,* Vol. 39, No. 02, pp. 25-31, February, 2006

Shaw, M & Garlan, D. (1996). Software Architecture: Perspective on an Emerging Discipline, Prentice Hall, pp. 242

Simos, M. et al. (1996). Organization Domain Modeling (ODM) Guidebook, Version 2.0, *Technical Report,* June, 1996, pp. 509

Sommerville, I. (2006). Software Engineering, pp. 840, Addison Wesley

Supply-Chain Council (1997). New group aims to improve supply chain integration, Purchasing, pp. 8-32, Vol. 123, No. 6

Svahnberg, M; van Gurp, J & Bosch, J. (2001). On the Notion of Variabilities in Software Product Lines, *Working IEEE/IFIP Conference on Software Architecture,* Amsterdam, Netherlands, August, 2001, pp. 45-54

Taulavuori, A.; Niemela, E. & Kallio, P; (2004). Component Documentation – a key issue in software product lines, *Journal Information and Software Technology,* Vol. 46, No. 08, pp. 535-546, June, 2004

Tracz, W. (1995). DSSA (Domain-Specific Software Architecture) Pedagogical Example, *ACM SIGSOFT Software Engineering Notes,* Vol. 20, No. 03, pp. 49-62, July, 1995

Weiss, D. M. & Lai, C. T. R. (1999). Software Product-Line Engineering: A Family-Based Software Development Process, Addison-Wesley, 1999, pp. 426

Winter, M.; Zeidler, C. & Stich, C. (2002). The PECOS Software Process, *Workshop on Component-based Software Development, 7th International Conference on Software Reuse*, pp. 07, Texas, USA, April, 2002, Austin

7

Return Policies and Coordination of Supply Chain

Mabel C. Chou
Department of Decision Sciences
National University of Singapore,
Singapore

1. Introduction

In many industries, manufacturers reply upon independent retailers to distribute their products as manufacturers can benefit from retailers' reputations, economies of scale, and knowledge about local markets. However, manufacturers and retailers being independent agents act in their own best interest. A supply chain composed of independent agents acting in their own best interests will generally not achieve system-wide efficiency, often due to some incongruence between incentives faced locally and the global optimization problem, a phenomenon known in the economics literature as "double marginalization". (Spengler 1950, Tirole 1988).

To efficiently use the supply chain one needs some mechanisms to coordinate supply chain to maximize total channel profit. Coordination of the supply chain implies that the profit of the supply chain is maximized, hence achieving system-wide efficiency. One way to coordinate channel decisions is by using a returns policy provided by the manufacturer.

This paper is organized as follows. Section 2 presents an overview of the benefits and costs of returns policies. Section 3 reviews the different kinds of returns policies that are required to coordinate the supply chain for the different types of products. Section 4 discusses how returns policies will be affected by demand uncertainty and Section 5 studies the impact that demand uncertainty together with retailing competition has on returns policies. Section 6 provides a conclusion of the paper.

2. Overview of return policies

A returns policy for excess inventory is a commitment made by a manufacturer, or distributor upstream when accepting products from a downstream channel member. The most generous policy promises to refund the full wholesale price for all returned products, while less generous policies offer credits against future orders. A partial returns policy gives only partial credits or refunds. (Padmanabhan & Png 1995)

The format of returns policies varies in and across industries. Manufacturers and distributors across a wide range of products have long allowed retailers to return excess inventory. Today, returns policies are common in the distribution of books, magazines and newspapers, recorded music, computer hardware and software, greeting cards and pharmaceuticals. (Padmanabhan & Png 1995)

2.1 Benefits of returns policies

At the most superficial level, returns policies will encourage retailers to carry larger stocks. From the manufacturer's standpoint, the more retailers sell, the higher the manufacturer's profit will be. Returns policies are thus beneficial to both the manufacturers and retailers. Below is a discussion of the benefits to the retailers.

2.1.1 Mitigate retailers' risk

Uncertain product demand is the reason for retailers' reluctance to carry more stock as they fear having excess inventory. By stocking conservatively, retailers limit the manufacturer's potential sales. The manufacturers can mitigate the retailers' risk of having excess inventory by offering returns policies which will encourage the retailers to stock more products and hence increase the sales of the manufacturers. (Padmanabhan & Png 1995)

2.1.2 Safeguard the brand

The institution of returns policies will discourage retailers from marking down the price of the product when it is nearing the expiration date and selling of stale products. Returns policies are a more attractive way of dealing with expiring products and will help to safeguard the manufacturer's brand. (Padmanabhan & Png 1995)

2.1.3 Support end-user returns policy

The institution of return policies will be beneficial to retailers as it allows them to return products to the manufacturers when they are pressured to accept returns from their own customers. End users want to be able to return products to retailers to safeguard against the risk that the product will not satisfy them. (Padmanabhan & Png 1995)

2.1.4 Facilitate distribution of new product information

Returns policies may aid information transmission between manufacturers and retailers more effectively. A manufacturer that is sure its products will sell well can afford to offer a generous returns policy, knowing that retailers will not return the items as explained by Chu (1993). Hence a returns policy is one way a new manufacturer can credibly signal information about expected market demand and product quality. (Padmanabhan & Png 1995)

Since returns policies also place the consequence of product failure on the manufacturers, the incentives of the manufacturers and retailers are thus aligned. As a result, a returns policy is also a way the manufacturer can commit to investments in advertising, promotion and other activities to enhance product sales. (Padmanabhan & Png 1995)

Returns policies also serve as forms of assurance to the retailers that a manufacturer will not bring out new products so quickly that the retailers' stock will be obsolete. If the manufacturer does so, retailers can return the superseded items, hence punishing the manufacturer. (Padmanabhan & Png 1995)

2.1.5 Structure competition

A returns policy strengthens a manufacturer's position relative to other competing brands as it reduces the cost of carrying excess inventory and tilts the balance in the retailers' mind towards carrying larger stocks of the manufacturer's product. As a result, the probability of a stock-out and of consumers switching to competing brands will be lower. Hence the institution of returns policies has strategic implications for manufacturer's profits by having an impact on competition among retailers and brands. (Padmanabhan & Png 1995)

2.2 Costs of returns policies

When some of the additional stocks are returned, a returns policy generates additional costs for the manufacturer which will be discussed below.

2.2.1 Logistics costs

One of the costs incurred by a returns policy is the logistic costs of organizing, packing, and shipping of products back and forth. Depreciation cost is another cost that the manufacturer incurred on returned items. Returned items depreciate as they may be damaged in shipment, decay physically or lose their marketability over time. (Padmanabhan & Png 1995)

2.2.2 Demand uncertainty

The demand for products such as new books, CDs, software and pharmaceuticals is uncertain. Since a returns policy transfers the cost of excess stocks from retailers to the manufacturer, the more uncertain demand is, the greater the cost of a returns policy to the manufacturer. (Padmanabhan & Png 1995)

2.2.3 Retailer incentives

By reducing the risk of losses due to excess inventory, a returns policy lessens some of the retailer's incentive to invest in efforts to promote retail sales by merchandising, doing point-of-sale advertising and providing attractive shelf space. (Padmanabhan & Png 1995) Hence, when manufacturer accepts the risk from the retailer, the retailer's incentive to invest in promotional efforts may be dulled and this is a cost to the manufacturer.

2.3 Implementation of returns policies

The partial returns policy which rebates only part of the wholesale price for return items is most widely implemented to address the retailers' incentive to overstock and avoid point-of-sale marketing efforts. In this way, the manufacturer and the retailer share the risk, hence providing some incentives for all parties to play their part. Partial risk gives the manufacturer an incentive to support the product and to select new introductions carefully while partial risk gives retailers the incentive to order conservatively and promote the product. Ultimately, the aim of the partial returns policy is to coordinate the supply chain to maximize total channel profit.

Some other kinds of partial returns policy are those with a time limit for returns as well as those which take into consideration a retailer's returns history into decisions on pricing, credit and even on whether to continue dealing with the retailer.

A time-consuming and expensive way for a manufacturer to devise returns policies for a mix of retailers that differ in risk aversion, competitiveness and skepticism is to tailor a returns policy for each retailer. Another way is to design a menu of alternative returns policies such as more generous returns policies with higher wholesale prices or strict limits on returns with lower wholesale prices and allow every retailer to choose among the options. (Padmanabhan & Png 1995)

3. Types of products

This section will investigate the different kinds of returns policies that are required to coordinate the supply chain for the different types of products, namely perishable commodities, style goods, catalogue style goods, experience goods, Internet sales and build-to-order products.

3.1 Perishable commodities

Perishable commodities are short life-cycle items with limited shelf or demand life such as newspapers, baked goods, periodicals and records.

An earlier investigation of employing a returns policy for perishable commodities to coordinate the distribution channel appeared in Pasternack (1985) where two extreme cases under the assumption of risk-neutral supply chain members are analyzed. It is impossible to achieve channel coordination if the manufacturer allows the retailer full credit for all unsold perishable goods or no returns for any unsold perishable goods.

On the other hand, if the manufacturer offers retailers full credit for a partial return of goods this may achieve channel coordination in a single-retailer environment. Since the optimal return allowance will be a function of retailer demand, such a returns policy cannot be optimal in a multi-retailer environment. In contrast, he shows that a returns policy in which a manufacturer offers a partial credit for all unsold goods can achieve channel coordination in a multi-retailer environment.

3.2 Style goods

Style goods are characterized by highly uncertain demand, long production lead time and have little or no salvage value at the end of their short selling seasons of a few weeks or months. They include products such as computer software, hardware, compact discs, fashion items, greeting cards, books and pharmaceuticals.

Mantrala and Raman (1999) study how the supplier's wholesale-buyback price policy influence the retailer's optimal order quantity decision with different levels of demand variability as well as study how demand uncertainty the profitability of both the suppliers and retailers. With demand variability remaining constant, supplier can induce the retailer to purchase more stock by increasing the buyback price and/or reducing the wholesale price.

When demand uncertainty increases, the supplier tends to drop his optimal buyback price even though the buyer is ready to increase her order quantity at any given wholesale price. This indicates that the supplier finds that his expected costs of returns at the higher buyback price are too high relative to his expected revenues. In addition, at any given wholesale price level, the retailer's total profits as well as the total system profits fall as demand variability increases.

On the other hand, at a very low wholesale price, the supplier does not find it optimal to accept returns at all and hence his optimal buyback price is zero. However, the buyer is still willing to order larger amount of stocks at higher levels of uncertainty even without a buyback price offer provided the wholesale price is sufficiently low. Hence, increasing demand uncertainty works completely in favor of the supplier in this situation and thereby improves his expected profits.

However, at the higher wholesale price, the supplier has to offer a large buyback price which significantly increases his expected costs of returns from the retailer, and thereby lowers his expected profits as demand uncertainty increases.

Finally, the supplier makes his largest profits at high wholesale prices accompanied by high buyback prices. With the right combination he can make the retailer purchase as much style goods as possible at a much lower wholesale price and no buyback price.

3.2.1 Catalogue style goods

Catalogue refers to a particular kind of style goods in which a retailer advertises an item at a particular price in a catalogue that is distributed to customers. Since cost considerations prohibit the frequent distribution of catalogues, the retailer must commit to a single price for an item for the entire selling season.

With a price-dependent demand model, Emmons and Gilbert (1998) show that under demand uncertainty, a supplier using a returns policy for catalogue style goods to repurchase excess stock from the retailer at the conclusion of the period can improve the profits if demand follows a uniform distribution. They also find that there exists at least a wholesale price where returns policy is a ‘win–win’ strategy for retailer and supplier. The offer to buy back excess stock tends to increase the total combined profits of the retailer and manufacturer; hence it is not a zero-sum game. A manufacturer can "buy" the loyalty of a retailer relatively cheaply by decreasing the wholesale price by a small amount. It is always in the retailer's interest to have the manufacturer buy back unsold items at the end of the season, and the manufacturer also benefits from such arrangements when wholesale prices are sufficiently large. However, when the manufacturer sets his wholesale price optimally, he gains more from the offer of a returns policy than does the retailer.

3.3 Experience goods

Experience goods are goods for which idiosyncratic valuations such as buyer’s remorse are possible only after purchase. This is because customers do not fully know their preferences for the products at the time of purchase, but after they gain some experience with them.

Returns policies for experience goods allow customers to defer their purchasing decisions until after gaining some experience with the products. A consumer who has learned that he does not like a product can cancel his purchase by simply returning it. These returns policies do not require customers to provide evidence or an explanation regarding the malfunction of the returned good. Instead, a customer not liking a product is often a sufficient reason for stores to accept the return. The "no-questions-asked" full refund policy is customary with many big retailers.

According to Che (1996), returns polices for experience goods insure consumers against ex post loss, which allow a monopoly seller to charge a higher price. This is because under the returns policy, consumers can return the good after learning their valuations, at zero cost. However, the seller cannot induce consumers to buy at a price above their ex post valuations with a no-returns policy. This is because under the no-returns policy, consumers bear the entire risk associated with their uncertain ex post valuation. As risk aversion increases, the seller must lower her price to compensate consumers for the risk.

Che (1996) also demonstrates that the returns policy is optimal if the consumers are sufficiently risk averse. This is because the returns policy eliminates a consumer's risk of paying more than his ex post valuation; hence the marginal consumer is not adversely affected by risk aversion. He also shows that returns policy is optimal if the retail costs are high. When the retail costs are high, the screening opportunity of the returns policy: seller can charge a high price and sell only to high-valuation consumers is relatively more valuable since the seller can maintain her profit margin by selling only to high-valuation consumers. Furthermore, superior risk sharing makes consumers strictly better off under the returns policy as the consumers are protected from any loss and so they receive strictly positive expected utility. However, the seller's failure to internalize this benefit leads to too little adoption of the returns policy in equilibrium.

3.4 Internet sales

The e-business revolution in recent time has brought an alternative model for the part of the supply chain from the manufacturer to the customer. More and more manufacturers are now attempting to sell directly to the customers bypassing the traditional distributor-wholesaler-retailer chain. Their motivation for this is to reduce the distribution cost and be more responsive to customers' requirements.

Customers also view Internet purchase as advantageous because it drastically reduces the search cost and is convenient due to the fact that the store is open 24 hours per day seven days a week. In an Internet direct sales supply chain, the customers buy direct from the manufacturer, hence sacrificing the benefits of physical inspection of the product. This increases the probability that the customers will be dissatisfied with the product and likely to return it. Hence, a common customer's concern is the lack of a proper returns policy for internet purchase and the complicated logistics for returning an item. As such, a clearly explained and generous returns policy will be welcomed by the customers and therefore will enhance demand.

From the manufacturer's point of view, a generous returns policy will increase customers' confidence and hence increase sales revenue by inducing more customers to buy. On the other hand, returns policies would increase the cost of business substantially due to increased likelihood of return. Hence, returns policy constitutes a tradeoff and an optimum policy would be one whereby the resultant profit would be maximized. (Mukhopadhyay and Setoputro 2004) The returns policy practice in e-business varies across industries and stores, and may range from unconditional money back guarantee to store credit only to no refund.

Mukhopadhyay and Setoputro (2004) find out that in a market where customers are less price-sensitive, the e-tailer can offer a more generous returns policy and at the same time will be able to charge higher price. Optimum pricing and returns policies will generate higher demand for the product but the e-tailer will also see higher return quantity. Overall, profit can be maintained at a higher level because the extra revenue from charging higher price and from increased sales outweighs the increase in cost due to increase in return quantity.

In addition, they also find out that in a market where customer's demand is increasingly more sensitive to the returns policy, the optimum price will increase and the e-tailer can offer more generous returns policy to increase sales. Although offering more generous returns policy will also increase return quantity, the e-tailer's profit will not be reduced. This is because when customer is more sensitive to returns policy, e-tailer can charge higher price to offset the cost increase due to offering more generous returns policy.

Results from the paper also show that when the customer is less sensitive to the rate of return parameter, offering generous or restrictive returns policy will not make much difference. Hence, the e-tailer can offer more generous returns policy without affecting profit level. This can be observed in the arts industry where the customers are less sensitive to the rate of return parameter and so sellers generally offer more generous returns policy. On the other hand, when customer is sensitive to the rate of return parameter, e-tailer tends to offer more restrictive returns policy because e-tailer is afraid that customers may abuse their returns policy. This can be seen in the electronic and apparel industries where the customers are widely known as sensitive to the rate of return parameter and so sellers often impose less generous returns policy on their customers.

3.5 Build-to-order products

A build-to-order product (BTO) is essentially built to customize the product to the requirement of the customers, hence increasing both lead time and cost of production. Firms are generally reluctant to offer a returns policy because the returned merchandise is practically useless as it was designed to meet the requirement of a particular customer. Hence returns policy will not be suitable in case of a BTO product with almost zero salvage value.

Mukhopadhyay and Setoputro (2005) propose the concept of modularity in the institution of returns policies for BTO products. When a BTO product with modular design is returned, the product can be dismantled very easily producing a number of components which are standard products keeping their full value. This returned product will give back a large salvage value to the firm, thereby cutting down its loss due to the return. Hence, the company will incur a lower cost of return from the returned product when the level of modularity is higher.

They also demonstrate that in the market where customer demand is more sensitive to the returns policy, the seller will offer more generous returns policy and simultaneously the optimum modularity level shall be increased. This is because since the BTO product is customized, offering a generous returns policy without increasing the modularity level will increase the cost of returning.

In addition, in the market where the demand is increasingly more sensitive to the modularity level, the optimum modularity level will increase and the seller shall simultaneously offer more generous returns policy. Moreover, when the seller can decrease their product development and design costs, the optimum modularity level will increase and at the same time the firm can offer more generous returns policy. Furthermore, when the seller can salvage more from the returned product, the optimal returns policy and the optimum modularity level offered will increase. Generous returns policy is favorable when the retailer can obtain high salvage value for returned merchandise.

4. Demand uncertainty

This section will discuss how return policies will be affected by demand uncertainty. Manufacturers whose products are subject to uncertain demand face a problem of inducing distributors to stock those products. A manufacturer may attempt to compensate its distributors by agreeing to accept returns of unsold goods for full or partial refunds of their purchase price.

4.1 Nature of demand uncertainty

Marvel and Peck (1995) show that the manufacturer's decision to accept returns depends crucially on the nature of demand uncertainty. Two cases of demand uncertainty are discussed in their paper. Valuation uncertainty occurs when firms are unsure about the willingness of customers to pay for the manufacturer's product while arrival uncertainty occurs when firms do not know how many customers will arrive. When uncertainty applies only to the valuation that consumers place on the manufacturer's product, but not to arrival uncertainty, returns policies distort pricing without offsetting the inventory efforts. Prices are forced up while the expected quantity sold remains low and hence returns policies are not employed in the case of valuation uncertainty. On the other hand, if the manufacturer and retailer have a good idea about the amount consumers will be willing to pay for the product, but do not know how many consumers will arrive at the marketplace, a high return allowance is attractive to the manufacturer.

In addition, the authors demonstrate that when arrival uncertainty is small relative to valuation uncertainty, the manufacturer chooses not to permit returns. However, increases in the arrival uncertainty parameter result in the institution of returns policies and the ratio of the returns to wholesale price rises rapidly with arrival uncertainty.

Marvel and Peck (1995) also point out that return allowances are far more widespread in Japan than in United States. The fragmented nature of Japanese distribution into smaller units as compared to those in the West is a consequence of legal obstacles to construction of large stores. Distributing sales over a larger number of outlets will increase the arrival uncertainty at each individual outlet relative to the sales at that outlet. However, valuation uncertainty is unlikely to be affected at any individual outlet. As a result, the liberal return policies of Japanese manufacturers are a profit-maximizing adaptation to the fragmented nature of Japanese distribution.

4.2 Multi-store retailer

Mantrala and Raman (1999) study the impact of demand uncertainty on supplier's returns policies for a multi-store style-good retailer. In their research the retailer, assumed to be a central department store, had two different retail outlets whose demand may or not have been correlated and the retail outlets are not in competition. For any given non-optimal wholesale price and returns value, they numerically study how demand variability affected suppliers' and retailers' profits and return credits. At any given wholesale price, the supplier tends to drop his optimal buyback price as demand uncertainty increases even though the buyer is ready to increase her order quantity. On the other hand, at a very low wholesale price, the supplier in fact does not find it optimal to offer to accept returns at all; hence his optimal buyback price is zero. However, the supplier makes his largest profits at high wholesale prices accompanied by high buyback prices.

4.3 Multi-item returns policy

Brown, Chou and Tang (2008) study a multi-item returns policy called "pooled" (or joint) returns policy under which the distributor can return any combination of the products up to R percent of the total purchases across all products. They analyze the distributor's optimal profit and order quantity under the pooled returns policy, and compare these operating characteristics to the case when a single-item "non-pooled" returns policy is instituted. Under the non-pooled returns policy, the distributor can only return on individual items using item-specific return limits.

Under the non-pooled policy, the distributor can return each product separately up to R percent of the purchase of that product. They show that the distributor will always achieve a higher profit under the pooled policy. They also show that the manufacturer could actually obtain a lower profit under the pooled policy due to a counter-intuitive result: the distributor may order less under the pooled policy even though the pooled policy offers more flexibility. This counter-intuitive result motivates them to determine the conditions under which the distributor would order less under the pooled policy. Finally, they develop a heuristic for determining the distributor's optimal order quantities associated with the n-product case under the pooled policy.

5. Demand uncertainty with retail competition

The papers by Marvel and Peck (1995) and Mantrala and Raman (1999) only consider the effect of demand uncertainty on returns policies. In this section, we will study the impact that demand uncertainty together with retailing competition have on returns policies.

Padmanabhan and Png (1997) point out that when retailing is competitive but demand is certain, a returns policy intensifies retail competition and reduces the retailer margins, hence benefiting the manufacturer. On the other hand, when retailing is a monopoly while demand is uncertain, a returns policy helps the manufacturer by intensifying retail distribution, but hurts by encouraging excessive stocking. They also demonstrate that when there are both competing retailers and demand is uncertain; there is a trade-off between benefits of more intense retail competition and the costs of excessive stocking of a returns policy. As a result, the manufacturer shall adopt a returns policy if the marginal production cost and the demand uncertainty are sufficiently low. The lower the marginal cost and demand uncertainty, the greater will be the benefit from more intensive retail competition and the smaller will be the manufacturer’s loss from excessive stocking. Since the marginal costs of production of books, CDs and computer software are small in comparison to their price, returns polices are common in the distribution of books, recorded music and software.

In addition, Yao, Wu and Lai (2005) study how demand variability and retailer’s competition interacts with manufacturers’ return prices in influencing retailer decisions on order size and retail price, and the implications for manufacturers’ policies and profitability of the parties in the supply chain.

They point out that a returns policy always benefits the manufacturer. If the demand uncertainty and wholesale price are very high, the best decision of manufacturer is to provide either a return credit or full returns policy. However, channel profits, and particularly the expected profits of the retailers, may not fare well under this policy. If the manufacturer provides a lower wholesale price, the optimal decision of the manufacturer is not to provide any returns credit.

They also demonstrate that the competing power of the retailer has an impact on the distribution channel members’ decisions. Intensifying the competition between two duopoly competing retailers will lead to a decline in both the retailers’ and the channel expected profits. In contrast, supplier’s expected profits are increasing as the supplier’s wholesale price increases with his setting the optimal returns price under all scenarios. The supplier’s action of setting high wholesale price leads to a high equilibrium retail price that in turn leads to a decline in market demand, thus resulting in a decrease in the retailers’ expected profits. Hence, although a returns policy can induce the retailer to order more so that the supplier’s profits and channel profits improve, the retailer’s expected profits are destroyed significantly. Furthermore, increasing the demand uncertainty works completely in favor of the supplier in this situation. On the other hand, in the higher uncertain demand situation, the retailers should adopt a risk-averse policy so that they can share more total profits.

Furthermore, they find out that the price sensitivity factor has a significant impact on the returns policy. With low price sensitivity, the manufacturer does not generally adopt a returns policy, particularly in low demand uncertainty. In contrast, with high price sensitivity, the manufacturer will like to adopt a returns policy to improve his profits. However, the returns policy requires the support of the high wholesale price, which leads to the severe cannibalization of the retailers’ profits.

6. Conclusion

We have come to the concluding section of this literature review. A returns policy provided by the manufacturer can coordinate the supply chain to maximize total channel profit. Any manufacturer may find that the benefits and costs of a returns policy pull in different directions and whether a manufacturer should accept returns depends on the balance between the benefits and costs.

We have also looked at the different kinds of returns policies that are required to coordinate the supply chain for the different types of products, namely perishable commodities, style goods, catalogue style goods, experience goods, Internet sales and build-to-order products. In addition, we also study the impact of these return policies on the optimal pricing of the manufacturer and optimal order quantity decision of the retailer.

We also discuss the impact that demand uncertainty alone have on return policies and extend our study to include the impact of both demand uncertainty and retailing competition on returns policies.

7. References

Brown, A., Chou, M.C., Tang, C.S., 2008. The Implications of Pooled Returns Policies. International Journal of Production Economics, 111, 129-146.

Che, Y.-K., 1996. Customer Return Policies for Experience Goods. Journal of Industrial Economics, 44, 17-24.

Chu, W., 1993. Demand Signaling and Screening in Channels of Distribution. Marketing Science, 11, 327-347.

Emmons, H., Gilbert, S.M., 1998. Note: The Role of Returns Policies in Pricing and Inventory Decisions for Catalogue Goods. Management Science, 44 (2), 276-283.

Lai, K.K., Wu, Y., Yao, Z., 2005. Demand uncertainty and manufacturer returns policies for style-good retailing competition. Production Planning & Control, 16(7), 691-700.

Mantrala, M.K., Raman, K., 1999. Demand uncertainty and supplier's returns policies for a multi-store style-good retailer. European Journal of Operational Research, 115, 270–284.

Marvel, H.P., Peck, J., 1995. Demand Uncertainty and Returns Policies. International Economic Review, 36(3), 691–714.

Mukhopadhyay, S.K., Setoputro, R., 2004. Reverse logistics in e-business optimal price and return policy. International Journal of Physical Distribution & Logistics Management, 34(1), 70-88.

Mukhopadhyay, S.K., Setoputro, R., 2005. Optimal return policy and modular design for build-to-order products. Journal of Operations Management, 23, 496-506.

Padmanabhan, V., Png, I.P.L., 1995. Returns policies: Make money by making good. Sloan Management Review, Fall, 65-72.

Padmanabhan, V., Png, I.P.L., 1997. Manufacturer's Returns Policies and Retail Competition. Marketing Science, 16 (1), 81-94.

Pasternack, B.A., 1985. Optimal Pricing and Return Policies for Perishable Commodities. Marketing Science, 4(2), 166–176.

Spengler, J. J. 1950. Vertical restraints and antitrust policy. J. Political Economics, 58, 347–352.

Tirole, J. 1988. The Theory of Industrial Organization. MIT Press, Cambridge, MA.

8

Operational Management of Supply Chains: A Hybrid Petri Net Approach

Mariagrazia Dotoli, Maria Pia Fanti and Agostino Marcello Mangini
Dipartimento di Elettrotecnica ed Elettronica, Politecnico di Bari
Bari,
Italy

1. Introduction

The emergence of Supply Chains (SCs) is an outcome of the recent advances in logistics and information technology. SCs are complex networks interconnecting different independent manufacturing and logistics companies integrated with material, information and financial flows (Viswanadham & Raghavan, 2000). Typically, SC management decisions are classified into three hierarchical levels according to the time horizon of decisions: strategic (longterm), tactical (medium-term), and operational (short-term, real-time) (Chopra & Meindl, 2001; Shapiro, 2001). Accordingly, different models have to be defined at each level of the decision hierarchy to describe the multiple aspects of the SC. While the development of formal models for SC design at strategic and tactical levels was addressed in the related literature (Dotoli et al., 2005; Dotoli et al., 2006; Gaonkar & Viswanadham, 2001; Luo et al., 2001; Vidal & Goetschalckx, 1997; Viswanadham & Gaonkar, 2003), research efforts are lagging behind in the subject of modeling and analyzing the SC operational performance.
At the operational level, SCs can be viewed as Discrete Event Dynamical Systems (DEDSs), whose dynamics depends on the interaction of discrete events, such as customer demands, departure of parts or products from entities, arrival of transporters at facilities, start of assembly operations at manufacturers, arrival of finished goods at customers etc (Viswanadham & Raghavan, 2000). Among the available DEDS analytical formalisms, Petri Nets (PNs) may be singled out as a graphical and mathematical technique to model systems concurrency and synchronization. Moreover, PNs are able to capture precedence relations and structural interactions and may be executed in standard engineering software packages simply implementing their dynamics via the corresponding matrix equations. However, most SC models based on PNs proposed in the related literature share the limitation that products are modelled by means of discrete quantities, called tokens (Desrochers et al., 2005; Dotoli & Fanti, 2005; Elmahi et al., 2003; Viswanadham & Raghavan, 2000; Von Mevius & Pibernik, 2004; Wu & O'Grady, 2005). This assumption is not realistic in large systems with a huge amount of material flow: the state space of the SC model generally turns out to be excessively large, so that inconveniences in the simulation and performance optimization often arise, leading to large computational efforts. Since SCs are DEDSs whose number of reachable states is very large, PN formalisms using fluid approximations provide an

aggregate formulation to reduce the state space dimension (Alla & David, 1998; Silva & Recalde, 2004). Hence, hybrid PNs may be employed to describe SCs efficiently and effectively at the operational level. In this context, First Order Hybrid Petri Nets (FOHPNs) (Balduzzi et al., 2000) are an emerging formalism.

This chapter focuses on SC operational management: we employ a FOHPN model recently proposed by the authors (Dotoli et al., 2008) and implement and compare two standard operational management strategies. The model is built by a modular approach based on the bottom-up methodology (Zhou & Venkatesh, 1998): manufacturers are described by continuous transitions, buffers are continuous places and products are continuous flows routing from manufacturers, buffers and transporters. Transporters are stochastic transitions with a triangular distribution for the transportation time. Moreover, discrete places and transitions describe the financial and information flows that enable, inhibit or change the material flow. Discrete exponential transitions model information about demands and occurrence of unpredictable events, e.g. the blocking of a supply or an accident in a transportation facility. The FOHPN model allows us to address two issues: system management and optimal mode of operation. The SC management is realized by PN structures that synthesize two well-known policies, namely Make-To-Stock and Make-To-Order, and standard inventory control rules. While the SC optimization models presented in the related literature determine the decision parameters off-line (e.g. see Gaonkar & Viswanadham, 2001; Vidal & Goetschalckx, 1997; Viswanadham, 1999; Viswanadham & Gaonkar, 2003) in order to design and manage the SC, the task of the considered model is selecting some operational SC parameters in a short time, based on knowledge of the system state and of the occasional uncontrollable events. More precisely, the optimal mode of operation is obtained by the computation of control variables, i.e., the instantaneous firing speeds of continuous transitions, solving a linear programming problem optimizing a chosen performance index. A representative SC example shows the technique effectiveness under the two operational management policies and by a standard inventory control rule, considering in each case a different optimal operative condition. Simulation results show that the selected formalism leads to an effective SC operational management, as well as to the possibility of choosing important system control parameters, e.g. the production rates. Future research includes implementing additional management policies and inventory rules in the chosen formalism.

2. The system description

2.1 The SC structure

The SC structure is typically described by a set of facilities with materials that flow from the sources of raw materials to subassembly producers and onwards to manufacturers and consumers of finished products. Moreover, feedback paths may be present if demanufacturers or recyclers are included in the SC. The SC facilities are connected by transporters of materials, semi-finished goods and finished products. More precisely, the entities of a SC are the following:

1. *Suppliers*: a supplier is a facility that provides raw materials, components and semifinished products to manufacturers that make use of them.
2. *Manufacturers and assemblers*: manufacturers and assemblers are facilities that transform input raw materials/components into desired output products.

3. *Distributors*: distributors are intermediate nodes of material flows representing agents with exclusive or shared rights for the marketing of an item.
4. *Retailers or customers*: retailers or customers are sink nodes of material flows.
5. *De-manufacturers or recyclers*: entities of the de-manufacturing stage feed recovered material, components or energy back to suitable upstream SC facilities.
6. *Logistics and transporters*: storage systems and transporters play a critical role in distributed manufacturing. The attributes of logistics facilities are storage and handling capacities, transportation times, as well as operation and inventory costs.

Here, part of the logistics, such as storage buffers, is considered pertaining to manufacturers, suppliers and customers. Moreover, transporters connect the different stages of the production process.

The SC dynamics is traced by the flow of products between facilities (i.e., entities of types 1-5) and transporters (i.e., entities of type 6). Because of the large amount of material flowing in the system, we model a SC as a hybrid system: the continuous dynamics models the flow of products in the SC, the manufacturing and the assembling of different products and its storage in appropriate buffers. Hence, resources with limited capacities are represented by continuous states describing the amount of fluid material that the resource stores.

Moreover, we consider also discrete events occurring stochastically in the system, such as:

a. the blocking of the raw material supply, e.g. modeling the occurrence of labour strikes, accidents or stops due to the shifts;
b. the blocking of the transport operations due to the shifts or to unpredictable events such as jamming of transportation routes, accidents, strikes of transporters etc.;
c. the start of a request from the retailers.

2.2 An example of SC

We describe an example of SC whose target product is a desktop computer system (Dotoli et al., 2008), inspired from a case study reported in (Dotoli et al., 2006; Luo et al., 2001). Figure 1 depicts the SC network, comprising three suppliers, two manufacturers, one distributor, two retailers and one de-manufacturer. Moreover, twelve transporters connect the facilities. Each edge represents the flow of the material and is labelled by the parts/products that are transported between the connected facilities: the Personal Computer or PC, the central processing unit or C, the hard disk driver or H, the keyboard or K and the monitor or M. In particular, the last four types of products are semi-finished products obtained by the suppliers S1, S2 and S3, while the PC is produced by manufacturer M1 (M2) with a bill of materials of C, H, M and K obtained from suppliers S1 and S2 (S3). Moreover, retailers R1 and R2 obtain from distributor D1 the product PC. In addition, the de-manufacturer DM1 obtains the finished product PC from the retailers and supplies manufacturer M1 (M2) with the semi-finished product H (C). Note that the SC scheme includes two inter-twined productive chains with a remarkable advantage: if a transportation link is temporarily unavailable the productive cycle does not stop.

2.3 SC management and inventory control rules

The operational SC dynamics depends on the considered planning and management methodology, which specifies the business model and determines the paths for the information and material flow in the SC, and on the corresponding inventory control rules governing each SC facility (Viswanadham & Raghavan, 2000).

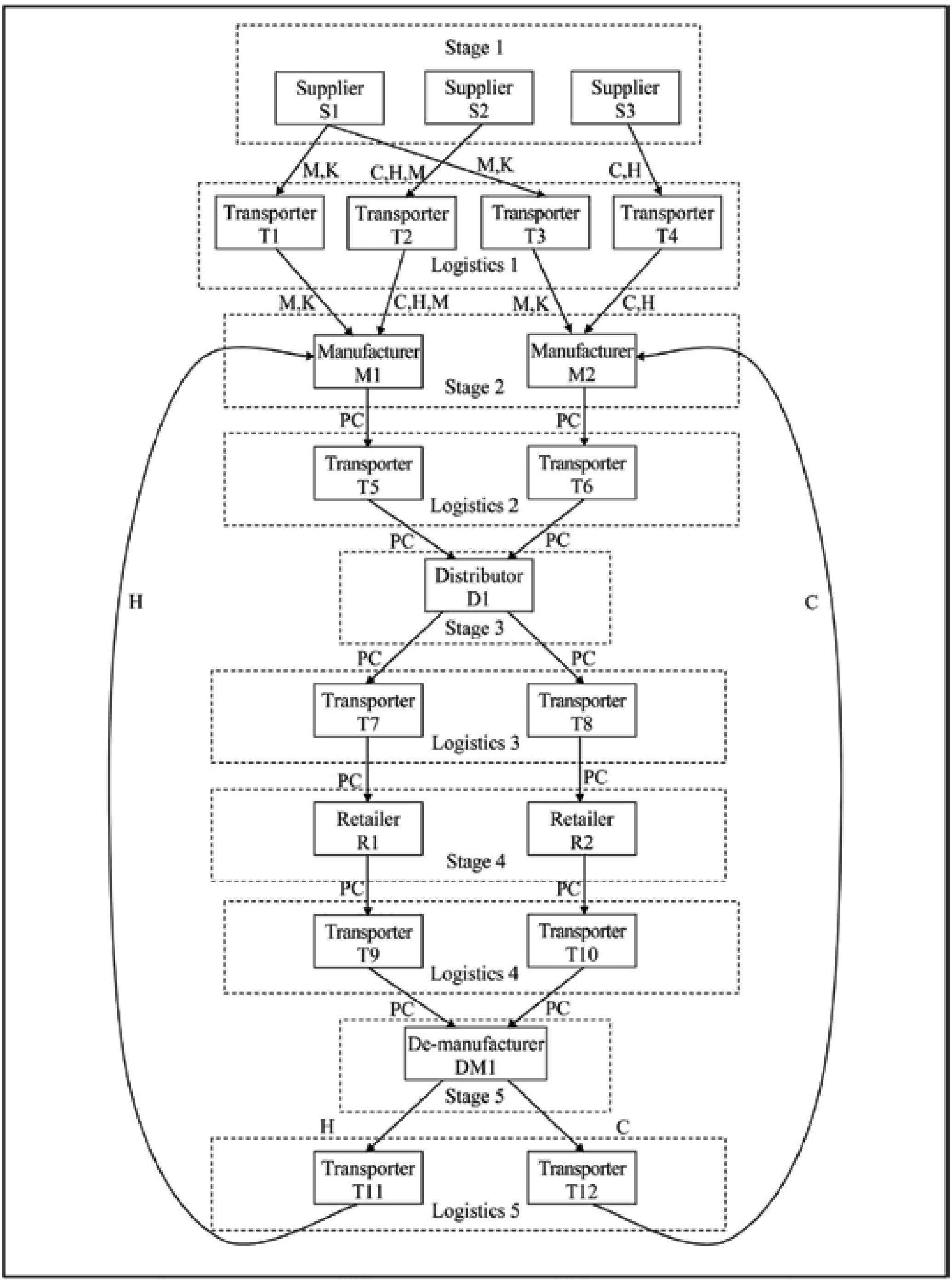


Fig. 1. The structure of the considered example SC.

According to the Wortmann classification (Wortmann, 1983), three SC managing policies are followed in practice: Make-To-Stock (MTS), Make-To-Order (MTO) and Assemble-To-Order (ATO) or Build-To-Order (BTO). In particular, in order to deliver on time the produced goods to end-users, the MTS strategy governs the system initiating production before the actual occurrence of demands, so that end customers are satisfied from stocks of inventory of finished goods. Often, under MTS management, stocks in the SC are governed by a reorder point based policy and inventory is replenished as soon as a preset reorder level is reached, so that the target level is maintained. On the other hand, in the MTO technique customer orders trigger the flow of materials and the requirements at each production stage of the SC. Hence, under such a management strategy, a lower level of material and product inventory is maintained. Furthermore, the ATO or BTO policy can be viewed as a hybrid of the former two strategies, basically applying MTS in the first stages of the SC and MTO in the last stages (Viswanadham & Raghavan, 2000). An additional production management choice is made between push and pull strategies (Hopp & Spearman, 2004): a pull production policy explicitly limits the amount of work-in-process that can be in the system, while in a push one no explicit limit on the amount of work-in-process is defined. Wellknown push systems are the Material Requirements Planning (MRP) approach and the Manufacturing Resources Planning technique (also known as MRP II) (Hopp & Spearman, 2004). Well-established pull strategies are the just-in-time technique based on Kanbans and the CONWIP (or CONstant Work-In-Process) strategy (Hopp & Spearman, 2004; Spearman et al., 1990), which is a generalized form of the Kanban policy: while the latter procedure establishes a fixed limit on work-in-process in a part of the system via the limited number of Kanban cards, the CONWIP strategy limits the total number of parts allowed in the whole system at any time, so that the SC is controlled at a constant level of work-in-process.

Together with the operational planning and management policy, inventory systems play a very important role in SC management. Inventory management addresses two fundamental issues: when a stock should replenish its inventory (order timing choice) and how much it should order from suppliers for each replenishment (order size choice) (Chen et al., 2005). These choices have to be adequately made in order to protect the SC from uncertainties, such as variations from their nominal values of demand quantity and mix, of production and transportation capacities, of quality and reliability of deliveries etc. (Rota et al., 2002). In particular, the numerous inventory management models proposed in the related literature may be mainly classified in four types, depending on order frequency and quantity. More precisely, order frequency may be either fixed, as in periodic review systems (T), or variable (R); similarly, order quantity may be either fixed, as in continuous review systems (Q), or variable (L). Accordingly, we may distinguish the following categories of inventory management rules (Vollmann et al., 2004):

i. (T,Q), in which orders are emitted with given frequency 1/T and ordered quantities are fixed and equal to Q, as in the well known economic order quantity model;
ii. (R,Q), where fixed quantities of parts that are Q in number are ordered any time the stock level drops below the reorder point R, as in the reorder point based rules;
iii. (T,L), where at every time step T a variable quantity of material is ordered so as to reach the preset desired level L, as in the reorder level rules;
iv. (R,L), where variable quantities are ordered to reach the preset level L each time the inventory level drops below the reorder point R.

Summing up, under the chosen SC management technique, a customer order for a product triggers a series of activities in the SC entities, and these have to be synchronized, so that the consumer demand and the selected inventory control rules are simultaneously satisfied.
This chapter focuses on SCs governed either by the MTS policy, which is typical of standardized products with high volumes (and reasonably accurate forecasts), or by the MTO strategy, which is characteristic of customized goods with low volumes. For the sake of brevity the ATO or BTO strategy is not considered in detail, although the presented SC model could be straightforwardly adapted to systems governed by such a strategy. Moreover, the (R,Q) inventory control rule is applied to manage the inventory of buffers that are governed as follows. Any time a withdrawal is made, a control system tracks the remaining inventory level of the buffer of products to determine whether it is time to reorder: in practice, thanks to automation and information systems, these reviews are continuous. At each review, the inventory level is compared with the pre-set reorder point R. In case the inventory level is higher than R, then no change in the inventory occurs. On the contrary, if the inventory level is lower than R, then a fixed quantity Q of products or lots of the considered items is ordered upstream, i.e., Q products or lots are manufactured if the considered stock level refers to an output product, or else they are ordered from an upstream facility in the SC.

3. First order hybrid Petri nets

In this section we briefly outline the basics of the FOHPN formalism (Balduzzi et al., 2000).

3.1 The FOHPN structure and marking

A FOHPN is a bipartite digraph described by the seven-tuple $PN=(P, T, \boldsymbol{Pre}, \boldsymbol{Post}, \Delta, F, RS)$. The set of places $P=P_d \cup P_c$ is partitioned into a set of discrete places P_d (represented by circles) and a set of continuous places (represented by double circles).
The set of transitions $T=T_d \cup T_c$ is partitioned into a set of discrete transitions T_d and a set of continuous transitions T_c (represented by double boxes). Moreover, the set of discrete transitions $T_d=T_I \cup T_S \cup T_D$ is further partitioned into a set of immediate transitions T_I (represented by bars), a set of stochastic transitions T_S (represented by boxes and including exponentially distributed transitions as well as transitions with triangular distribution) and a set of deterministic timed transitions T_D (represented by black boxes). We also denote $T_t=T_S \cup T_D$, indicating the set of timed transitions.
The matrices $\boldsymbol{Pre}$ and $\boldsymbol{Post}$ are the pre-incidence and the post-incidence matrices, respectively, of dimension $|P|\times|T|$. Note that the symbol $|A|$ denotes the cardinality of set A. Such matrices specify the net digraph arcs and are defined as follows:

$$\boldsymbol{Pre}, \boldsymbol{Post} : \begin{cases} P_c \times T \rightarrow \mathbb{R}^+ \\ P_d \times T \rightarrow \mathbb{N} \end{cases}.$$

We require that for all $t \in T_c$ and for all $p \in P_d$ it holds $\boldsymbol{Pre}(p,t)=\boldsymbol{Post}(p,t)$ (*well-formed nets*).
The function Δ: $T_t \rightarrow \mathbb{R}^+$ specifies the timing associated to timed transition. In particular, we associate to each $t_j \in T_S$ the average firing delay $\Delta(t_j)=\delta_j=1/\lambda_j$, where λ_j is the average firing rate of the transition. More precisely, in case the transition is exponential δ_j represents the expected value of the associated distribution, while in case it is triangular δ_j represents the

modal value of such a distribution and we assume that the minimum and maximum values of the range in which the firing delay varies equal respectively $d_{\delta j}$=0.8 δ_j and $D_{\delta j}$=1.2 δ_j. In addition, each $t_j \in T_D$ is associated the constant firing delay $\Delta(t_j)= \delta_j$. Moreover, the function F: $T_c \rightarrow \mathbb{R}^+ \times \mathbb{R}^+_\infty$ specifies the firing speeds associated to continuous transitions (we denote $\mathbb{R}^+_\infty = \mathbb{R}^+ \cup \{+\infty\}$). For any continuous transition $t_j \in T_c$ we let $F(t_j)=(V_{mj},V_{Mj})$, with $V_{mj} \leq V_{Mj}$, where V_{mj} represents the minimum firing speed and V_{Mj} the maximum firing speed of the generic continuous transition. Finally, the function RS: $T_d \rightarrow \mathbb{R}^+$ associates a probability value called random switch to conflicting discrete transitions.

Given a FOHPN and a transition $t \in T$, the following place sets may be defined: $\bullet t=\{p \in P: Pre(p,t)>0\}$ (pre-set of t); $t\bullet=\{p \in P: Post(p,t)>0\}$ (post-set of t). Moreover, the corresponding restrictions to continuous or discrete places are defined as ${}^{(d)}t=\bullet t \cap P_d$ or ${}^{(c)}t=\bullet t \cap P_c$. Similar notations may be used for pre-sets and post-sets of places. The incidence matrix of the net is defined as $\boldsymbol{C}=\boldsymbol{Post}-\boldsymbol{Pre}$. The restriction of $\boldsymbol{C}$ to P_X and T_Y (with X,Y$\in\{c, d\}$) is denoted by $\boldsymbol{C}_{XY}$.

A marking $\mathbf{m}: \begin{cases} P_d \rightarrow \mathbb{N} \\ P_c \rightarrow \mathbb{R}^+ \end{cases}$ is a function that assigns to each discrete place a non-negative number of tokens, represented by black dots, and to each continuous place a fluid volume; m_i denotes the marking of place p_i. The value of a marking at time τ is denoted by $\mathbf{m}(\tau)$. The restrictions of $\mathbf{m}$ to P_d and to P_c are denoted by $\mathbf{m}_d$ and $\mathbf{m}_c$, respectively. A FOHPN system $\langle PN,\mathbf{m}(\tau_0) \rangle$ is a FOHPN with initial marking $\mathbf{m}(\tau_0)$.

The following statements rule the firing of continuous and discrete transitions:

1. a discrete transition $t \in T_d$ is enabled at $\mathbf{m}$ if for all $p_i \in \bullet t$, $m_i > \boldsymbol{Pre}(p_i,t)$;
2. a continuous transition $t \in T_c$ is enabled at $\mathbf{m}$ if for all $p_i \in {}^{(d)}t$, $m_i > \boldsymbol{Pre}(p_i,t)$.

Moreover, we say that an enabled transition $t \in T_c$ is strongly enabled at $\mathbf{m}$ if for all places $p_i \in {}^{(c)}t$, $m_i>0$; we say that transition $t \in T_c$ is weakly enabled at $\mathbf{m}$ if for some $p_i \in {}^{(c)}t$, $m_i=0$.

In addition, for any continuous transition $t_j \in T_c$ its IFS is indicated by v_j and it holds:

1. if t_j is not enabled then $v_i=0$;
2. if t_i is strongly enabled, then it may fire with any firing speed $v_j \in [V_{mj},V_{Mj}]$;
3. if t_j is weakly enabled, then it may fire with any firing speed $v_j \in [V_{mj},V_j]$, where $V_j \leq V_{Mj}$ depends on the amount of fluid entering the empty input continuous place of t_i.

We denote by $\boldsymbol{v}(\tau)=[v_1(\tau)\ v_2(\tau) \ldots\ v_{|Tc|}(\tau)]^T$ the IFS vector at time τ. Hence, any admissible IFS vector $\boldsymbol{v}$ at $\mathbf{m}$ is a feasible solution of the following set of linear constraints:

$$\begin{array}{ll} V_{Mj} - v_j \geq 0 & \forall t_j \in T_\varepsilon(\mathbf{m}) \\ v_j - V_{mj} \geq 0 & \forall t_j \in T_\varepsilon(\mathbf{m}) \\ v_j = 0 & \forall t_j \in T_\upsilon(\mathbf{m}) \\ \sum_{t_j \in T_\varepsilon(\mathbf{m})} C(p,t_j) v_j \geq 0 & \forall p \in P_\varepsilon(\mathbf{m}), \end{array} \tag{1}$$

where $T_\varepsilon(\mathbf{m}) \subset T_c$ ($T_\upsilon(\mathbf{m}) \subset T_c$) is the subset of continuous transitions that are enabled (not enabled) at $\mathbf{m}$ and $P_\varepsilon(\mathbf{m}) = \{p_i \in P_c \mid m_i=0\}$ is the subset of empty continuous places. In particular, the first three constraints in (1) follow from the firing rules of continuous transitions, while the last constraint in (1) imposes that if a continuous place is empty then its fluid content does not become negative.

The set of all feasible solutions of (1) is denoted as $S(PN,\mathbf{m})$.

3.2 The FOHPN dynamics

The dynamics of the hybrid net combines both time-driven and event-driven dynamics. We define *macro-events* the events that occur when:

i. a discrete transition fires or the enabling/disabling of a continuous transition takes place;
ii. a continuous place becomes empty.

The equation that governs the time-driven evolution of the marking of a place $p_i \in P_c$ is:

$$\dot{m}_i(\tau) = \sum_{t_j \in T_c} C(p_i, t_j) v_j(\tau) \tag{2}$$

Now, if τ_k and τ_{k+1} are the occurrence times of two subsequent macro-events, we assume that within the time interval $[\tau_k,\tau_{k+1})$ (macro-period) the IFS vector $v(\tau_k)$ is constant. Then the continuous behavior of an FOHPN for $\tau \in [\tau_k,\tau_{k+1})$ is described by:

$$\begin{aligned} \mathbf{m}^{\mathbf{c}}(\tau) &= \mathbf{m}^{\mathbf{c}}(\tau_k) + \boldsymbol{C}_{\mathbf{cc}} v(\tau_k)(\tau - \tau_k) \\ \mathbf{m}^{\mathbf{d}}(\tau) &= \mathbf{m}^{\mathbf{d}}(\tau_k) \end{aligned} \tag{3}$$

The evolution of the net at the firing of a discrete transition $t_j \in T_d$ at $\mathbf{m}(\tau_k\text{-})$ yields the following marking:

$$\begin{aligned} \mathbf{m}^{\mathbf{c}}(\tau_k) &= \mathbf{m}^{\mathbf{c}}(\tau_k^-) + \boldsymbol{C}_{\mathbf{cd}} \boldsymbol{\sigma}(\tau_k) \\ \mathbf{m}^{\mathbf{d}}(\tau_k) &= \mathbf{m}^{\mathbf{d}}(\tau_k^-) + \boldsymbol{C}_{\mathbf{dd}} \boldsymbol{\sigma}(\tau_k), \end{aligned} \tag{4}$$

where $\sigma(\tau_k)$ is the firing count vector associated to the firing of transition t_j at time τ_k.
Moreover, we associate to each timed transition $t_j \in T_t$ a timer v_j and we call $\mathbf{v}(\tau_k)$ the vector of timers associated to timed transitions at time τ_k. Hence, the timer evolution within the macro-period $[\tau_k,\tau_{k+1})$ for each transition $t_j \in T_t$ is as follows:

$$\mathrm{v}_j(\tau_{k+1}) = \begin{cases} \mathrm{v}_j(\tau_k) = 0 & \text{if } t_j \text{ is not enabled,} \\ \mathrm{v}_j(\tau_k) + (\tau - \tau_k) & \text{if } t_j \text{ is enabled,} \end{cases} \quad \text{for } j=1,\dots,|T_t|\,. \tag{5}$$

Whenever t_j is disabled or it fires, its timer is reset to zero.
Equations (3)-(4)-(5) describe the dynamics of the FOHPN model. The overall state of the system at time τ_k is given by the marking of all places and by the values of all timers and is hence indicated by $\mathbf{x}(\tau_k) = \begin{bmatrix} \mathbf{m}^{\mathbf{c}}(\tau_k) \\ \mathbf{m}^{\mathbf{d}}(\tau_k) \\ \mathbf{v}(\tau_k) \end{bmatrix}$.

Moreover, the system input is vector $\mathbf{u}(\tau_k) = \begin{bmatrix} \tau_{k+1} - \tau_k \\ \boldsymbol{\sigma}(\tau_{k+1}) \end{bmatrix}$, collecting the length of the current macro-period and the transition (if any) that will fire at the end of such macro-period. Note that this input vector depends on the current state vector $\mathbf{x}(\tau_k)$ and on the next macro-event occurring at the end of the current macro-period. Consequently, a FOHPN system (3)-(4)-(5) can be described by a linear discrete-time time-varying state variable model, so that an efficient simulation algorithm determining the state vector at the beginning of each macroperiod, given the initial state $\mathbf{x}(\tau_k)$, may be straightforwardly derived.

3.3 The FOHPN control

Once the set of all admissible IFS vectors has been defined, a procedure is required to select one among them and thus let equation (3) of the net dynamics be univocally determined. In other words, each IFS vector $v \in S(PN,\mathbf{m})$ solving (1) represents a particular mode of operation of the system described by the FOHPN. Consequently, the designer may choose the best operative condition according to a given objective and solving the corresponding optimization problem with the constraint set (1). In this chapter the following two cases are considered.

1) Maximize flows. We may consider as optimal the solution v^* of (1) that maximizes the performance index $J=\mathbf{1}^T \cdot v$, which is intended to maximize the sum of all the flow rates. In the manufacturing domain, this corresponds to maximizing resource utilization.

2) Minimize stored fluid. We may choose as optimal the vector v^* solving (1) that minimizes the derivative of the marking of each place $p \in P_c$. This can be done by minimizing the performance index $J=\mathbf{c}^T \cdot v$ where $c_j=\mathbf{C}(p,t_j)$ if $t_j \in p^{(c)} \cup {}^{(c)}p$ and 0 otherwise. In the manufacturing domain, this corresponds to minimizing the work-in-process.

3.4 An example of FOHPN

In this section we describe an example of FOHPN in order to clarify its dynamics.

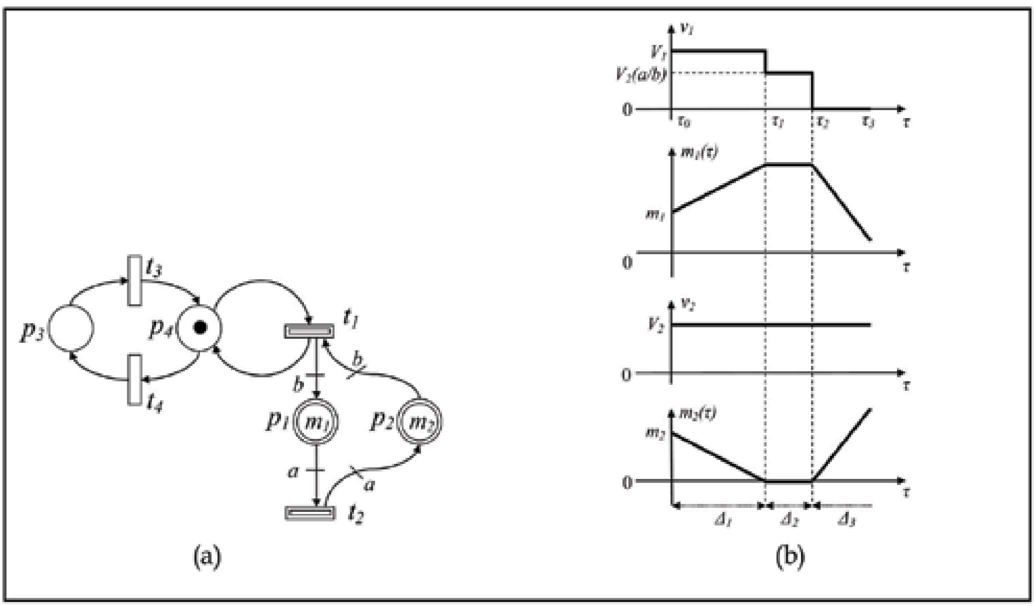


Fig. 2. An example of FOHPN (a) and its evolution (b).

Consider the net in Fig. 2a. Places p_1 and p_2 are continuous places and places p_3 and p_4 are discrete places. Transitions t_1 and t_2 are continuous transitions with firing speeds $v_1 \in [0, V_1]$ and $v_2 \in [0, V_2]$, respectively. We assume $V_1 b > V_2 a$ (here a and b are the arc weights in Fig. 2a). In addition, the discrete transitions t_3 and t_4 are exponentially distributed timed transitions with average firing rates λ_3 and λ_4, respectively.

The net dynamics, depicted in Fig. 2b, is described as follows. Since place p_4 is initially marked, transition t_1 is enabled. Moreover, the initial markings of the continuous places are $m_1(\tau_0)>0$ and $m_2(\tau_0)>0$ so that transitions t_1 and t_2 are both strongly enabled and may fire according to the set of constraints (1):

$$\begin{cases} V_1 - v_1 \geq 0 \\ V_2 - v_2 \geq 0 \\ v_1, v_2 \geq 0. \end{cases} \tag{6}$$

We assume $v_1=V_1$ and $v_2=V_2$. By (3), the continuous marking of the net during this first macro-period Δ_1 is $\boldsymbol{m}^c(\tau)=\begin{cases} m_1(\tau)=m_1(\tau_0)-(V_2a-V_1b)(\tau-\tau_0) \\ m_2(\tau)=m_2(\tau_0)-(V_1b-V_2a)(\tau-\tau_0) \end{cases}$ for $\tau>\tau_0$ until the subsequent macro-event. Moreover, by (5) the timer vector is $\boldsymbol{\nu}(\tau)=[0\ \tau-\tau_0]^T$ for $\tau>\tau_0$, since t_3 is disabled and t_4 is enabled. Figure 2b shows the corresponding marking evolution and the IFSs of the net continuous transitions. In particular, we remark that the marking m_1 increases while m_2 decreases since it holds V_1b>V_2a.

At time τ_1 a macro-event occurs because place p_2 becomes empty. Consequently, t_1 becomes weakly enabled and the set of constraints (1) has to be re-written as follows:

$$\begin{cases} V_1 - v_1 \geq 0 \\ V_2 - v_2 \geq 0 \\ v_1, v_2 \geq 0 \\ v_2 a - v_1 b \geq 0. \end{cases} \tag{7}$$

Since t_2 remains strongly enabled, its firing speed is assumed $v_2=V_2$. On the other hand, we choose the firing speed of t_1 as $v_1=V_2$(a/b). Therefore, during the macro-period Δ_2, by (3) the continuous marking is expressed by $\boldsymbol{m}^c(\tau)=\begin{cases} m_1(\tau)=m_1(\tau_1) \\ m_2(\tau)=0 \end{cases}$ for $\tau>\tau_1$ until the subsequent macro-event (see Fig. 2b). Moreover, by (5) it holds $\boldsymbol{\nu}(\tau)=[0\ \tau-\tau_1]^T$ for $\tau>\tau_1$.

Next, suppose that at time τ_2 transition t_4 fires and the macro-event updates the discrete markings to $m_3(\tau_2)=1$ and $m_4(\tau_2)=0$. Hence, t_1 is disabled, i.e., $v_1=0$, while t_2 remains strongly enabled and we assume $v_2=V_2$. Then, during the macro-period Δ_3 the marking is given, as in (5), by $\boldsymbol{m}^c(\tau)=\begin{cases} m_1(\tau)=m_1(\tau_2)-V_2a(\tau-\tau_2) \\ m_2(\tau)=m_2(\tau_2)+V_2a(\tau-\tau_2) \end{cases}$ (see Fig. 2b). Moreover, by (5) it holds $\nu(\tau)=[\tau-\tau_2\ 0]^T$ for $\tau>\tau_2$.

4. The SC model

Based on the idea of the bottom-up approach (Zhou & Venkatesh, 1998), this section reviews a modular FOHPN model to describe a SC previously proposed by the authors (Dotoli et al., 2008). Such a method can be summarized in two steps: decomposition and composition. Decomposition consists in partitioning a system into several subsystems. In SCs this subdivision can be performed based on the determination of distributed system entities (i.e., suppliers, manufacturers, distributors, customers and transporters). All these subsystems are modelled by FOHPNs modules. On the other hand, composition involves the interconnections of these sub-models into a complete model, representing the whole SC.

In particular, manufacturers are described by continuous transitions, buffers are continuous places and products are represented by continuous flows (fluids) routing from manufacturers, buffers and transporters. Moreover, transporters are described by discrete stochastic transitions with a triangular distribution and the customers demand is modelled by exponential transitions. In addition, exponential transitions model discrete events occurring stochastically in the system, such as the blocking of a raw material supply or of a transport operation due to unpredictable events. Hence, the state of the SC model at the beginning of each macro-period is a vector $\mathbf{x}(\tau_k)$ that includes the following sub-vectors:

- the sub-vector $\mathbf{m}^c(\tau_k)$, collecting the markings of the continuous places, i.e., the buffer places and the associated capacity places (absent for infinite capacity buffers);
- the sub-vector $\mathbf{m}^d(\tau_k)$, collecting the markings of the discrete places, i.e., the places modeling choices, constraints and the operative states of entities;
- the timers vector $\mathbf{v}(\tau_k)$, collecting the values of the timers of discrete timed transitions, i.e.,
- the transitions associated to customer demands or transporters and the transitions modeling
- the blockings of supplies or transports due to unpredictable or external events.

The following FOHPN modules model the individual subsystems composing the SC (Dotoli et al., 2008).

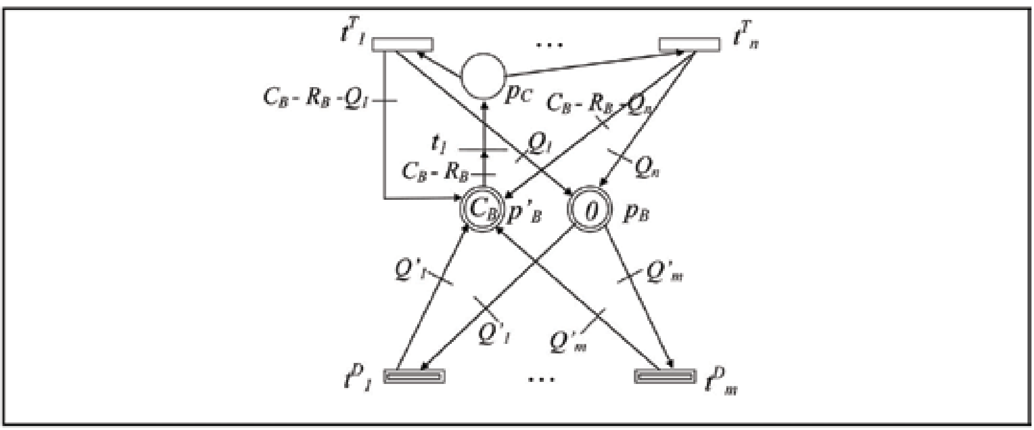


Fig. 3. The FOHPN modeling the input buffers.

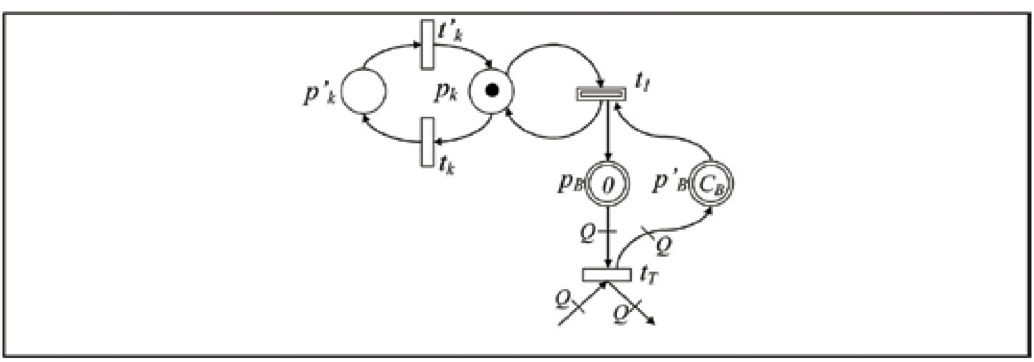


Fig. 4. The FOHPN modeling the suppliers.

4.1 The inventory management model of the input buffers

In this section we describe the model of the input buffers of manufacturers and distributors managed by the (R,Q) policy. On the other hand, the output buffers are not managed by the (R,Q) policy since they are devoted just to providing the requested material. The basic quantities of the (R,Q) inventory management strategy are: the *fixed order quantity* Q; the *lead time*, i.e., the time between placing an order and receiving the goods in stock; the *demand* D, i.e., the number of units to be supplied from stock in a given time period; the *reorder level* R, i.e., the new orders take place whenever the stock level falls to R.

Figure 3 shows the FOHPN model for the input buffers managed by the (R,Q) policy (Furcas et al., 2001). The continuous place p_B denotes the input buffer of finite capacity C_B. The complementary place p'_B models the available buffer space so that at each time instant it holds $m_B+m'_B=C_B$. Here and in the following models the assumed initial marking corresponds to empty buffers. We assume that the buffer can receive demands from different facilities and can require the goods from different transporters. Hence, each demand is modelled by a continuous transition t^D_i with $i=1,\dots,m$ so that the demand to be fulfilled is $D_i=v^D_iQ'_i$. When $m_B>0$ a transition t^D_i with $i\in\{1,\dots,m\}$ may fire at the firing speed v_{Di}, reducing the marking of the place p_B with a constant slope $v^D_iQ'_i$. As soon as m_B falls below the level R_B (or, equivalently, the marking m'_B goes over C_B-R_B), the immediate transition t_1 is enabled. When t_1 fires, the choice place $p_C\in P_d$ becomes marked and performs the choice of the input facility. Hence, new materials/products are requested by enabling one of the transitions t^T_i according to the value of the random switches RS(t^T_i) with $i=1,\dots,n$. If a particular transition t_{Ti} with $i\in\{1,\dots,n\}$ is selected and fires after the lead time of average$\Delta(t^T_i)=\ \delta_i=1/\ \lambda_i$, Q_i products are received in the buffer and $C_B-R_B-Q_i$ units are restored in the buffer capacity. Typically, transitions t_{Ti} can represent a transport operation.

4.2 The inventory management model of the SC entities

The supplier model. Suppliers are modelled as a continuous transition and two continuous places (see Fig. 4). The continuous place p_B represents the raw material output buffer of finite capacity C_B and the complementary place p'_B represents the available buffer space. Moreover, the continuous transition t_1 models the arrival of raw material into the system. In addition, we consider the possibility that the providing of raw material is blocked. This situation is represented by a discrete event modelled by two exponentially distributed transitions and two discrete places. In particular, place $p_k\in P_d$ models the operative state of the supplier and $p'_k\in P_d$ is the non-operative state. The blocking and the restoration of the raw material supply correspond to the firing of exponential transitions t_k and t'_k, respectively. For the sake of clarity, Fig. 4 depicts the transition $t_T\in T_S$ that, as discussed later, models the transport operation. Here and in the following models the initial marking assumes that the entity is operative.

The manufacturer and assembler model. Manufacturers and assemblers are modelled by the FOHPN shown by Fig. 5. More precisely, the continuous places p_{Bi} and p'_{Bi} with $i=2,\dots n$ describe the input buffers and the corresponding available capacities, respectively. Each buffer stores the input goods of a particular type. Analogously, the continuous places p_{B1} and p'_{B1} model the output buffer and its capacity, respectively. The production rate of the facility is modelled by the continuous transition t_1 with the assigned firing speed $v_1\in[V_{m1},V_1]$.

The transporter model. The transporters connecting the different facilities are modelled by a set of timed transitions t^T_i with triangular distributions for $i=1,\dots,n$ (see Fig. 6), according to

(Kelton et al., 2004; Law & Kelton, 2000). Each transition describes the transport of items of a particular type from an upstream facility to a downstream one in an average time interval $\Delta(t^T_i)=\delta_i=1/\lambda_i$. The control places $p_{Ci}\in P_d$ with i=1,…,n determine the choice of only one type of material to transport among the available set by the firing of the corresponding immediate transition t_i with i=1,…,n modeling the replenishment request to the transporter by a downstream SC entity. In addition, place $p_1\in P_d$ disables the remaining transitions. Moreover, the random stop of the material transport is represented by two places p_k, $p'_k\in P_d$ and two exponentially distributed transitions $t_k, t'_k\in T_S$. The transporter capacity is Q and places $p_{Bi}\in P_b$ and $p'_{Bi}\in P'_b$ with i=1,…n in Fig. 4 describe the n input buffers of the downstream facility (e.g., a manufacturer, a distributor, a retailer) and the corresponding available capacities, respectively. The initial marking shown assumes that no material has yet been selected for transportation.

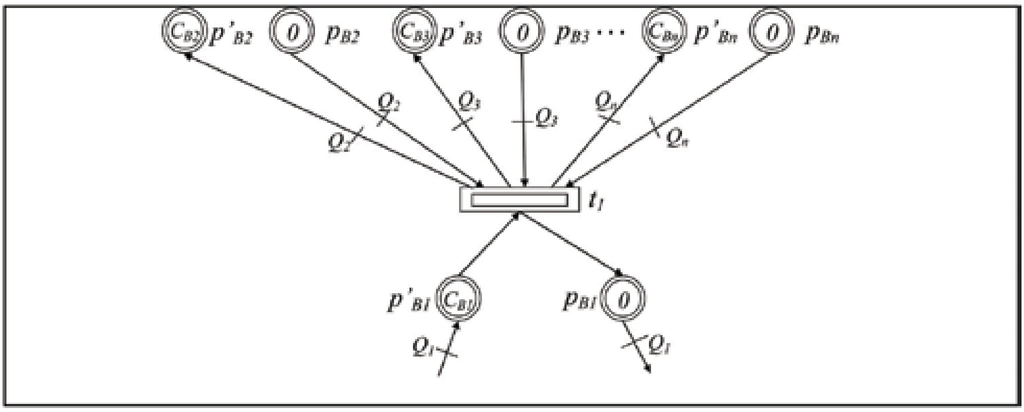


Fig. 5. The FOHPN modeling manufacturers and assemblers.

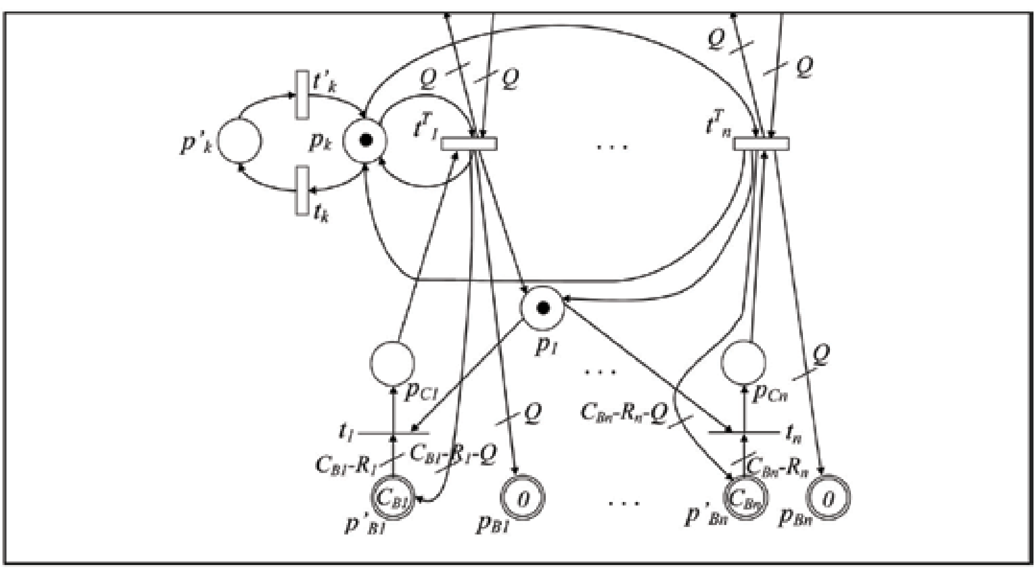


Fig. 6. The FOHPN modeling the transporters.

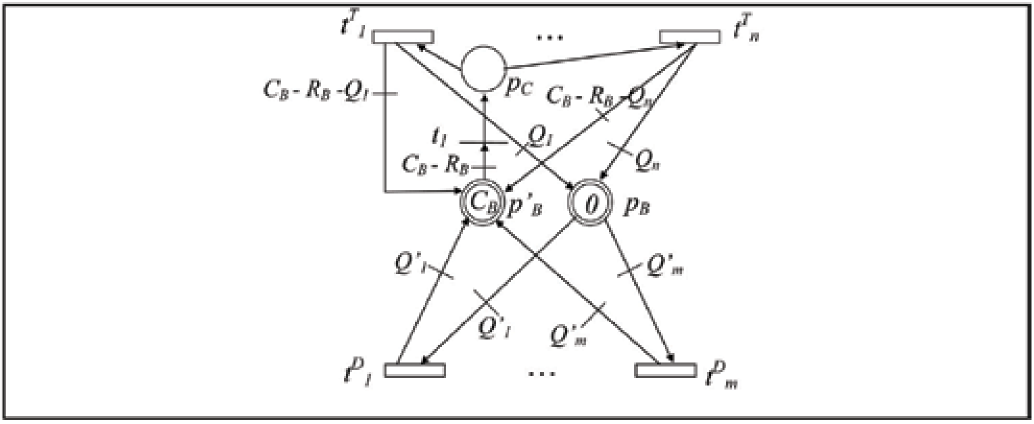


Fig. 7. The FOHPN modeling the distributors.

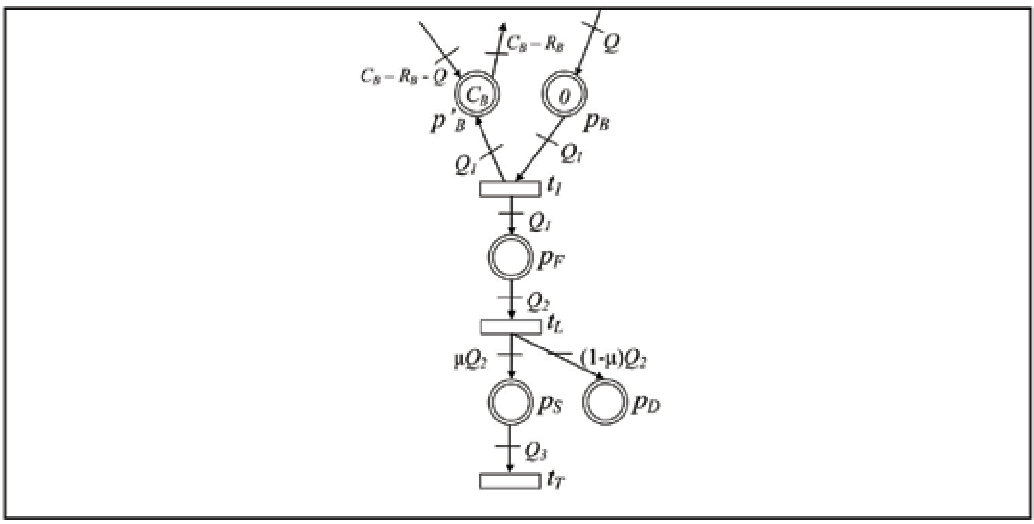


Fig. 8. The FOHPN modeling the retailers under the MTS strategy.

The distributor model. The model of the distributors is represented by an input buffer managed by the (R,Q) inventory control rule. Hence, the model is similar to the FOHPN represented in Fig. 3, where each downstream continuous transition t^D_i with i=1,…,m is substituted by a stochastic timed transition representing a transport operation (see Fig. 7).

The retailer model. The retailer model is different in the two cases of the MTS and MTO management. If the MTS strategy is used, the retailer is represented by an input buffer managed by the (R,Q) policy with a finite lead time and stochastic demand (see Fig. 8). Hence, the model is similar to the FOHPN represented in Fig. 3 where all the downstream continuous transitions are substituted by one or more exponential transition (such as t_1) modeling the stochastic demand of the consumers. Moreover, the continuous place p_F

collects all the products obtained by the retailer, i.e., it corresponds to an infinite capacity buffer. In addition, the discrete transition with triangular distribution t_L models the deterioration of the finished products used by the customer that are stored in the infinite capacity buffers p_S and p_D. In particular, p_S collects the μQ_2 products to be de-manufactured with $\mu \in [0,1]$, and p_D the $(1-\mu)Q_2$ goods to be discarded. In addition, transition t_T represents the transport operation transferring products to the de-manufacturer.

Similarly, under the MTO management the retailer (see Fig. 9) is described by an output buffer place $p_B \in P_c$ and a stochastic transition t_1 modeling the rate with which the consumer withdraws products, that in this case have been previously ordered. The infinite capacity place $p_F \in P_c$ collects all the products obtained by the retailer. In addition, transition t_L models the products deterioration, place p_S denotes the system output and transition t_T represents the transport operation. The customer demand, just as under the MTS policy, is represented by one (or more) exponential transition, such as t_2. On the other hand, differently than the MTS case, when fired transition t_2 stores an order of Q products in the infinite capacity buffer place p_O. Hence, its marking m_o enables transition t_3 modeling the production of the upstream manufacturer. When the Q orders are satisfied, marking m_o becomes zero, t_3 is weakly enabled and its firing speed is equal to zero until the upstream transition t_2 fires again.

The de-manufacturer model. De-manufacturers are modelled by the FOHPN shown by Fig. 10, which is the reverse of the manufacturers model reported in Fig. 5. More precisely, p_{B1} and p'_{B1} model the input buffer and its capacity, respectively, and the continuous places p_{Bi} and p'_{Bi} with i=2,…n describe the output buffers and the corresponding available capacities, respectively. The continuous transition t_1 models the production (or, more properly, the disassembly) rate of the facility.

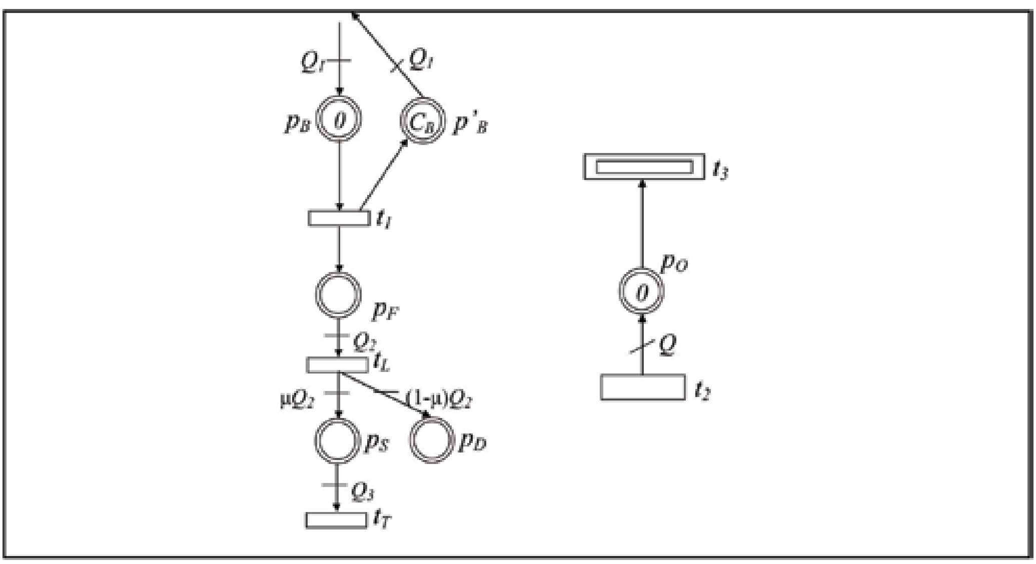


Fig. 9. The FOHPN modeling the retailers under the MTO strategy.

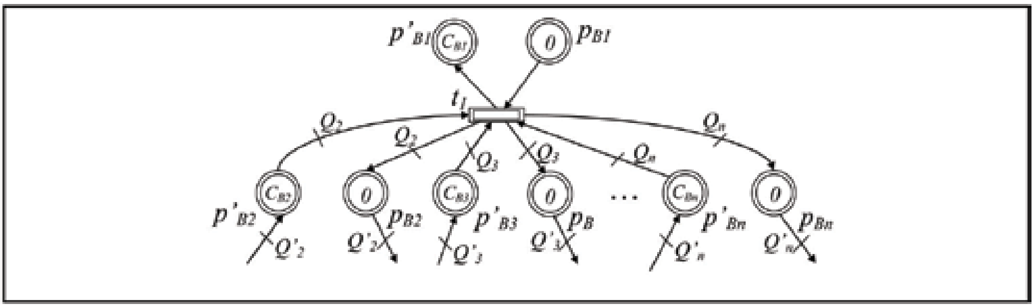


Fig. 10. The FOHPN modeling the de-manufacturers.

5. A case study

We consider the SC of Fig. 1 described in Section 2.B. To implement and compare the system managed under the MTS and MTO strategies, we model the whole SC properly merging the elementary modules described in the previous section. Figures 11 and 12 show the FOHPN modeling the SC under MTS and MTO, respectively. The dashed rectangles depict in the two figures the correspondence between each module and the entities of Fig. 1. We remark that for the sake of simplicity in this chapter we consider only single product SCs, since our ultimate aim is to show the effectiveness of a FOHPN formalism for operational management of SCs. However, common multi-product SCs may straightforwardly be considered within the proposed formalism thanks to its simplicity and modularity.

5.1 The SC under the MTS strategy

Figure 11 shows the SC model of the case study managed by the MTS policy. The production is determined by the firing of the continuous transitions $t_1, \ldots, t_7$ (modules S1, S2 and S3) that describe the input of the raw materials that can be interrupted by stochastic events only. Consequently, under this control technique, each input buffer is managed by the (R,Q) strategy.

Moreover, if the input buffer of manufacturer M1 (M2) requires a particular product, a request has to be sent to the corresponding transporter. Hence, places p_{60}, p_{63}, p_{66} and p_{67} (modules T1 and T2) (p_{72}, p_{74}, p_{75} and p_{77} (modules T3 and T4)) are introduced to select a particular transporter. For example, if place p_{60} (module T1) is marked then the transport modelled by t_{43} (module T1) is enabled. In addition, transitions t_{56} and t_{58} and place p_{63} (modules T1 and T2) are introduced since the buffer of M1 storing the semi-finished products monitors (denoted by p_{17} and p'_{17}) can require material from S1 by T1 or from S2 by T2. Consequently, place p_{63} with transitions t_{56} and t_{58} model the choice.

Finally, according to the SC scheme of Fig. 1, in the model of Fig 11 the supply of some semifinished products at the manufacturers (i.e., H at place p_{21} in M1 and C at place p_{27} in M2) may be obtained via two different paths, either by a supplier or by the de-manufacturer. The corresponding choice of the replenishment transition to enable (i.e., t_{60} of T2 or t_{82} of T11 for M1 and t_{63} of T4 or t_{83} of T12 for M2) is modelled via a random switch that associates a 100% probability to the less costly supply obtained by the de-manufacturer (i.e., to the enabling of t_{82} and t_{83}): the corresponding transition, however, is enabled via the respective arc weights Q_{13} and Q_{14} only when the corresponding semi-finished product output buffer of the demanufacturer (i.e., place p_{96} or p_{97} of module DM1) contains sufficient material.

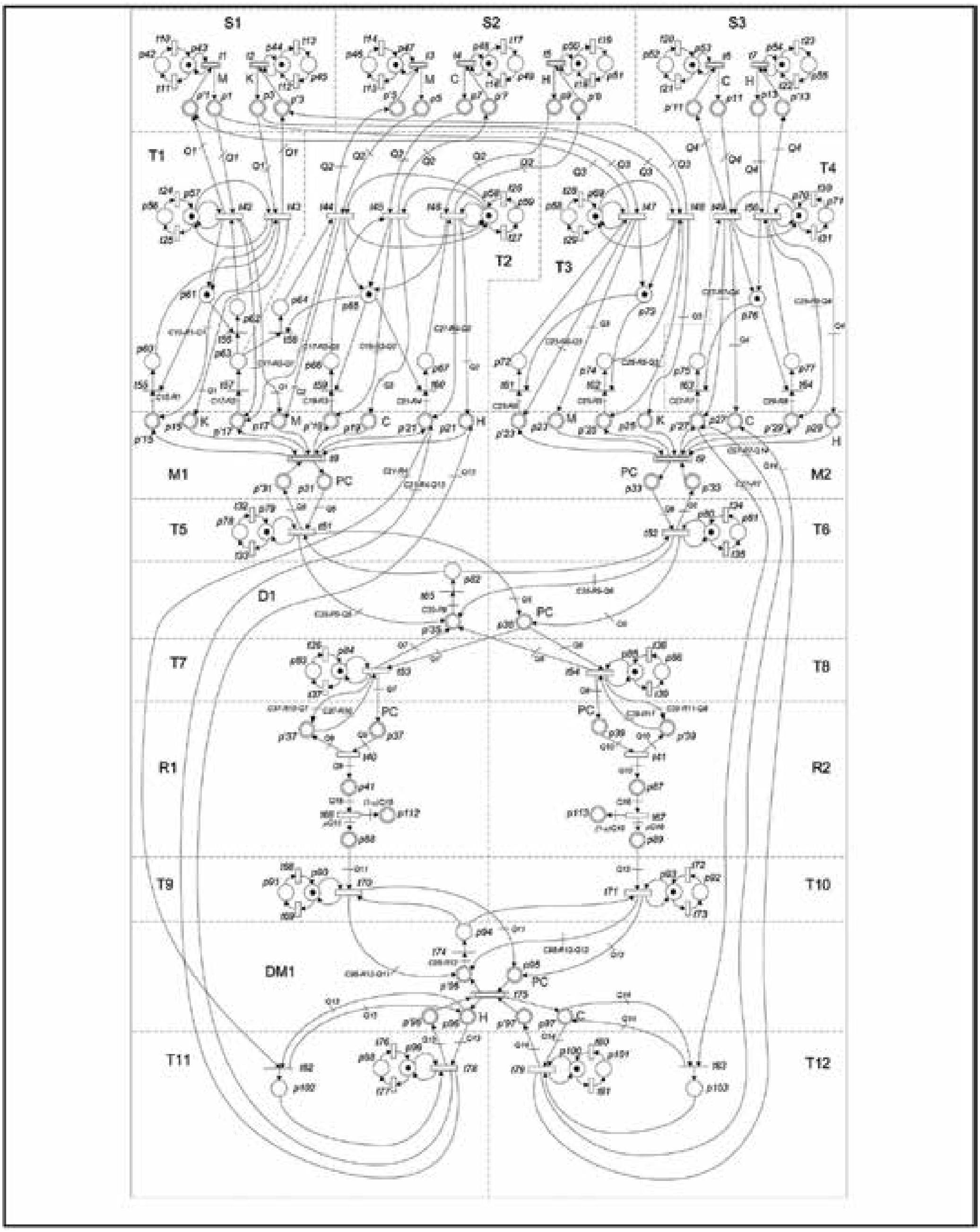


Fig. 11. The FOHPN modeling the case study under the MTS policy.

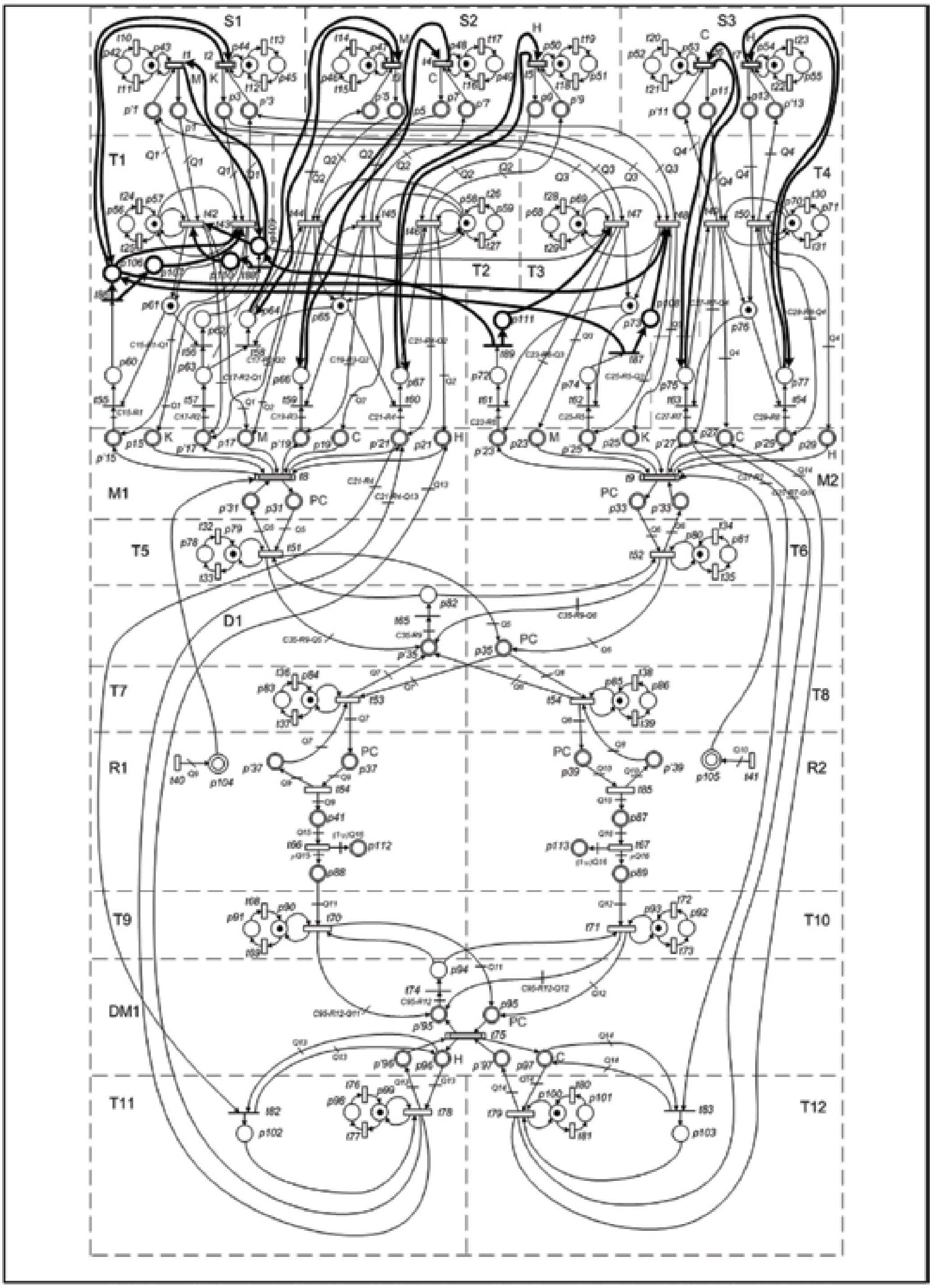


Fig. 12. The FOHPN modeling the case study under the MTO policy.

5.2 The SC under the MTO strategy

Figure 12 shows the SC model of the case study managed by the MTO policy. The model is similar to the model shown in Fig. 11 with two exceptions. First, some additional edges, places and transitions are introduced and drawn in bold. Second, the retailer modules are different than those employed under the MTS strategy, as detailed in Section 5.B. Hence, the actual assembling of a finished product is triggered by the demand of a consumer that in the MTO policy is modelled by the discrete exponential transitions t_{40} and t_{41} (modules R1 and R2, respectively) and by the places p_{104} and p_{105} (modules R1 and R2, respectively). For example, if a request is present for R1 (i.e., $m_{104}>0$), then t_8 (module M1) is enabled and fires. Consequently, the markings of p_{15}, p_{17}, p_{19} and p_{21} (module M1) decrease. In such a condition, if there is material stored in each input buffers, no raw material is requested to the suppliers. Indeed, the markings $m_{106}=m_{109}=m_{64}=m_{66}=m_{67}=0$ (modules T1 and T2) disable transitions t_1, t_2, …, t_5 (modules S1 and S2). On the contrary, if for example the input buffer p_{19} (module M1) is at a low level (i.e., it holds $m_{19}\leq R_3$ and $m'_{19}\geq C_{19}-R_3$), then transition t_{59} (module T2) is enabled. If such an immediate transition fires, it holds $m_{66}=1$ and the continuous transition t_4 (module S2) fires to provide the raw material C in the buffer p_7 (module S2). When it holds $m_7\geq Q_2$, transition t_{45} (module T2) modeling the transport is enabled and fires after $\Delta(t_{45})$ time units on average. After transition t_{45} fires, it holds $m_{66}=0$ so that t_4 is disabled.

We remark that the replenishment requests of a manufacturer input buffer might influence the replenishment of other manufacturer input buffers. Hence, in order to avoid such a condition, we suitably introduce places p_{106}, p_{107}, p_{108}, p_{109}, p_{110} and p_{111} and transitions t_{86}, t_{87}, t_{88} and t_{89} (modules T1, T2 and T3, drawn in bold). As an example, a token in p_{106} (module T1) may be determined by the firing of transition t_{86} (module T1) or t_{87} (module T3). In both cases, the marking $m_{106}=1$ enables transition t_2 (module S1). The subsequent enabling of the transport transition is managed by the added places p_{107} (module T1) and p_{108} (module T3) that respectively enable the transports from supplier S1 to manufacturers M1 and M2 modelled by t_{43} (module T1) and t_{48} (module T3).

Finally, note that the supply of the semi-finished products at the manufacturers from suppliers or de-manufacturers is governed in the same way as in the previously detailed SC model managed by MTS.

5.3 The simulation specification

The SC dynamics under the MTS and MTO management strategies is analyzed via numerical simulation using the data reported in Table 1. This table shows the manufacturer production rates and the average firing delays of discrete stochastic transitions. In addition, Table 2 shows further data necessary to fully describe and simulate the system, namely the buffer capacities for the inventories of each stage and the initial markings of odd continuous places (those of the other continuous places are complementary with respect to the capacities reported in the table and hence omitted). Furthermore, the initial markings of discrete places and the values of the edge weights are depicted in Figs. 11 and 12. Moreover, Table 3 reports the reorder levels and fixed order quantities. Note that the reorder levels are set to zero when the MTO policy is used, since the production is triggered by the customers demand only. In addition, we assign a value of 0.5 to each random switch. Moreover, the fraction of consumed goods to be recycled is set equal to $\mu=0.5$ in both retailers (see modules R1 and R2 of Figs. 11 and 12).

Continuous transitions		Discrete transitions			
	$[V_{min}, V_{max}]$	Exponential	Average firing delay (hours)	Triangular	Average firing delay (hours)
$t_1\ t_5\ t_7$	[0, 4]	$t_{22}\ t_{40}\ t_{72}\ t_{84}$	2	t_{53}	1
$t_2\ t_3\ t_4$	[0, 5]	$t_{16}\ t_{26}\ t_{34}$	3	$t_{42}\ t_{43}\ t_{70}$	2
t_6	[0, 6]	$t_{10}\ t_{14}\ t_{18}\ t_{76}$	4	$t_{47}\ t_{48}\ t_{78}$	2
t_8	[0,7]	$t_{24}\ t_{28}\ t_{32}$	4	$t_{52}\ t_{54}\ t_{79}$	2
t_9	[0,6]	$t_{36}\ t_{38}\ t_{80}$	4	$t_{44}\ t_{45}\ t_{71}$	3
t_{75}	[0,7]	$t_{20}\ t_{30}\ t_{41}\ t_{85}$	5	$t_{46}\ t_{49}$	3
		$t_{12}\ t_{68}$	6	$t_{50}\ t_{51}$	3
		$t_{13}\ t_{69}$	18	t_{66}	60
		$t_{21}\ t_{31}$	19	t_{67}	60
		$t_{11}\ t_{15}\ t_{19}$	20		
		$t_{25}\ t_{29}\ t_{33}\ t_{81}$	20		
		$t_{37}\ t_{39}\ t_{77}$	20		
		$t_{17}\ t_{27}\ t_{35}$	21		
		$t_{23}\ t_{73}$	22		

Table 1. Firing speed and average firing delay of continuous and discrete transitions.

Initial marking	Parts	Capacities	Parts
$m_1(0)$, $m_5(0)$, $m_{11}(0)$, $m_{15}(0)$	20	C_1, C_5, C_{11}, C_{15}	100
$m_{23}(0)$, $m_{25}(0)$	20	C_{23}, C_{25}	100
$m_{31}(0)$	20	C_{31}	150
$m_{37}(0)$, $m_{39}(0)$	20	C_{37}, C_{39}	70
$m_3(0)$, $m_9(0)$, $m_{13}(0)$	15	C_3, C_9, C_{13}	100
$m_7(0)$, $m_{27}(0)$	25	C_7, C_{27}	100
$m_{17}(0)$, $m_{19}(0)$, $m_{29}(0)$	30	C_{17}, C_{19}, C_{29}	100
$m_{33}(0)$	30	C_{33}	150
$m_{21}(0)$	35	C_{21}	100
$m_{35}(0)$	0	C_{35}	120
$m_{95}(0)$, $m_{96}(0)$, $m_{97}(0)$	0	C_{95}, C_{96}, C_{97}	80
$m_{41}(0)$, $m_{87}(0)$, $m_{88}(0)$	0	/	/
$m_{89}(0)$, $m_{104}(0)$, $m_{105}(0)$	0	/	/

Table 2. Initial marking of odd continuous places, capacities and edge weights.

To analyze the SC, the following performance indices are selected (Viswanadham, 1999):

i. the system throughput T, i.e., the average number of products obtained in a time unit;
ii. the average system inventory SI, i.e., the average amount of products stored in all the system buffers during the run time TP;
iii. the average lead time LT=SI/T that is a measure of the time spent by the SC to convert the raw material in final products.

Note that in the considered simulation experiments the SI performance index (and, consequently, the LT value) is calculated taking into account only the buffers that are upstream with respect to the retailers.

The FOHPN models of the case study under the MTS and MTO policies are implemented and simulated in the Matlab environment (The Mathworks 2006), which is ideally suited when dealing with modular, numerical matrix-based large-scale systems. All the indices are evaluated by a simulation run of 600 time units with a transient period of 100 time units, so that the run time equals 500 hours if we associate one time unit to one hour. The estimates of the performance indices are deduced by 100 independent replications with a 95% confidence interval. Besides, we evaluate the percentage value of the confidence interval half width to assess the accuracy of the performance index estimation: the half width of the confidence interval, being 3.3% in the worst case, confirms the sufficient accuracy of the performance indices estimation.

Reorder levels (parts) for the MTS case	Reorder levels (parts) for the MTO case	Fixed order quantities (parts)
R_1=18	R_1=0	Q_1=50
R_2=25	R_2=0	Q_2=45
R_3=25	R_3=0	Q_3=55
R_4=25	R_4=0	Q_4=40
R_5=15	R_5=0	Q_5=60
R_6=15	R_6=0	Q_6=50
R_7=20	R_7=0	Q_7=30
R_8=20	R_8=0	Q_8=25
R_9=30	R_9=0	Q_9=2
R_{10}=10	R_{10}=0	Q_{10}=5
R_{11}=10	R_{11}=0	Q_{11}=35
R_{12}=10	R_{12}=0	Q_{12}=45
		Q_{13}=40
		Q_{14}=40
		Q_{15}=25
		Q_{16}=35

Table 3. Reorder levels and fixed order quantities.

The simulation study is performed in three different operative conditions, respectively denoted *OCi* with i=1,2,3, and each operative condition *OCi* corresponds to a different choice of the IFS vector within the set of admissible solutions of (1).
Operative Condition 1 (OC1). At each macro-period the IFS vector $\boldsymbol{v}$ is selected so as to maximize the sum of all flow rates:

$$\begin{aligned} &\max\ (\mathbf{1}^{\mathrm{T}}\cdot\boldsymbol{v}) \\ &\quad \text{s.t. } \boldsymbol{v}\in S(PN,\boldsymbol{m}). \end{aligned} \tag{8}$$

Operative Condition 2 (OC2). At each macro-period the IFS vector $\boldsymbol{v}$ is selected so as to minimize the sum of all the stored materials:

$$\begin{aligned} &\min\ (\mathbf{c}^{\mathrm{T}}\cdot\boldsymbol{v}) \\ &\quad \text{s.t. } \boldsymbol{v}\in S(PN,\boldsymbol{m}). \end{aligned} \tag{9}$$

with $c_j=\boldsymbol{C}(p,t_j)$ if $t_j\in p^{(c)}\cup{}^{(c)}p$ and c_j=0 otherwise.
Operative Condition 3 (OC3). At each macro-period the IFS vector $\boldsymbol{v}$ is selected so as to maximize the sum of all flow rates as in (8) while setting all the capacities, re-order levels and fixed order quantities of the SC equal to 3/5 of the nominal values reported in Tables 2 and 3.
We remark that the first operative condition allows us to estimate the maximum level of performance of the SC with respect to the production capacity. Obviously, if we wish to maximize the manufacturing of a sub-set of products only, then (8) may be accordingly modified with a suitable objective function. Similarly, in the second operative condition the

stored materials in the SC buffers are minimized and, if we wish to minimize the stocks of inventory of a sub-set of products, then (9) may be accordingly modified by a suitable vector $\mathbf{c}^T$. Moreover, the third operative condition aims at imposing an *almost* constant work-inprocess in each buffer as a fraction of each capacity, similar to the previously described CONWIP management technique (Spearman et al. 1990). In this case the storage costs are a priori limited and at the same time the SC does not evolve at its maximum productivity level.

5.4 The simulation results

Figures 13, 14 and 15 report the selected SC performance indices, i.e., throughput, system inventory and lead time, respectively, obtained under the two management strategies and in the three operative conditions. In particular, Fig. 13 shows that in the two conditions *OC1* and *OC2* the system throughput values obtained under the MTS strategy are always greater than the corresponding values obtained with the MTO policy, since in the latter case the production is triggered by orders only. On the other hand, the reduced productivity imposed to the SC under *OC3* leads to equivalent throughput values in such a condition under MTS and MTO. Moreover, the throughput values obtained under *OC1* for a given management strategy (i.e., MTS or MTO) are greater than the ones corresponding to *OC2* and these are in turn bigger than those obtained in *OC3*. Indeed, solving (8) corresponds to maximizing the flow rates of the SC while the objective of (9) is to minimize the SC inventory and finally the third condition corresponds to a decreased level of the SC overall maximum productivity. In addition, it is noteworthy that the greatest (smallest) average throughput value, i.e. the highest (lowest) productivity level, is obtained when the SC is governed by the MTS (MTS or MTO) policy and in *OC1* (*OC3*).

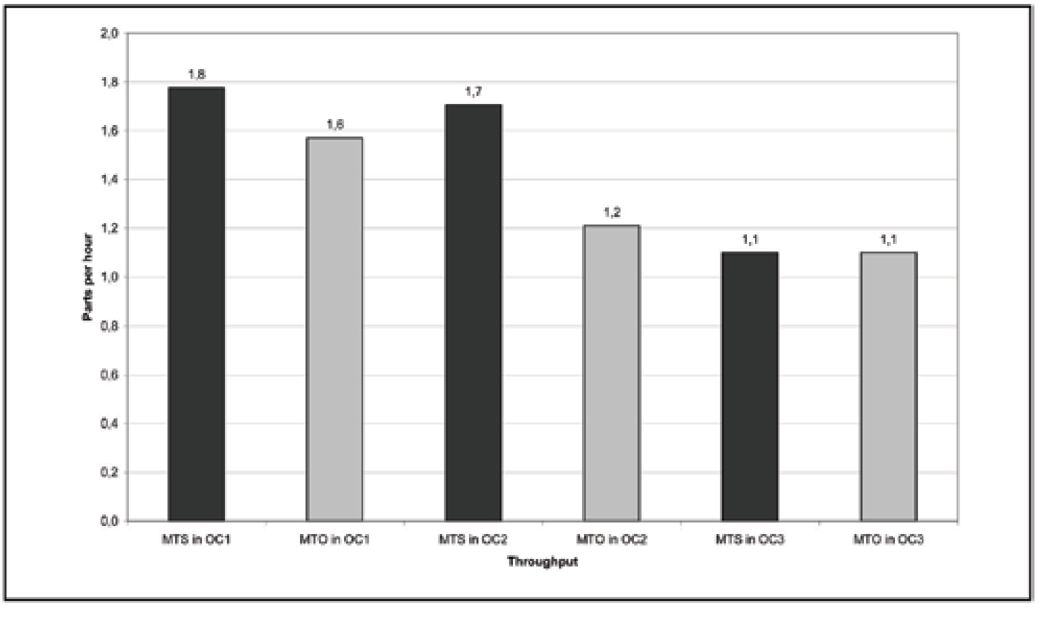


Fig. 13. Throughput (parts per hour) for different operative conditions under MTS and MTO.

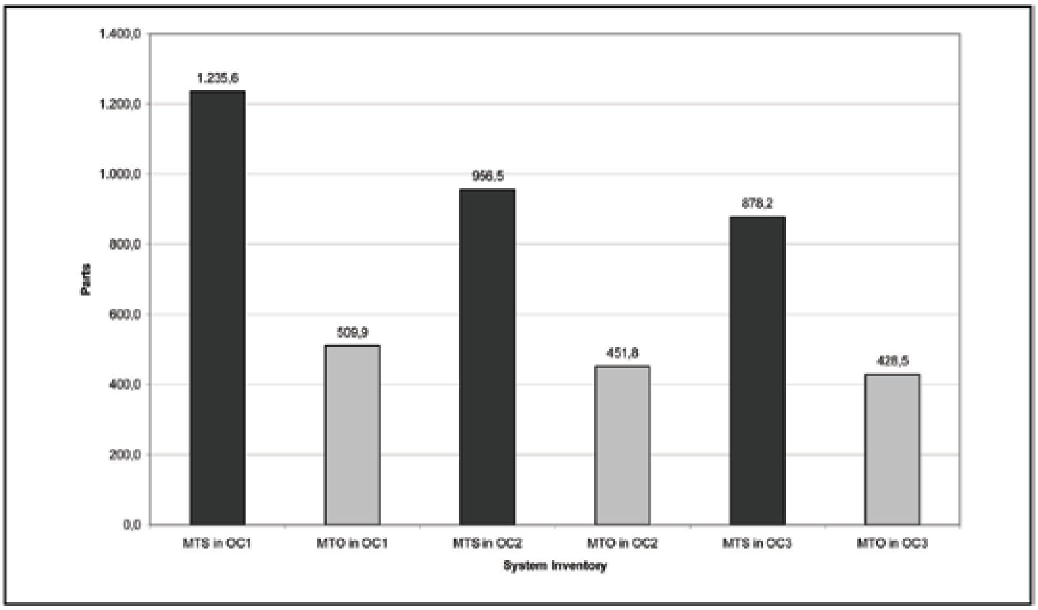


Fig. 14. System inventory (parts) for different operative conditions under MTS and MTO.

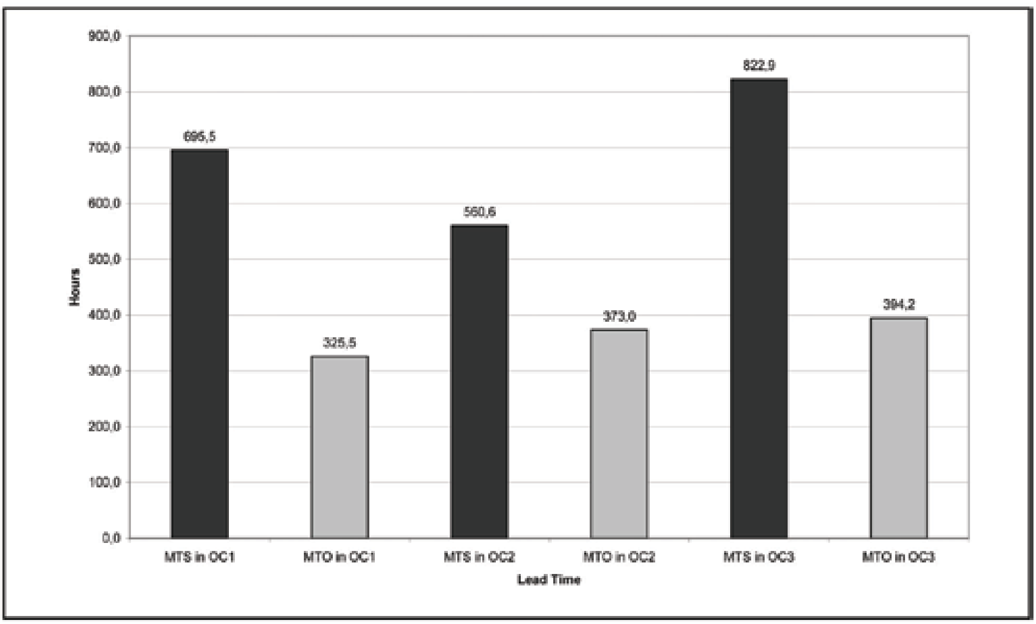


Fig. 15. Lead time (hours) for different operative conditions under MTS and MTO.

On the other hand, Fig. 14 shows that in the three cases the system inventory values evaluated under the MTS policy are always much greater than the corresponding values obtained with the MTO strategy. Moreover, the values of index SI obtained in *OC1* under a given policy (i.e. MTS or MTO) are greater than the ones corresponding to *OC2* for the same

strategy and these are in turn bigger than those obtained in *OC3* with that policy. The greatest (smallest) average system inventory value, i.e. the highest (lowest) storage level, is obtained when the SC is governed by the MTS (MTO) policy and in *OC1* (*OC3*).
Besides, Fig. 15 shows that in any operative condition the lead time values obtained under MTS are always greater than those obtained with the MTO policy, since in the former case the higher production corresponds on average to a longer permanence of materials and products in the system. Furthermore, Fig. 15 shows that under a the MTS operational management policy the LT values obtained in *OC1* are greater than the corresponding ones obtained in *OC2*, since the former case corresponds to a higher productivity. On the other hand, in *OC3* the value of LT under MTS (MTO) is greater than the corresponding index obtained in *OC1* under the same policy, since in the former condition the considerable throughput diminishment counterbalances the less significant system inventory decrease. Conversely, the MTO strategy lead to similar LT values in the three operative conditions.
Summing up, the SC managed under MTS is more productive than the system using MTO. On the other hand, using the latter policy the stocks in the considered SC are reduced and the same applies to the lead time values. Moreover, a different choice of the production rates and inventory management (as in *OC1*, *OC2* and *OC3*) lets the designer further find a tradeoff between different key performance indicators of the SC. For instance, comparing the system under the MTS policy in *OC1* and *OC2*, it is apparent that in these conditions a similar productivity level is attained (compare the corresponding throughput values in Fig. 13) but in the latter case the system inventory is quite smaller (as enlightened by the SI values in Fig. 14), leading to more sustainable storage costs than in the former case, with a lower lead time, as well (see the LT values in Fig. 15). Moreover, *OC3* under the MTS or MTO policy is characterized by the lowest value of average throughput and highest value of lead time compared to the other operative conditions with the same management policy (see Fig. 13 and Fig. 15) while offering the lowest average system inventory (see Fig. 14). On the other hand, under a given management strategy *OC1* is characterized by a lower value of LT with respect to *OC3* (see Fig. 15) but *OC1* involves a higher SI value than *OC3*, while leading to a better (or at most equal) throughput (see Fig. 13 and Fig. 14). With regard to the MTO strategy, the obtained LT values in the three operative conditions are quite similar (see Fig. 15), with *OC3* leading to the lowest SI index (see Fig. 14) at the price of a low value of T (see Fig. 13).

6. Conclusion

The chapter focuses on the problem of managing at the operational level Supply Chains (SCs), new emerging company networks, very complex to describe and manage. The SC system is described by a modular model based on the first order hybrid Petri net formalism, previously proposed by the authors: a fluid approximation of material and products is considered and discrete unpredictable events occurring stochastically (i.e., blocking of suppliers, manufacturers, transporters, etc.) are singled out by the discrete event dynamics. The model can effectively describe the operational management policies and the inventory control rules, and enables the designer to impose an optimal SC dynamics according to given objective functions.
To show the effectiveness and simplicity of the modeling technique, a SC case study is modelled and simulated under the alternative Make-To-Stock and Make-To-Order policies with a reorder point based inventory rule. The simulation results show that the fluid

approximation leads to an effective management policy implementation and to the possibility of choosing some important control parameters of the system, such as the production rates. Perspectives on future research include investigating the optimal decoupling point position, as well as implementing additional inventory control rules.

7. References

Alla, H. & David, R. (1998). Continuous and hybrid Petri nets. *Journal of Circuits, Systems and Computers*, Vol. 8, No. 1, 159-188, ISSN 0218-1266.

Balduzzi, F.; Giua, A. & Menga, G. (2000). Modelling and control with first order hybrid Petri nets. *IEEE Transactions on Robotics and Automation*, Vol. 4, No. 16, 382-399, ISSN 1042-296X.

Chen, H.; Amodeo, L.; Chu, F. & Labadi, K. (2005). Modeling and performance evaluation of supply chains using batch deterministic and stochastic Petri nets. *IEEE Transactions on Automation Science and Engineering*, Vol. 2, No. 2, 132-144, ISSN 1545-5955.

Chopra, S. & Meindl, P. (2001). *Supply chain management: strategy, planning, operation*. Prentice Hall, ISBN 0130264652, Upper Saddle River.

Desrochers, A.; Deal, T.J. & Fanti, M.P. (2005). Complex-valued token Petri nets. *IEEE Transactions on Automation Science and Engineering*, Vol. 2, No. 4, 309-318, ISSN 1545-5955.

Dotoli, M. & Fanti, M.P. (2005). A generalized stochastic Petri net model for distributed manufacturing system management. *Proceedings of CDC-ECC'05, 44th IEEE Conference on Decision and Control and European Control Conference*, pp. 2125-2130, ISBN 0-7803-9568-9, Seville, Spain, December 12-15, 2005, IEEE, Seville, Spain.

Dotoli, M.; Fanti, M.P.; Iacobellis, G. & Mangini, A.M. (2008). A first order hybrid Petri net model for supply chain management, to appear on the *IEEE Transactions on Automation Science and Engineering,* ISSN 1545-5955.

Dotoli, M.; Fanti, M.P.; Meloni, C. & Zhou, M.C. (2005). A multi-level approach for network design of integrated supply chains. *International Journal of Production Research*, Vol. 43, No. 20, 4267-4287, ISSN 0020-7543.

Dotoli, M.; Fanti, M.P.; Meloni, C. & Zhou, M.C. (2006). Design and optimization of integrated e-supply chain for agile and environmentally conscious manufacturing. *IEEE Transactions on Systems Man and Cybernetics, part A*, Vol. 36, No. 1, 62-75, ISSN 1083-4427.

Elmahi, I.; Grunder, O. & Elmoudni, A. (2003). A max plus algebra approach for modelling and control a supply chain. *Proceedings of IEEE-CCA'2003, 2003 IEEE International Conference on Control Applications*, pp. 1425-1430, ISBN 0-7803-7729-X , Istanbul, Turkey, June 23-25, 2003, IEEE, Istanbul.

Furcas, R.; Giua, A.; Piccaluga, A. & Seatzu, C. (2001). Hybrid Petri net modelling of inventory management systems, *European Journal of Automation APII-JESA*, Vol. 35, No. 4, 417-434, ISSN 0296-1598.

Gaonkar, R.S. & Viswanadham, N. (2001). Collaboration and information sharing in global contract manufacturing networks. *IEEE/ASME Transactions on Mechatronics*, Vol. 6, No. 4, 366-376, ISSN 10834435.

Hopp, W.J. & Spearman, M.L. (2004). To pull or not to pull: what is the question? *Manufacturing and Service Operations Management*, Vol. 6, No. 2, 133-148, ISSN 1523-4614.

Kelton W.D.; Sadowski R.P. & Sturrock D.T. (2004), *Simulation with Arena, Third edition*, McGraw-Hill, ISBN 9780072919813, Boston, MA.

Law A.M. & Kelton W.D. (2000). *Simulation modeling and analysis, Third edition*. McGraw-Hill, ISBN 978-0070366985, Boston, MA.

Luo, Y.; Zhou, M.C. & Caudill, R.J. (2001). An integrated e-supply chain model for agile and environmentally conscious manufacturing. *IEEE/ASME Transactions on Mechatronics*, Vol. 6, No. 4, 377-386, ISSN 10834435.

Rota, K.; Thierry, C. & Bel, G. (2002). Supply chain management: a supplier perspective. *Production Planning and Control*, Vol. 13, No. 4, 370-380, ISSN 0953-7287.

Shapiro, J.F. (2001). *Modeling the supply chain*. Duxbury Press, ISBN 978-0-534-37741-0, Pacific Grove.

Silva, M., Recalde, L. (2004). On the fluidification of Petri nets: from discrete to hybrid and continuous models, *Annual Reviews in Control*, Vol. 28, 253-266, ISSN 1367-5788.

Spearman, M.L.; Woodruff, D.L. & Hopp, W.J. 1990. CONWIP: a pull alternative to kanban. *International Journal of Production Research*, Vol. 28, No. 5, 879-894, ISSN 0020-7543.

The Mathworks (2006). *MATLAB Release Notes for Release 14*. The Mathworks, Natick, MA, USA.

Vidal, C. & Goetschalckx, M. (1997). Strategic production-distribution models: a critical review with emphasis on global supply chain models. *European Journal of Operational Research*, Vol. 98, No. 1, 1-18, ISSN 0377-2217.

Viswanadham, N. (1999). Analysis of manufacturing enterprises: an approach to leveraging value Delivery processes for competitive advantage. Kluwer Academic, ISBN 1-85233-186-0, Boston.

Viswanadham, N. & Gaonkar, R.S. (2003). Partner selection and synchronized planning in dynamic manufacturing networks. *IEEE Transactions on Robotics and Automation*, Vol. 19, No. 1, 117-130, ISSN 1042-296X.

Viswanadham, N. & Raghavan, S. (2000). Performance analysis and design of supply chain: a Petri net approach. *Journal of the Operational Research Society*, Vol. 51, No. 10, 1158-1169, ISSN 0160-5682.

Vollmann, T.E.; Berry, W.L.; Whybark, D.C. & Jacobs, F.R. (2004). *Manufacturing planning and control systems for supply chain management*. Irwin/Mc Graw-Hill, ISBN 978-0072299908, New York.

Von Mevius, M. & Pibernik, R. (2004). Process management in supply chains - a new Petrinet based approach. *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, pp. 71-80, Big Island, Hawaii, USA, January 5-8 2004, IEEE, Big Island, Hawaii, USA.

Wortmann, J.C. (1983). A classification scheme for master production schedule. In: *Efficiency of manufacturing systems*, Berg, C., French, D. Wilson, B., eds, Plenum Press, ISBN 978-0306412561, New York.

Wu, T. & O'Grady, P. (2005). A network-based approach to integrated supply chain design. *Production Planning and Control*, Vol. 16, No. 5, 444-453, ISSN 0953-7287.

Zhou, M.C. & Venkatesh, K. (1998). Modeling, simulation and control of flexible manufacturing systems. A Petri net approach. World Scientific, ISBN 981-02-3029-X, Singapore.

9

A Physics Approach to Supply Chain Oscillations and Their Control

Ken Dozier and David Chang
Western Region Application Center
University of Southern California
U.S.A.

1. Introduction

Virtually all manufacturing involves supplies chains, in which value is added by each entity in the chain until a finished product emerges at the end. A serious problem that plagues all supply chains is unwanted fluctuation in the inventories along the chain. This disrupts the product output, is costly in capital, and can result in considerable disruption and hardship in personal lives. J.D. Sterman and his colleagues at MIT (Sterman & Fiddaman, 1993) have performed a useful service by providing business schools with a widely used simulation game that demonstrates for a beer distribution supply chain how easily the fluctuations can arise, and how difficult they can be to control. Their simulation results indicate that the oscillations are due to the over-reaction to input fluctuations by the individual entities in the chain, and the results strongly suggest that the situation could be improved by information technology that enables real-time feedback.

The purpose of this chapter is to describe a statistical physics approach to understanding the supply chain oscillations. Standard techniques from statistical physics can lead to insights on the nature of the oscillations and on means for their control. Much of the supply chain work has been reported earlier, primarily at a number of conferences (Dozier a& Chang, 2004a, 2004b, 2005, 2006a, 2006b, 2007), but this chapter is the first time that the general approach together with its applications have been assembled in one place, along with a number of possible extensions.

Section 2 provides some background on the application of statistical physics techniques to manufacturing issues. The focus in this section is on a pseudo-thermodynamics description of a manufacturing sector, and on the nature of effective intervention forces that can be uniquely derived from this thermodynamics.

Section 3 demonstrates how the statistical physics approach leads directly to the existence of supply chain normal mode oscillations. It is shown that the nature of the normal modes depends on the extent of the information exchange between the entities in the supply chain. When the information exchange occurs only between an entity and the two entities immediately below and above it in the chain, the normal mode oscillation frequencies depend strongly on the way the oscillation amplitudes change along the chain: the normal modes resemble sound waves. On the other hand, when the information exchange occurs between an entity and all the other entities in the chain, the normal mode oscillation

frequencies depend more weakly on the way the oscillation amplitudes change along the chain: the normal modes resemble plasma oscillations.
Section 4 discusses possible interventions that can increase the production rate of a supply chain, utilizing the effective pseudo-thermodynamic forces discussed in Section 2. It demonstrates that both the time scale and the position focus of the intervention are important for determining the effectiveness of the intervention. An example is given of a useful application to supply chains of a quasilinear approximation technique that is often used in plasma physics problems.
Section 5 summarizes the approach and results to date, and suggests several possible extensions.

2. Background on statistical physics modeling in manufacturing

There has been a long-term association between economics and one particular aspect of statistical physics: thermodynamics. This association can be found in both neoclassical economics and modern new growth economics. For example, Krugman (1995) points out that economics is based on physics, and one of his favorite examples is that of the thermodynamics of economics. Even systems far from economic equilibrium can be treated by (open system) thermodynamics (Thorne & London, 2000). Costanza, et. al. (1997) have noted that biology has also played an important role, emphasizing that both thermodynamics and biology drove the development of new growth economic models: Smith and Foley (2002) have remarked that both neoclassical economics and classical thermodynamics seek to describe natural systems in terms of solutions to constrained optimization problems.
The foregoing observations about the role of thermodynamics in economics are in the quasi-static realm: the variables involved are relatively slowly changing in time. However, statistical physics also treats dynamic (time-dependent) phenomena, and can be used to understand important. time-dependent processes in economics. Koehler (2001, 2003) has been studying the interactions between the private and public sectors, and has emphasized the need to understand how best to conduct interactions between sectors that can have quite drastically different intrinsic time scales.

2.1 Rationale for application of statistical physics

Why does statistical physics provide a good framework for analyzing static (time-independent) and dynamic (time-dependent) aspects of cybernetic-based management of manufacturing? There are two reasons:

1. For the time-independent aspects, the formalism is designed to display the implications of systematically focusing on a few macroscopic variables of interest, an advantage for understanding complex systems such as manufacturing.
2. For time-dependent aspects, the formalism naturally takes into account the impact of time-dependent collective (cooperative) effects on the systems, of importance for designing effective cybernetic management control practices for complex situations.

These two reasons are considered further below.

2.1(i). Time-independent aspects: systematic focus on macroscopic variables

The systematic focusing on a few macroscopic variables of interest is important for understanding the relationships that can exist between the variables, and the associated implications for how the associated microscopic variables are distributed in an industrial sector. For example, suppose the objectives of policies and management actions on a particular industrial sector are to increase the productivity and the sales for the sector. Statistical physics can explicitly exhibit the relationships between the two variables and show how the impact of policies and management on the two are reflected by changes in the distribution of productivities and sales of the companies in the sector.

The systematic focusing on a few macroscopic variables of interest is accomplished by the constrained optimization approach of statistical physics. The approach is "constrained" by the requirement that specific measured values can be assigned to the few chosen relevant variables, and it is "optimized" by deriving the probability distribution that corresponds to the maximum possible number of possible arrangements of the entities comprising the system. This approach leads to well-defined "management forces" that can change the distribution of the variables within an industrial sector.

2.1(ii). Time-dependent aspects: collective effects in a sector

An understanding of the collective effects in an industrial sector is important for optimizing the focus and timing of both policy and management efforts. In particular, it is well known that companies in supply chains exhibit inventory oscillations that can be larger than the oscillations in market demands. This is wasteful of resources, and impacts costs and profits in a negative manner. Properly focused policy and cybernetic-based management can reduce these oscillations, and statistical physics can provide a systematic guide on how best to focus and time these efforts to perform this cost-saving reduction.

2.2 Effective pseudo-thermodynamic forces

In this chapter we would like to consider the impact of interventional control on reducing the resource-wasteful inventory oscillations in supply chains. This involves the application of effective pseudo-thermodynamic forces to the entities in the supply chains. These forces can be most easily understood by considering a thermodynamic description of the quasistatic state of a manufacturing sector. In this subsection, the thermodynamics approach – with its definition of an effective force - is briefly summarized, and some validating examples are given of its successful application to real data.

2.2(i). Conservation laws for information transfer involving effective "information force"

Thermodynamics can be applied to manufacturing in a number of ways. As an illustration of a general approach, let us suppose that a particular industrial sector has an objective of decreasing the average unit cost of production, and wishes to do this by implementing information technology transfer from an external source. (For example, this external source could be the government to which the companies in the sector have provided real-time data.) The statistical physics approach will produce a concrete measurable value-added for that sector due to the information transfer.

This objective can be posed as a straightforward constrained optimization problem in statistical physics, and the problem can be solved by maximizing a probability distribution subject to taking into account the constraints by the use of Lagrange multipliers. As an example, suppose that the total costs incurred in some length of time, C(total) = $\sum$ C(i)n(i), is known and that the total number of units produced in that time, N(total) = $\sum$ n(i), is known, where in both expressions the summation is from i = 0 to i = N_T, the total number of companies in the sector. With this knowledge, we then ask what the most likely distribution of production is over i Since the total number of allowable ways that N(total) can be arranged subject to an assumed distribution n(i) is P[N(total), n(i)] = N(total)! / [n(1)!n(2)! … $n(N_T!]$, the most likely distribution n(i)can be determined by maximizing ln P subject to the constraints that N(total) and C(total) are known. Introducing the Lagrange multipliers α and β to form F(n(i)) = ln {P[N(total), n(i)]}- α [$\sum$ n(i) - N(total)] - β[$\sum$ C(i)n(i) – C(total)], we find on setting dF(n(i)) /dn(i) = 0 and using the Stirling large number approximation for a factorial,- ln{n(i)}-1 - α - β C(i) = 0, i.e. n(i) = A exp[-βC(i)], the familiar Maxwell-Boltzmann distribution. (A more detailed derivation of the relevant equations is contained in Dozier and Chang (2004).)

Given the Maxwell-Boltzmann distribution function, it is straightforward to define quantities that are analogous to the conventional thermodynamic quantities of thermodynamics. These are shown in Table 1:

Variable	Physics	Economics
State (i)	Hamiltonian eigenfunction	Production site
Energy	Hamiltonian eigenvalue E_i	Unit production cost C_i
Occupation number	Number in state N_i	Sales output Ni
Partition function Z	$\sum \exp[-(1/k_B T)E_i]$	$\sum \exp[-\beta C_i]$
Free energy F	$k_B T \ln Z$	$(1/\beta) \ln Z$
Generalized force f_ξ	$\partial F/\partial \xi$	$\partial F/\partial \xi$
Example	ξ = Pressure	ξ = Information technology
Example	ξ= Electric field x charge	ξ = Mfg. technology
Entropy (randomness)	$- \partial F / \partial T$	$k_B \beta^2 \partial F/\partial \beta$

Table 1. Comparison of statistical formalism in physics and in economics (based on Dozier and Chang (2004a)

In this Table, the statistical physics formalism has naturally replaced the product of Boltzmann's constant k_B and the temperature T in the physics realm by the inverse of a "bureaucratic factor β" in the economic realm.

Note that the constrained optimization approach leads directly to a generalized pressure-like force corresponding to information technology transfer. (A second example is also provided in Table 1 of an effective force due to the introduction of new manufacturing technology.) For the purposes of this chapter, the important result is that an effective "force" can be derived from a partition function or associated free energy provided by a constrained optimization pseudo-thermodynamics of manufacturing.

As a further illustration of the result of applying a statistical physics technique to model a manufacturing sector, a conservation equation relating pseudo-thermodynamic quantities can be obtained by differentiating the expression for the total cost of production. This is shown in Table 2.

We note that the conservation equation (that corresponds to the First Law of Thermodynamics) involves the effective force f_ξ that is uniquely determined by the "free energy function" F. It is also interesting to note that if the two terms in the conservation equation have opposite signs, then the effect of information technology transfer is not as large as it could be. On the other hand, if the two terms have the same sign, then the impact of information technology transfer on costs for the sector are maximized.

Total cost of production

$$C = \Sigma\, C(\xi, i)\exp[-\beta C(\xi, i) - F(\xi)]$$

Effect of a change $d\xi$ in a parameter ξ in the system and a change $d\beta$ in bureaucratic factor

$$dC = -\langle f_\xi \rangle d\xi + \beta(\partial^2 F/\partial\beta\partial\xi)d\xi + (\partial^2\{\beta F\}/\partial\beta^2)d\beta$$

which can be rewritten

$$dC = -\langle f_\xi \rangle d\xi + TdS$$

Significance:

First term on the RHS describes lowering of unit cost of production
Second term on RHS describes increase in entropy (with a temperature
that is defined to be the inverse of β

Table 2. Conservation law for information technology transfer (based on Dozier and Chang (2004a)

As a preliminary data comparison, these results were applied to two particular sectors (the semiconductor sector and the heavy springs sector) in the U.S. Economic Census Data for the Los Angeles Metropolitan Statistical Area for 1992 and 1997 (Dozier & Chang, 2006b) The per capita information technology investments were known for both of these sectors: they were high for the semiconductor sector and low for the heavy springs sector.

The semiconductor sector experienced both a decrease in the average unit cost of production) and a decrease in the entropy term (a lower effective temperature). In this case, both terms in the conservation equation were operating synergistically. The sector also displayed a marked increase in output.

The heavy spring sector (the information-poor sector), on the other hand, displayed little change in output. Its entropy term increased and its average unit cost of production also increased. In this case, both terms acted in the wrong direction.

Although this preliminary examination of data is not directly applicable to information transfer in general, and although it is not possible to attribute all the changes to information technology investments since several other factors were operating, the examples suggest that both the force and entropy terms (due to all the relevant factors) in the conservation equation can act synergistically in the desired direction.

Accordingly, this suggests that information technology transfer can act to both decrease the effective temperature of a sector and to reduce costs. The former might be achieved by focusing information technology transfer efforts on companies that have unit costs that are far out of line with the average unit costs. It might also be achieved by focusing information technology transfer efforts on companies that have the best unit costs, letting the other companies drop by the wayside. Another possible method of decreasing entropy might be to provide market incentives for “heat flow” from the sector to sectors (including the government itself) that have lower temperatures.

2.2(ii). Static phenomena: Impact of information on productivity in the Los Angeles region

As a second example of the application of the constrained optimization technique of statistical physics, the impact of information technology on the improvement of productivity (i.e. on improving the output per employee) was considered. A detailed account of this example is contained in Dozier and Chang (2006b) and is described in the proceedings of the CITSA 04 conference (Dozier & Chang 2004a).

The constrained optimization formalism is applied here also, except that now the focus is not on the unit cost of production but rather on the shipments per company. In this example, the technique is used to derive the most likely distribution of the number of companies having a particular value of shipments per company vs. the shipments per company.

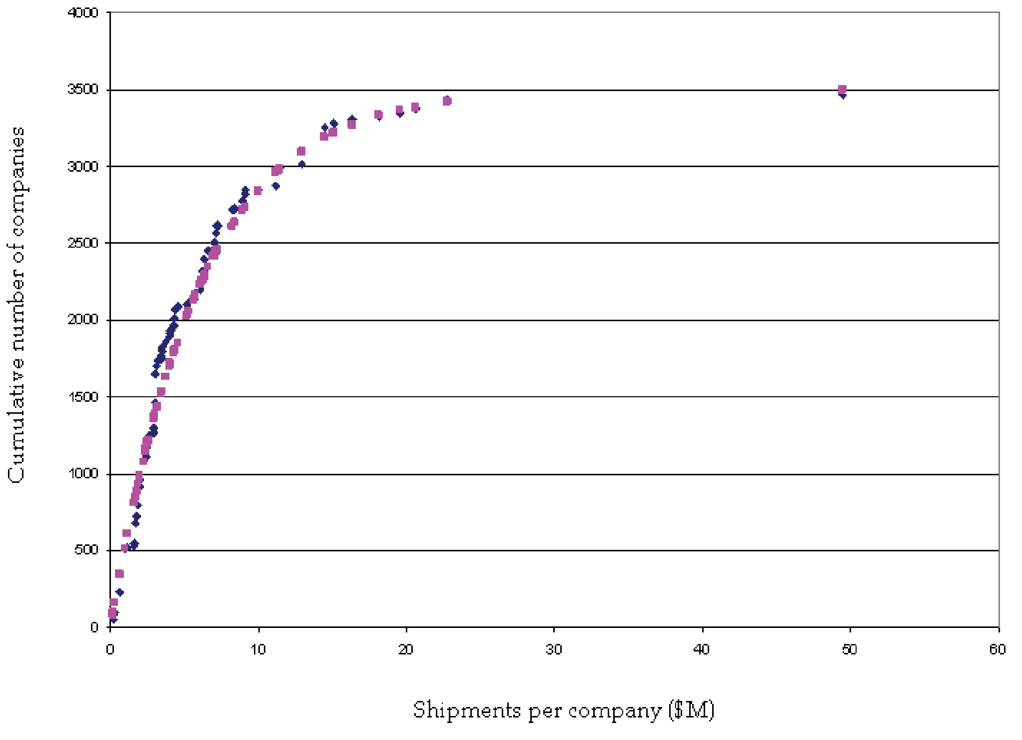


Fig. 1. Comparison of U.S. economic census cumulative number of companies vs. shipments/company (diamond points) in LACMSA in 1992 and the statistical physics cumulative distribution curve (square points) with β = 0.167 per $\$10^6$

The figure above (taken from Dozier and Chang, 2006b) compares the predicted cumulative distribution for a value of "inverse temperature" β = 0.167 per million dollars with actual data for the manufacturing sectors in the Los Angeles Metropolitan Statistical Area for 1992. As can be seen from the Figure, the comparison is quite good, giving good validation for the statistical physics constrained optimization applicability to an economics problem.

In this preliminary study, the interest was in the impact of information technology (IT) on various statistical physics parameters of companies as a function of company size. Table 3 shows the ratios of the 1997 statistical physics characterization parameters to their values in

1992 for large, intermediate, and small companies in the Los Angeles manufacturing sectors. Shown are the ratios for the number of companies in a size segment (#). the number of employees in the corresponding segment (E), the average number of employees per company (E/comp). the shipments of the sector (Sh), the shipments per employee (Sh/E), and a normalized "inverse temperature (β)

Company size:	Large	Intermediate	Small
IT rank	59	70	81
#	0.86	1.0	0.90
E(1000s)	0.78	0.98	1.08
E/company	0.91	1.0	1.21
Sh ($million)	1.53	1.24	1.42
Sh/E ($1000)	1.66	1.34	1.35
β	1.11	0.90	0.99

Table 3. Ratio ("97/'92) of the statistical physics characterization parameters (from Dozier & Chang , 2006b)

Without going into detail, it can be seen from Table 3 that the impact of information technology investments on the statistical physics characterization parameters does seem to be correlated with the company size. Whether or not the same type of size-dependency exists in the technology transfer realm bears analysis.

To summarize, Section 2 has shown that the application of the constrained optimization technique of statistical physics to (quasi) time-independent economic phenomena is supported by preliminary comparisons with U.S. Economic Census Data for the Los Angeles Metropolitan Statistical Area. The application of the technique appears to be justified both to systematically analyze the data and to provide a comprehensive and believable framework for presenting the results. For our immediate purpose of analyzing oscillatory phenomena in supply chains, the successes also provide confidence in the concept of an effective pseudo-thermodynamic-derived "information force".

3. Normal mode oscillations in supply chains

In this section, we describe a simple statistical physics model of inventory oscillations in a supply chain. It is based on material presented in three conferences (Dozier & Chang, 2005, 2006a, 2007), and follows especially closely the presentation in Dozier & Chang (2006a)

Section 3a describes a simple time-dependent conservation equation for the production units flowing through the chain. This equation is the supply-chain equivalent of the Liouville or Vlasov equations of statistical physics.

Section 3b applies this equation to derive results for a chain in which each entity only exchanges information with the two entities immediately above and below it in the chain.

Section 3c derives the oscillation dispersion relation for a supply chain in which every entity in the chain exchanges information with all the other entities in the chain.

In order to focus on the bare essentials of the approach, only a linear supply chain will be considered in the following. In addition, end effects will be ignored. It should be quite apparent how to generalize the results to a situation in which the number of companies at each level of the supply chain varies, and in which effects associated with the finite length of a chain are included.

3.1 Conservation equation for supply chain

In order to replace difference equations with more familiar differential equations, instead of designating each level in the chain by a discrete label n, the position in a chain will be designated by a continuum variable x. (Dozier & Chang (2005) shows the modifications resulting from dealing with the discrete label n instead of the continuous variable x, and demonstrates that for long supply chains, the basic features of both treatments are essentially the same.) The rate of flow of a production unit through each position x in the chain can be characterized by a velocity variable v.

Introduce a distribution function f(x,v,t) that depends on position, production flow rate velocity, and time. In statistical physics parlance, x and v denote the phase space for the problem, and f(x,v,t)dxdv denotes the number of production units in the intervals dx and dv at a given x and v at the time t.

From its definition, this distribution function satisfies a conservation equation in the phase space of x and v, if it is assumed that no production units are destroyed at each level of the chain:

$$\partial f/\partial t + \partial[f dx/dt]/\partial x + \partial[f dv/dt]/\partial v = 0 \tag{1}$$

This equation simply states that the change of fdxdv is due only to the divergence of the flow into the phase space volume dxdv. In a perfectly operating supply chain, we would expect that there would be no divergence in the flow. By permitting a divergence in the flow – i.e. permitting the flow into a volume element dxdv to be different than the flow out, the possible existence of local inventory fluctuations is allowed.

By introducing a force F that influences the velocity of the production rate v, this equation can be rewritten

$$\partial f/\partial t + \partial[f v]/\partial x + \partial[f F]/\partial v = 0 \tag{2}$$

We assume that the force F is of the type discussed earlier in Section 2b(i), i.e. that it can be uniquely derived from a pseudo-thermodynamic partition function for the chain entities.

Since position x and velocity of the production rate v are independent variables, $\partial v/\partial x = 0$. If, moreover, the force F does not depend on the flow rate v, $\partial F/\partial v = 0$, so that Eq. (2) becomes

$$\partial f/\partial t + v\partial f/\partial x + F\partial f/\partial v = 0 \tag{3}$$

This is similar to the Liouville equation of classical mechanics, and has the familiar form of the Vlasov equation for collisionless plasmas (Spitzer, 1956).

3.2 Supply chain with local exchange of information

Now assume that F at the position x is determined only by the level of the inventories of the production units immediately above and below x in the chain. As explained below, it is reasonable to assume that the fractional change in the time rate of change of velocity (1/v)dv/dt is proportional to the fractional change in the gradient of the density N(x,t):

$$(1/v)dv/dt \propto - (1/N)dN/dx \tag{4}$$

Where $N(x,t) = \int dv f(x,v,t)$, and where the negative sign is explained below.

For local information exchange with the levels immediately above and below the level of interest, the change in the density is observed over only the interval dx $=2l$, where l is the spacing between levels in the supply chain. Thus, we can further write

$$(1/v)dv/dt \propto -(2l/N)dN/dx \tag{5}$$

The rationale for this expression is that when the inventory of the level below the level of interest is less than normal, the production rate (v) will be diminished because of the smaller number of production units being introduced to that level. At the same time, when the inventory of the level above the level of interest is larger than normal, the production rate will also be diminished because the upper level will demand less input so that it can "catch up" in its production through-put. Both effects give production rate changes proportional to the negative of the gradient of N. It is reasonable also that the fractional changes are related rather than the changes themselves, since deviations are always made from the inventories at hand.

We note in passing that the quantity l is somewhat arbitrary, and reflects an equally arbitrary choice of a scale factor that relates the continuous variable x and the discrete level variable n.

A time scale for the response is missing from Eq. (5). We know that a firm must make decisions on how to react to the flow of production units into the firm. Assume that the time scale of response $\tau_{response}$ is given by $\tau_{response} = (1/\xi)\tau_{processing}$, where $\tau_{processing}$ is the processing time for a unit as it passes through the firm, and for simplification we are assuming ξ and $\tau_{processing}$ are constant throughout the chain. Because of a natural inertia associated with cautious decision-making, it is likely that ξ will be less than unity, corresponding to response times being longer than processing times.

Then Eq. (5) becomes

$$(1/v)dv/dt = -(2\xi l/\tau_{processing}\, N)dN/dx \tag{6}$$

Since by definition, the steady state production rate velocity is given by $V_0 \approx l/\tau_{processing}$, this gives finally for the effective internal force that changes production flow rates:

$$F = dv/dt = -2\xi\, V_0^2(1/N)dN/dx \tag{7}$$

Insertion of this expression into Eq. (3) then yields

$$\partial f/\partial t + v\partial f/\partial x - 2\xi\, V_0^2(1/N)(dN/dx)\, \partial f/\partial v = 0 \tag{8}$$

In the steady state, the equation is satisfied by $f(x,v,t) = f_0(v)$, i.e. by a distribution function that is independent of position and time: In this desired steady state, production units flow smoothly through the line without bottlenecks. For a smoothly operating supply chain, $f_0(v)$ will be centered about the steady state flow velocity V_0, a fact that we shall make use of later.

Now suppose there is a (normal mode) perturbation of the form $\exp[i(\omega t - kx)]$, i.e.

$$f(x,v,t) = f_0(v) + f_1(v) \exp[-i(\omega t - kx)] \tag{9}$$

On linearizing eq. (8) with this $f(x,v,t)$, we find that $f_1(v)$ satisfies

$$-i(\omega - kv)f_1 - ik\, 2\xi\, V_0^2(1/N_0)N_1\partial f_0/\partial v = 0 \tag{10}$$

Solving for f_1:

$$f_1 = -2\xi k(N_1/N_0)\, V_0^2\partial f_0/\partial v(\omega - kv)^{-1} \tag{11}$$

On integrating this equation with respect to v, we get the statistical physics dispersion relation relating ω and k:

$$1+ 2\xi kV_0^2(1/N_0) \int dv\partial f_0/\partial v(\omega\text{-}kv)^{-1} =0 \tag{12}$$

This equation contains a singularity at $\omega=kv$. This singularity occurs where the phase velocity ω/k becomes equal to the velocity of flow v. There are well-defined methods for the treatment of singularities: Following the Landau prescription (Landau, 1946; Stix,1992)

$$\int dv\partial f_0/\partial v(\omega\text{-}kv)^{-1} = PP\int dv\partial f_0/\partial v(\omega\text{-}kv)^{-1} - i\pi(1/k)\partial f_0(\omega/k) /\partial v \tag{13}$$

where PP denotes the principal part of the integral, i.e. the value of the integral ignoring the contribution of the singularity.
To evaluate the principal part, assume that for most v, $\omega>>kv$. Then approximately

$$PP\int dv\partial f_0/\partial v(\omega\text{-}kv)^{-1}\approx \int dvkv\partial f_0/\partial v(1/\omega^2) \approx - kN_0/\omega^2 \tag{14}$$

This gives the sound-wave-like dispersion relation

$$\omega \approx (2\xi)^{1/2}\, kV_0 \tag{15}$$

Addition to this of the small contribution from the imaginary part yields

$$\omega = (2\xi)^{1/2}\, kV_0+ (k/2)(1/N_0)i\pi(\omega/k)^2\partial f_0(\omega/k) /\partial v \tag{16}$$

or, on using the approximate relationship of Equation [21] for the ω's in the second term on the RHS

$$\omega = (2\xi)^{1/2}\, kV_0 + i(\pi/2)(k/N_0)\, (2\xi)^{3/2}\, V_0{}^3\partial f_0((2\xi)^{1/2}\, V_0\,) /\partial v] \tag{17}$$

For the fast response times made possible by first order rapid information exchange, $\xi = O(1)$. Thus, with $f_0(v)$ peaked around V_0, $\partial f_0(4\xi V_0\,) /\partial v <0$.
Accordingly, the imaginary part of ω is less than zero, and this corresponds to a damping of the normal mode oscillation. It is interesting to note that since $(2\xi)^{1/2}V_0 >> V_0$ (where the distribution is peaked), the derivative will be small, however, and the damping will be correspondingly small.
We note in passing that the discrete level variable is used instead of the continuous variable x, the dispersion relation is the same as Eq. (10) for small k, but when $kl \rightarrow 1$, the dispersion relation resembles that of an acoustic wave in a solid (Dozier & Chang, 2004, and Kittel, 1996).
To summarize, this sub-section has shown that when an entity in this linear supply chain exchanges information only with the two entities immediately above and below it in the chain, a slightly damped sound-wave-like normal mode results. Inventory disturbances in such a chain tend to propagate forwards and backwards in the chain at a constant flow velocity that is related to the desired steady-state production unit flow velocity through the chain.

3.3 Supply chain with universal exchange of information

Consider next what happens if the exchange of information is not just local. (Suppose that information is shared equally between all participants in a supply chain such as in the use of

grid computing.) In this case, the force F in Eq. (3) is not just dependent on the levels above and below the level of interest, but on the f(x,v,t) at all x.
Let us assume that the effect of f(x,v,t) on a level is independent of what the value of x is. This can be described by introducing a potential function Φ that depends on f(x,v,t,) by the relation

$$\partial^2\Phi/\partial x^2 = -[C/N_0]\int dv\, f(x,v,t) \tag{18}$$

from which the force F is obtained as $F = -\partial\Phi/\partial x$. (That this is so can be seen by the form of the 1-dimensional solution to Poisson's equation for electrostatics: the corresponding field from a source is independent of the source position.)
The constant C can be determined by having F reduce approximately to the expression of Eq. (7) when f(x,v,t) is non zero only for the levels immediately above and below the level x_0 of interest in the chain. For that case, take $N(x+l) = N(x_0) + dN/dx\, l$ and $N(x-l) = N(x_0) - dN/dx\, l$, and $N(x)$ zero elsewhere. Then

$$F = -\partial\Phi/\partial x = -[C/N_0](dN/dx)\, 2l^2 \tag{19}$$

On comparing this with the F of Eq. (7), $F = -2\xi v^2(1/N)dN/dx$, we find (since the distribution function is peaked at V_0) that we can write $C = \xi V_0^2 / l^2$.
Accordingly,

$$\partial^2\Phi/\partial x^2 = -[\xi V_0^2/N_0 l^2]\int dv\, f(x,v,t) \tag{20}$$

With these relations, F from the same value of f(x,v,t) at all x above the level of interest is the same, and F from the same value of f(x,v,t) at all x below the level of interest is the same but of opposite sign.
This is the desired generalization from local information exchange to universal information exchange.
It is interesting to see what change this makes in the dispersion relation. Eq. (3) now becomes

$$\partial f/\partial t + v\partial f/\partial x - \partial\Phi/\partial x\, \partial f/\partial v = 0 \tag{21}$$

and again the dispersion relation can be obtained from this equation by introducing a perturbation of the form of Equation (15) and assuming that Φ is of first order in the perturbation. This gives

$$-i(\omega-kv)f_1 = ik\Phi_1\partial f_0/\partial v \tag{22}$$

i.e.,

$$f_1 = -k\Phi_1\partial f_0/\partial v\, (\omega-kv)^{-1} \tag{23}$$

Since Eq. (20) implies

$$\Phi_1 = (1/k^2)\, [\xi V_0^2/N_0 l^2]\int dv\, f_1(v) \tag{24}$$

we get on integrating Eq. (23) over v:

$$1 + (1/k)\, [\xi V_0^2/N_0 l^2]\int dv\partial f_0/\partial v\, (\omega-kv)^{-1} = 0 \tag{25}$$

Once again a singularity appears in the integral, so we write

$$\int dv\partial f_0/\partial v\, (\omega-kv)^{-1} = PP\int dv\partial f_0/\partial v\, (\omega-kv)^{-1} - i\pi(1/k)\partial f_0(\omega/k)/\partial v \tag{26}$$

Evaluate the principal part by moving into the frame of reference moving at V_0, and in that frame assume that $kv/\omega<<1$:

$$PP\int dv\partial f_0/\partial v\,(\omega-kv)^{-1} \approx \int dv\partial f_0/\partial v\,(1/\omega)[1+(kv/\omega)]$$

$$= -kN_0/\omega^2 \tag{27}$$

Moving back into the frame where the supply chain is stationary,

$$PP\int dv\partial f_0/\partial v\,(\omega-kv)^{-1} \approx -kN_0/(\omega-kV_0)^2 \tag{28}$$

This gives the approximate dispersion relation

$$1 - (1/k)\,[\xi V_0^2/N_0l^2]\,kN_0/(\omega-kV_0)^2 \approx 0 \tag{29}$$

i.e.

$$\omega = kV_0 + \xi^{1/2}\,V_0/l \quad \text{or} \quad \omega = kV_0 - \xi^{1/2}\,V_0/l \tag{30}$$

To describe a forward moving disturbance, we take $\omega>0$ as $k\to0$, discarding the minus solution for this case.
Now add the small imaginary part to the integral:

$$1+ (1/k)\,[\xi V_0^2/N_0l^2][\,-kN_0/(\omega-kV_0)^2 - i\pi(1/k)\partial f_0(\omega/k)/\partial v]= 0 \tag{31}$$

On iteration, this yields

$$\omega \approx kV_0 + \xi^{1/2}(V_0/l)\,[1 + i\,\{\pi\xi V_0^2/(2k^2l^2N_0)\}\partial f_0/\partial v\,] \tag{32}$$

where $\partial f_0/\partial v$ is evaluated at $v = \omega/k \approx V_0 + (\xi^{1/2}V_0/kl)$.
Since for velocities greater than V_0, $\partial f_0/\partial v< 0$, we see that the oscillation is damped.
Moreover, the derivative $\partial f_0/\partial v$ is evaluated at a velocity close to V_0, the flow velocity where the distribution is maximum. Since the distribution function is larger there, the damping can be large. (We note here that the expression of Eq. (32) differs a little from that in Dozier & Chang (2006a), due to an algebraic error in the latter.)
To summarize, Section 3 has shown that universal information exchange results both in changing the form of the supply chain oscillation to a plasma-like oscillation, and in the suppression of the resulting oscillation. Specifically, it has been shown that for universal information exchange, the dispersion relation resembles that for a plasma oscillation. Instead of the frequency being proportional to the wave number, as in the local information exchange case, the frequency now contains a component which is independent of wave number. The plasma-like oscillations for the universal information exchange case are always damped. As the wave number k becomes large, the damping (which is proportional to $\partial f_0(\omega/k)/\partial v$) can become large as the phase velocity approaches closer to the flow velocity V_0. This supports Sterman and Fiddaman's conjecture that IT will have beneficial effects on supply chains.

4. External interventions that can increase supply chain production rates

In Section 3, we have seen that universal information exchange among all the entities in a supply chain can result in damping of the undesirable supply chain oscillations. In this section, we change our focus to see if external interactions with the oscillations can be used to advantage to increase the average production rate of a supply chain.

A quasilinear approximation technique has been used in plasma physics to demonstrate that the damping of normal mode oscillations can result in changes in the steady state distribution function of a plasma. In this section, this same technique will be used to demonstrate that the resonant interactions of externally applied pseudo-thermodynamic forces with the supply chain oscillations also result in a change in the steady state distribution function describing the chain, with the consequence that production rates can be increased.

This approach will be demonstrated by using a simple fluid flow model of the supply chain, in which the passage of the production units through the supply chain will be regarded as fluid flowing through a pipe. This model also gives sound-like normal mode waves, and shows that the general approach is tolerant of variations in the specific features of the supply chain model used. A more detailed treatment of this problem is available at Dozier and Chang (2007).

4.1 Moment equations and normal modes

The starting point is again the conservation equation, Eq. (5), for the distribution function that was derived in Section 3a. To obtain a fluid flow model of the supply chain, it will be useful to take various moments of the distribution function:

Thus, the number of production units in the interval dx and x at time t, is given by the v^0 moment, $N(x, t) = \int dv f(x,v,t)$; and the average flow fluid flow velocity is given by the v^1 moment $V(x,t) = (1/N)\int v dv f(x,v,t)$.

By taking the v^0 and v^1 moments of Eq. (3) – see, e.g. Spitzer (2006) - we find

$$\partial N/\partial t + \partial[NV]/\partial x = 0 \tag{33}$$

and

$$\partial V/\partial t + V\partial V/\partial x = F_1 - \partial P/\partial x \tag{34}$$

where $F_1(x,t)$ is the total force F acting per unit dx and P is a "pressure" defined by taking the second moment of the dispersion of the velocities v about the average velocity V: $P(x,t) = \int dv(v-V)^2 f(x,v,t)$

We can write the pressure P in the form

$$P(x,t) = \int dv(v-V)^2 f(x,v,t) = N(x,t)\,(\Delta v)^2 \tag{35}$$

where

$$(\Delta v)^2 = \int dv(v-V)^2 f(x,v,t)/N(x,t) \tag{36}$$

This is a convenient form, since we it shall assume for simplicity that the velocity dispersion $(\Delta v)^2$ is independent of level x and time t. In that case, Eq. (34) can be rewritten as

$$\partial V/\partial t + V\partial V/\partial x = F_1 - (\Delta v)^2\,\partial N/\partial x \tag{37}$$

This implies the change in velocity flow is impacted by the primary forcing function and the gradients of the number density of production units. Equations (33) and (37) are the basic equations that we shall use in the remainder to describe temporal phenomena in this simple fluid-flow supply chain model.

Before considering the effect of externally applied pseudo-thermodynamic forces, we derive the normal modes for the fluid flow model. Accordingly, introduce the expansions $N(x,t) = N_0 + N_1(x,t)$ and $V(x,t) = V_0 + V_1(x,t)$ about the level- and time-independent steady state density N_0 and velocity V_0. (We can take the steady state quantities to be independent of the level in the supply chain, since again we are considering long supply chains in the approximation that end effects can be neglected.)
Upon substitution of these expressions for N(x,t) and V(x,t) into Eqs. (33) and (37), we see that the lowest order equations (for N_0 and V_0) are automatically satisfied, and that the first order quantities satisfy

$$\partial N_1 / \partial t + V_0 \, \partial N_1 / \partial x + N_0 \, \partial V_1 / \partial x = 0 \tag{38}$$

and

$$\partial V_1 / \partial t + V_0 \, \partial V_1 / \partial x = F_1(x,t) - (\Delta v)^2 \partial N_1 / \partial x \tag{39}$$

where $F_1(x,t)$ is regarded as a first order quantity.
As before,. the normal modes are propagating waves:

$$N_1(x,t) = \mathit{N_1} \exp[i(\omega t - kx)] \tag{40}$$

$$V_1(x,t) = \mathit{V_1} \exp[i(\omega t - kx)] \tag{41}$$

With these forms, Eqs. (38) and (39) become

$$i\,(\omega - kV_0)\mathit{N_1} + N_0\, ik\mathit{V_1} = 0 \tag{42}$$

$$i\, N_0\, (\omega - kV_0)\mathit{V_1} = -ik\, (\Delta v)^2 \mathit{N_1} \tag{43}$$

In order to have nonzero values for N_1 and V_1, these two equations require that

$$(\omega - kV_0)^2 = k^2(\Delta v)^2 \tag{44}$$

Equation (44) has two possible solutions

$$\omega_+ = k\,(V_0 + \Delta v) \tag{44a}$$

$$\omega_- = k\,(V_0 - \Delta v) \tag{44b}$$

The first corresponds to a propagating supply chain wave that has a propagation velocity equal to the sum of the steady state velocity V_0 plus the dispersion velocity width Δv. The second corresponds to a slower propagation velocity equal to the difference of the steady state velocity V_0 and the dispersion velocity width Δv. Both have the form of a sound wave: if there were no fluid flow ($V_0 = 0$), ω_+ would describe a wave traveling up the chain, whereas ω_- would describe a wave traveling down the chain. When $V_0 \neq 0$, this is still true in the frame moving with V_0

4.2 Resonant interactions resulting in an increased production rate

As indicated earlier, our focus in this section is on the effect of external interactions (such as government actions) on the rate at which an evolving product moves along the supply chain. This interaction occurs in the equations through an effective pseudo-thermodynamic

force $F_1(x,t)$ that acts to accelerate the rate. From the discussion of Section 3, we expect that this force will be most effective when it has a component that coincides with the form of a normal mode, since then a resonant interaction can occur.
To see this resonance effect, it is useful to present the force F in its Fourier decomposition

$$F_1(x,t) = (1/2\pi)\iint d\omega dk \boldsymbol{F_1}(\omega,k)\exp[i(\omega t-kx)] \quad (45)$$

where

$$\boldsymbol{F_1}(\omega,k) = (1/2\pi)\iint dxdt F_1(x,t)\exp[-i(\omega t-kx)] \quad (46)$$

With this Fourier decomposition, each component has the form of a propagating wave, and it would be expected that these propagating waves are the most appropriate quantities for interacting with the normal modes of the supply chain.
Our interest is in the change that F_1 can bring to V_0, the velocity of product flow that is independent of x. By contrast, F_1 changes V_1 directly, but each wave component causes an oscillatory change in V_1 both in time and with supply chain level, with no net (average) change.
To obtain a net change in V, we shall go to one higher order in the expansion of V(x,t):

$$V(x,t) = V_0 + V_1(x,t)+ V_2(x,t) \quad (47)$$

On substitution of this expression into Eq. (37), we find the equation for $V_2(x,t)$ to be

$$N_0(\partial V_2/ \partial t + V_0\partial V_2/\partial x) + N_1(\partial V_1/ \partial t + V_0\partial V_1/\partial x) + N_0 V_1\partial V_1/\partial x = - (\Delta v)^2 \partial N_2 /\partial x \quad (48)$$

This equation can be Fourier analyzed, using for the product terms the convolution expression:

$$\iint dxdt \exp[-i(\omega t-kx)] f(x,t)g(x,t) = \iint d\Omega dK \mathit{f}(-\Omega+\omega, -K+k)\mathit{g}(\Omega,K) \quad (49)$$

where

$$\mathit{f}(\Omega,K) = \iint dxdt \exp[-i(\Omega t-Kx)]f(x,t) \quad (50a)$$

$$\mathit{g}(\Omega,K) = \iint dxdt \exp[-i(\Omega t-Kx)]g(x,t) \quad (50b)$$

Since we are interested in the net changes in V_2 – i.e. in the changes brought about by F_1 that do not oscillate to give a zero average, we need only look at the expression for the time rate of change of the ω=0, k=0 component, $V_2(\omega=0, k=0)$.
From Eq. (48), we see that the equation for $\partial V_2(\omega=0, k=0)/\partial t$ requires knowing N_1 and V_1. When $\boldsymbol{F_1}(\omega,k)$ is present, then Eqs. (42) and (43) for the normal modes are replaced by

$$i (\omega-kV_0)N_1(\omega,k) + N_0 ikV_1(\omega,k) = 0 \quad (51)$$

$$i N_0 (\omega-kV_0)V_1(\omega,k) = -ik (\Delta v)^2 N_1(\omega,k) + \boldsymbol{F_1}(\omega,k) \quad (52)$$

These have the solutions

$$N_1(\omega,k) = -ik \boldsymbol{F_1}(\omega,k) [(\omega-kV_0)^2 - k^2 (\Delta v)^2]^{-1} \quad (53)$$

$$V_1(\omega,k) = - i \{\boldsymbol{F_1}(\omega,k)/N_0\}(\omega-kV_0) [(\omega-kV_0)^2 - k^2 (\Delta v)^2]^{-1} \quad (54)$$

Substitution of these expressions into the ω=0, k=0 component of the Fourier transform of Eq. (48) gives directly

$$\partial V_2(0,0)/\partial t = \iint d\omega dk(ik/N_0^2)\,(\omega-kV_0)^2\,[(\omega-kV_0)^2 - k^2\,(\Delta v)^2]^{-2}\,\boldsymbol{F_1}(-\omega,k)\,\boldsymbol{F_1}(-\omega,k) \qquad (55)$$

This resembles the quasilinear equation that has long been used in plasma physics to describe the evolution of a background distribution of electrons subjected to Landau acceleration [Drummond & Pines (1962)].

As anticipated, a resonance occurs at the normal mode frequencies of the supply chain, i.e. when

$$(\omega-kV_0)^2 - k^2\,(\Delta v)^2 = 0 \qquad (56)$$

First consider the integral over ω from $\omega = -\infty$ to $\omega = \infty$. The integration is uneventful except in the vicinity of the resonance condition where the integrand has a singularity. As before, the prescription of Eq. (13) can be used to evaluate the contribution of the singularity.

For Eq. (55), we find that when the bulk of the spectrum of $F_1(x,t)$ is distant from the singularities, the principal part of the integral is approximately zero, where the principal part is the portion of the integral when ω is not close to the singularities at $\omega = k(V_0 \pm \Delta v)$. This leaves only the singularities that contribute to $\partial V_2(0,0)/\partial t$.

The result is the simple expression:

$$\partial V_2(0,0)/\partial t = \pi/(N_0^2\Delta v)\int dk(1/k)\,[\,\boldsymbol{F_1}(-k(V_0-\Delta v, -k)\boldsymbol{F_1}(k(V_0-\Delta v),k) -$$

$$(-k(V_0+\Delta v, -k)F_1(k(V_0+\Delta v),k)] \qquad (57)$$

Equation (57) suggests that any net change in the rate of production in the entire supply chain is due to the Fourier components of the effective statistical physics force describing the external interactions with the supply chain, that resonate with the normal modes of the supply chain. In a sense, the resonant interaction results in the conversion of the "energy" in the normal mode fluctuations to useful increased production flow rates. This is very similar to physical phenomena in which an effective way to cause growth of a system parameter is to apply an external force that is in resonance with the normal modes of the system.

To summarize, Section 4 has shown that the application of the quasilinear approximation of statistical physics to a simple fluid-flow model of a supply chain, demonstrates how external interactions with the normal modes of the chain can result in an increased production rate in the chain. The most effective form of external interaction is that which has Fourier components that strongly match the normally occurring propagating waves in the chain.

5. Discussion and possible extensions

In the foregoing, some simple applications of statistical physics techniques to supply chains have been described.

Section 2 briefly summarized the application of the constrained optimization technique of statistical physics to (quasi) time-independent economic phenomena. It showed some preliminary comparisons with U.S. Economic Census Data for the Los Angeles Metropolitan Statistical Area, that supported the approach as a good means of systematically analyzing the data and providing a comprehensive and believable framework for presenting the results. It also introduced the concept of an effective pseudo-thermodynamic-derived "information force" that was used later in the discussion of supply chain oscillations.

Section 3 discussed supply chain oscillations using a statistical physics normal modes approach.

It was shown that the form of the dispersion relation for the normal mode depends on the extent of information exchange in the chain. For a chain in which each entity only interacts with the two entities immediately below and above it in the chain, the normal more dispersion relation resembles that of a sound wave. For a chain in which each entity exchanges information with all of the other entities in the chain, the dispersion relation resembles that of a plasma oscillation. The Landau damping in the latter could be seen to be larger than in the limited information exchange case, pointing up the desirability of universal information exchange to reduce undesirable inventory fluctuations.

Section 4 applied the quasilinear approximation of statistical physics to a simple fluid-flow model of a supply chain, to demonstrate how external interactions with the normal modes of the chain can result in an increased production rate in the chain. The most effective external interactions are those with spectra that strongly match the normally occurring propagating waves in the chain.

The foregoing results are suggestive. Nevertheless, the supply chain models that were used in the foregoing were quite crude: Only a linear uniform chain was considered, and end effects were ignored.

There are several ways to improve the application of statistical physics techniques to increase our understanding of supply chains. Possibilities include (1) the allowance of a variable number of entities at each stage of the chain, (2) relaxation of the uniformity assumption in the chain, (3) a more comprehensive examination of the effects of the time scales of interventions, (4) a systematic treatment of normal mode interactions, (5) treatment of end effects for chains of finite length,(6) consideration of supply chains for services as well as manufactured goods, and (7) actual simulations of the predictions. We can briefly anticipate what each of these extensions would produce.

Variable number of entities at each level Equations similar to those in Sections 3 and 4 would be anticipated. However, in the equations, the produced units at each level would now refer to those produced by all the organizations at that particular level. The significance is that the inventory fluctuation amplitudes calculated in the foregoing refer to the contributions of all the organizations in a level, with the consequence that the fluctuations in the individual organization would be inversely proportional to the number of entities in that level. Thus, organizations in levels containing few producing organizations would be expected to experience larger inventory fluctuations.

Nonuniform chains In Sections 3 and 4, it was assumed that parameters characterizing the processing at each level (such as processing times) were uniform throughout the chain. This could very well be unrealistic: for example, some processing times at some stages could be substantially longer than those at other stages. And in addition, the organizations within a given stage could very well have different processing parameters. This would be expected both to introduce dispersion, and to cause a change in the form of the normal modes.

As a simple example, suppose the processing times in a change increased (or decreased) linearly with the level in the chain. The terms of the normal mode equation would now no longer have coefficients that were independent of the level variable x. For a linear dependence on x, the normal modes change from Fourier traveling waves to combinations of Bessel functions, i.e. the normal mode form for a traveling wave is now a Hankel function. The significance of this is that the inventory fluctuation amplitudes become level-dependent: A disturbance introduced at one level in the chain could produce a much larger (smaller) fluctuation amplitude at another level.

Time scales of interventions Since inventory fluctuations in a supply chain are disruptive and wasteful of resources, some form of cybernetic control (intervention) to dampen the fluctuations would be desirable. In Section 4, it was suggested that interventions that resonate with the normal modes are most effective in damping the fluctuations and converting the "energy" in the fluctuations to useful increased production rates. Koehler (2001, 2002) has emphasized, however, that often the time scales of intervention are quite different from those of the system whose output it is desired to change.

A systematic means of analyzing the effects of interventions with time scales markedly different from those of the supply chain is available with standard statistical physics techniques:

For example, if the intervention occurs with a time scale much longer than the time scales of the chain's normal modes, then the adiabatic approximation can be made in describing the interactions. The intervention can be regarded as resulting in slowly changing parameters (as a function of both level and time). Eikonal equations (Weinberg 1962) can then be constructed for the chain disturbances, which now can be regarded as the motion of "particles" comprising wave packets formed from the normal modes.

At the other extreme, suppose the intervention occurs with time scales much less than the time scales of the chain's normal modes. When the intervention occurs at random times, the conservation equation (Eq. 3) can be modified by Fokker-Planck terms (Chandrasekhar, 1943). The resulting equation describes a noisy chain, in which a smooth production flow can be disrupted.

Normal mode interactions The beer distribution simulation (Sterman & Fiddaman, 1993) has shown that the amplitudes of the inventory oscillations in a supply chain can become quite large. The normal mode derivation in Sections 3 and 4 assumed that the amplitudes were small, so that only the first order terms in the fluctuation amplitudes needed to be kept in the equations. When higher order terms are kept, then the normal modes can be seen to interact with one another. This "wave-wave" interaction itself can be expected to result in temporal and spatial changes of the supply chain inventory fluctuation amplitudes.

End effects of finite chains The finite length of a supply chain has been ignored in the calculations of this chapter, i.e. end effects of the chain have been ignored. As in physical systems, the boundaries at the ends can be expected to introduce both reflections and absorption of the normal mode waves described. These can lead to standing waves, and the position and time focus of optimal means of intervention might be expected to be modified as a result.

Supply chains for services as well as manufactured products In the foregoing, we have been thinking in terms of a supply chain for a manufactured product. This supply chain can involve several different companies, or – in the case of a vertically integrated company – it could comprise several different organizations within the company itself. The service sector in the economy is growing ever bigger, and supply chains can also be identified, especially when the service performed is complex. The networks involved in service supply chains can have different architectures than those for manufacturing supply chains, and it will be interesting to examine the consequences of this difference. The same type of statistical physics approach should prove useful in this case as well.

Numerical simulations The statistical physics approach to understanding supply chain oscillations can lead to many types of predicted effects, ranging from the form and frequencies of the inventory fluctuations to the control and conversion of the fluctuations. Computer simulations would be useful in developing an increased understanding of the predictions. This is especially true when the amplitudes of the oscillations are large, since then the predictions based on small-amplitude approximations would be suspect.

The application of statistical physics techniques to understand and control supply chain fluctuations may prove to be very useful. The initial results reported here suggest that further efforts are justified.

6. References

Bogolyubov, N.N., and Zubarev, D.N..(1955) An Asymptotic Approximation Method for a System with Rotating Phases and its Application to the Motion of a Charged Particle in a Magnetic Field, *Ukrainian Math. Journal VI,I* 5, ISSN 0041-5995

Chandrasekhar, S. (1943) Stochastic Processes in Physics and Astronomy, *Revs. Modern Physics 15*, 1-89, ISSN 1539-0756

Chang, David B. (1964) Landau Damping and Related Phenomena, *Phys. Fluids 7*, 1980-1986, ISSN 1070-6631

Costanza, R., Cumberland, J., Daly, H., Goodland, R., Norgaard, R. (1997) *An Introduction to Ecological Economics. ISBN 1884015727, Saint Lucie Press, Boca Raton, Florida*

Dozier, K., and Chang, D. (2004a) Thermodynamics of Productivity: *Framework for Impacts of Information/Communication Investments, International Conference on Cybernetics and Information Technologies, Systems and Applications (CITSA 2004),* ISBN 980-6560-19-1, Orlando, Florida, July 21-25, 2004, International Institute of Informatics and Systemics, Winter Garden, Florida

Dozier, K., and Chang, D. (2004b) A Thermodynamic Model for Technology Transfer, presentation at *Technology Transfer Society (T2S) Annual Conference: Emerging Issues in Technology Transfer*, RPI , September 29-October 1, 2004, Albany , NY, available at www.wesrac.usc.edu

Dozier, K., and Chang, D. (2005) Cybernetic control in a supply chain: wave propagation and resonance. *The 11th International Conference on Information Systems Analysis and Synthesis. (ISAS 2005) and the 2nd International Conferenc4 on Cybernetics and Information Technologies, Systems and Applications (CITSA 2005).* July 14-17, 2005, Orlando, Florida, ISBN 980-6560-41-8 and 980-6560-42-6, International Institute of Informatics and Systemics, Winter Garden, Florida

Dozier, K., and Chang, D. (2006a) Role of information exchange in damping supply chain oscillations. *Information Resources Management Association (IRMA) conference*, ISBN 1-59904-020-4, Washington, D.C., May 21-24, 2006, Information Resources Management Association, Hershey, PA

Dozier, K., and Chang, D. (2006b) The effect of company size on the productivity impact of information technology investments. *Journal of Information Technology Theory and Application (JITTA) 8*:1, ISSN 1532-4516

Dozier, K., and Chang, D. (2007) The impact of information technology on the temporal optimization of supply chain performance. *Proceedings of the Hawaii International Conference on System Sciences HICSS-40, 2007*, ISBN 0-7695-2755-8, January 3-6, 2007, Hilton Waikoloa Village, Big Island, Hawaii, IEEE Computer Society, Washington, D.C..

Drummond, W.E. and Pines, D. (1962) Non-linear stability of plasma oscillations, *Nucl. Fusion, Suppl, 3,* 1049-1057, ISSN 0029-5515

Kittel, Charles (1996) *Introduction to Solid State Physics*. ISBN 047111813, John Wiley & Sons, New York

Koehler, Gus (2003) Time, Complex Systems, and Public Policy: A Theoretical Foundation for Adaptive Policy Making, *Nonlinear Dynamics, Psychology and the Life Science*s *7*, 1, 99, January, 2003, ISSN 1090- 0578

Koehler, Gus, (2001) A Framework for Visualizing the Chronocomplexity of Politically Regulated Time- Ecologies, presentation at *International Society for the Study of Time 2001 Conference*, Gorgonza, Italy, July 8-22, 2001, available at *www.timestructures.com/ downloads/ISST-Arrowhead%20PowerPt.pdf*

Krugman, Paul (1995) *Development, Geography, and Economic Theor.* ISBN 0585003246, The MIT Press , Cambridge, Mass.

Landau, L.D. (1946) On the Vibrations of the Electronic Plasma. *J. Phys*. (U.S.S.R.) 10, 25-34, ISSN 0368-3400

Smith, E. and Foley, D. (2002) *Classical thermodynamics and economic general equilibrium theory*. 32 page working paper, New School for Social Research, New York, available at Duncan K. Foley homepage

Spitzer, L. (2006). *Physics of Fully Ionized Gases*. ISBN 0486449823, Dover Publications, Mineola, New York

Sterman, J.D. & Fiddaman, T. (1993). *The Beer Distribution Game Flight Simulator, Software (for Macintosh) and Briefing Book.* Sloan School of Management, MIT E53-351 Cambridge, MA 02142

Stix, T.H. (1992) *Waves in Plasmas,* ISBN 0883188597, American Institute of Physics, New York

Thome, F. A., and London, S. (2000). Disequilibrium economics and development. *Computing in Economics and Finance 2000,* paper 377, repec.sce.scecf0, Society for Computational Economics, Barcelona, Spain

Weinberg, S. (1962) Eikonal method in magnetohydrodynamics. *Phys. Rev. 126,* 1899-1909, ISSN 1550-2376

10

Utilizing IT as an Enabler for Leveraging the Agility of SCM

Mehdi Fasanghari and S. K. Chaharsooghi
Iran Telecommunication Research Center (ITRC) & Tarbiat Modares University (TMU)
Iran

1. Introduction

Supply chain management (SCM) is the 21st century operations strategy for achieving organizational competitiveness. Companies are attempting to find ways to improve their flexibility, responsiveness, and competitiveness by changing their operations strategy, methods, and technologies that include the implementation of SCM paradigm and Information Technology (IT).
The use of IT is considered as a prerequisite for the effective control of today's complex supply chains. Indeed, a recent study is increasingly dependent on the benefits brought about by IT to: improve supply chain agility, reduce cycle time, achieve higher efficiency, and deliver products to customers in a timely manner (Radjou, 2003).
However, IT investment in the supply chain process does not guarantee a stronger organizational performance. The debate on the ''IT-productivity'' paradox and other anecdotal evidence suggests that the impact of IT on firm performance remains unclear (Lucas & Spitler, 1999). In fact, the adoption of a particular technology is easily duplicated by other firms, and it often does not provide a sustained competitive advantage for the adopting firms (Powell & Dent-Micallef, 1997).
The implementation of IT in the SCM can enable a firm to develop and accumulate knowledge stores about its customers, suppliers, and market demands, which in turn influences firm performance (Tippins & Sohi, 2003).
The main objective of this paper is to provide a framework that enhances the agility of SCM with IT.
The rest of this article is organized as follows. IT systems and Supply Chain Management will be described in the next sections. Therefore we begin with a brief review of the IT and SCM. Definitions for agility–as key subjects in this article- are ambiguous. Then, leveraging the agility of SCM is argued and the framework is represented. This is ended by conclusion.

2. IT systems

As for IT systems, when discussing the use of IT in SCM, we refer to the use of interorganizational systems that are used for information sharing and/or processing across organizational boundaries. Thus, besides internal IT systems such as Enterprise Resource Planning systems we also consider identification technologies such as RFID from the scope of this study (Auramo et al., 2005).

3. Supply chain management

A business network is defined as a set of two or more connected business relationships in which exchange in one relationship is contingent on (non-) exchange in another (Campbell & Wilson, 1996). Stevens (1989) defines SCM as 'a series of interconnected activities which are concerned with planning, coordinating and controlling materials, parts, and finished goods from supplier to customer. A supply chain typically consists of the geographically distributed facilities and transportation links connecting these facilities. In manufacturing industry this supply chain is the linkage which defines the physical movement of raw materials (from suppliers), processing by the manufacturing units, and their storage and final delivery as finished goods for the customers. In services such as retail stores or a delivery service like UPS or Federal Express, the supply chain reduces to problem if distribution logistics, where the start point is the finished product that has to be delivered to the client in a timely, manner. For a pure service operation, such as a financial services firm or a consulting operation, the supply chain is principally the information flow (Bowersox & Closs, 1996).

SCM and logistics definitions entail a supply chain perspective from first supplier to end-user and a process approach, but the main difference between them is that Logistics is a subset of SCM. Companies have realized that all business processes along with logistics process cut across supply chains (Lambert & Cooper, 1998). According to that, SCM ideally embraces all business processes cutting across all organizations within the supply chain, from initial point of supply to the ultimate point of consumption (Lambert & Cooper, 1998). For, SCM embraces the business processes identified by the International Center for Competitive Excellence (see Fig. 1).

4. IT and supply chain management

Recently with development of information technologies that include electronic data interchange (EDI), the Internet and World Wide Web (WWW), the concepts of supply chain design and management have become a popular operations paradigm. The complexity of SCM has also forced companies to go for online communication systems. For example, the Internet increases the richness of communications through greater interactivity between the firm and the customer (Walton & Gupta, 1999). Armstrong & Hagel (1996) argue that there is beginning of an evolution in supply chain towards online business communities.

Supply chain management emphasizes the long-term benefit of all parties on the chain through cooperation and information sharing. This signifies the importance of communication and the application of IT in SCM. This is largely caused by variability of ordering (Yu et al., 2001).

There have been an increasing number of studies of IT's effect on supply chain and interorganizational relationships (Grover et al., 2002). In this article, IT appears to be an important factor for collaborative relationships. A popular belief is that IT can increases the information processing capabilities of a relationship, thereby enabling or supporting greater interfirm cooperation in addition to reducing uncertainty (Subramani, 2004). IT decreases transaction costs between buyers and suppliers and creates a more relational/cooperative governance structure, leads to closer buyer-supplier relationships (Bakos & Brynjyoolfsson, 1993), may decrease trust-based interorganizational partnerships and removes a human element in buyer-supplier interaction, while trust is built on human interaction (Carr &

Smeltzer, 2002). A new challenge of marketing is occurred with combination of e-business and SCM. IT allows suppliers to interact with customers and receive enormous volumes of information for data mining and knowledge extraction; this knowledge help suppliers for better relationship with their customers (Zhang, 2007). Network Integration in e-business environment increase the flexibility and link the suppliers and customers electronically based on three basic components (Poirier & Bauer, 2000): e-network (for satisfying the customer demands through a seamless supply chain), responses (for integrating inter-enterprise solutions and responses and customer based supply chain strategy), and technology (for supporting the goals of the supply chain).

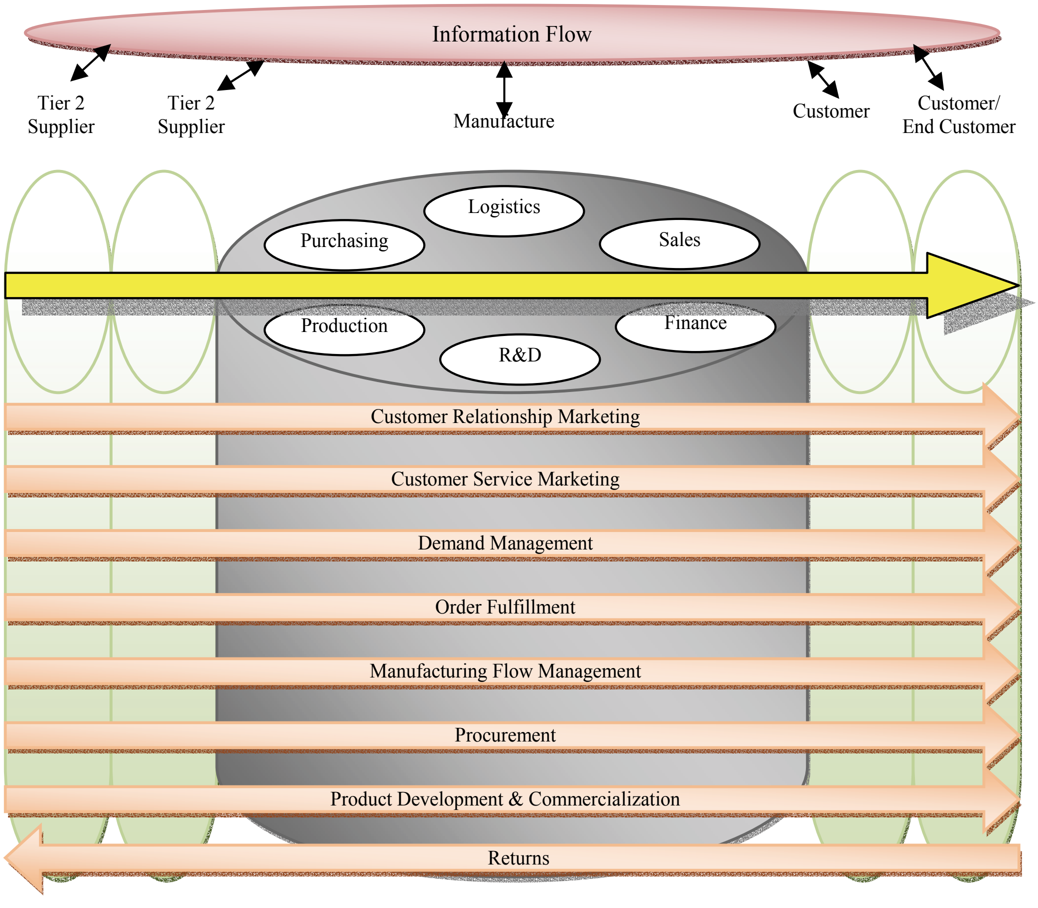


Fig. 1. Supply Chain Management (Lambert & Cooper, 1998)

As late description, in next section a main framework will be represented to illustrate the impact of IT on SCM.

5. Definition of agility

Agility is a business-wide capability that embraces organizational structures, information systems, logistics processes, and, in particular, mindsets. A key characteristic of an agile organization is flexibility.

Initially, it was thought that the route to manufacturing flexibility was through automation to enable rapid change (i.e., reduced set-up times) and, thus, a greater responsiveness to changes in product mix or volume. Later, this idea of manufacturing flexibility was extended into the wider business context (Powell & Dent-Micallef, 1997) and the concept of agility as an organizational orientation was born.

Agility should not be confused with leanness. Lean is about doing more with less. The term is often used in connection with lean manufacturing (Womack et al., 1990) to imply a "zero inventory" just-in-time approach. Paradoxically, many companies that have adopted lean manufacturing as a business practice are anything but agile in their supply chain. The car industry, in many ways, illustrates this conundrum. The origins of lean manufacturing can be traced to the Toyota Production System (TPS) (Ohno, 1988), with its focus on the reduction and elimination of waste.

Provided that reaction of supply chain increased for responding the real demands, the agility of SCM grows. Emersion of IT and its application in SCM cause to virtual SCM emerges which is more information-based than inventory-based. So, collaboration along buyers, suppliers, and the firm enhances the agility of SCM.

6. The framework of leveraging the agility of SCM by embedding it

The research revealed that the most impact of IT on SCM is on procurement, logistic, firm, vendor relationship management and CRM described in follows and illustrated in Fig. 2.

The final and perhaps most important prerequisite is the need for a high level of "connectivity" between the firm and its strategic suppliers and customers. This implies not just the exchange of information on demand and inventory levels, but multiple, collaborative working relationships across the organizations at all levels. It is increasingly common today for companies to create supplier development teams that are cross-functional and, as such, are intended to interface with the equivalent customer's management team within the supplying organization (Lewis, 1995). Through using of IT in the supplier and customer area of SCM, the agility of SCM could be leveraged (Fasanghari et al., 2007, Fasanghari et al., 2008).

6.1 IT & procurement

The use of the IT in managing purchasing in the supply chains has developed rapidly over the last 10 years. The research demonstrates that the IT is utilized in a variety of procurement applications including the communication with vendors, checking vendor price quotes, and making purchases from vendor catalogs. Vendor negotiation has also been streamlined through the use of the IT. Face-to-face negotiations are not used as frequently because the negotiations can conducted through the IT. This includes the bargaining, renegotiation, price, and term agreements (Olsen & Ellram, 1997). The receipt of queries from vendors, providing vendors with information, and the processing of returns and damaged goods were all handled by the IT.

The other more popular use of the IT in supply chains is in order processing applications. The most frequent use of the IT here is in order placement and order status. Over half of the firms use the IT for this purpose. This has dramatically reduced the costs of order processing. The use of the IT in order processing has reduced the error rate involved in order processing. Errors now can be detected more easily and corrected more quickly.

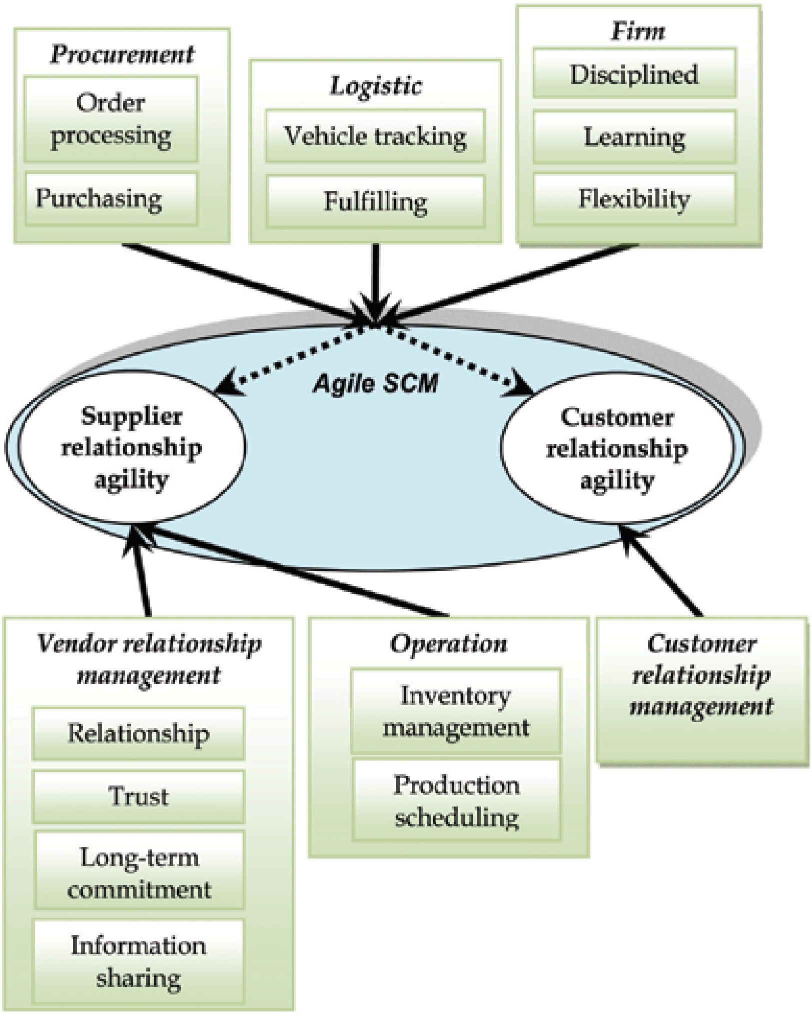


Fig. 2. Framework for impact of IT on SCM

6.2 IT & operation

1) One of the most costly aspects of supply chains is the management of inventory. The research has shown that the most popular use of the IT in this area is the communication of stock outs by customers to vendors, or the notification of stock outs by companies to their customers. The IT has enabled companies to more quickly institute EDI information programs with their customers. The IT has affected inventory management most dramatically in the ability of firms to be proactive in the management of inventory systems. This is demonstrated in the ability of firms to notify customers of order shipping delays and inventory emergencies, in order to decrease the delivery lead time and inventory.

2) Production scheduling has traditionally been the most difficult aspect of SCM. The IT has enabled firms to minimize the difficulty in their production scheduling by improving the communication between vendors, firms, and customers. The research showed that some of

the firms in the study use the IT to coordinate their JIT programs with vendors. In addition, some of the firms are beginning to use the IT to coordinate their production schedules with their vendors.

6.3 IT & firm

1) To keep costs down, an organization must have a high level of discipline based on the size of the firm: each person knows what needs to be done, knows how to do it, and does it quickly and efficiently. To do this requires a discipline of change which encourages innovation, and yet retains the stability of existing procedures until innovations are ready for wide-spread adoption. IT could overcome this problem.
2) The need for continued learning is acute in today's competitive environment. As new teams are formed, individuals must be able to learn rapidly what is needed to deal with a new set of issues. As new knowledge is developed, it must be made available to other members of the team and to individuals in other parts of the larger organization, that IT has the main impact on improving this process.
3) An organization must be "tight" at the same time that it is "loose". By light, we mean the need to have a lean, disciplined operation, in which there is a strong and ceaseless attention to keeping costs down and providing quality service at the same time. By loose, we mean the need to be innovative, to be responsive to customers' needs, to be flexible and adaptive to changing conditions and changing customer needs in each local situation. This flexibility is the other area that IT has critical impact on firm in the SCM.

6.4 IT & logistic

1) The research showed that the monitoring of pickups at regional distribution centers by carriers is the most popular application of the IT in this area. This is particularly important for a company, since tracking shipments to regional depots provides the firm with data on the reliability performance of the carriers it is using. This enables transportation managers to make sure that the motor carriers they use are meeting their promised arrival times.
2) In production and logistics, many parties are involved in coordinating all the processes that are involved in fulfilling a customer's order: manufacturer, suppliers of parts and subassemblies, material managers, logistics managers, transportation carriers, customer service representatives, quality assurance staffs, and others. The goals are to reduce the cycle time to fill a customer's order, reduce the inventory of parts, work in process, and finished goods in the pipeline, increase the accuracy and completeness of filling a customer's order and of billing him for it, and accelerate the payment for the delivered items to put cash in the bank as soon as possible. To achieve this degree of Order Cycle Integration, manufacturers, merchandisers, and their trading partners are using IT.

6.5 IT & customer relationships

Many management experts argue that, by focusing on total customer satisfaction, a company can improve its processes to deliver better service at a lower cost. Customer-satisfaction driven is often described as the next step beyond TQM, total quality management: the objective is not simply to deliver some abstract definition of quality, but to deliver total satisfaction to the customer, of which the delivery of quality is only a part.

Meanwhile, in the past, customer information could not be fully utilized in setting processes of firms' conditions. With recent increase in the speed of the IT, it has provided firms with the ability to offer their customers another way to contact the firm regarding service issues and integrate customer information and firm information to bring great benefits to both customer and firm. The research shows that some of the companies use the IT to receive customer complaints, while the other utilizes it for emergency notifications.

6.6 IT & vendor relationships

1) For IT in general, Auramo et al. (2005) propose that IT deployment in supply chains leads to closer buyer-supplier relationships. Stump & Sriram (1997) provide empirical evidence that the use of IT is associated with the overall closeness of buyer-supplier relationships. Subramani (2004) reports a positive relationship between an IT-based supply chain and organizational benefits. Lewis (1995) suggest that the decision to use IT within the dyad could encourage the commitment to establishing relational behavior. Their results show that IT decreases transaction costs between buyers and suppliers and creates a more relational /cooperative governance structure.

2) Trust plays a key role in any organizational relationship that IT facilitates it. Trust exists when a party believes that its partner is reliable and benevolent (Heikkilä, 2002). There has been a noticeable increase in the importance of trust in different forms of interorganizational relationships in management literature. The need for trust between partners has been identified as an essential element of buyer-supplier relationships.

3) Studies recognize long-term orientation commitment as a predictor for successful interorganizational relationships (Bensaou & Anderson, 1999). Long-term orientation refers to parties' willingness to exert effort in developing long-term relationships. Such willingness is frequently demonstrated by committing resources to the relationship, which may occur in the form of an organization's time, money, facilities, etc. Productivity gains in the supply chains are possible when firms are willing to make transaction or relation-specific investments, an important indication of commitment that was increased by IT.

4) Several studies suggest that successful buyer-supplier relationships are associated with high levels of information sharing. Information sharing (quality and quantity) refers to the extent to which critical and proprietary information is communicated to one's supply chain partner. IT caused to open and collaborative information sharing lead to positive effects on interfirm relationship.

7. Conclusion

In this article, at first was presented the definition of IT and SCM and afterward the impact of IT on SCM was illustrated in a framework. It is important that, the impact of IT on SCM is much larger as it facilitates inter-organizational communication and in turn reduces cycle times and develops collaborative work. IT provides opportunities for an organization to expand their markets worldwide. IT opens up the communication and enlarges the networking opportunities. IT supports seamless integration of partnering firms. This facilitates an increase in agility and a reduction in cost. Also, IT enhanced teamwork and CRM for designing new products and receiving feedback from customers and being proactive on responding to change market requirements. Considering the recent trend in IT,

more and more companies are attempting to use IT in producing and selling their products/services. Reduction of manual work and costs, improvement of information quality, speeding up of information transfer, and volume of transactions were found to be the drivers for the transaction processing role of IT in SCM.
One set of strategies for gaining competitive advantage is based on a simple principle: use IT to enhance the ways in which people work. To improve the communication between customers and suppliers, IT would be useful in exchanging the information about products and services. Many companies lack knowledge and skills about IT. This could be due to lack of understanding of the implications of IT and lack of fund for IT investment. These require education and training and also government support to facilitate easy access to the Internet service and development of web site for use of IT in SCM. As a main deduction, IT is a major source to enhance the competitive advantages of the SCM.
The implementation of IT in the SCM can enable a firm to develop and accumulate knowledge stores about its customers, suppliers, and market demands, which in turn influences firm performance. The key to survival in this changed condition is through agility in particular by the use of IT in the important segment of SCM. Moreover, the investment of IT for leveraging the agility of SCM can be optimized as the proposed framework for the affected dimensions of the SCM through IT support organizations for use of IT in SCM according to their goal and resources.

8. References

Armstrong, A. & Hagel, J. (1996). The real value of online communities. *Harvard Business Review*, 134-140

Auramo, J., Kauremaa, J. & Tanskanen, K. (2005). Benefits of IT in supply chain management - an explorative study of progressive companies. *International Journal of Physical Distribution & Logistics Management,* Vol. 35, No. 2, 82-100

Bakos, J. Y. & Brynjyoolfsson, E. (1993). From vendors to partners: Information technology and incomplete contracts in buyer-supplier relationships. *J. Org. Comput.,* Vol. 3, No. 3, 301-329

Bensaou, M. & Anderson, E. (1999). Buyer-supplier relations in industrial markets: When do buyers risk making idiosyncratic investments? *Org. Sci.,* Vol. 10, No. 4, 460-481

Bowersox, D. J. & Closs, D. J. (1996). *Logistical Management-The Integrated Supply Chain Process*, McGraw-Hill Companies, New York

Campbell, A. J. & Wilson, D. (1996). Managed Networks: Creating Strategic Advantage, *Proceeding of Networks in Marketing*, London, England, London, England

Carr, A. S. & Smeltzer, L. R. (2002). The relationship between information technology use and buyer-supplier relationships: An exploratory analysis of the buying firm's perspective. IEEE Trans. Eng. Manage., Vol. 49, No. 3, 293-304

Fasanghari, M., Mohammadi, S., Khodaei, M., Abdollahi, A. & Roudsari, F. H. (2007). A Conceptual Framework for Impact of Information Technology on Supply Chain Management, *Proceeding of Second International Conference on Convergence Information*

Technology (ICCIT '07), pp. 72-76, Gyeongju, South Korea, IEEE Computer Society, Gyeongju, South Korea

Fasanghari, M., Roudsari, F. H. & Chaharsooghi, S. K. (2008). Assessing the Impact of Information Technology on Supply Chain Management. *World Applied Sciences Journal,* Vol. 4, No. 1, 87-93

Grover, V., Teng, J. & Fiedler, K. (2002). Investigating the role of information technology in building buyer-supplier relationships. *J. Association Information Systems,* Vol. 3, 217-245

Heikkilä, J. (2002). From supply to demand chain management: Efficiency and customer satisfaction. *Journal of Operation Management,* Vol. 20, 747-767

Lambert, D. M. & Cooper, M. C. (1998). Supply Chain Management: Implementation issues and research opportunities. *The International Journal of Logistics Management,* Vol. 9, No. 2, 1-19

Lewis, J. (1995). *The Connected Corporation*, Free Press, New York

Lucas, J. H. C. & Spitler, V. K. (1999). Technology use and performance: A field study of broker workstations. *Decision Sciences,* Vol. 30, No. 2, 291- 311

Ohno, T. (1988). *The Toyota Production System Beyond Large Scale Production*, Productivity Press, Portland, Oregon

Olsen, R. F. & Ellram, L. M. (1997). A Portfolio Approach to Supplier Relationships. *Industrial Marketing Management,* Vol. 26, 101-113

Poirier, C. C. & Bauer, M. J. (2000). *E-supply chain: Using the Internet to revolutionize you business*, Berrett-Koehler, San Francisco

Powell, T. C. & Dent-Micallef, A. (1997). Information technology as competitive advantage: The role of human business, and technology resources. *Strategic Management Journal,* Vol. 18, No. 5, 375- 405

Radjou, N. (2003). U.S. manufacturers' supply chain mandate. *World Trade,* Vol. 16, No. 12, 42-46

Stevens, G. C. (1989). Integrating the Supply Chain. *International Journal of Physical Distribution & Materials Management,* Vol. 19, 3-8

Stump, R. L. & Sriram, V. (1997). Employing information technology in purchasing: Buyer-supplier relationships and size of the supplier base. *Industrial Marketing Management,* Vol. 26, No. 2, 127-136

Subramani, M. R. (2004). How do suppliers benefit from IT use in supply chain relationships. *MIS Quarterly,* Vol. 28, No. 1, 50-75

Tippins, M. J. & Sohi, R. S. (2003). IT competency and firm performance: Is organizational learning a missing link? *Strategic Management Journal,* Vol. 24, No. 8, 745- 761

Walton, S. & Gupta, N. D. (1999). Electronic data interchange for process change in an integrated supply chain. *International Journal of Operations & Production Management,* Vol. 19, No. 4, 372-388

Womack, J., Jones, D. & Roos, D. (1990). *The Machine that Changed the World*, Macmillan, New York

Yu, Z., Yan, H. & Cheng, T. C. E. (2001). Benefits of information sharing with supply chain partnerships. *Industrial Management & Data Systems,* Vol. 101, No. 3, 114-119

Zhang, Q. (2007). *E-Supply Chain Technologies and Management*, InformatIon scIence reference, Hershey

11

Development and Evolution of the Tiancalli Project

Macías Galindo Daniel, Vilariño Ayala Darnes and López y López Fabiola
Benemerita Universidad Autonoma de Puebla, Puebla, Mexico

1. Introduction

Every year since 2003, a group of universities and research centers compete against each other to prove mechanisms for supply chain situations, in a simulation known as The Trading Agent Competition: Supply Chain Management Game (TAC SCM). Several common situations such as customer and supplier care, and storage management are considered as important but should always be well-balanced in order to obtain the highest amount amongst other five agents in competence. With the purpose of participating on the TAC SCM, the Tiancalli Project was created since 2004. Some experiments were conducted in order to develop the first agent –Tiancalli 2005- which participated in the current SCM competition, and since then, evolutions and changes over the structure of the agent have been done to create even better agents for fast decision-making.

In this paper, we would like to present the course of this research, detailing it as much as possible. Thus a comparison between the three main versions of the Tiancalli agent, from 2005 to 2007 will be presented, in order to prove if the current agent *plays* better than the previous ones. In order to demonstrate the improvements obtained with every agent, an experiment with 100 games -25 for each version- has been designed and presented on the latter pages of this chapter. An analysis for each of the three agents is performed, by measuring the results obtained from each simulation.

It is a fact that Tiancalli 2005 is more *reactive* than the 2007 version, which is presented as an *intelligent* agent. However, this fact will be discussed and even proved on this paper. This analysis will also serve to suggest new implementations for the new version of the agent, which will be participating on the 2009 game –due to the 2008 competition has already taken place.

The metrics proposed and generated and the strategies and algorithms developed to make Tiancalli an agent capable of winning a higher amount of orders and a null generation of late deliveries are exposed. These both points are considered the most important on developing a successful SCM player agent.

2. The trading agent competition and the Tiancalli project

Supply chain is the part of management which is "…concerned with planning and coordinating bidding, production, sourcing and procurement activities across the multiple organizations involved in the delivery of one or more products (Arunachalam & Sadeh,

2004)". With the purpose of analyzing the behaviour of dynamic markets based on supply chains, the Trading Agent Research Group has developed the TAC SCM platform. On this simulation there are three main entities: Customers, Suppliers and Agents, the latter also own an assembly factory and a bank account.
Supply chains are the group of processes and activities required in order to transform raw materials to final products. On TAC SCM, a software agent should be developed by a participant institute. This agent should buy components, assemble them to create and sell computers. These activities are performed on a changing environment, against five other agents. The strategies implemented on each agent will allow it to gain customer orders and supplier orders. Finally, this will help the agent to get the highest amount through 220-days simulations.
The use of agent technology is encouraged because of the following features:

- Agents must be enabled to adapt themselves to their environment, and sometimes, with their actions, modify it.
- Agents must be proactive; this means that depending on the see-through capacity of the agent, they must be able to modify "intelligently" their behaviour before an expected situation arrives.
- Sometimes, agents must act as a "reaction" coming from the changes on the environment.
- Finally, the agents must be able to cooperate and collaborate to offer accurate answers, in other words, to work as a society.

2.1 The trading agent competition: supply chain management game description

The simulation consists on developing an autonomous agent capable of performing common activities on a supply chain process. The activities performed are, in general:

- To attend a group of customers interested on acquiring a product, for this game it will be computers. The agent must propose them an affordable price that should be more interesting than the offers made by the rivals.
- To request a group of suppliers to obtain the raw materials to assemble a computer. For this game, the components to produce a PC are: processor, motherboard, memory and hard disk. The agent should be able to request the materials and offer an interesting price to the producers –better than prices offered by the rivals-, in order to acquire components.
- To organize and plan the production and delivery of computers on time to avoid penalties for late deliveries or for storing too many components or computers.
- All the entities of the game are limited in resources: the suppliers have a limited production, and the agent owns a factory that can produce less than 500 computers per day. The participants are then competing against them to dominate the market.
- The market is not a static environment, so the agent must be able to adapt to the current conditions, for example: unavailability of components, preference over certain type of computer, and other conditions which may be changed by the other agents on the competition.

The TAC SCM scenario is formed by six agents capable of producing PCs, a group of customers and eight different suppliers, two for each type of component. The simulation longs 55 minutes that represent 220 TAC days when the agents have to make important decisions about their tasks, in order to obtain the highest incomes and become the winner.

The agent starts without money in its bank account. The common competition is to go from negative amount to positive during the whole game. Then, it is important to obtain components, sell them and avoid the penalties that can be applied for both late deliveries and component storage.

The components available are listed as follows: (a) Processors of 2 or 5 GHz, supplied by Pintel or IMD –so the processors are the only element that can be considered as four different; (b) Motherboard for Pintel or IMD processors, supplied by Basus and Macrostar; (c) Memory of 1 or 2 GB, supplied by MEC and Queenmax; and (d) Hard disks of 300 or 500 GB, supplied by Mintor and Watergate. The total combination gives 16 types of computers and 10 different components to combine.

This is a brief competition of the TAC SCM game. For more details about it, please check (Arunachalam et al, 2003).

As it is mentioned, many universities and research centres have tried to approach the problem from different perspectives. Some of the most remarked techniques include fuzzy logic, stochastic methods, regression trees and prediction techniques, all of them applied on Multi-Agent and Sub-Agent systems.

With the purpose of creating an agent that could participate on the TAC SCM tournament, and following a philosophy of gradual improvements by knowing the basics of the competition, the Project TIANCALLI was created on 2004. The project consisted on presenting an agent for the following SCM games, by a process of reading, comprehension and implementation of several learning techniques. Several experiments which allowed understanding the behaviour of the simulated markets were run before obtaining the first agent, Tiancalli 2005. There were other implementations on the following years -2006 and 2007-, and we postponed the presentation of the 2008 agent because of lack of time, for 2009.

With each year, the agent presented many improvements, especially on areas such as:

- The relationship with the customers.
- The relationship with the suppliers.
- The use of the full capacity of the factory during the whole simulation.
- The decisions taken in order to obtain more incomes than expenses.

In this chapter, the full notes of each of the performed Tiancalli agents are presented. It is organized as follows: On section 3, the nine non-so-called-Tiancalli agents are presented as a background of the first analysis made about the TAC SCM game. On section 4, the "Tiancalli 2005" agent and its features are discussed. Then the sections 5 and 6 present the background and the final implementations for the agents "Tiancalli06" and "Tiancalli07", where a subagent perspective was approached. On section 7, a mathematical comparison between the obtained results for these three agents is presented and discussed. Finally, section 8 and the conclusions present the issues not corrected on the previous versions of the agent in order to obtain the principles to construct the Tiancalli09 agent.

3. The agents before Tiancalli

In order to know how to approach to the SCM game, several experiments were conducted to know how the platform functioned. These agents served to know which strategy could work better with the found limitations of our first design, this is why they are mainly divided on two groups:

- Those agents which purchase components after the orders from the customers arrive, which are known as the "Loco_Avorazado" agents –which can be translated as "Mad and voracious". They were implemented by considering the features of the voracious algorithms. They can be easily described as agents who want to "have customers first, and then find how to satisfy them".

- Those agents which purchase components before the orders from the customers arrive. They are better known as the "Previsor" agents –which can be translated as "Foreseer". Their philosophy is based on the idea of "selling what we already have".

Both approaches served to know how the platform behaved. These agents used essentialy the same commands but allowed to understand how much does a supplier delayed on delivering the items to the agent, and how to control the storage and the assembly factory of the agent.

3.1 Loco_Avorazado agents

A. Loco_Avorazado agent. Two versions of these agents were constructed. The first version had implemented the following behaviour:

- The agent selects its customers by taking the due date of each request. This date must be of eight days after the current TAC day.
- The agent offers to customers at 95% of the customer suggested price.
- The agent requests its components after the offers to the customers have been made.
- The agent chooses its suppliers at random.

The results obtained for this agent are shown in Figure 1. Some observations were made and are described next.

- The agent takes 10% of total requests sent in a TAC day.
- The factory of the agent is used in a 24%.
- The customer chooses this agent in 96% of the offers made by it.
- The agent delivers an average of 80% of the total orders obtained.
- The incomes for this agent are between $4 and $8.2 M, and it always obtains the last place.

B. Loco_Avorazado_2 agent. The second version, named Loco_Avorazado_2, was implemented long after the first version. It implemented several features of the Previsor agents –which will be described next- and it served to discard or continue implementing these agents. Its behaviour can be described on the following quotes:

- Considering the delivery problem of the previous version, this agent chooses customers by taking the delivery date of the request of over 10 days after the current date.
- The agent offers to the customers a static price of 95% of their suggested price.
- The agent also makes requests of components in the same day that the offers for customers are sent.
- The agent chooses suppliers at random. It means that the preference concept is not implemented here.

After implementing the agent and obtaining results, some other facts obtained of the performance of this agent are discussed next.

- Loco_Avorazado_2 takes almost the 25% of total requests made daily by customers, and gains all these requests in a static price environment.
- The agent uses over 60% of factory duty cycles.
- The customers prefer the agent offers in 96% of the cases.
- The agent has incomes from $4 to $37 M and can reach first place in most of the simulations.

3.2 Previsor agents

A. Previsor agent. The agents that belong to the Previsor Series implement the strategy of "first buy, and then sell what was bought". The first version of this agent works as follows:

- The method for handling customer requests chooses them by defining a function value that considers the suggested price and the quantity of computers that the customer wants to purchase. This function is established on (1):

$$f = \frac{reservePrice + (quantity * 100)}{2} \tag{1}$$

Where:

- *reservePrice* is the price suggested by the customer (top price)
- *quantity* is the amount of computers requested.

All those request that exceed a *f* value of 1500 are considered to be attended by the agent.

- The assembly of computers is made when the offers are sent by the agent to the customer.
- The agent obtains its inventory a priori based on static daily amounts. Then it has to optimally assign the inventory and the free duty cycles of the factory. When the inventory and the free cycles are calculated for the current day, an optimization problem is organized and intended to be solved. This problem is formulated on (2).

$$\begin{aligned} &Maximize\ \ n_1x_1 + n_2x_2 + n_3x_3 + \ldots + n_jx_j \\ &adjusted\ to: \quad k_1x_1 + k_2x_2 + k_3x_3 + \ldots + k_jx_j \le 2000 \\ &\qquad\qquad\qquad c_1^1x_1 + c_2^1x_2 + c_3^1x_3 + \ldots + c_j^1x_j \le 100 \\ &\qquad\qquad\qquad c_1^2x_1 + c_2^2x_2 + c_3^2x_3 + \ldots + c_j^2x_j \le 100 \\ &\qquad\qquad\qquad \ldots \\ &\qquad\qquad\qquad c_1^ix_1 + c_2^ix_2 + c_3^ix_3 + \ldots + c_j^ix_j \le 100 \end{aligned} \tag{2}$$

Where:

- x_j can be binary {0,1}
- *j* is the amount of requests that satisfy the conditions to be considered
- *n* is the product of the quantity and the unit price of the request
- *k* is the product of the quantity and duty cycles
- c_j^i is the quantity of component *i* used in request *j*

This problem is best known in literature as the Multidimensional Knapsack Problem. It can be solved by using different types of algorithms. For this strategy, we solved the problem using a Monte Carlo Algorithm. This algorithm randomly selects the requests that could be satisfied with agent limitations.

With this mechanism as the essential core of this and the following Previsor agents, this version achieved the following results:

- The agent has a customer acceptance rate of 96%, mainly because the agent offered the 95% of the price suggested by the customer.
- Monte Carlo Algorithm is a selection method that is little effective for so many request in a very few seconds. Some days, the algorithm never returns a solution, so the agent cannot take any request. Then, the agent loses a lot of money because of the a priori purchasing system.

- There is an important stock of components and assembled computers at the end of the round, basically because the objective is to always have items on the store. Then a mechanism for managing the inventory is required
- The Previsor agent always makes 6th on any round.

B. Previsor_2 agent. This agent is an improvement of the Previsor agent. It was intended to correct the absence of response by using the Monte Carlo algorithm, by implementing a selective algorithm. Its functionalities are described next.

- The agent chooses its customers by using the same function than Previsor as the first filter to choose customers.
- The agent obtains its inventory by purchasing a constant amount of each component on each TAC day.
- Its customer requests selection works with a selective algorithm instead of the Monte Carlo algorithm of Previsor agent. In this case the agent orders upwards the requests by considering the f value of the first filter, and decides which ones could be satisfied with both the inventory and the free duty cycles limitations.
- The orders for assembling computers are made once the offers are sent to the customer.
- The results of one simulation with this agent are shown in Figure 3. Some remarks on this agent are shown next.
- The customer accepts 99% of the offers made by the agent. So, a feared trouble about the inventory is partially corrected on a static environment –because the Dummy agents do not change their prices.
- The factory of the agent has a remaining inventory –but quite lower than the Previsor agent stock.
- The incomes for this agent are from under zero to $14 M, and sometimes it gets from 3rd to last place. This lack of consistence in the results is not enough to win to a Dummy agent.

C. Previsor_3 agent. This agent became the first winner agent of this project, working over static environments with Dummy agents. Its features are enlisted next.

- The Previsor_3 agent uses the same function to handle customer orders. However, it implements a new component purchase systems that defines certain quotes for purchasing. These quotes are established as follows:
  - A Normal Quote is applied when there are between 500 and 2000 components on stock.
  - An Extended Quote is applied when there are less than 500 components in stock.
  - A Minimal Quote is applied when there are more than 2000 components in stock.

These quotes are applied for each TAC day, excepting the last 10 days when the agent applies only Minimal Quotes.

There is also an important rule about the quotes. The quotes for purchasing processors are always half the quotes for purchasing motherboards, memories and hard disks, because there are four types of processors and only the two types of the other three components. The objective of this rule is to have less stock at the end of the round.

- Computer production is made in the same way than every previous Previsor agent.
- The agent limits its components purchase. This is made in two ways: the agent offers computers until three days before the game ends. Also, the agent considers all the customer requests before the game ends. This is done in order to empty the inventory.
- The results of Previsor_3 agent in a game are shown in Figure 4. Other results are discussed next.

- The agent offers prices in two ways. They could be static at 90.7% or non-static in a random range between 75 and 90.7%. The agent always wins the customer orders, in a static pricing environment.
- There is a remaining inventory of components and computers, but lower than on Previsor_2 agent.
- The agent gets incomes from $29 to $40 M in the rounds. This performance makes that agent to get the first place in all simulations competing against Dummy agents.
- All orders received by the agent are delivered on time; there are very few penalties which do not interfere with the final result in all tests.

D. Previsor_4 agent. This was the latest Previsor implementation, which allowed starting with the development of the first "Tiancalli 2005" agent. Its features are presented next:

- Its behavior is similar to Previsor_3 agent except that this agent has methods to sell the assembled computers in stock to other customers; it means, in case that a customer does not accept the agent offer, the assigned computers for this sale are reassigned to other requests, and so until they are out of the stock. This was implemented after developing a prototype of Vendedor_Inmediato agent, a hybrid that will be discussed next. • The agent also offers non-static prices, depending on the behaviours of the other agents in the current game. As this was the latest agent of the platform research phase, it was intended to give this agent the chance to change the prices given to the customer. This was done by considering the acceptance and reject of the offers sent by the agent, and the price was moved from 75 to 95% of the customer suggested price.
- The results of the performance of the Previsor_4/4x agent are shown in Figure 6. Some other interesting facts of this agent are shown below.
- This is the first agent that implements a system of non-static prices which is not random –as on the hybrid agents shown next.
- Almost all deliveries are on time. There are very few penalties which do not affect the final result in all tests.
- There is an inventory remaining of components and computers. It is the smallest of any other implemented agent at this point of the work.
- Previsor_4 gets incomes from $19 to $30 M in test rounds against Dummy agents.
- This is the first agent that was tested to compete against the other ones implemented in this work, and always reaches first place.

3.3 Other agents

Some other hybrid agents were created, in order to test other possibilities for analyzing the platform. They are described below.

A. Vendedor_Inmediato agent. This agent implements the same function for the suggested price and computers quantity that Previsor series. Also it is a testing agent for computer inventory handling, which leads to the development of Previsor_4/4x agent (described previously).

This agent purchases the components a priori. When the components are delivered, it orders the factory a constant number of computers that must be assembled. Then the restrictions for this agent are not the components nor duty cycles, but the assembled computers currently in disposal.

This agent could have been the best agent in the phase, but due to the lack of time to develop it, it was left as a reference intend for the agents developed on the following years.

It was a good step to avoid the dependence on the factory cycles, but the main situation was that a higher penalty is applied for computers than for components. So in a real competition, to risk this situation is very difficult.

B. Previsor_Avorazado agent. It was a combination of both Previsor and Loco_Avorazado strategies. It takes the early purchase system implemented in Previsor agent and its filters to consider requests and assembling, and the ability to search for more incomes with the due date mechanism of the Loco_Avorazado agent.

- The agent chooses customers limited by components in store and duty cycles. It also chooses extra customers that their due date is over 10 days, both of them must previously satisfy function value.
- The offers made by the agent have static price about 90.7% of the suggested price for customers –this value was considered from previous agents.
- As the agent chooses the selected requests, an extra amount of no more than 10% of the total production is sent to the "long-term production". This means that the agent should ask for these extra components and try to assemble these computers on the following five days at least. However, as the factory is mostly used at full capacity, it was not clearly defined how to obtain this daily 10 percent required for extra orders. It was mostly obtained from rejected offers.

The performance of Previsor_Avorazado agent in a game is shown in Figure 9. Some other observations are made below.

- The agent gets all the orders from the customers, because of its static prices of 90.7%, which can not be improved by a Dummy agent.
- The factory uses over 60% of its daily duty cycles.
- Some orders which should wait for components are sent before the components are received. Then some orders are delivered late or never delivered, or the buy ratio is smaller.
- There are more outcomes than incomes. However, in some simulations, the agent gets first place and in the worst case, the 4th place.

4. The Tiancalli 2005 agent

The presented background is quite consistent to provide tools to develop the first version of the "Tiancalli 2005" agent. As it was explained before, the agent is based on the implementation of Previsor_4x agent, but it was modified to consider the market price changes which occur through the competition.

As a first effort to get a well-defined structure of the agent, it was partitioned on three submodules, which persisted over the design of the following versions of the agent. These modules are explained next.

A. Customer Selection System (CSS). The submodule included all the behaviors to choose and give a price to an expected customer. To choose a customer, the agent has two established filters. On the entire incoming request, first the agent has to consider that the customer must offer a higher amount than the production cost. This cost is calculated with the function seen on (1) where, as previously seen, *reservePrice* is the customer suggested price, and *quantity* is the amount of computers requested by the customer. The f parameter is a numerical value which was determined through several calculations as the average cost for any computer – not exactly the real one, however. Thus, if the request exceeded an f value of 1500, it is considered as winnable, which means that the agent may get an income for delivering this

order in time. Then the request passes to the second filter. All the approved requests are ordered from top to down considering the *reservePrice* offered.
The second filter is applied once the Inventory System has determined the current inventory. Then, the entire ordered requests are subject to the availability of components in stock. If there is enough stock to satisfy the complete order, either by:

- using already-assembled computers -and if that decision could not affect another previously made orders; or,
- using the stock components –which will require duty cycles from the factory.

Those requests that pass the second filter are then considered as *offers*, and the next step is to define the price for the customer. However, to use already-assembled computers, the agent must have considered this order *delayed*, in other words, that the customer is no more expecting the requested computers. In most of the cases, the agent only offers computers that still should be assembled.
As the offered price has to be at least equal or less to the customer suggested price, this price can be calculated in two different ways:
a) Mirror Pricing. When the *reservePrice* is higher than the base price defined by the agent during the simulation, Tiancalli uses the Mirror Pricing Strategy. It consists on applying the function (3) to the *reservePrice* defined previously, as shown.

$$off\,\mathrm{Pr}ice = \left(2 - \frac{reserve\,\mathrm{Pr}ice}{base\,\mathrm{Pr}ice}\right) * reserve\,\mathrm{Pr}ice \tag{3}$$

The basePrice applied is the static price of a computer by considering the initial price of the components required to assemble it. Commonly, the obtained *offPrice* will give a lower price than the customer suggested price. Otherwise, then the next strategy is used.
b) Factor Discount. In the special cases where the *reservePrice* is smaller than the base price, the Mirror Pricing strategy could give a higher price than the customer price. Then another strategy should be applied. A factor discount for each type of computer –which is stored in the Inventory System- is taken and applied to the price obtained with the previous strategy. It is certain that always the price will be lower than the original customer price, however if it fails, the Inventory System can apply another discount –this will be discussed on next. The function used then in this strategy is represented on (4):

$$offPrice = suggestedPrice * discountFactor \tag{4}$$

This discount factor is modified for each computer type at the end of each TAC day, using the following algorithm:

```
For each computer i[1...16] and this i computer discountFactor
df[0...1]
              Calculate Factor = orders received / offers sent
      If Factor value goes:
                   under 30%: lower df by 5%
           between 30 and 90%: lower df by 1%
           over 90%: wait for 5 days with the same df value; if the
           condition is maintained, increase df by 1%.
Next i
```

The orders received come from the offers made on the previous TAC day –so the offers must be saved on the agent IS. If just a few orders –or none- have been received with this discount factor, it means the offered price is too high, so it must be highly decreased. If most of the offers become orders, a trigger is thrown in order to expect good results with this price and then, if this condition is maintained on the following four days, the discount factor is increased.

There is an additional strategy that is normally applied after the TAC day 200. It is applied to decrease the amount of components in stock. However, this strategy can be applied if the agent has not gained any order on five straight TAC days. With this strategy, the agent fixes the discount factor under 50%, as an emergency method to gain customers.

B. Supplier Purchase System (SPS). The platform had a bug which allowed agents to make big purchases at the zero day of the competition –the configuration day- so as it was unknown to us once the competition started, a very primitive strategy was implemented, in order to get components for five non-consecutive TAC days. After the "zero day", and through the first ten TAC days, Tiancalli tries to purchase components for making an initial stock. So the agent does not offer computers until TAC day 11. Then the purchase of components becomes somehow difficult and limited because of the limited production due to the zero day bug.

To calculate an accurate price for the supplier, a similar idea to the customer pricing was implemented. In this case, instead of a percentage, a binary parameter was used –if the request became an offer from the supplier or not. On the first case, a trigger to wait if the price is accepted on the following three TAC days is thrown; then the offered price may go down a percent unit. On the latter case, the factor for the component must be increased 2% and expect the next day result.

As it is shown, the agent only makes one request per day for each component. This was procured in order to maintain certain organization with the requests; however it is insufficient to reach the half of the game with many components in stock to keep production levels high.

The SPS communicates with the IS as it holds the component prices and the pricing factor committed previously. Depending on the level of stock maintained for each component, the agent may request a different amount of components, which is explained next, as the quota system was explained before in the paper:

- A Normal Quote requires 100 components.
- An Extended Quote requires 250 components.
- A Minimal Quote requires 10 components –this quote is commonly used to poll the market.

The proposed quota must be modified as each TAC day passes, for example: a Normal Quote applies for 500 to 2000 components on the early competition days. However, the quote is reduced on 10 to 100 as the simulation comes close to the end –approx. on day 200.

C. Inventory System (IS). The IS works more as a database than as a system, because it stores most of the factors, prices and numerical considerations applied on the performance of the agent. This system offers limits for prices –for example, no discount factor for computers may go under 40%, however it is a possible situation for any game.

The IS is modified by the methods applied within both CSS and SPS, considering the customer acceptance ratio –the amount of orders against the one of offers sent by the agent-, and the supplier acceptance value –if it was accepted or rejected.

First the Tiancalli agent was able to choose a discount factor of just one value for all the computers and another for all the components. It was necessary to implement an array of sixteen values for each computer and ten for each component. These arrays -which represent discount factor for each component and computer-, were used to give the agent an individual perspective of each market element during the simulation.

D. Issues related to the performance of Tiancalli 2005. The Tiancalli 2005 agent had a desirable performance, by achieving the mid-table during most of the competition and reaching the Quarter Finals. The main issues found with this agent are enlisted next:

- The final implementation of the agent was created during the competition, so many modules and methods which were programmed on the agent do the same.
- The methods for modifying the discount factors are very restrictive and are commonly modified during a competition. The agent goes from 95% to 40% and resets to 95% in no more than twenty days, so the prices do not remain static in most of the time.
- The agent does not take automatically any learning from each game, as the values are reset on each simulation, and if some learning is acquired, it is applied by the programmer and not the agent itself.
- The simulation was not completely understood, and several facts, such as the factory usage, the punctual deliveries and the disposition of components, did not work properly, delivering several penalties to the final results of the agent in each game.

These are some of the issues that are expected to be corrected on the following version for 2006. However, for a better comprehension and review of the previous agents and the performance of Tiancalli 2005, we encourage you to read (Macías et al, 2005).

5. Development of the agent Tiancalli06

Tiancalli06 is a software agent that develops real-time techniques such as case-event learning. This allows the agent to improve to the Tiancalli 2005 agent on the following tasks:

- It has an improved mechanism for managing the component purchase, in order to become the preferred customer to the suppliers. The agent offers a better price based on the current situation of the game, in order to dispose of components during the whole game.
- The price selector for customers is improved by considering several other factors such as the current day and acceptance in the market for each computer.
- The factory is better organized and the agent intends to keep the lowest stock possible to satisfy orders on the following days.

This agent is organized on a better architecture which has the same three modules –defined from now on as sub-systems- of the previous agent, but well defined and separated. Then, if a modification is required, the programmer may direct easily to the required class and update it without affecting the normal performance of the agent. There are other classes included on the architecture of the agent that keep the game information inside the agent, and allow displaying the relevant information of a game on output files. The architecture is displayed on Figure 1.

A. The Customer Selection System version 2(CSSv2). For the analysis of the customer requests, the agent considers and makes the following activities:

- At the beginning of the game, the agent applies an initial discount factor for each type of computer –this discount factor works as in the previous version.

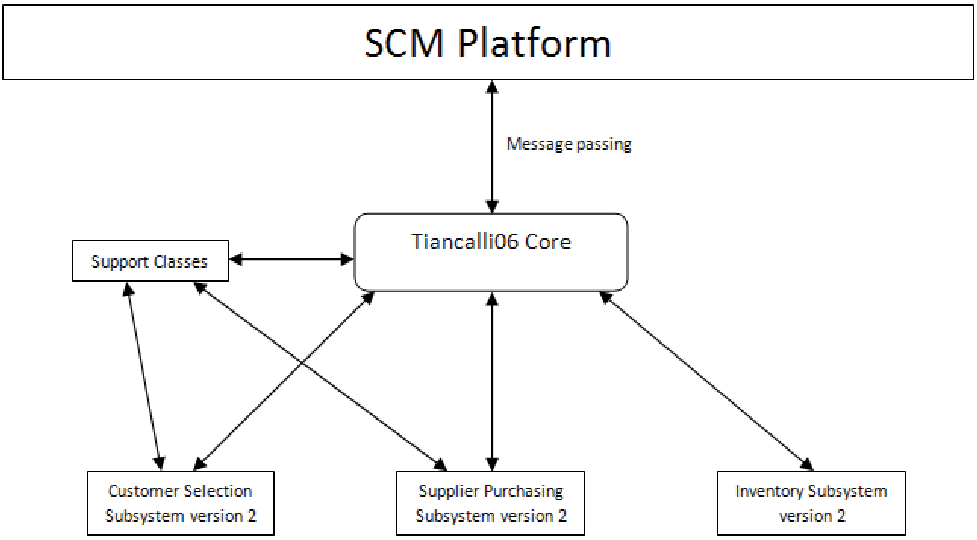


Fig. 1. The architecture of the agent Tiancalli06.

- From the TAC day 11 and until day 216, the agent analyzes the customer requests –the final day was decided because this day is the latest when the agent could delivery an order because of the process that will be explained later.
- The analysis of requests is done considering the available inventory for the next TAC day and if the request may produce higher incomes. The latter is determined by comparing the offered price with a calculated average price, which is described on function 5.

$$csp \times dsf_{pcType} \geq mp_{pcType} \tag{5}$$

Where *csp* is the price that the customer is willing to pay; *dsf* is the discount that the agent offers to the customer for this PC type (*pcType* goes from 1 to 16); and *mp* is the lowest price for a computer determined by the agent. This value is calculated by adding the average price of the four components that form any computer. Hence the first filter considered for the previous agent was changed for this process on this version. The entire group of requests that satisfy this requirement will pass to the following filter.

- The second filter works as on the previous version, any request for there is either enough both components and duty cycles, or enough free computers in stock, will be considered as offers. The determined price for these offers will be these which were calculated on the first filter –then, to save time the agent saves the price when the first filter is passed so when the offer is sent, it is sent with the stored price.
- The discount factor can never be more than 100% -in fact, it is never over 95% for most of the games-, or less than 42%. This is done in order to offer an optimal price to the customer that guarantees an even little income.
- The mechanism for changing the discount factor was slightly changed from the previous version. The calculation is the same; however more situations to move the prices are different, with this a better movement of the prices is allowed, and the agent

may even try with high prices for the customer. Then the function to determine the discount movement is shown next.

```
factor = (orders from the current day) / (offers from the previous day)
If Factor is:
- Zero, then to lower the discount by 10%
- under 20%, then to lower the discount by 2%
- between 20 and 50%, then lower the discount by 1%
- between 75 and 90%, then keep the discount
- over 90%, then increase the discount by 1%.
```

- The complete routine to deliver a PC to a customer consists on three TAC days, as following:

DAY d: The agent receives the request and sends an offer. Then the agent plans the factory usage for the following day –if required, only if the computers are not still assembled. This behaviour is approximated to the one implemented on the Previsor agents; however the computers are assembled before the agent receives any response from the customer.

DAY d+1: The agent receives the orders from the customers; if not, the rejected computers will be, on the following day, disposed for other requests. If the orders are received, the agent plans the delivery of the computers for the following day –because the computers are being assembled on the current day.

DAY d+2: The request is satisfied, and the agent expects the incomes on the following days –for the agreed date.

With this new structure of the customer attention, the agent is pretended to deliver on time all of its orders. It is expected that the agent increases its acceptance amongst the other agents, by evaluating simpler parameters for giving prices to the requests received.

B. Supplier Purchase System version 2 (SPSv2). For the interaction with the suppliers, the second version of this agent implemented the following activities:

- An improved purchase for the zero day. With this, the agent gets approximately the 35% of the total purchase of the game on the first day. The agent requests the components on determined days which are considered as crucial. The days are 9, 25, 50 and 100.
- The pricing for the suppliers is proposed on the same way that the previous version of the agent, with a specific discount factor for each type of component –not so for each seller. Each TAC day, the agent requests no more than 100 components to each supplier.
- The price is influenced by the acceptance or rejection of the price of the previous days, and each day that the price is accepted, the system receives and stores this price and calculates an average with the previous accepted prices, in order to propose the computer base price that is used on the CSSv2 system. Then the discount factor movement is determined with the following pseudo code:

```
If request is accepted
  then decrease discountFactor by 1%
Else
  increase discountFactor by 1.5%
```

- The system pretends to offer an initial purchase volume of approximately 70,000 computers per game –this was a calculation employed to determine how many computers an agent can assemble using the total factory cycles and components.
- Obviously this quantity can be affected by the market preference –maybe one type of computer is more purchased than the other-, so the system must be able to determine when a computer is preferred, and intend to offer more of these computers. Then, the system fixes at the start of a game, an initial quote of 70,000 components. When this quantity is almost to be reached and there are more TAC days following, this quote is modified as follows:

```
for id = 1 to 16, do:
  if (initialQuote (componentid) * 0.9) • soldComponents(id))
    then initialQuote(componentid) is increased by 500 components
```

- It must be noticed that the count that matters is the quantity of components sold and not those purchased. This feature is intended to decrease the amount of stored components at the end of the game.

C. Inventory System version 2 (ISv2). On the new version of the IS, a more precise calculation for stored components and computers is obtained daily. Before deciding if most components must be purchased, the real quantity available of components available in stock is determined with the function (6).

$$realInventoryForNextDay = componentAcquiredYesterday - componentSoldYesterday \qquad (6)$$

This formula marks a significant difference against the agent of 2005, because it only considered raw components that were purchased and not those which are still waiting to be assembled. Also, this version considers computers in stock that have not been sold. The agent has on disposal such components –for they need duty cycles to become computers- and computers to be offered to the clients. However, this system is centralized and offers the inventory and the calculations to perform the operations of the other two subsystems.

D. The performance of Tiancalli06. The main issue that Tiancalli06 deals with is the price estimation. The changing conditions of the market and the changes that several agents try to impose on the competition are the most difficult challenges for a very sensitive mechanism of pricing. However, the performance of the agent was acceptable, especially on the Seeding phase of the competition.

The agent achieved a constant acquisition of orders through most of the games; however the agent had difficulties in most of the games to get orders due to this pricing mechanism. Anyway, the most important achievement of the agent was the avoidance of late deliveries. The agent deliveries all of its orders in time, and just one game had penalties because of a connection issue at the end of the game. Finally, the storing prices were reduced because of the new quotes mechanism for purchasing components that were described on the SPSv2.

For second year, the agent lasted until Quarter Finals and achieved a final 18th place. This year, the competition presented more agents and the improved versions of the previous participants.

Although the advances were significant, the next version of the agent should offer better pricing systems. Some of the other agents implemented fuzzy logic or prediction heuristics to determine the prices. Then the next effort for Tiancalli07 should include an intelligent mechanism for pricing. The results obtained with the agent can be consulted on (Macías et al, 2006, 2) and an interesting comparison about the results of both agents Tiancalli 2005 and Tiancalli06 can be found on (Macías et al, 2006, 1).

6. Development of the Tiancalli07 agent

In the effort of improving the performance on the activities that the agent does on the competition, the agent Tiancalli07 was developed. This agent features new prediction techniques to give prices to both customers and suppliers. The obtained prices are stored in external files, which are updated during the current game, and used for the upcoming games. Since the beginning of the participation of the Tiancalli agents, any of the developed agents took experiences from the previous games and stored the information automatically, so it is the first agent which can be considered as "evolutionary". As it is the last agent developed by the same team project, it should offer a brand new and detailed architecture, conformed by three subagents –as defined on (He et al, 2006)- with specific functionalities, and two extra modules for information control. The general architecture is presented on the following figure.

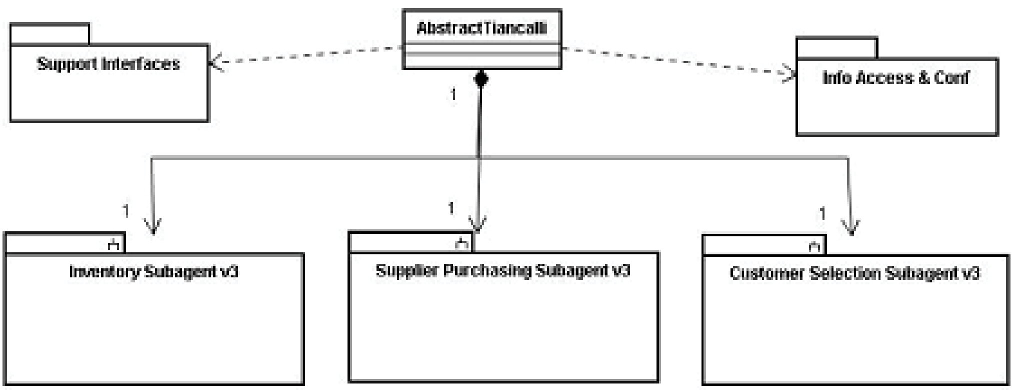


Fig. 2. Architecture of the Tiancalli07 agent.

A. Customer Selection Subagent version 3 (CSSv3). The general problem of attending the customers is subdivided in several areas. This allows the recognition of the main activities in order to correct any significant behavior. Some of these areas are described with their current developed activities:

- Customer selection. This is the least modified area of the general problem. It uses both filters applied since the CSSv2 on Tiancalli06. Just as a reminder, the first filter checks all those customer requests for which their customer suggested price –multiplied by the current discount ratio- satisfy the base price for assembling the current computer type. The second filter checks the availability of both components and duty cycles, or free computers that remain in stock. Those orders that satisfy both filters are considered as reasonable and the subagent then intends to give them a good price.
- Customer pricing. This are, on the contrary, is the most modified of the agent. A regression tree has been designed to have a better statistic of pricing movement. The tree is controlled by three main variables, which are described as follows:
    - Requested quantity of computers. It is required to order each type of computer, from top to down, as all the requests have arrived to the agent during the past

days. For the classification, fuzzy values of maximum, high, medium and low are applied to determine this variable.

- Price ratio. This is determined by considering three prices: pmin which is the minimal price that offered the agent on the previous day and became order; pens that is the base price to assemble the requested type of computer; and psug that is the suggested price by the customer. However only three cases are considered – because the others are illogical:

$$p_{sug} > p_{ens} > p_{\min} \rightarrow case\ 1$$
$$p_{ens} > p_{sug} > p_{\min} \rightarrow case\ 2$$
$$p_{ens} > p_{\min} > p_{sug} \rightarrow case\ 3$$

- Range of days. It considers the day when the computer is requested. The whole competition is divided then in subintervals of ten days, which give a total of 22 different ranges.

Once the concepts for pricing are presented, the structure of the trees is suggested. Sixteen trees, which represent each type of computer, were built to represent the required structure. Each tree has four branches for the requested quantity of computers, then three sub-branches for price ratio, and finally 22 leaves for each sub-branch. This leaves a total of 243 leaves, which include each a range of values –maximum and minimum- that is used to determine the customer price. The leaves also include the acceptance ratio of each leaf during the current game.

The acceptance ratio modifies the range of values as following:

- If the prices are accepted, the tendency is to close the intervals in order to find a convergence on the prices. If the optimal is found, the optimal should tend to increase its value in order to search for more incomes.
- If the prices are rejected, the tendency is to open the intervals, in order to reduce the minimal price and improve the prices against the other agents in the competition.

The trees are updated with each TAC game and stored on external files. These files are loaded at the zero day.

B. Supplier Purchasing Subagent version 3 (SPSv3). The subagent is in charge of performing the following activities:

- Price calculations. In order to calculate the base price that is considered on the first filter of the CSSv3, this subagent must determine a price for each component. This is done by applying the following pseudo code:

```
Variables:
Input: p as the offered Price, q as the quantity offered
Intermediate: qPurchased as the previously purchased amount of
components, avgP as the previous average price, temp as a temporary
variable.

AT THE BEGINNING OF THE SIMULATION:
qPurchased ←q, avgP ← p
EVERY DAY WHEN THE AGENT RECEIVES A SUPPLIER OFFER:
temp ← qPurchased + q
```

$$avgp \leftarrow \frac{qPurchased \times p}{temp} + \frac{q \times p}{temp}$$

```
qPurchased ←temp
```

This calculation generates a base price for each component, which once it is added to the remaining components, can be applied for the base price of each computer. This price is better approached than the previous versions, because the price is determined by the amount of computers which have been already acquired.

- Component purchasing. The subagent recognizes two ways to purchase components: one at the beginning of the game with the zero day purchase, and the other during the competition. The first purchase is intended to get an initial stock for the first orders; the latter purchase is to maintain this stock. As the zero day purchase was corrected by the game developers –you can still buy components but the prices are higher and consider a general statistic for all the requester agents- the amount of components acquired on these days has been reduced; hence the strategy must be improved to acquire the items during the whole competition.

The pricing system implements a single list for each component –ten lists in total- with 22 spaces. These spaces should be filled with the maximum and minimum ranges of discount factor during each 10 TAC days. As it can be seen the structure is similar but simpler than the trees implemented for the customer pricing. The ranges are modified by following these rules:

- If the supplier accepts the discount, the ranges are closed, tending to find an optimal. Once an optimal value is found, decrease this value to find a minimum value.
- If the supplier rejects the price, the ranges are opened by increasing the maximum price, in order to improve the prices and make them competitive.

C. Inventory Subagent version 3 (ISv3). The ISv3 is in charge of organizing production and delivery of computers to the customer. Its main goals are the following:

- Avoid the loss of orders by production delays, and reduce at most the delayed deliveries, if possible nullify them.
- Bring an accurate use of the factory and stock of components and computers, by delivering real statistics of the current situation to the other subagents. Sometimes this statistic must foresee the availability of components that will arrive on the following days.

D. Support packages. The whole operational environment of the Tiancalli07 agent includes two support packages that will be described next.

- Information Access and Configuration Package (IA&CP). This package is in charge of handling the most important data about the current simulation, in order to bring these data fast to any requiring entity. Also several methods for reading and writing external files are included to save the knowledge generated during the simulation. Finally, a configuration file is included, which commands the whole system the files that will be used to generate the initial trees and lists for the current simulation. The configuration file and the files for the storage of the structures are overwritten or substituted at the end of a simulation.
- Support Interfaces Package (SIP). The current package is only used for developing purposes. It implements classes for graphic interfaces to manage the information generated during a game. The programmers can so recognize the behavior of the current simulation with more detail than with the sole Agentware interface –the

original SCM interface included with the downloadable test agent. The developed interfaces are to display information such as: (a) maximal and minimal customer acceptance prices; (b) amount of requests against amount of orders, and (c) demand representation for the customers.

E. Implementation of the knowledge repositories. The files for configuring the agent are stored on a subfolder named "playbooks". The first read configuration file is named "init.tcf" and includes, mainly, the names of the files that include both the trees –p####.cus, where # is a serial number- and the lists –s####.sup. The serial number is updated once a simulation is finished.

The general structure of a customer tree file is the following:

```
p0003    //file name
16    //number of trees included
4    //possible values for the first parameter of the tree
3    // possible values for the second parameter of the tree
22    // possible values for the third parameter of the tree
a0a 90000 100000 10  19
a0b 90000 100000 15  16
a0c 90000 100000 24  30
...
```

The first five lines of the file are explained in the structure itself. The following lines are the ones that represent the tree structure as following: the first letter (a) could have three other values (b, c or d) and represents the demand of the current computer type –where a is *maximum* and d is *low*, as explained before-; then a number (0) can take other two values (1 or 2) and represents the price relationship; and the last letter (a) represents the segment of time on the competition where the computer is requested –it may have 22 values from a to v, where a is the segment from day 1 to day 10 and v the segment from day 211 to 220.

The following two quantities -90,000 and 100,000- represent the percentage applied for the discount factor. They are not decimal values because of the expense to store a *double* value, and the capacity to represent exact quantities which can not be represented with a wide double variable. This is explained as follows:

Considering the minimum amount that can be applied to the percentage discount so that you can discount one unit of money, the obtained number is 0.00012. So if for example, the range is between 90,000 and 90,012, the price may vary in just one unit, maybe the price will be 1500 or 1501 with this range. The value of 0.00012 is considered as the less significant number in the price calculation. This is why an integer –formed with the double value and multiplied by one hundred thousand- is required to store this factor.

The last couple of numbers represent the ratio of orders against the offers proposed. So in the first element of the tree –a0a- it can be observed that this leaf has generated 10 orders of 19 offers to the customer. Then the efficiency of the leaf can be also obtained.

This structure of the file and the tree allows the system to obtain easily the leaves, so the time required to seek the leaf is minimum. But the structure to store the prices for suppliers is not so different. The list is stored on a file named "s###.sup". Here an example is presented:

```
s0003
1
16
22
```

```
aa 74450 74453 11  13
ab 69550 72864 8  10
```

The file stores the name of it, which is s0003. It will propose the construction of only one list of 16 x 22 elements (the sixteen components divided by supplier and the 22 ranges of TAC days, as explained previously. The list is represented now with two elements –aa, ab- which represent, respectively, the component type and the range of days.

F. The performance of the Tiancalli07 agent.

The results obtained with this agent were compared with the previous two agents by separate on (Macías et al, 2007). However, the results in the competition were not as good as expected: the agent achieved again the 17th place. The comments about this result and several remarks on what issues should be corrected on the next version of the agent are discussed on the conclusions of this paper.

7. Experiment

In the previous lines, all the Tiancalli agents that have been developed until now have been described with detail. As it has been discussed, the intention of building an intelligent agent that can participate on the TAC SCM game has been reached with Tiancalli07. How the intended mechanisms worked in order to obtain this intelligent behavior, is one of the goals of the current paper. To prove the thesis that Tiancalli07 works in a more intelligent way than the two previous versions of reactive agents, the following experiment has been developed.

Over a local computer with the SCM platform, the three agents have been installed and configured to play one against the others on fifty games. The other three agents participating are *dummy* agents, in order to not alter the basic results –it has been tested that when many agents are running on a local platform, the messages are not well received by the agents.

The average results are presented on Table 1.

The following lines are a discussion of the obtained results on the experiment.

- The agent that receives more money from its customers is Tiancalli 2005, closely followed by Tiancalli07. However, Tiancalli 2005 is the agent which expenses more money to assemble its computers. This can be explained as follows: the agent of 2007 has improved its mechanisms to purchase components with lower prices, and sell computers to higher prices. This difference can be explained by consulting the parameter “Ratio S/P” that is intended as the amount of incomes divided against the amount of expenses, or in single words, the benefit that the agent receives for each $1 that it invests. Curiously, this is the only important parameter where Tiancalli06 is over Tiancalli 2005.
- The agent Tiancalli07 remains more time on the competition with a positive balance. This is also reflected on the final result of the competition, where the same agent also takes the lead.
- The amount for storing components and computers is the highest on the agent of 2007. It may be explained by considering that the agent risks more the prices to the customers, and many more computers remain in the current inventory. Also a huge problem for this agent is that when purchasing components, the market tends to offer more components of one type instead of another –for example, there are more processors than motherboards. The absence of one type of component will lead to have a big inventory during days –the days that will take to the agent to get the absent items-, and so the storage costs will be higher.

Agent	Tiancalli 2005	Tiancalli06	Tiancalli07
Revenue	91,535,200.85	33,321,284.62	85,679,710.08
Interest	142,610.15	82,597.46	342,191.54
Material	78,282,031.77	25,593,433.46	56,140,179.23
Storage	1,181,238.23	647,656.08	1,841,481.00
Penalty	288,055.54	0.00	0.00
Penalty%	0.4%	0	0
Ratio S/P	115%	127%	148%
Result	11,926,485	7,162,793	28,040,241
Orders	3,976	1,833	5,122
Utilization	73%	24%	56%
Delayed	27	0	0
Missed	8	0	0
Del. Performance	99%	100%	100%

Table 1. Results of the experiment of fifty simulations between Tiancalli and *dummy* agents.

- The only agent that receives penalties for late or missed deliveries is the 2005 agent. As it is shown, the amount of penalties represents less than 1% of the total outcomes that the agent makes during the game. It is not an important quantity, however, while it is kept as low or null as possible.
- The difference on the final result of an average game of Tiancalli07 is twice and a bit more of the result of Tiancalli 2005. The mechanisms implemented on Tiancalli06 allow it to obtain better prices than the agent of 2005; however the amount of components acquired is not enough to produce computers.
- The agent of 2007 obtains more orders, but uses the factory less than Tiancalli 2005. This is possible by consulting the formula that Tiancalli 2005 uses to determine the potential customers; it tends to accept customers which require a lot of computers –but also with a high amount of money to be paid. Then the agent will require assembling more computers. The case with Tiancalli07 is that the agent takes all the possible requests which may generate an income, so the quantity requested is not so important; then the agent may satisfy more orders without caring the requested quantity of computers.
- It may be discussed that "Tiancalli06 was an evolution over Tiancalli 2005" due to the poor obtained results. However, an important argument to defend this hypothesis is the ratio of sells against purchases, where it is clearly noticeable that Tiancalli06 obtains more money per invested unit than the agent from 2005. The mechanisms for purchasing components were not substantially different on Tiancalli06, but it is possible that maybe installing SPSv3 on Tiancalli06 the results obtained would be better.
- On the experiment, it is important to notice that in most of the simulations, Tiancalli07 obtained the first place; however, just in four simulations, the agent descended to second or even fourth place. This argument can be justified with the slow learning curve that the agent performs for new prices and conditions of the market. Tiancalli 2005 obtained only one second, two fourth places and fifth place in the simulations, and Tiancalli06 always obtained sixth place.

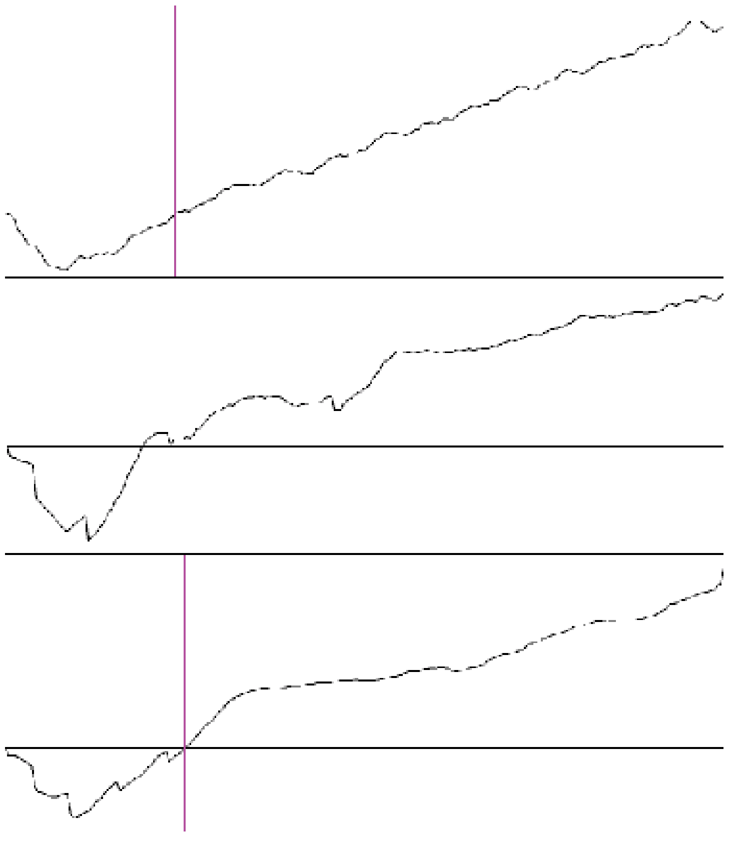


Fig. 3. The balance of the bank account of Tiancalli 2005 (top graph), Tiancalli06 (middle graph) and Tiancalli07 (lower graph). The vertical lines show the moment when the agent has a positive balance.

- On Figure 3, an example of a simulation is shown, and it displays the balance during the simulation time. The vertical line shows the moment when the agent passes the negative balance, and surprisingly, the agent that achieves this first is Tiancalli06, on day 42. Both Tiancalli 2005 and Tiancalli07 obtain the positive balance over 50 days of competition. However, in the graph can be seen that the agent from 2007 has a more constant line than the other agents, and Tiancalli06 has the most erratic line, due to the lack of components during most of the competition.

8. Conclusions and future work

This paper presents the deepest analysis about the construction of the Tiancalli agents since 2005. It intends to describe all the effort that has been developed and both the evolution and experience acquired during the three years participating on this competition, the TAC SCM.
It is remarkable that the agent concepts have evolved from Tiancalli 2005 to Tiancalli07. Certainly, the results achieved with the latest version of the agent are far from the previous results; however during the last competition these results could not be reflected on the reached place of the competition. Several considerations and facts that are planned for the following versions of the Tiancalli agent are the following:

- The proposed learning curve for the learning structures –but specially the tree- is too slow, and many more simulations must be done in order to obtain an agent with enough knowledge. If new conditions are added and the structure changes, it only affects to the leaves that participated on the price proposal, the others remain static. So, in order to promote the learning capabilities of the agent, several modifications about the knowledge update must be conducted.
- There is a lot of work done on the pricing system for customers; however the prices for suppliers are the most common issue and the penalties for storing too many useless components –referring to "useless" because they can not be used to assemble any computer due to absence of another component- are still high. There must be an improvement on the IS structure that intends to get the missing components as soon as possible in order to avoid important penalties.
- There must be an incentive to produce more computers on the factory. It is used the half of the total capacity.
- Finally, these improvements should be noticed on the competition, reaching even better results.
- There is still a lot of work to do about the TAC SCM. It is expected that this work serves to promote and invite researchers and universities to participate on the competition. To review more information and results of the Trading Agent Competition and all the contests, it is suggested to check the website: http://www.sics.se/tac

9. References

Arunachalam, R., & Sadeh, N., (2004). The supply chain trading agent competition. *Electronic Commerce Research and Applications 4, 2005*. pp. 63–81. Revised and extended version of paper at *AAMAS-04 Workshop on Trading Agent Design and Analysis*, New York, 2004.

Arunachalam, R., et al (2003), Design of the Supply Chain Trading Competition, *IJCAI-03 Workshop on "Trading Agent Design and Analysis"*, Mexico, August 2003.

He, M., et al, (2006). Designing a successful trading agent for supply chain management. In: *Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-06)*, pp. 1159–1166, Hakodate, Japan, 2006.

Macías, D., et al (2005). Tiancalli: An Agent for the Supply Chain Management Game 2005. *La Computación en Puebla en el Siglo XXI*, pp. 11-16, ISBN, Puebla, México, November 2005. BUAP FCC, Puebla.

Macías, D., et al (2006,1) Statistic Analysis for the Tiancalli Agents on TAC SCM 2005 and 2006. *In Proceedings of the 15th International Conference on Computing (CIC'2006)*, pp. 161-166, Mexico City, Mexico, November 2006, ISBN, IEEE Mexico,

Macías, D., et al (2006,2). Tiancalli06: An Agent for the Supply Chain Management Game 2006, *Proceedings of the International Conference on Computational Intelligence for Modelling, Control & Automation -CIMCA 2006*, published on CD-ROM, ISBN, NSW, Australia, November 2006., IEEE, Sydney, Australia.

Macías, D., et al (2007) Desarrollo de Agentes Inteligentes para la TAC SCM 2007 (Development of Intelligent Agents for the TAC SCM 2007), *Actas de la XII Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA-TTIA 2007)*, pp. 81-90, ISBN, Salamanca, Spain, November 2007, University of Salamanca, Spain.

12

A Framework and Key Techniques for Supply Chain Integration

Yanfang Huo[1], Xinyue Jiang[1], Fu Jia[2] and Bingguang Li[3]
[1]School of Management, Tianjin University
[2]Cranfield School of Management, Cranfield University,
[3]Byrd School of Business, Shenandoah University,
[1]P.R. China
[2]UK
[3]USA

1. Introduction

Supply chain integration is a new kind of organizational model, taking dynamic alliance of supply chain as a subject, to realize global resource integration through interactive collaborate operation of supply chain. Different from vertical integration, this integration focuses on the seamless connection of firms to improve the whole supply chain competitiveness by establishing and maintaining a long-term strategic partnership based on information integration, function and business reengineering, organization integration, cultural adaptation and strategic resources reorganization, etc (Chen & Ma, 2006).

Erengue (1999) brought forward four valuable research fields of supply chain, among them 3 are relevant to supply chain integration, i.e. integrated approaches to managing inventory decisions at all stages of the supply chain, the use of information sharing in a multi-partner supply chain, and analytical and simulation models that integrate the three major stages of supply chains which he thought to be is an important future direction of research in the area of supply chain. The literature suggests different theoretical models to describe and operationalise integration. Hammel & Kopczak (1993) pointed out that supply chain life cycle is a whole process from the construction to disaggregation of supply chain and defined the processes of construction, operation and disaggregation from the perspective of focal firm. Fisher (1997) and Nissen (2001) investigated supply chain integration from the aspect of product and agent-based separately. More recently, Lalwani & Mason (2004) and Mason & Lalwani (2006) use a model to characterize the extent of integration in relation to TPL providers. Tang & Qian (2008) established a PLM framework to enable supplier integration and partnership management in the automotive development process.

In connection with a survey of the relation between supply chain collaboration and logistics service performance (Stank et al., 2001), a framework for establishing the degree of internal and external collaboration is set up. This framework is further developed by Gimenez & Ventura (2005) in order to study internal and external integration as well as the influence of integration on performance. Again, the framework is appealing, but does not include a more systematic and detailed description of the specific tasks and processes involved in the cooperation. An operationalisation of the integration concept requires the identification of both the most essential tasks to be solved in connection with supply chain management and the underlying activities to be carried out to accomplish these tasks (Mortensen et al, 2008).

In this paper, investigating the methods of supply chain integration for manufacturing industry in the background of China, a three-echelon theoretical framework for supply chain integration based on Thorn's model (Thorn, 2002) is established. And then the relative techniques are presented in each level, which will be illustrated as following.

2. The three-echelon theoretical framework for supply chain integration

In order to integrate supply chain effectively, based on the comprehensive hierarchical planning framework by Thorn, we establish a theoretical framework for supply chain integration. According to the framework, the key relative techniques can be sorted into three echelons based on the rules from entity objectives to relative objectives and from basic capabilities to advanced capabilities- the basic operations management level, the planning and controlling level and the strategic management level.

The integration in operations management level. The supply chain operations level involves the whole process from the material acquisition to order fulfillment, which is the physical level and basic elements of supply chain. The supply chain integration must begin with the integration of this level, which is the basis of collaboration between all the firms.

According to the organizations and the functions, the integration in the operational management level includes internal integration of the focal manufacturer, the supplier integration, the distributor integration and the customer integration.

The integration in planning and controlling level. The excellent operation needs the support of integrated planning and performance evaluation, which involves the utilization of multiply techniques to plan, control, assess and improve performance. The integration in planning and controlling level coordinates all the business processes, esp. source, make, order fulfillment and inventory replenishment by information utilization and coordination. The core competencies in this level involve: databases, which enable the members to share necessary information; transaction system, which can initiate and deal with inventory replenishment and customer order fulfillment. Besides, it is vital to form the capability relevant to internal communication and collaborative operation. IT-based CPFR strategy is helpful to forming the core competencies. Meanwhile, it is also vital to monitor the business process through performance evaluation and improve the integrated performance continuously. Therefore, the key elements in planning and controlling level can be summed up as IT-based CPFR strategy and performance evaluation.

The integration in strategic management level. Out of question, successful supply chain integration needs partnership and management skills to maintain the partnership, while the skills always come from unique organization culture which is the basis of partner selection and maintaining. Thereby partnership maintaining and cultural adaptation are the two kernel elements in strategic management level.

Fig. 1 shows the framework and the key elements in every level for supply chain integration. On the basis of core competencies are integrated well with support capacities, the focal firm can coordinate four kinds of flows- product/service flow, knowledge flow, information flow and fund flow – to produce value.

The product/service flow refers to the serials of value-adding activities of product/service from material acquisition to end customers. The values add to the product when the product

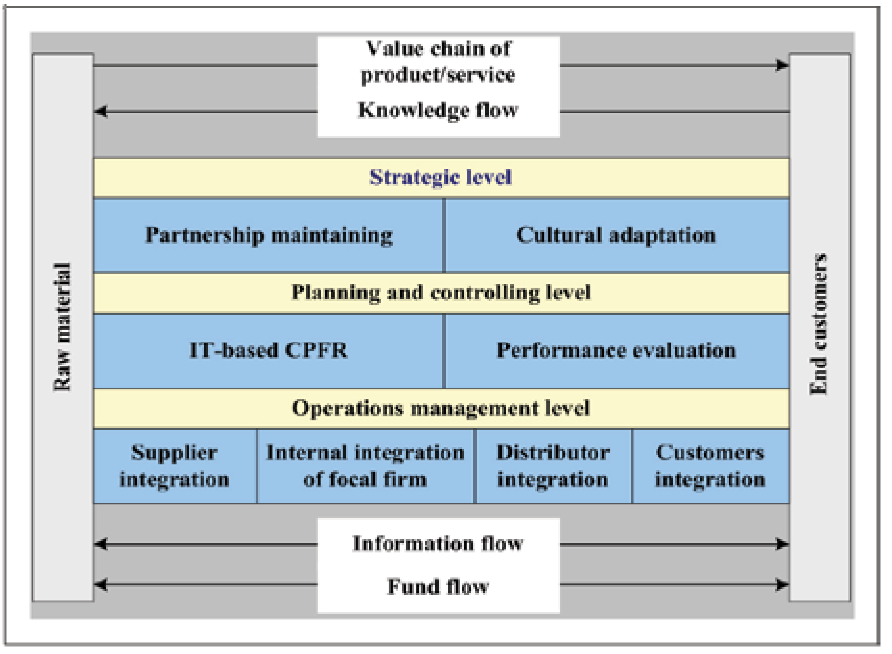


Fig. 1. The three-echelon theoretical framework for supply chain integration

flows along the supply chain and experiences physical changes, packaging, launch, customization, service support and other relative activities until meet the needs of end customers.

The knowledge flow is a reverse flow from end customers to suppliers. It always involves the exchange information about sales mode and product description, such as the customization requests, POS data, consumption information of end customers, warehouse and shipment information, etc. The information is vital to supply chain planning because it is helpful for the members to know the sales status well and then reach consistent understanding on customer requirements and consumption status. On the basis, better plans are formulated and the supply chain can work collaboratively.

The Information flow is a kind of inter-communicational flow between supply chain members. The information always includes forecast information, promotional plan, purchase order, order confirmation, shipment and inventory information, invoice, payment and replenishing requirements, etc. The information exchange can trigger, control and record the flow of product/service. With the ICT development, more and more information are exchanged by EDI and network instead of paperwork.

The fund flow always moves against the value-adding activities. Capital turnover and return on assets are the two main financial metrics which are of great importance to the supply chain performance.

The four flows exist always even if no coordination in the supply chain. However, under the circumstance of low integration and bad coordination, the flows move unsmooth which will result in delay, redundancy and inefficiency. While, integrated supply chain will accelerate the flows, with which the supply chain can produce maximum customer value and keep in a good condition meanwhile.

3. Key elements in the operational management level

The operation management level is the level aiming at synchronous operation of supply chain. The key issue in this level is how to balance and coordinate the restrictions, such as resources, information, capacity and time, through integration and coordination within each firm and between firms. The integration in operations level of supply chain can be illustrated in two dimensions: internal integration and external collaboration, shown in Figure 2. In the following, we will illuminate them and the key elements separately.

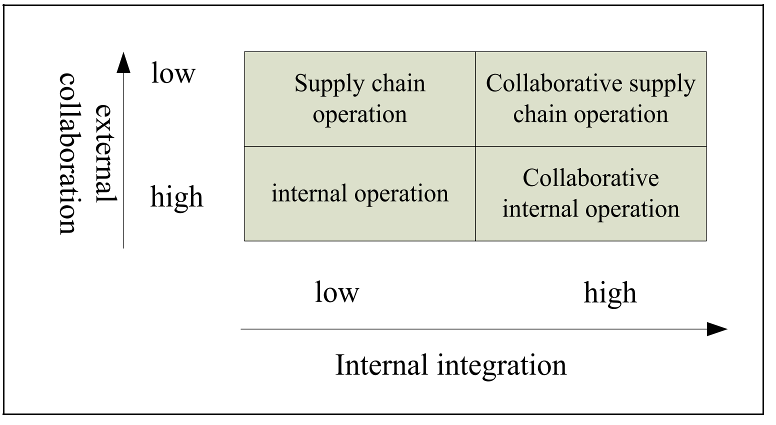


Fig. 2. The operational integration matrix

3.1 Internal integration and collaborative internal operation of within focal firm

The internal integration, the function to function integration within the focal firm, is the first step of operations integration, and also the basis of to success of supply chain integration. High internal integration can reach a level of "collaborative internal operation", with which the whole firm works like an integrated system that results in better performance and better interdepartmental effectiveness, such as cycle time reduction, better in-stock performance, increased product availability levels, and improvement in order-to delivery lead times (Harrison et al, 2008). Moreover, high internal integration is also the foundation of high external integration. A study of Spanish food manufacturers by Gimenez (2006) shows that the highest levels of external integration are achieved by firms which have already achieved the highest levels of internal integration between logistics, production and marketing.

For the internal integration is process-oriented, the firms need to come across the border of functions to build a borderless flat organization through BPR (business process reengineering) combined with ICT-based advanced production modes, such as MRPII (manufacturing resources planning), ERP (enterprise resources planning), lean production, agile manufacturing, concurrent engineering, etc.

BPR is the fundamental approach for internal integration within the focal firm, which emphasizes '... the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical contemporary measures of performance, such as cost, quality, service, and speed.' To implement BRP effectively, Hammer (1990) presents seven principles for BPR: organize around outcomes, not tasks; have those who use the output of the process perform the process; subsume information-processing work into the real work that produces the information; treat geographically dispersed resources as though they were centralize; link parallel activities instead of integrating their results; put the decision point where the work is performed, and build control into the process; capture the information once and at the source.

Base on these principles, cross-functional team, first presented by Toyota in lean production, can be one of the rational reference models for internal integration. While before the actual efforts of new model, the firms should define some mechanisms and actions to monitor and evaluate the status of collaboration, and then improve the initiative planning process. 'Alignment Compass' (van Hoek & Mitchell, 2006), the alignment analysis tool shown as Figure 3, illustrates four areas where alignment improvement efforts could be focused: in the interactions with peers from other functions, in interactions with their bosses and the Board, in interactions with their teams, and in their own day-to-day behavior.

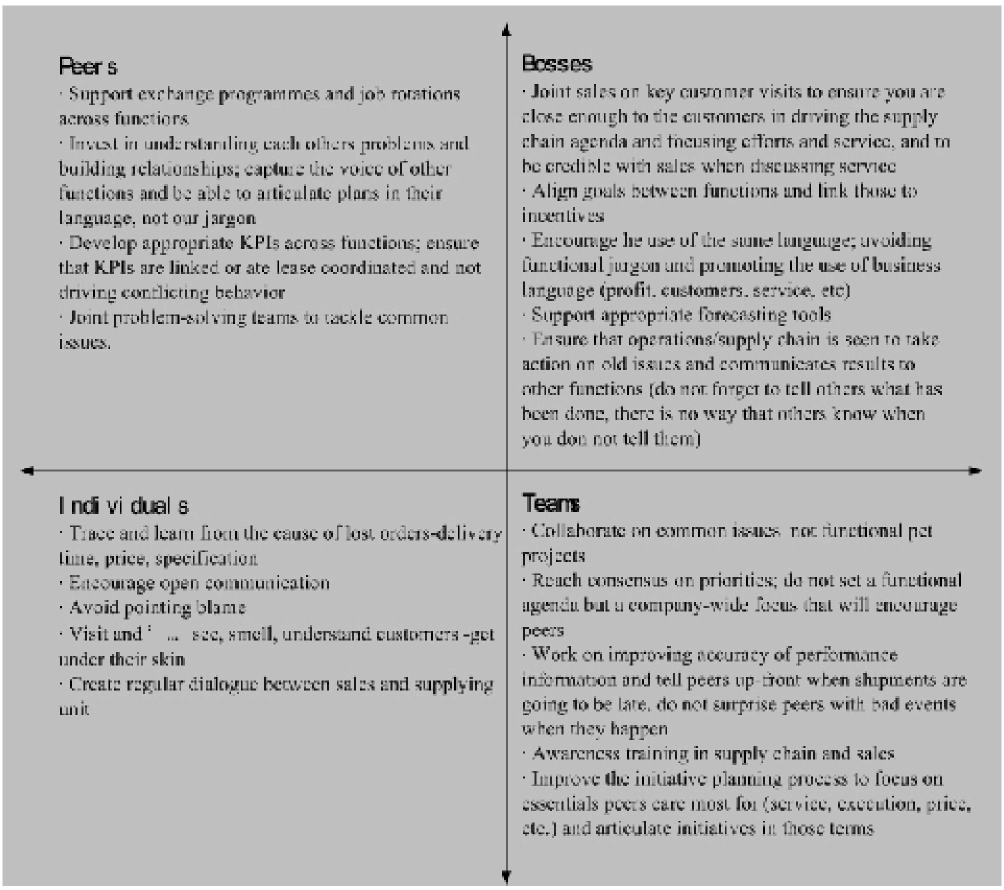


Fig. 3. Alfa Laval's alignment compass (van Hoek and Mitchell, 2006)

When the focal firm finished internal integration, it must follow some rules as below to carry out the new business model successfully:

Rule 1 Transfer from functional management to process management. The new model is process-oriented, and put the decision point where the work is performed, therefore response to market and customer will be improved through the shortened communication channel and time.

Rule 2 Focus on systematic philosophy about whole process optimization. Reengineer and optimize the business process to delete the useless activities or non-value-added activities, meanwhile, make each activity add the maximum value to customers. Note that all of these are base on global optimization, not local optimization, aiming at eliminating the selfish departmentalism and advantages equalitarianism.

Rule 3 Build a flat organization. Design processes first, and then build the organization based on processes. Remove the middle-level managers as could as possible.

Rule 4 Have everybody play his important role in the whole business process. Each person who processes the business should have comprehensive qualification and teamwork spirit. At the same time, the organization should build a new mechanism for self-learning.

Rule 5 Integrate business processes oriented to customers and suppliers. In the age of competition between supply chains, the firm should consider not only the collaboration between internal business processes, but the redesign of the interfaces between the focal firm with its customers and suppliers, when implementing BPR.

Rule 6 Resolve the conflicts between dispersed business and centralized management using ICT. When designing and improving the business processes, the firm should make the most use of ICT to process and share information as far as possible, convert sequencing processes into synchronous ones, resolve the conflicts between separated businesses with centralized management.

All in all, the firm can provide the right products with lowest costs and accurate amounts at right time and right place through BPR and integration of internal core businesses. Furthermore, high internal integration will improve the firm's decision-making capability dramatically, so the firm can capture the opportunities and win the competition in the fierce market.

3.2 External collaboration and collaborative supply chain operation

The second dimension of operational integration is called external integration, or inter company integration, referring to the cross-border operational integration in the supply chain which can place customer and supplier processes closer together. Compared with internal integration, external integration is a relative new concept, which integrates a firm's logistics with external logistics of suppliers and customers by the excellent collaboration between the partners. High external integration has some features like: increased logistics transactions with suppliers and customers; increased logistics collaboration between the focal firm with their customers and suppliers; more indistinct organizational boundary between partners in logistics collaboration. External integration makes the supply chain operate like a real physical entity to gain more powerful competitive advantage.

High external integration can be divided into 'supply chain operation' and 'collaborative supply chain operation' based on the internal integration level of each firms. The former one is high external integration with low internal external integration (In fact, it rarely exists). The latter one is a real high integration type base on high internal integration and high external integration. High-integration supply chain operates in a form of virtual organization, which is like a physical entity with high competency.

External integration can be divided into three basic types according to the partner along the material flow - supplier integration, distributor integration and customer integration, which will be explained in detail as following.

3.2.1 Supplier integration

Supplier integration plays a very important role in the operational integration of supply chain. One of the keys to increased responsiveness in the supply chain is a high-integration with upstream suppliers. Therefore the focal firm should pay more attentions on supplier development and integration to build partnership with the suppliers, which can increase the firm's performance or capabilities and meet short-term or long-term supply needs of buyers. Based on PDCA (Plan-do-check-action) cycle, we build a model– supplier integration cycle to support the activity of supplier integration. The model can be divided into 5 stages around with the goal of supplier development and excellent performance. At the initial stage, the firm should set up clearly and consistent objectives, which must communicate with the suppliers carefully. At the second stage, the firm should select qualified suppliers and establish a perfect evaluation system and assess the supplier performance based on the evaluation system. At the third stage, feed back the actual performance to the suppliers. While at the fourth state, the suppliers analyze the process to find the performance gap, and then formulate plans to improve their performance. Finally, the focal firm admits suppliers' performance and gives them the relative treatments according to their performance to support joint development. The supplier integration cycle is shown as Figure 4. In the following we will illustrate how the cycle operates based on John Deere's practices.

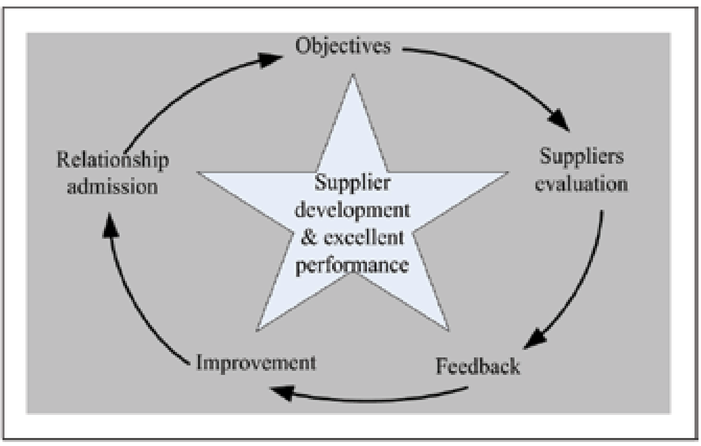


Fig. 4. The supplier integration cycle

A. Objective setting. The first step of supplier integration is to set the consistent objectives and strategies, on the basis, integrate information, processes and resources to realize quick response to customer needs. The objectives of supplier integration always involve:

- Build a win-win relationship with supplier, share the resources each other, and achieve continuous improvement (CI) to win the competitive advantage in the marketplace.
- Control the total cost by quality improvement combined with cost management.
- Build effective performance measurement system to lead and encourage CI, and communicate with suppliers on their performance timely and accurately. Meanwhile, provide standards of supplier admission for their excellent performance. On the basis, carry out strategic sourcing.
- Encourage the suppliers to involve in all the core processes and make full use of their technical supports, innovations and experiences to improve the capability and competitiveness of the whole chain.

B. Supplier evaluation. Often, the suppliers need to be evaluated from such aspects as quality, delivery, cost, technical support and collaboration, to determine their relationship with the focal firm. The contents of these metrics are illustrated as below.

- *Quality,* the metrics on the capability that suppliers' quality management system and material provided meet the requirements, expectation and material quality of the focal firm. This metrics, involving quality PPM and quality effects, provides suppliers with the statistical information on their products/service quality.
- *Delivery.* Delivery evaluation provides the statistical basis for suppliers with their order fulfilling capability based on order amounts and delivery date. The delivery level is denoted by PPM (Delivery PPM = number of defects/supplied amount×10^6), which is calculated based on the amounts of early delivery, late delivery or over delivery.
- *Technical support,* the metrics on suppliers' knowledge and their capability of technology application, to determine whether they can support strongly the focal firm's needs of product development and manufacturing. The performances of suppliers in such fields as the product techniques and innovations, the delivery process, the manufacturing process, the time of manufacturing critical path, and the warranty are involved in the evaluation system.
- *Collaboration,* comprehensive analysis of suppliers' initiative, attitude, responsiveness, communication, detail-concerned, and safety performance. The evaluation fields always include information share, problem-solving, responsiveness to customer requirements, business relations, consistency of processes, and the quality and smoothness of electric commercial.
- *Cost*, the metrics on the capabilities of supplier in price competitiveness, cost control and reduction. The evaluation objects include the cost reduction planning, the net cost reduction performance, the enthusiasm in cost control and the performance during product delivery.

As for the collaborative relationship, according to the supplier performances, it can be divided into four grades as following.

- *Partners*, the suppliers whose performances have topped the performance measurements of the focal firm and reach the world class level, and played an important role to customers satisfaction of the focal firm.
- *Key suppliers*, the suppliers whose performance have topped the lowest level of the focal firm and are go further towards the world class level continuous.
- *Qualified suppliers*, the suppliers whose performances reach the lowest level of the focal firm, but no action on continuous improvement.
- *In questioned suppliers*, the suppliers whose performances are lower than the lowest level of the focal firm and may be rejected from the supplier group.

The final level is determined by the lowest one of any field and the supplier performance in each field is assessed by cross-functional team. The evaluation system is shown in Table 1.

C. Feedback and improvement. After evaluation, the focal firm will give the suppliers feedback on their performances. Besides the information of past and current performances, the status of other key suppliers is also presented in the report to encourage the supplier's improvement. Moreover, it is prescribed that the suppliers those achieve a 50% improvement in some field (quality, delivery, etc) than the early year can be upgraded to the upper one level (e.g. from qualified supplier to key supplier, but never from key supplier to partner). It is a measure to encourage the CI activities of suppliers.

	Partner 92%-100%	Key supplier 80%-91%	Qualified supplier 70%-79%	In questioned <70%
Quality	Excellent quality, topped in all metrics; high reliable product/service; well-recorded and filed quality plan & improvement measures; excellent outcomes in continuous improvement	Excellent quality plan; timely response to quality problems; reliable product/service; better outcomes than measurements consistently	Quality meets the expectation; passive response to quality problems; substantially satisfied by inner customer; approved but not implemented collaboration agreement	Ordinary service; inadequate product quality; may not keep the business relations with this supplier.
Delivery	Always deliver product/service on-time; actively respond to the short-time order; seamless link of delivery process to pricing process.	Can deliver product/service on-time and respond to the short-time order; a few pricing problems.	Always be reminded of product/service delivery; sometimes need to be followed on the problems of delivery and/or price; respond passively with no expectation	Need to be traced; always failed to due date; very long lead time; and no measures are taken to shorten lead time.
Technical support	Respond to technical problems and service problems promptly; explore and implement innovation in techniques; provide advanced techniques for customers; service representatives are fully-trained and are the expert of their field.	Rarely go wrong when introduce new or existing product/ service; service representatives are able to solve most techniques or service problems; make efforts to implement innovation in techniques under proper directions.	Reply to technical change passively (not actively); sometimes new products or new services are unusable; need to be followed to implement innovation in technique sufficient technical support.	No technical support; and no response to changes.
Collabo-ration	Care for customers' experiences; make great efforts to continuous improvement actively; provide accurate information over expectation; be good at teamwork and communication; clear, open and frank business relations	Support its staff and worker fully; outcome-oriented to meet the expectations all the time; com in on problem-solving actively.	Need to be traced; respond to customers requirements passively with no prediction; share information sufficiently	Difficult to file and share information; low reliability; need to clarify and trace the information all the time.
Cost	Have set goals of continuous cost reduction and monitor its operations; focus on inner cost, make great efforts to reduce the cost both internally and externally with suppliers; keep the most competitive price always.	Make efforts to reduce cost and gain some opportunities internally and external internal from suppliers; the price is competitive	Make few efforts to control cost control; need to be encouraged to implement cost reduction activities.	No cost reduction action; uncompetitive price

Table 1. Supplier evaluation system

D. Relationship admission. Supplier relationship admission refers to the plans and activities of supplier development after determining the supplier relationships by evaluation, feedback and improvement. According to the different relationship, different development plans are formulated. For instance, the partners can be involved in the new product design and bid for other businesses. Besides, they can join the training plans, attend the meetings of supplier management, and so on. The key suppliers are also qualified to join the training plans and the meetings of supply management. But they are just considered to be involved in the business processes of new product design and new business development. As for qualified suppliers, they are qualified to some specific training plans, and may be invited to participate in the meeting of supply management, and also have the possibility to be thought of involving in the business processes of new product design or new business development. However, the suppliers in questioned are not admitted at all.

The relationship admission can enhanced the relationship between focal firm and its suppliers to upgrade the supplier integration level step by step.

3.2.2 Distributor integration

The downstream of the focal firm is the distributors. It has been suggested for a long time that manufacturers, esp. those produce industrial product, should treat their distributors as partners (Narus et al, 1986), which means that the manufacturers should admire the value of distributors and provide necessary support to the distributors to win the competition in marketplace. In fact, the distributors always possess lots of information about the customer requirements, which the manufacturers will need when they want to develop new products and production line successfully. Moreover, integrating with the distributors can share the skills between the distributors to meet the end customers’ needs much better.

The distributor integration can be realized mainly in two aspects.

First is sharing inventories within the alliance of the focal firm and all the distributors to protect the downtime from emergency orders. Traditional distributor management fulfilled the emergency orders by increased inventory. On the contrary, distributor integration can decrease the inventory by sharing inventory information between all the distributors. Every distributor can check others’ inventory records to determine its own. And in some cases, distributors have the contract-type duty to exchange parts at a consentaneous price. Sure, it needs the support of advanced information system.

Second is to improve each distributor’s capacity in explicit skills and capability of response to non-routine requirements. In this type of alliance, different distributors can cultivate their skills in different fields. And a specific customer requirement can be inducted to the most skillful distributor. Otra is a good example of this skills integration and collaboration.

There are two issues which need to be paid much attention in distributor integration. First, the distributors may doubt about their returns of taking part in such a system. They will feel upset when they think of they are providing some inventory control skills to their inexperience partners, esp. when some distributors are more powerful and holding more inventory. Second, some distributors have to rely on other distributors to help them improve the customer service. However, sometimes the distributors who are relied on may not know what will happen. Third, the new type of collaboration will easily lead to a status that a certain responsibilities and skills may be transferred from some distributors to a certain new distributors, which upsets some distributors. All these problems illuminate that distributor integration is a very difficult task in SCI which need the focal firm to devote lots of resources and efforts to get the trust of its distributors (Simchi-Levi et al, 2003)

3.2.3 Customer integration

Ken Burnett (2002) pointed out that the capability of identifying, understanding and meeting the needs of important customers can always be seen from most of successful companies. Developing a long-term partnership with these important customers is key to successful customer integration. Only if the partnership can be maintained well, can a firm realize lean and integrate with the customers. The process of partnership development and customer integration is shown as below (Zhou, 2006):

Step 1 Requirements analysis. In this more and more stinging market, the company must perceive the rapid change of customers needs (not only the explicit requirements, but also the tacit requirements), and give quick responses to them. In order to reach the target, the company should understand them from multiple perspectives – industrial perspective, dynamic perspective and purchase perspective, to determine its targeted customers based on mutual complement, compatibility and win-win premise, and then attract the customers depending on its business strengths.

Step 2 Value positioning. Value positioning is the kernel of customer partnership. Whether a company can establish long-term customer relationship depends more on its capability of creating durative customer value. Value identification, value selection and value supply compose the value creation process. Indentifying the factors of affecting customer value judgment, finding the features the customer most concerned, and thinking over the value positioning of competitors combining with its own advantages, will help the company to find new breaks of value innovation. Then the company can configure the resources based on the new breaks and the total product concept as well to provide conceived value for customers and then to realize value position.

S*tep 3 Strategies matching*. It means that the strategic orientation, competitive strategy and strategic resources input to the focal firm must match its customers' needs to enhance the collaboration. Based on the analysis of the enterprise environment and its own resources, the focal firm should formulate its strategy aligning with the selected value position and matching it to its partners.

Step 4 Process improvements. The focal firm should transfer its main business process into strategic capability to provide valuable service to customers. That means the firm should enhance its core business process and make it un-imitable to customer.

Step 5 Partnership maintenance. In order to make full use of customer partnership, the company must deal with the relationship carefully, including effective innovation and strict risk control. Rational partnership innovation combined with effectual control is key elements to customer partnership success.

4. Key techniques in planning and controlling level

4.1 IT-based CPFR strategy

Information plays a very important role in the operational process of supply chain. Low level of information transparency and visibility will result in the unnecessary 'transit' cost. While the CPFR strategy based on more active information sharing mechanism and more effective inventory forecasting and replenishing measures will ensure a more smooth supply chain. In this section, we will provides a practical method to realize CPFR by forecasting demand with time series analysis and the adoption of a Push/Pull integrated inventory management system.

4.1.1 Supply chain visibility and bullwhip effect

Supply chain visibility was defined as 'the ability to see clearly from one end of the supply chain to another and, in particular, to share information on supply and demand issues across corporate boundaries' (Christopher & Gattorna, 2005). Lack of visibility will cause the uncertainty of demand and the most obvious phenomenon led by uncertainty is the bullwhip effect. When a supplier from a certain level makes decision only based on the demand of its next distributor, the data distortion will move upstream along the supply chain therefore intensify level by level. As a result, when reaching the headstream supplier, the accumulated distortion will vary the demand quite a lot from the real condition of the ultimate market (Zhang et al, 2003). Figure 5 shows typical order patterns faced by each node in a supply chain that consists of a manufacturer, a distributor, a wholesaler and a retailer. The retailer's orders to the wholesaler display greater variability than the end-consumer sales; the wholesaler's orders to the manufacturer show even more oscillations; and, finally, the manufacturer's orders to its supplier are the most volatile. Even a slight change in consumer sales ripples backward in the form of magnified oscillations upstream, resembling the result of a flick of a bullwhip handle.

4.1.2 CPFR-base logistics and supply chain management

To decrease the bullwhip effect and enhance supply chain visibility, there is a growing recognition of the importance of shared information within the supply chain. For example, the adoption of 'Collaborative Planning Forecasting and Replenishment' (CPFR) is beginning to make a difference. CPFR involves the joint determination of forecasts through pooled knowledge and information. Based on this agreed view of demand over the forecast horizon, the supplier takes responsibility for the replenishment of supplies based upon the actual rate of sale or usage (Ireland, 2000). The realization of CPFR requires the establishment of an information-sharing platform, which pools the information of customer demand, order adjustment, transportation plans etc., among manufacturer, distributor, wholesaler and retailer. In this way, the visibility of supply chain could be enhanced whilst suppliers and distributors could clearly know the sales volume and the products inventory at the retail store, distribution center and manufacturing sites, so that they can better forecast the demand.

A. Demand Forecast by Time Series Analysis

The first step to realize CPFR is to accurately forecast demand in the future. Therefore, a time series analysis could be a practical method.

Hypothesis 1 Influence factors of demand. A time series can be defined as chronologically ordered data that may contain several components of demand: trend, seasonal, cyclical, autocorrelation and random. In practice, it is relatively easy to identify the trend and the seasonal component, but more difficult to identify the cycles, autocorrelation and random components. We hypothesize that demand contains both seasonal and trend effects in here.

Hypothesis 2 The type of demand variation: multiplicative seasonal variation. The multiplicative seasonal variation is the usual experience. Figure 6 shows the seasonal variation increasing as the trend increases because its size depends on the trend. Forecast including trend and seasonal= Trend × Seasonal factor.

Hypothesis 3 Seasonal Factor. A seasonal factor is the amount of correction needed in a time series to adjust for the season of the year.

Align with these hypothesizes, we take A firm as an example to explain how to forecast demand of the future. A firm's sales for a product line during the 12 quarters of the past three years were as follows (Table 2). Our objective is to forecast demand for the four quarters of the fourth year using Least Squares Regression to analyze the Time Series.

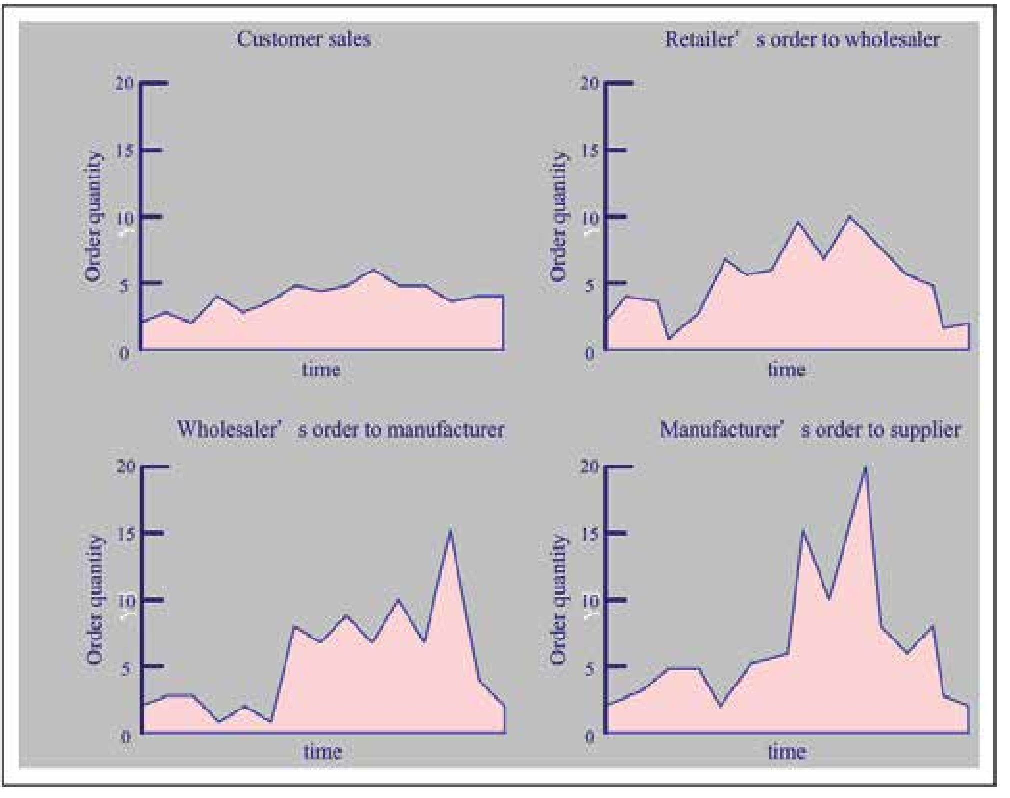


Fig. 5. Increasing variability of orders up the supply chain

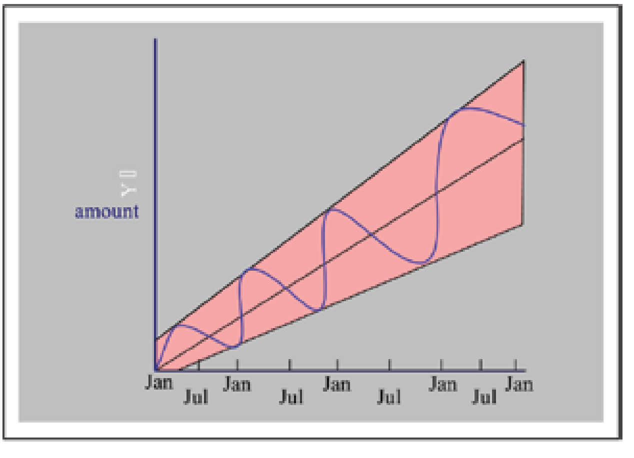


Fig. 6. Multiplicative Seasonal

Step 1. Determine the seasonal factor. Table II summarizes the calculations needed. Column 4 develops an average for the same quarters in the three-year period. For example, the first quarters of the three years are added together and divided by three. A seasonal factor is then derived by dividing that average by general average for all 12 quarters (33,350/12 or 2,779.1667). These are entered in column 5. The seasonal factors are identical for similar quarters in each year.

Qtr	sales	Qtr	sales
1	600	7	2600
2	1550	8	2900
3	1500	9	3800
4	1500	10	4500
5	2400	11	4000
6	3100	12	4900

Table 2. A firm's sales during the past three years

Step 2. Deseasonalize the original data. To remove the seasonal effect on the data, we divide the original data by the seasonal factor. The result is shown in column 6 of Table 3.

(1) Prd (x)	(2) Qtr	(3) Actual Demand (y)	(4) Average of the Same Quarters of Each Year	(5) Seasonal Factor	(6) Deseasonalized Demand(y_d) Col.(3) ÷Col.(5)	(7) x^2 Col.(1)^2	(8) $x \times y_d$ Col.(1) ×Col.(6)
1	I	600	(600+2,400+3,800)/3 =2,266.7	0.82	735.7	1	735.7
2	II	1,550	(1,550+3,100+4,500)/3 =2,700	1.10	1,4122.4	4	2,824.7
3	III	1,500	(1,500+2,900+4,900)/3 =3,100	0.97	1,544.0	9	4,631.9
4	IV	1,500		1.12	1,344.8	16	5,379.0
5	I	2,400		0.82	2,942.6	25	14,713.2
6	II	3,100		1.10	2,824.7	36	16,948.4
7	III	2,600		0.97	2,676.2	49	18,733.6
8	IV	2,900		1.12	2,599.9	64	20,798.9
9	I	3,800		0.82	4,659.2	81	41,932.7
10	II	4,500		1.10	4,100.4	100	41,004.1
11	III	4,000		0.97	4,117.3	121	45,290.1
12	IV	4900		1.12	4,392.9	144	52,714.5
78		33,350		12.03	33,350.1	650	265,706.9

$$\bar{x} = \frac{78}{12} = 6.5 \qquad b = \frac{\sum xy_d - n\overline{xy}_d}{\sum x^2 - n\bar{x}^2} = \frac{265{,}706.9 - 12(6.5)2{,}779.2}{650 - 12(6.5)^2} = 342.2$$

$$\bar{y}_d = 33{,}350/12 = 2{,}779.2 \qquad a = \bar{y}_d - b\bar{x} = 2{,}779.2 - 342.2(6.5) = 554.9$$

Therefore, $Y = a + bx = 554.9 + 342.2x$

Table 3. Deseasonalized demand

Step 3. Develop a least squares regression line for the deseasonalized data. The purpose here is to develop an equation for the trend line Y, which we then modify with the seasonal factor. The least squares equation for linear regression:

$$Y = a + bx \tag{1}$$

Where,
Y_d = Deseasonalized demand (See Table 3)
x= Quarter
Y= Demand computed using the regression equation (1)
a = Y intercept
b= Slope of the line
The least squares calculations using columns 1, 7, and 8 of Table II are shown in the lower section of the table. The final deseasonalized equation for our data is $Y = 554.9 + 342.2x$.
Step 4. Project the regression line through the period to be forecast. Our purpose is to forecast periods 13 through 16. We start by solving the equation for Y at each of these periods (shown in Table III, column 3)

Prd	Qtr	*Y* from regression line	seasonal factor	forecast (*Y*×seasonal factor)
13	1	5,003.5	0.82	4,102.87
14	2	5,345.7	1.10	5,880.27
15	3	5,687.9	0.97	5,517.26
16	4	6,030.1	1.12	6,753.71

Table 4. Final forecast adjusted by the seasonal factor

Step 5. Create the final forecast by adjusting the regression line by the seasonal factor. Multiply the quarterly data we derived by the seasonal factor for that quarter and we can get the final forecast (shown in Table 4, column 5).
B. Push/Pull Integrated Inventory Management System
The second step to realize CPFR is to replenish inventory quickly and reasonably. According to compacting multistage response management mode, not only does an organization have to turn out the right amount of products, but also has to response quickly to the market. It is not enough to fulfill the management demand when we use push or pull operational system singly. So a push and pull integrated inventory management system is brought out.
The upstream companies will carry out processes of purchase, manufacture and replenishment according to the Push Operational System, which basically belong to the mode of make to stock (MTS). They can also use the forecast method we mentioned above to determine the production lot sizes, stock level and reorder point based on the history orders of the downstream companies. The key point of the upstream operational process is JIT supply – to fulfill the downstream companies' producing demand while making them kept a low inventory level.
The downstream companies will carry out processes according to the Pull Operational System, which basically belong to the mode of make to order (MTO). They make the decision of orders to the upstream companies based on the history orders of their customer. The key point of the downstream operational process is quick response (QR), which means to perfectly fulfill the customer's purchasing demand through close cooperation in accelerating the logistics flow between retailer and supplier.

Push/Pull integrated operational system requires the upstream to prepare the spare parts aptly for the downstream or to finish the assembly to a certain degree, so that the downstream can accomplish the final process of assembly quickly when the customer's order arrived. The companies in both upstream and downstream have to build a quick response network which integrates each other's needs and techniques. In this way, more efficient logistics and information flows can be adopted, which ensure a better realization of the shared goal to provide higher quality product in shorter time. Push/Pull will realize economies of scale in the upstream of supply chain whilst satisfy the customization needs in the downstream.

4.2 Performance evaluation

Performance evaluation is the basis of supply chain management. SCOR (Supply Chain Operational Reference) model (SCC, 2008), which is founded on five distinct management processes and viewed in terms of overlapping management processes – source, make, deliver and return – within an integrated framework that encompasses all of the organizations in the chain, provides a good method to assess the supply chain performance, a process-based method of supply chain evaluation. From the model, the firm can obtain the benchmarking data and judge their internal operational performance.
There are three levels to the SCOR model: Level 1 defines the scope and content for the Supply Chain Operations Reference-model. Here basis of competition performance targets are set. Level 2 configures a company's supply chain from core "process categories" (e.g., make to order, make to stock). Level 3 defines a company's ability to compete successfully in its chosen markets. Detailed performance metrics are set in this level.
Table 5 shows 10 metrics at level 1 in SCOR model, which is taken from the SCOR website. Then the model provides a breakdown of level 2 and level 3 subcomponents of level 1 metrics to identify the opportunities of improvement. An individual company should not attempt to be "best in class" in all areas. Rather, a given company should targets its strength I four to six selected areas to create differentiation in the marketplace, and ensure that it stays competitive in the other areas. The Supply Chain Council also gives the benchmarking information for participating companies.

Level 1 Metrics	Performance Attributes				
	Customer-Facing			Internal-Facing	
	Reliability	Responsiveness	Agility	Cost	Assets
Perfect Order Fulfillment	√				
Order Fulfillment Cycle time		√			
Upside Supply Chain Flexibility			√		
Upside Supply Chain Adaptability			√		
Downside Supply Chain Adaptability			√		
Supply Chain Management Cost				√	
Cost of Goods Sold				√	
Cash-to-Cash Cycle Time					√
Return on Supply Chain Fixed Assets					√
Return on Working Capital					√

Table 5. Supply chain performance and the measurements that can be benchmarked

A company must measure and then enhance its capability of planning, sourcing, making and delivering illustrated in level 2 and level 3 (shown as Table 6), to meet the needs of customers. The degree of these metrics reach will determine its operational competency. Order and order fulfillment by XX, JIT delivery, inventory turnover, and cash-to-cash cycle time are the most important metrics of operational measurement. Some activities are linked to the metrics, e.g. the inventory management will affects the fulfillments of order and order by XX highly. So the company should collect and evaluate all the four processes operational information to make decisions. If it was found that the operations have some biases with the targeted performance, the company should investigate the relative businesses processes from the three types attributes shown as level 3, which are also called level 3 consultant metrics. Table 6 shows that the consultant metrics can be used not only to analyze the complexity and structure of supply chain, but to investigate the concrete business operations. By effective data collection and performance evaluation, the companies can find its weakness and then make plans to improve their performance and competency gradually.

	Level 2	Level 3		
	Performance metrics	Complexity metrics	Structure metrics	Practices matrix
Plan	Planning cost Financial cost Inventory days of supply	Rate of order change Number of SKUs held Throughput Inventory holding cost	Products amount of different channels Number of channels Number of supply chain location	Planning period Forecast accuracy Obsolete stock on-hand
Source	Material acquisition cost Source cycle time Material acquisition time	Number of Suppliers Rate of source from long distance	Materials from long distance Rate of source from long distance	Supplier delivery performance Payment period Percentage of each purchased items by lead time
Make	Product defects or number of customer complaints Make cycle time Order fulfillment Product quality	Number of SKUs Flexibility with the increasing demand	Make process steps by area Utilization	Rate of value-added Rate of order fulfillment Inventory turnover Rate of change order by internal problems WIP
Deliver	Order fill rates Order management cost Order fulfillment lead times Rate of product return	Order numbers by channels Items and shipments by channel Rate of items return	Distribution location by area Number of Channels	Delivery lead time Percentage of invoices with wrong bill Methods of order entry

Table 6. Supply chain performance metrics and diagnosis metrics

5. Critical contents in strategic management level

5.1 Partnership maintaining

5.1.1 The types of enterprises relationship

The key to supply chain success is good relationship with partners and excellent collaboration in product design, manufacture and competitive strategy between them. What type relationship that the supply chain selected finally relies on the degree of knowledge reliance and information sharing between all the members of supply chain (General Administration of Quality Supervision, Inspection and Quarantine of PRC, 2001). Figure 7 shows the main five types of supply chain relationship. According to Figure7, contract and outsourcing is the basic types of collaboration. In these two types, the reliance is just accepted to a certain extent, and only a little information (of operations) is shared. Moreover, the relationships can exist just in a certain period.

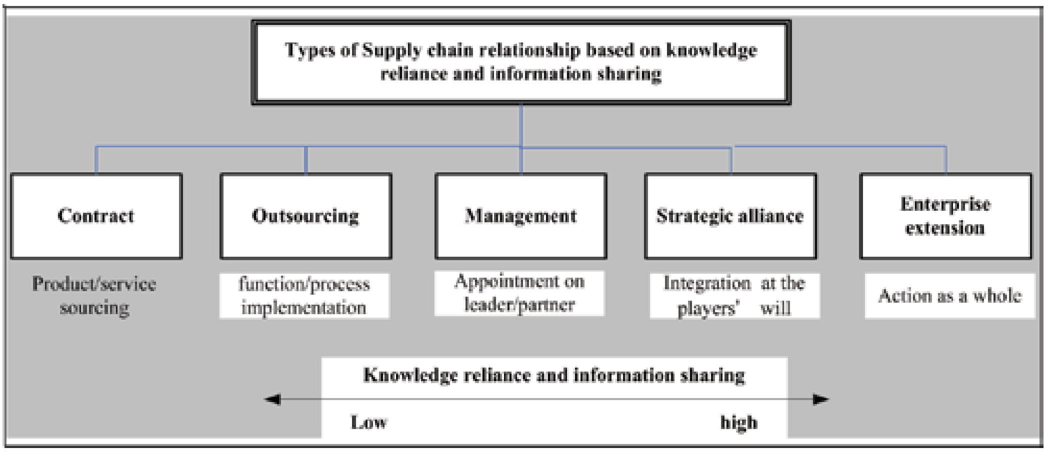


Fig. 7. The types of supply chain relationship

While with the management relationship, the focal firm usually takes the role of leadership and is in charge of looking for better collaboration with trade partners and service providers. Not only the operational information is shared, but some strategic information is shared as well. The relationships can last for a comparatively long time.

Strategic alliance and enterprise extension are two types of partnership, with which the firms require collaboration and are willing to cooperate rationally in an integrated way. They reach consistency automatically to integrate human, finance, operations and techniques to provide greater customer value with higher efficiency. And an extended collaboration planning aiming at maintaining this relationship is included as well. Enterprise extension, which is across the border of single firm, is the end of knowledge reliance and information sharing. By total information and planning share, enterprise extension can increase the operational efficiency and enhance the relationship. Moreover, it presents a more simple way of CPFR which we have discussed in section 4.1.

5.1.2 Strategic partnership and the key to partnership maintaining

There is no doubt that Strategic partnership based collaboration can increase the cooperation and communication between functions and firms so as to balance production, synchronize logistics, at the same time, shorten the time to market of new product remarkably.

Furthermore, the partnership strengthens the flexibility and agility in the fierce market by a production mode of modularization, simplification and standardization oriented to high customization. Virtual manufacturing and dynamic alliance are typical forms of strategic partnership, which enhance the effect of outsourcing.

However, it is not easy to establish and maintain the relationship. The main reason is that all firms are always concern with their own benefits. So the depth and scope usually limited, even for the strategic partners. When the internal or external environments change, the firms may be suffer great disasters because their partners' mistakes or abandonment. Ericsson Corp. lost its competitive advantage in mobile phone market and declined generally from March 2000, when the Philip Co., a supplier of Ericsson, fired unexpected in a plant which resulted in a downtime in Ericsson for the lack of key components.

What firms can be chosen as the partners? In order to seamlessly cooperate, the partners should have consistent cultures, uniform strategic insights and inter-supported operational philosophy, which can ensure their core competencies are complementary to each other.

Then, how to maintain a long-term partnership? It relies on three aspects: common strategy and operational vision; bi-directional performance evaluation metrics, and the formal and informal feedback mechanism.

First, define the strategy and operational goals all together, and then trace, evaluate and update the goals often to achieve long-term improvements. For example, if the focal firm develops a new product, it should decide the common goals with its customers about the product market orientation. Also, the goals should take the retailers' key role in the process into consideration.

Second, transfer the strategic and operational goals into detailed traceable performance measurements. The focal firm and their partners should decide the metrics and measure frequency together. Meanwhile, the metrics should be bidirectional. Generally, the metrics between manufacture and their suppliers focus more on the suppliers' performance, for example, JIT delivery and quality. A research on strategic alliance developed a successful united metrics - total systematical inventory. The research point out that it is the inventory decrease of two sides that is really important to improve the whole supply chain performance, not only the inventory decrease of manufacturers.

Third, evaluate performance, feedback and improve formally and informally. Annually assessment is the most popular formal method, which is usually done by top managers aiming at checking and updating strategy goals. While quarterly check or monthly check are two kinds of informal method focusing more on tracing and evaluating the operational performance, which is usually top manager excluded. When informal checks implement, the alliance can change their operational practice to create good conditions for improve the planning and reach strategic goals. Weekly/daily checks are also informal activities, which are carried out by coordinators to solve routine problems and find the opportunities to improve. Although they are informal, they have a detailed mechanism to solve problems. In a word, they are vital for collision avoidance and are good for establishing close relationship between coordinators.

5.2 Cultural integration and cultural adaptation

Organizational culture is the common cultural values growing up with the development of an enterprise, which has been accepted by all the staff and workers, including the vision, the management philosophy, the tradition, the behavior regulation, the management system, the relevant enterprise spirit, and so on. Generally, the culture will affect the behavior of firms and the culture consistency of supply chain, and then enhance the cohesion and competence of supply chain.

Cultural integration and adaptation is on the top of supply chain integration, which can be divided into integration within focal firm and adaptation between focal firm and their partners. Within the focal firm, the cultural integration should more strength the firm features, e.g. the values, the spirits, and the philosophy. And for the whole supply chain, the cultures should be adapted from the following three aspects.

Strategies consistency, the integration on the macro-level. The focal firm should confirm and enhance its core competence, while outsource the non-core competitive business by establishing strategic partnership with its supplier; on the basis, integrate the visions, competitive strategies and development tacit to reach their common goals.

Philosophies or values adaptation, the kernel and difficulty of cultural adaptation. The organizational philosophies and values, comprised of vision, philosophy, spirit, concepts of benefit, service, quality, etc., are the special standards on operational behaviors selection and evaluation shaping from its long-time operational process and the essential part of organizational culture. Values adaptation along the supply chain needs to separate the good from the bad, and then strengthen the good while delete the bad. Meanwhile, promote the values of focal firm and then form the common values accepted by all members generally.

Management models integration. In this operations level, different management models in supply chain should be analyzed to find, integrate and develop suitable spirits and souls for supply chain integration. By integrating management models, the improvement of employees' quality is linked with the improvement of supply chain competency, and a new incentive mechanism and supply chain culture that employees' fate is connected closely with the status of supply chain will be formed finally.

6. Conclusion

It is the good choice for a firm try to enter the global operation system to enlarge its market share and raise efficiency while its business develops strong enough. Generally, the process can be divided into three stages: international trade, branches establishment abroad and globalization. At the third stage, the firm can develop its business across the boundary in the international market. Therefore, it can improve the operational efficiency from three aspects at least: implement strategic supply of raw material and components; gain profit from low price labor by making and delivering in developing country; gain more profits from the tax preferential policy which makes the value-added model more attractive.

Up-to-date, most Chinese firms are still at the first stage. They are still relying on the low manufacturing cost from low price labor. However, just like the book *Supply chain management: the practices in Hong Kong Li & Fung Group* said: The production cost of a 4-yuan product in American market is only 1 Yuan. More important, it has almost reached the lowest and is difficult to decrease. So the firms must turn to the other 3 Yuan to make profits, i.e., make money by cost reduction in the whole supply chain processes, including product design, material supply, transportation, wholesaler and retailer, information and management. From the supply chain perspective, there are still lots of opportunities to decrease cost (Li & Fung Research Centre, 2003). ZARA and H&M have made good examples. Although no plant is established in China, they have entered Chinese market successfully with low cost and high profit through global supply chain integration. Chinese firms should pay more attention on this trend and try to enhance the competitive advantage from supply chain advantage.

The paper explores a framework for supply chain integration, and explained the relative methods from three aspects- operational management, planning and controlling, and strategic management. By effective supply chain integration, Chinese firms will find a more competitive way to increase the capability of soft 3-Yuan and compete in world market.

7. References

Burnett, K. (Translated by Liu, R. et al). (2002). *The Handbook of Key Customer Relationship Management*, Publishing House of Electronics Industry, Beijing

Chen, J. & Ma, S. (2006). Implementable Mechanisms and Technical Solution for Supply Chain Integration Management, *Industrial Engineering and Management*. Vol.8 No.1, 23~31

Chien T. W. (1993). Determining profit-maximizing production/shipping policies in a one-to-one direct shipping, stochastic environment. *European Journal of Operational Research*, No.64: 83~100

Christopher, M. & Gattorna, J. (2005). Supply chain cost management and value-based pricing. *Industrial Marketing Management*, (34): 115~121

Christopher, M. (2000). *Marketing Logistics*. Simon & Schuster (Asia) Pte Ltd.

Cohen M.A. & Lee H.L. (1988). Strategic analysis of integrated production-distribution systems: models and methods. *Operational Research*, 36(2): 216~228

Erengue S. S., Simpson N. C., Vakharia A. J. (1999). Integrated production distribution planning in supply chain: an invited review. *European Journal of Operational Research*, No.115, 219~236

Fisher M. L. (1997). What is the right supply chain for your product? *Harvard Business Review*, No.2, 105~116

General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. (2001). *National Standards of the People's Republic of China: terms of logistics*, Standards Press of China, Beijing

Gimenez, C. (2006) . Logistics integration process in the food industrial. *International Journal of Physical Distribution and Logistics Management*, Vol. 36, No.3, 231-45.

Gimenez, C. & Ventura, E. (2005). Logistics-production, logistics marketing and external integration: their impact on performance. *International Journal of Operations & Production Management*, Vol. 25 No. 1, 20-38.

Hammer M. (1990). Reengineering Work: Don't Automate, Obliterate. *Harvard Business Review*, July-Auguste, 104~120

Hammel T.D. & Kopczak L.R. (1993). Tighting the Supply Chain, *Production and Inventory Management*. Second Quarter, 63~69

Harrison, A. &van Hoek, R (2008). *Logistics Management and Strategy: Competing through the supply chain (3rd edition).* Prentice Hall Press.

Ireland, R. & Bruce R. (2000). CPFR only the beginning of collaboration. *Supply Chain Management Review*. No.5, 80~88

Lalwani, C. & Mason, R. (2004). Transport integration tools for supply chain management *Proceedings of the 9th Annual Logistics Research Network Conference*, pp. 346-355, University College Dublin, Quinn School of Business, Li & Fung Research Centre. (2003). *Supply Chain Management: the practices in Hong Kong Li & Fung Group*, China People's University Press, Beijing

Mason, R. & Lalwani, C. (2006). Transport integration tools for supply chain management, *International Journal of Logistics: Research and Applications*, Vol. 9 No. 1, 57-74.

Mortensen, O. & Lemoine l. (2008). Integration between Manufacturers and third party logistics providers. *International Journal of Operations & Production Management*, Vol. 28 No. 4, 331-359

Muekstadt, J.A. & Roundy, R.O. (1993). *Logistics of production and inventory*. NortHolland Press

Narus, J. & Anderson, J. (1986). Turn Your Industrial Distributors into Partners. *Harvard Business Review*, No.2: 66~71

Nissen, M.E. (2001). Agent-Based supply chain integration. *Information Technology and Management*, No.2, 289~312

Yilmaz, P. & Catay B. (2005). Strategic level three-stage production distribution Planning. *Computers and Industrial Engineering,* Vol. 48 No. 4, 799~809

Pyke, D.E. & Cohen M.A. (1993). Performance characteristics of stochastic integrated production-distribution systems. *European Journal of Operational Research*, No. 68, 23~48

Pyke, D.E. & Cohen M.A. (1994). Multiproduct integrated production-distribution systems. *European Joumal of Operational Research*, No. 74, 18~49

Simchi-Levi, D., Kaminsky, P., et al. (2003). *Designing and Managing the Supply Chain (2nd Ed)*. McGraw-Hill, New York

Stank, T.P., Keller, S.B. and Daugherty, P.J. (2001). Supply chain collaboration and logistical service performance, *Journal of Business Logistics*, Vol. 22 No. 1, 29-48

Supply Chain Council (SCC). (2008). *SCOR 9.0 Overview*. www.supply-chain.org

Tang, D. & Qian, X. (2008). Product lifecycle management for automotive development focusing on supplier integration. *Computers in Industry*, No.59, 288–295.

Thorn, J. (2002). *Taktisches Supply Chain Planning*, 22, Frankfurt am Main: Peter Lang

van Hoek, R. & Mitchell, A. (2006). Why supply chain efforts fail: the crisis of misalignment, International Journal of Logistics, *Research and Applications*, Vol. 9, No.3, pp269-281

Zhang, L., Pan, R. et al. (2003). The Bullwhip Effect and the Visibility in Supply Chain. *Commercial Research*, No.24, 179-181.

Zhou, W. (2006). *Strategy of key account loyalty: Partnership-based marketing,* Hunan People's Press: Changsha

13

New Approaches for Modeling and Evaluating Agility in Integrated Supply Chains

Vipul Jain and Lyes Benyoucef
INRIA Nancy Grand Est, COSTEAM Project, ISGMP, Bat. A, Ile du Saulcy 570000, Metz, France

1. Introduction

In the exisiting hotly competitive environment, companies/enterprises/organizations are interesting by the following question: How to provide the desired products and/or services to customers faster, cheaper, and better than the competitors?. Managers have come to realize that they cannot do it alone; rather, they must work on a cooperative basis with the best organizations in their supply chains in order to succeed. Moreover, the emerging global economy and the advent of IC technologies have significantly modified the business organisation of enterprises and the way of doing business. New forms of organisations such as extended enterprises, virtual enterprises, long supply chains etc. appeared and are quickly adopted by most leading enterprises. It is more and more noticed that "Competition in the future will not be between individual organizations but between competing supply chains" (Christopher, 2004). More and more business opportunities are captured by groups of enterprises in the same supply chains. The main reason for this change is the global competition that force enterprises to focus on their core competences (i.e. to be what you do the best and let others do the rest). According to the visionary report of Manufacturing Challenges 2020 conducted in USA, this trend will continue and one of the six grand challenges of this visionary report is to ability to reconfigure manufacturing enterprises rapidly in response to changing needs and opportunities.
While alliances like supply chains represent tremendous business opportunities, they also make related enterprises face greater uncertainties and risks. First supply chains are subject to market volatility and will have to be modified or dissolved once the business opportunities evolve or disappear. Changes or major perturbations at one enterprise will propagate through the supply chains to other enterprises and hence adversely influence the overall performance of the supply chains/networks. These issues are particularly important for SMEs. SMEs have to be part of some supply chains for business opportunities but they are not strong enough to face high uncertainties and risks, which are very common in today's dynamic and volatile markets. The capabilities to evaluate agility, benefits, performances, risks, etc. of supply chains are crucial for the long term efficiency and thus need serious research attentions.
Existing in both service and manufacturing activity sectors, generally speaking, a supply chain includes the transition and transportation of material from raw form through several

stages of manufacturing, assembly and distribution to a finished product delivered to the retailers and/or the end customers (Jain *et al.,* 2006). In addition to the material flows, it also includes the flows of information and finance. Each stage of material transformation or distribution may involve inputs coming from several suppliers and outputs going to several intermediate customers. Each stage will also involve information and material flows coming from immediate and distant preceding and succeeding stages.

Supply chains in general and integrated supply chains in particular are complex systems and their modeling, analysis and optimization requires carefully defined approaches /methodologies. Also, the complexities may vary greatly from industry to industry and from enterprise to enterprise. Since technological complexity has increased, supply chains have become more dynamic and complex to manage. *Consequently, it is easy to get lost in details and spend a large amount of efforts for analyzing the supply chain. On the other hand, it is also possible to execute too simplistic analysis and miss critical issues, particularly using tools that do not take into account agility, uncertainties, risks, etc.*

It is important to recognize that supply chain power has shifted from manufacturer to retailer, and finally to consumer (Blackwell & Blackwell, 2001). Most of the supply chain researchers and practitioners have agreed that there is a real need to develop integrated supply chains significantly more flexible, responsive and agile than existing traditional supply chains. It is essential that supply chains continually re-examine how they can compete and agility is one of the underlying paradigms to enable them to re-invent the content and processes of their competitive strategies. The main objectives of this chapter is to discuss ***two new approaches for modeling and evaluating agility in dynamic integrated supply chains.*** The rest of the chapter is organized is as follows: Section 2 deals with the complexities of integrated supply chains. Section 3 discusses the need for agile integrated supply chains. Section 4 presents the two novel approaches. Finally, section 5 concludes the chapter with some perspectives.

2. Integrated supply chains complexities

The key to genuine business growth is to emphasize the creation of an effective supply chain with trading partners, while at the same time maintaining a focus on the customer. Today, instead of simply focusing on reducing cost and improving operational efficiency, more efforts are put on customer satisfaction and the enhancement of relationships between supply chain partners. Traditional supply chain management (structural and operational strategies) are more incompetent and integration between all supply chain partners is essential for the reliability and durability of the chain. Therefore, more and more companies in different sectors like automotive, textile, grocery, petrochemical etc. are giving much more emphasizes on the integration of all their supply chain partners.

Integrated supply chains are dynamic complex processes, which involves the continuous flow of information, materials, and funds across multiple functional areas both within and between chain members. Each member of the integrated supply chain is connected to other parts of the integrated chain by the flow of materials in one direction, the flow of information and money in the other direction. Changes in any one of these integrated chain members usually creates waves of influence that propogate throughout the integrated supply chains. These waves of influence are reflected in prices (both for raw materials, labor, parts, and finished product), flow of materials and product (within a single facility or between facilities within the supply chain), and inventories (of parts, labor capacity, and finished product). Besides its effectiveness, integrated supply chain management is a difficult process because of the stochastic and dynamic nature, multi-criterion and ever-increasing complexity of integrated supply chains. Due to highly complex nature of integrated supply chains, designing, analyzing and re-engineering of integrated supply chain processes using formal and quantitative approaches seems to be very difficult (Jain *et al.*, 2006, Ding *et al.,* 2006).

Several researchers, such as Evans *et al.*, 1995, Vander Aalst, 1998, Lin and Shaw 1998, etc. have developed some frameworks and models to design and analyze the supply chain processes. These models are either oversimplified or just qualitatively described (some of them are based on simulation study (Bhaskaran, 1998) and are difficult to apply for evaluating real supply chains with quantitative analysis and decisions. Because today's manufacturing enterprises are more strongly coupled in terms of material, information and service flows, there exists a strong urge for a process-oriented approach to address the issues of integrated modeling and analysis (Ding *et al.*, 2006, Jain *et al.* 2006, 2007a). Many of the past studies neglected significant impacts of such integration issues because of dramatic increase in modeling complexity. Therefore, models from past studies are confined in their capability and applicability to analyze real supply chain processes. An integrated formal and quantitative model, addressing the above mentioned issues that allows supply chain managers to quickly evaluate various design and operation alternatives with satisfactory accuracy, has become imperative (Jain *et al.,* 2007b).

Moreover, the need for agility for competitiveness has traditionally been associated with the integrated supply chains that provide and manufacture innovative products, such as high-technology industry products characterized by shortened life-cycles, a high degree of market volatility, uncertainty in demand, and unreliability in supply. Similarly, traditional, more slow moving industries face such challenges in terms of requirements for speed, flexibility, increased product diversity and customization. The next section discusses more in detail why the need for agile integrated supply chain?

3. Why agile integrated supply chain?

Agility – namely, the ability of a supply chain to rapidly respond to changes in market and customer demands – is regarded as the bearer of competitive advantage in today's business world (Yusuf *et al.,* 2004, Christopher & Towill, 2001, Gunasekaran, 1999). Based on a survey of past decade management literature, van Hoek (2001) identify the two most significant lessons for achieving competitive advantage in the modern business environment. The first lesson is that companies have to be aligned with suppliers, the suppliers' of the suppliers, customers and the customers' of the customers, even with the competitors, so as to streamline operations (Simchi-Levi *et al.*, 2003). As a result, individual companies no longer compete solely as autonomous entities; rather, the competition is between rival supply chains, or more like closely coordinated, cooperative business networks (Christopher, 1998, Lambert *et al.*, 1998). The second lesson is that within the supply chain, companies should work together to achieve a level of agility beyond the reach of individual companies. *All*

companies, suppliers, manufacturers, distributors, and even customers, may have to be involved in the process of achieving an agile supply chain (Christopher, 2000, Christopher and Towill, 2001).
Furthermore, "Agility" includes "Leanness" because a high stock or spare capacity method of providing flexibility to changing customer demands or adversity is not a viable financial option. *Since, agile manufacturing incorporates all the elements of lean manufacturing and thus lean and agile supply chains have commonality of characteristics except that the latter ascribes to additional principles and practices, which enhances its capability to balance both predictable and unpredictable changes in market demands* (Yusuf *et al.,* 2004). In a changing competitive environment, there is a need to develop supply chains and facilities significantly more flexible and responsive than existing ones. It is essential that supply chains continually re-examine how they can compete and agility is one of the underlying paradigms to enable them to re-invent the content and processes of their competitive strategy. In agility, therefore, lies the capability to survive and prosper by reacting quickly and effectively to changing markets. As a result, more recently, the agile manufacturing paradigm has been highlighted as an alternative to, and possibly an improvement on, leannessAn agile supply chain is seen as a dominant competitive advantage in today's business; however, the ability to build an agile supply chain has developed more slowly than anticipated (Lin *et al*., 2006).

Based on a survey of past decade management literature, van Hoek (2001) identify the two most significant lessons for achieving competitive advantage in the modern business environment. One lesson is that companies have to be aligned with suppliers, the suppliers' of the suppliers, customers and the customers' of the customers, even with the competitors, so as to streamline operations (Simchi-Levi *et al.,* 2003). As a result, individual companies no longer compete solely as autonomous entities; rather, the competition is between rival supply chains, or more like closely coordinated, cooperative business networks (Christopher 1998, Lambert *et al.* 1998). Another lesson is that within the supply chain, companies should work together to achieve a level of agility beyond the reach of individual companies (van Hoek, 2001). All companies, suppliers, manufacturers, distributors, and even customers, may have to be involved in the process of achieving an integrated agile supply chain (Christopher, 2000, Christopher & Towill, 2001).

The need for agility for competitiveness has traditionally been associated with the supply chains that provide and manufacture innovative products, such as high-technology industry products characterized by shortened life-cycles, a high degree of market volatility, uncertainty in demand, and unreliability in supply. Similarly, traditional, more slow moving industries face such challenges in terms of requirements for speed, flexibility, increased product diversity and customization. Consequently, the need for agility is becoming more prevalent. These demands come, typically, from further down the supply chain in the finishing sector, or from end customers (Gunasekaran & Ngai, 2004). Some traditional companies have already elements of agility because the realities of a competitive environment dictate these changes (e.g. in sectors such as automobiles, food, textiles, chemicals, precision engineering and general engineering) (Christian *et al.*, 2001). According to Christian *et al.* (2001), this is, however, usually outside any strategic vision and is approached in an ad-hoc fashion. *The lack of a systematic approach to agility does not allow companies to develop the necessary proficiency in change, a prerequisite for agility (Lin et al., 2006).*

Kidd (1994) stated that Supply Chain Management (SCM) is a fairly well defined topic, but agility is not so well defined. Agility can be something that companies achieve without realizing it, or it can relate to issues that are difficult to quantify. The nature of the competencies implied by agility is such that they would be better considered as intangibles, similar to intellectual property, company specific knowledge, skills, expertise, etc. *In summary, SCM and agility combined are significant sources of competitiveness in the business world. Thus, it is no surprise that they are favored research areas in the academic research world (Yusuf et al., 2004, Swafford et al., 2006).*

The fact that agile attributes are necessary but not sufficient conditions for agility points to a major research issue to be addressed (Yusuf and Burns, 1999). It is essential that the attributes are transformed into strategic competitive bases of speed, flexibility, proactivity, innovation, cost, quality, profitability and robustness. More importantly, these attributes are of very little significance to practitioners unless there is a way of deploying them. In addition, the changing nature of the market requirements suggests the need for a dynamic deployment tool for evaluating agility. Integrated supply chains have realized that agility is essential for their survival and competitiveness. Consequently, there is no generally accepted method by researchers and practitioners for designing, operating and evaluating agile supply chains. Moreover, *the ability to build agile supply chain has developed more slowly than anticipated, because technology for managing agile supply chain is still being developed.*

Based on a synthesis of the literature (Sharp *et al.,* 1999, Yusuf *et al.,* 1999, Jharkaria and Shankar, 2005) and interviews of several industrial partners in the EU-I*Proms project (www.Iproms.org), the following critical questions and extracted motivations form the basis of this research work:

Some critical questions

Question 1: What precisely is agility/leanness and how it can be measured?

Question 2: How to develop an integrated agile/lean supply chain?

Question 3: How will lean and agile supply chains know what they have it, as there are no simple metrics or indexes available?

Question 4: How and to what degree does the integrated lean and agile supply chain attributes affect supply chains business performance?

Question 5: How to compare agility/leanness with competitiveness?

Question 6: How can the integrated supply chains identify the principal obstacles to improvement, if a supply chain wants to improve agility and leanness?

Question 7: How to assist in achieving agility/leanness effectively?

Some extracted motivations

Motivation 1: All companies, suppliers, manufacturers, distributors, and even customers, may have to be involved in the process of achieving an agile supply chain (Christopher, 2000, Christopher & Towill, 2001).

Motivation 2: The lack of a systematic approach to agility does not allow companies to develop the necessary proficiency in change, a prerequisite for agility (Lin *et al.,* 2006).

Motivation 3: SCM and agility combined are significant sources of competitiveness in the business world. Thus, it is no surprise that they are favored research areas in the academic research world (Yusuf *et al.,* 2004, Swafford *et al.,* 2006).

Motivation 4: Most agility measurements are described subjectively by linguistic terms, which are characterized by ambiguity and multi-possibility. Thus, the scoring of the existing techniques can always be criticized, because the scale used to score the agility capabilities has limitations (Lin *et al.,* 2006).

Motivation 5: The fact that agile attributes are necessary but not sufficient conditions for agility points to a major research issue to be addressed (Yusuf & Burns, 1999). It is essential that the attributes are transformed into strategic competitive bases of speed, flexibility, proactivity, innovation, cost, quality, profitability and robustness.

Motivation 6: There is no methodology and tools for introducing and implementing such a complex and dynamic interactive system which incorporate both quantitative and qualitative attributes as agile supply chains (Lin *et al.,* 2006).

Motivation 7: Recently, the use of intelligent agents for supply chain management has received great attention as agent technology is the preferable technology for enabling a flexible and dynamic coordination of spatially distributed entities in integrated supply chains (Swaminathan *et al.*, 1998).

Motivation 8: Fuzzy logic provides a useful tool to deal with problems in which the attributes and phenomena are imprecise and vague (Zadeh, 1965).

Motivation 9: Relational databases have been widely used in support of business operations, and there the size of database has grown rapidly, for the agility of decision making and market prediction for varying degree of importance for agility evaluation, knowledge discovery from a database is very important for sustaining essential information to a business (Berry & Linoff, 1997).

Motivation 10: Association rules are one of the ways of representing knowledge, having been applied to scrutinize market baskets to help managers and decision makers understand which item/ratings are likely to be preferred at the same time (Han *et al.,* 2000).

4. New approaches

Motivated by the above extracted motivations and to find the answers to the aforementioned questions, which are critical to the practitioners and to the theory of integrated agile supply chains design, in this section, we will discuss two novel approaches for modeling and evaluating agility in dynamic integrated supply chains (Jain *et al.,* 2008a,b).

4.1 Fuzzy intelligent based approach

In this section, we discuss a novel approach to model agility (which includes leanness) and introduce *Dynamic Agility Index* through fuzzy intelligent agents. Generally, it is difficult to emulate human decision making if the recommendations of the agents are provided as crisp, numerical values. The multiple intelligent agents used in this study communicate their recommendation as fuzzy numbers to accommodate ambiguity in the opinion and the data used for modeling agility attributes for integrated supply chains. Moreover, when agents operate based on different criteria pertaining to agility like flexibility, profitability, quality, innovativeness, pro-activity, speed of response, cost, robustness etc for integrated supply chains, the ranking and aggregation of these fuzzy opinions to arrive at a consensus is complex. The proposed fuzzy intelligent agents approach provides a unique and unprecedented attempt to determine consensus in these fuzzy opinions and effectively model dynamic agility.

As producers, wholesalers and retailers seek more effective ways of marketing their products, they increasingly examine their supply chains for ways to reduce costs. Strategic planning of performance improvement is gaining attention in all areas of manufacturing. The reason for that is that it takes into account the long-term interest of the company in determining suitable business and operational policies. The agility in supply chains is determined by certain time variables, which we refer to here as 'agility characteristics'. These characteristics evolve in time and determine the entire behavior of the supply chains, refer Figure 1. The rate of change of these characteristics is a function of the current values of all the attributes as well as some suitable 'input' variables, like the size and numbers of teams, refereed as team formation, the level of integration of the database.

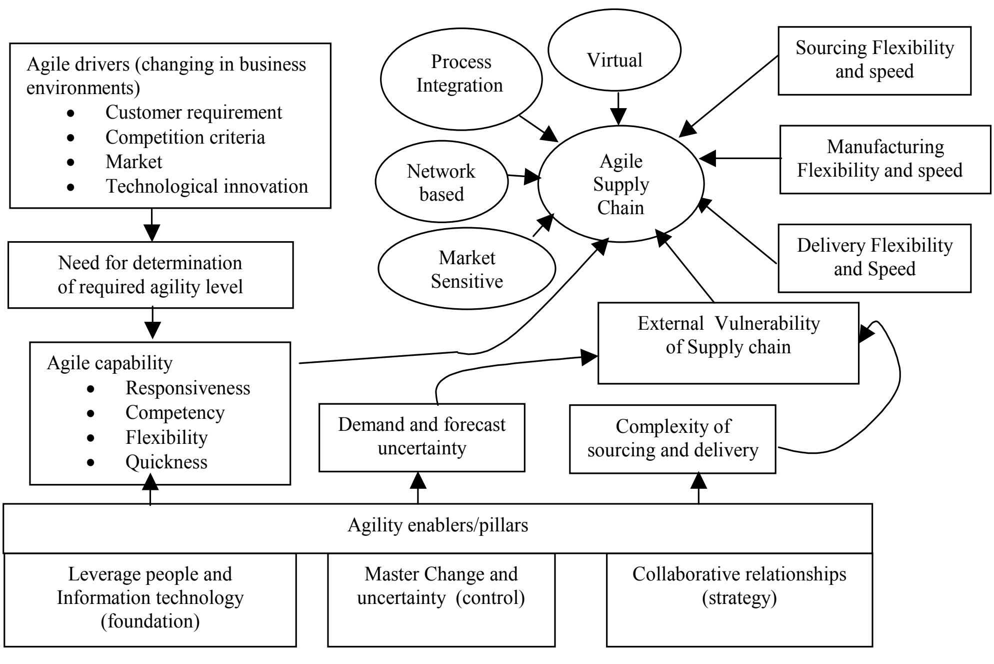


Fig. 1. The conceptual model for agile supply chains

The proposed dynamic agility index (DA_{Li}) of an integrated supply chain can be given a numerical value calculated as the sum of the products of suitable 'economical bases', i.e.

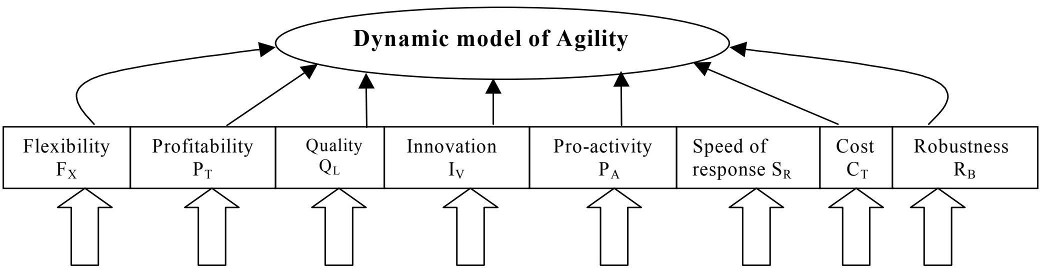
$$DA_{Li} = W_1 \times F_X + W_2 \times P_T + W_3 \times Q_L + W_4 \times I_V + W_5 \times P_R + W_6 \times S_R + W_7 \times C_T + W_8 \times R_B$$

Where:

- F_X is a measure of Flexibility, and W_1 is a weight assumed constant but time varying in general,
- P_T is a measure of Profitability, and W_2 is a weight assumed constant but time varying in general,
- Q_L is a measure of Quality, and W_3 is a weight assumed constant but time varying in general,
- I_V is a measure of Innovation, and W_4 is a weight assumed constant but time varying in general,

- P_R is a measure of Profitability, and W_5 is a weight assumed constant but time varying in general,
- S_R is a measure of Speed of response, and W_6 is a weight assumed constant but time varying in general,
- C_T is a measure of Cost, and W_7 is a weight assumed constant but time varying in general,
- R_B is a measure of Robustness, and W_8 is a weight assumed constant but time varying in general,

The dynamic agility index model considered in this research is shown in Figure 2.



Agility enablers			
Dimensions	Related attributes	Dimensions	Related attributes
Integration	Concurrent execution of activities Enterprise information Information accessible to employees	Team building	De-centralized decision making Empowered individuals working in teams Cross-functional team Teams across company borders
Change	Culture of change Continuous improvement	Education	Response to changing market requirements New product introduction Customer driven innovations Customer satisfaction
Competence	Business practice and structure are difficult to replicate Multi-venturing capabilities	Technology	Technology awareness Leadership in the use of current technology Skill and knowledge enhancing technologies Flexible production technology
Partnership	Trust-based relationship with customers/suppliers Rapid partnership formation Strategic relationships with customers Close relationship with suppliers	Market	Response to changing market requirements New product introduction Customer-driven innovations Customer satisfaction
Welfare	Employee satisfaction	Quality	Quality over product life Products with substantial added value First-time right design Short development cycle times

Fig. 2. The proposed dynamic model for agile supply chains

The mathematical model developed is based on dynamical systems theory and recognizes that the integrated supply chains attributes have evolutionary approaches. Therefore, a new generation tools should be developed and the existing tools significantly enhanced to support decision-making processes and to deliver required solutions to extended businesses.

Now, we present the various steps of the proposed Fuzzy Intelligent agent based approach to study and model agility for integrated supply chains. More details of the proposed approach can be found in (Jain *et al.,* 2008a).

Step 1: Select criteria for evaluation. We have listed several important criteria including: Flexibility (F_X), Profitability (P_T), Quality (Q_L), Innovation (I_V), Pro-activity (P_R), Speed of response (S_R), Cost (C_T), Robustness (R_B).

"These selected eight criteria's and their possible combinations abbreviated as (C_0, C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8) are listed in Table 1. The agility of integrated supply chains can be given a numerical value calculated as the sum of the products of the aforementioned criteria and their possible combinations as given in Table 1. The eight criteria's listed above are by no means exhaustive and therefore new factors may be added depending on the product, industry and market characteristics."

Step 2: Determine the appropriate linguistic scale to assess the performance ratings and importance weights of the agility capabilities.

"Noteworthy, many popular linguistic terms and corresponding membership functions have been proposed for linguistic assessment. In addition, the linguistic variables selected to assess the importance weights of the agility capabilities are {Very High (VH), High (HG), Fairly High (FH), Medium (M), Fairly Low (FL), Low (L), Very Low (VL)}."

Step 3: Measure the importance and the performance of agility capabilities using linguistic terms.

"Once the linguistic variables for evaluating the performance ratings and the importance weights of the agility capabilities are defined, according to the supply chains policy and strategy, profile, characteristics, business changes and practices, marketing competition information, the agents can directly use the linguistic terms above to assess the rating which characterizes the degree of the performance of various agility capabilities. The results, integrated performance ratings and integrated importance weights of agility capabilities measured by linguistics variables, are shown in Table 2."

Step 4: Approximate the linguistic terms by fuzzy numbers.

"We perform trapezoidal approximations of fuzzy numbers. Tapping the properties of trapezoidal fuzzy numbers, a set of fuzzy numbers for approximating linguistic variable values was developed as shown in Table 3."

Step 5: Cumulate fuzzy opinions with fuzzy weights.

"Several aggregation techniques require that the fuzzy opinions have some intersection so that they are not entirely out of agreement. In case, the opinions do not have some agreement, the agents negotiate until they can arrive at a consensus. However, these methods will not be considered, as agents assumed in this research may intentionally have disparate recommendations due to their diverge viewpoints for supply chain management. Weighted linear interpolation is used to aggregate the opinions for every alternative, incase, there is no common interaction between agent opinions."

Combination C_0 of criteria	Combination C_1 of criteria	Combination C_2 of criteria	Combination C_3 of criteria	Combination C_4 of criteria	Combination C_5 of criteria	Combination C_6 of criteria	Combination C_7 of criteria	Combination C_8 of criteria
Flexibility (F_X) Profitability (P_T) Quality (Q_L) Innovation (I_V) Pro-activity (P_R) Speed & response (S_R) Cost (C_T) Robustness (R_B)	$F_X P_T$ $F_X Q_L$ $F_X I_V$ $F_X P_R$ $F_X S_R$ $F_X C_T$ $F_X R_B$	$P_T Q_L$ $P_T I_V$ $P_T P_R$ $P_T S_R$ $P_T C_T$ $P_T R_B$	$Q_L I_V$ $Q_L P_R$ $Q_L S_R$ $Q_L C_T$ $Q_L R_B$	$I_V P_R$ $I_V S_R$ $I_V C_T$ $I_V R_B$	$P_R S_R$ $P_R C_T$ $P_R R_B$	$S_R C_T$ $S_R R_B$	$C_T R_B$	$F_X P_T Q_L I_V P_R S_R C_T R_B$

Table 1. Criteria's for modeling dynamic agility

Criteria									Weight									Rank
C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	W_0	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	$R_{0\text{-}8}$
F_X	$F_X P_T$	$P_T Q_L$	$Q_L I_V$	$I_V P_R$	$P_R S_R$	$S_R C_T$	$C_T R_B$	$F_X P_T Q_L I_V P_R S_R C_T R_B$	VH	VH	VH	FH	H	VH	FH	H	VH	EP
P_T	$F_X Q_L$	$P_T I_V$	$Q_L P_R$	$I_V S_R$	$P_R C_T$	$S_R R_B$			H	H	VH	H	FH	H	VH			VG
Q_L	$F_X I_V$	$P_T P_R$	$Q_L S_R$	$I_V C_T$	$P_R R_B$				VH	VH	H	VH	VH	VH				GD
I_V	$F_X P_R$	$P_T S_R$	$Q_L C_T$	$I_V R_B$					H	FH	FH	VH	FH					FR
P_R	$F_X S_R$	$P_T C_T$	$Q_L R_B$						FH	VH	H	FH						GD
S_R	$F_X C_T$	$P_T R_B$							H	M	VH							FH
C_T	$F_X R_B$								VH	FH								VG
R_B									FH									GD

Table 2. Aggregated performance rating with aggregated important weight for selected agility criteria

Criteria									Rank
C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	$R_{0\text{-}8}$
F_X	$F_X P_T$	$P_T Q_L$	$Q_L I_V$	$I_V P_R$	$P_R S_R$	$S_R C_T$	$C_T R_B$	$F_X P_T Q_L I_V P_R S_R C_T R_B$	(7, 8, 9, 10)
P_T	$F_X Q_L$	$P_T I_V$	$Q_L P_R$	$I_V S_R$	$P_R C_T$	$S_R R_B$			(5, 6.04, 7, 8)
Q_L	$F_X I_V$	$P_T P_R$	$Q_L S_R$	$I_V C_T$	$P_R R_B$				(3.49, 4.51, 5.50, 6.52)
I_V	$F_X P_R$	$P_T S_R$	$Q_L C_T$	$I_V R_B$					(2.52, 3.50, 4.50, 5.56)
P_R	$F_X S_R$	$P_T C_T$	$Q_L R_B$						(3.50, 4.50, 5.50, 6.50)
S_R	$F_X C_T$	$P_T R_B$							(5, 6, 7, 8)
C_T	$F_X R_B$								(3.52, 4.50, 5.48, 6.25)
R_B									(5, 6, 7, 8)

Table 4. Ranks of dynamic agility index for selected agility criteria

Performance rating		Importance weighting	
Linguistic variable	Fuzzy number	Linguistic variable	Fuzzy number
Worst (WT)	(0, 0.05, 0.25, 1.25)	Very Low (VL)	(0, 0.005, 0.025, 0.125)
Very Poor (VP)	(1, 2, 3, 4)	Low (LW)	(0.1, 0.2, 0.3, 0.4)
Poor (PR)	(1.5, 2.5, 3.5, 4.5)	Fairly Low (FL)	(0.15, 0.25, 0.35, 0.45)
Fair (FR)	(2.5, 3.5, 4.5, 5.5)	Medium (MD)	(0.25, 0.35, 0.45, 0.55)
Good (GD)	(3.5, 4.5, 5.5, 6.5)	Fairly High (FH)	(0.35, 0.45, 0.55, 0.65)
Very Good (VG)	(5, 6, 7, 8)	High (HG)	(0.5, 0.6, 0.7, 0.8)
Exceptional (EP)	(7, 8, 9, 10)	Very High (VH)	(0.7, 0.8, 0.9, 1.0)

Table 3. Fuzzy numbers for approximating linguistic variables for selected agility criteria

Each agent, ξ, is assigned a rating, ψ_ξ. The most crucial agent is specified a rating of 1 and the others are given ratings less than 1, in relation to their significance. To the ratings the following properties holds:

Maximum ($\psi_1, \psi_2, \psi_3 \ldots, \psi_\delta$) = 1

Minimum ($\psi_1, \psi_2, \psi_3 \ldots, \psi_\delta$) < 1

The degree of significance (DOS) is defined as:

$$\mathrm{DOS} = \Pi_\xi = \frac{\psi_\xi}{\sum_{\xi=1}^{\delta} \psi_\xi} \qquad \xi = 1, 2, 3, \ldots, \delta \tag{1}$$

The cumulated fuzzy opinion for alternative η is formed as a Trapezoidal fuzzy number (TFN) tuple ($\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3, \bar{\lambda}_4$) using formulas:

$$\begin{cases} \bar{\lambda}_1 = \sum_{\xi=1}^{\delta} \Pi_\xi \lambda_{1\xi}, \ \bar{\lambda}_2 = \sum_{\xi=1}^{\delta} \Pi_\xi \lambda_{2\xi}, \\ \bar{\lambda}_3 = \sum_{\xi=1}^{\delta} \Pi_\xi \lambda_{3\xi}, \bar{\lambda}_4 = \sum_{\xi=1}^{\delta} \Pi_\xi \lambda_{4\xi} \end{cases} \tag{2}$$

where: δ is the number of agents with opinions on alternatives η, Π_ξ corresponds to the degree of significance of agent ξ and $(\lambda_{1\xi}, \lambda_{2\xi}, \lambda_{3\xi}, \lambda_{4\xi})$ symbolizes TFN opinion of agent ξ for alternative η. The resulting inferred aggregated opinion ($\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3, \bar{\lambda}_4$) can be represented as:

$$\left(RI_A\right)^* = \sum_{\xi=1}^{\delta} \Pi_\xi (\circ) R^* \tag{3}$$

where R^* = ($\lambda_{1\xi}, \lambda_{2\xi}, \lambda_{3\xi}, \lambda_{4\xi}$) and $(\circ)$ is the fuzzy multiplication operator.

Thus, the trapezoidal fuzzy membership function is used to determine the agility level and the required fuzzy index of the selected criteria can be calculated using equation (3).

$$R_0=\frac{\begin{bmatrix}(7,8,9,10)\otimes(0.7,0.8,0.9,1.0)\oplus(7,8,9,10)\otimes(0.7,0.8,0.9,1.0)\\ \oplus(7,8,9,10)\otimes(0.7,0.8,0.9,1.0)\oplus(7,8,9,10)\otimes(0.35,0.45,0.55,0.65)\\ \oplus(7,8,9,10)\otimes(0.5,0.6,0.7,0.8)\oplus(7,8,9,10)\otimes(0.7,0.8,0.9,1.0)\\ \oplus(7,8,9,10)\otimes(0.35,0.45,0.55,0.65)\oplus(7,8,9,10)\otimes(0.5,0.6,0.7,0.8)\end{bmatrix}}{\begin{bmatrix}(0.7,0.8,0.9,1.0)\oplus(0.7,0.8,0.9,1.0)\oplus(0.7,0.8,0.9,1.0)\\ \oplus(0.35,0.45,0.55,0.65)\oplus(0.5,0.6,0.7,0.8)\oplus(0.7,0.8,0.9,1.0)\\ \oplus(0.35,0.45,0.55,0.65)\oplus(0.5,0.6,0.7,0.8)\end{bmatrix}}=(7,8,9,10)$$

Applying the same equation the other fuzzy indexes of agility criteria are obtained as listed in Table 4. Finally, applying the same equation again, we calculate the proposed Dynamic Agility level index (DA_{Li}) for modeling agility for integrated supply chains with the taken 8 criteria and their all possible combinations is evaluated as:

$$DA_{Li}=\frac{\begin{bmatrix}(7,8,9,10)\otimes(0.7,0.8,0.9,1.0)\\ \oplus(5,6.04,7,8)\otimes(0.5,0.6,0.7,0.8)\\ \oplus(3.49,4.51,5.5,6.52)\otimes(0.7,0.8,0.9,1.0)\\ \oplus(2.52,3.5,4.5,5.56)\otimes(0.5,0.6,0.7,0.8)\\ \oplus(3.5,4.5,5.5,6.5)\otimes(0.35,0.45,0.55,0.65)\\ \oplus(5,6,7,8)\otimes(0.5,0.6,0.7,0.8)\\ \oplus(3.52,4.5,5.48,6.25)\otimes(0.7,0.8,0.9,1.0)\\ \oplus(5,6,7,8)\otimes(0.35,0.45,0.55,0.65)\end{bmatrix}}{\begin{bmatrix}(0.7,0.8,0.9,1.0)\oplus(0.5,0.6,0.7,0.8)\\ \oplus(0.7,0.8,0.9,1.0)\oplus(0.5,0.6,0.7,0.8)\\ \oplus(0.35,0.45,0.55,0.65)\oplus(0.5,0.6,0.7,0.8)\\ \oplus(0.7,0.8,0.9,1.0)\oplus(0.35,0.45,0.55,0.65)\end{bmatrix}}=(4.544,5.486,6.352,6.982)$$

Step 6: Rank the fuzzy opinions.

"The superior alternative must be chosen, once the opinions of the agents have been aggregated to produce a consensus opinion for each alternative. The findings of Nakamura (1986) emphasize a fuzzy preference function that outline a comparison index, which compares opinions k_i and k_j that accounts for the hamming distance of every fuzzy number to the fuzzy minimum and the fuzzified best and worst states."

The FFCF is defined as:

$$\mu_p(K_i,K_j)=\begin{cases}\dfrac{1}{\varpi_\beta}\begin{bmatrix}\beta\chi\left(K_{i*},K_{i*}\wedge K_{j*}\right)\\ +\left(1-\beta\right)\chi\left(K_i^*,K_i^*\wedge K_j^*\right)\end{bmatrix} & if\ \ \varpi_\beta\neq 0\\ \dfrac{1}{2} & if\ \ \varpi_\beta=0\end{cases}\tag{4}$$

where :

$$\varpi_\beta=\beta\left[\chi\left(K_{i*},K_{i*}\wedge K_{j*}\right)+\chi\left(K_{j*},K_{i*}\wedge K_{j*}\right)\right]+(1-\beta)\left[\chi\left(K_i^*,K_i^*\wedge K_j^*\right)+\chi\left(K_j^*,K_i^*\wedge K_j^*\right)\right]$$

$$\mu_{K^*}(\phi) = \underset{\{\theta/\theta \geq \phi\}}{Sup} \mu_K(\theta) \qquad \forall \phi \in V \tag{5}$$

Further, K_* is the highest upper set of K defined by:

$$\mu_{K_*}(\phi) = \underset{\{\theta/\theta \leq \phi\}}{Sup} \mu_K(\theta) \qquad \forall \phi \in V \tag{6}$$

$K_i \wedge K_j$ is the extended minimum defined by:

$$\mu_{K_i \wedge K_j}(\sigma) = \underset{\{\theta,\phi/\theta \wedge \phi = \sigma\}}{Sup} [\mu_{K_i}(\theta) \wedge \mu_{K_j}(\phi)] \forall \sigma = V \tag{7}$$

and the Hamming distance between K_i and K_j is given by $\chi(K_i, K_j)$, which is

$$\chi(K_i, K_j) = \int_{\varepsilon} \left| \mu_{K_i}(\theta) - \mu_{K_j}(\theta) \right| d\theta \tag{8}$$

Theoretically, $\chi\left(K_{i*}, K_{i*} \wedge K_{j*}\right)$ and $\chi\left(K_i^*, K_i^* \wedge K_j^*\right)$ signifies the advantages of K_i over K_j with respect to the fuzzified worst states and the fuzzified best states. The fraction of the weighted combination of the advantages of K_i and K_j over the worst states and the above the best states, to the sum of such weighted combinations of K_i's and s's is represented by the fuzzy first choice function (FFCF), $\mu_p(K_i, K_j)$.

In this chapter, the fuzzy first choice function compares every fuzzy opinion to a "Standard" fuzzy number, which demonstrates the case where the opinion is "Most Likely". Hence, the difficulty with existing methods suffers when comparing fuzzy numbers with identical modes and symmetric spreads is eliminated. Also, in this chapter, the fuzzy opinions are not only judge against "Most Likely" fuzzy numbers but also are already ranked in contrast to this value, thus eliminating the procedure of determining the ranking based on pairwise comparison. The result of every fuzzy first choice calculation for every node presents its ranking. The FFCF evaluating opinion K_i and the most likely mode, M, substitutes the second fuzzy opinion with M and is defined as:

$$\mu_p(K_i, M) = \begin{cases} \dfrac{1}{\varpi_\beta} \begin{bmatrix} \beta\chi\left(K_{i*}, K_{i*} \wedge K_{j*}\right) \\ +(1-\beta)\chi\left(K_i^*, K_i^* \wedge M^*\right) \end{bmatrix} & if \ \ \varpi_\beta \neq 0 \\ \dfrac{1}{2} & if \ \ \varpi_\beta = 0 \end{cases} \tag{9}$$

The FFCF can be simplified by showing that $\chi\left(K_i^*, K_i^* \wedge M^*\right) = 0$, when M is a TFN defined as $(\lambda_1, \lambda_2, 1, 1)$. Thus, if M is signified by $(\lambda_1, \lambda_2, 1, 1)$, the modified fuzzy first choice function used to evaluate opinion K_i with the most likely mode, M, is defined as:

$$\mu_p(K_i, M) = \begin{cases} \dfrac{1}{\varpi_\beta} \beta\chi\left(K_{i*}, K_{i*} \wedge M^*\right) & if \ \ \varpi_\beta \neq 0 \\ \dfrac{1}{2} & if \ \ \varpi_\beta = 0 \end{cases} \tag{10}$$

where $\varpi_\beta = \beta\left[\chi(K_{i^*}, K_{i^*} \wedge M^*) + \chi(M_*, K_{i^*} \wedge M^*)\right] + (1-\beta)\chi(M_*, K_{i^*} \wedge M^*)$

This fuzzy first choice function is able to distinguish between fuzzy numbers with identical modes and symmetric spreads while reducing the computational complexity.

Step 7: Match the fuzzy opinions with an appropriate agility level.

"In this case the natural language expression set selected is given as: Exceedingly Agile (EA), Very Agile (VA), Agile (AG), Fairly Agile (FA), Most Likely Agile (MLA), Slowly Agile (SA), No Agile (NA). "

The Euclidean distance ED is calculated by using the Euclidean distance formula as given in Equation (11) below:

$$ED(AG_L, F_N) = \left(\sum_{x \in P}\left(f_{AG_L}(x) - f_{F_L}(x)\right)^2\right)^{\frac{1}{2}} \tag{11}$$

Where $P = \{x_0, x_1, ..., x_m\} \subset [0, 10]$ so that $0 = x_0 < x_1 < ... < x_m = 10$.

The ED for the selected set of natural expression set is given as: ED (EA)= 1.2364, ED(VA)= 0.0424, ED(AG)= 1.0241, ED(FA)= 1.1462, ED(MLA)= 1.5321, ED(SA)= 1.6422 and ED(NA)= 1.8041.Thus, by matching a linguistic label with the minimum ED, dynamic agility can be modeled with the given criteria's. From the numerical example given in (Jain *et al.,* 2008a), it can be seen that the selected eight criteria (F_X, P_T, Q_L, I_V, P_R, S_R, C_T, R_B), the supply chain falls under the Very Agile (VA) category. Depending on the selected criteria, for any supply chains, the proposed approach will help the decision makers and analysts in quantifying agility.

Step 8: Analyze and classify the main obstacles to improvement.

"Modeling agility not only measures how agile is integrated supply chain, but also most importantly helps supply chain decision makers and practitioners to assess distinctive competencies and identify the principal obstacles for implementing appropriate improvement measures. In supply chain network, the factual environment of the problem engrosses statistics, which is repeatedly fuzzy and indefinite. This is primarily owing to its imprecise interfaces and its real-world character, where uncertainties in activities starting raw material procurement to the end consumer make the supply chain unfocused. As customer's demands are always uncertain, manufacturers tend to manage their suppliers in different ways leading to a supplier-supplier development, supplier evaluation, supplier selection, supplier association, supplier coordination etc."

However, it is difficult to emulate human decision making if the recommendations of the agents are provided as crisp, numerical values. Intelligent agents must express their opinions in similar terms to emulate human experts. Moreover at times, the agents make their recommendations based upon incomplete or unreliable data. A second problem arises when intelligent agents base their opinions on different viewpoints. The proposed approach provides an overall picture about the possibly agility of an integrated supply chain. Although, the dynamic agility index is conveyed in a range of values, the proposed approach ensures that the decision made in the selection using the fuzzy intelligent agents will not be biased.

4.2 Fuzzy association rules mining based approach

As a second approach, we present a Fuzzy Association Rule Mining based approach to support the decision makers by enhancing the flexibility in making decisions for evaluating agility with both tangibles and intangibles attributes/criteria such as Flexibility, Profitability, Quality, Innovativeness, Pro-activity, Speed of response, Cost and Robustness. Also, by checking the fuzzy classification rules, the goal of knowledge acquisition can be achieved in a framework in which evaluation of agility could be established without constraints, and consequently checked and compared in several details. More details of the proposed approach can be found in (Jain *et al.,* 2008b).

Mining association rules is one of the most important research problems in data mining. Many organizations have devoted a tremendous amount of resources to the construction and maintenance of large information databases over recent decades, including the development of large scale data warehouses. Frequently the data cannot be analyzed by standard statistical methods, either because there are numerous missing records, or because the data are in the form of qualitative rather than quantitative measures.

In many cases, the information contained in these databases is undervalued and underutilized because the data cannot be easily accessed or analyzed. Some databases have grown so large that even the system administrators do not always know what information might be represented or how relevant it might be to the questions at hand. Data sets commonly contain some an uncertain, particularly incompleteness and inconsistency. One example is a distributed information environment, where data sets are generated and collected from different sources, and each source may have different constraints. This can lead to different interrelationships among the items, thus imposing vagueness on the data set. Recent years have witnessed many efforts on discovering fuzzy associations, aimed at coping with fuzziness in knowledge representation and decision support process. Therefore, the necessity of applying Fuzzy Logic in data mining is due to the following:

- One is that fuzziness is inherent in many problems of knowledge representation, and the other is that high-level managers or complex decision processes often deal with generalized concepts and linguistic expressions, which are generally fuzzy in nature.
- Moreover fuzziness may prevail in many other association cases in which impression, matching, similarity, implication, partial truth or the like is present.
- The modeling of imprecise and qualitative knowledge, as well as the transmission and handling of uncertainty at various stages are possible through the use of fuzzy sets.
- Fuzzy logic is capable of supporting to a reasonable extent, human type reasoning in natural form.

A method to find the large itemsets and also an apriori algorithm is proposed in the literature (Agarwal *et al.,* 1996). However, to find the large itemsets, these algorithms should scan the database several times. Also, while they generated a candidate itemset, the apriori-gen function must have exhausted a good deal of time to confirm, if its subsets are large or not. Further, the well known methods viz. Partial completeness (Srikant and Agarwal 1996), Optimized association rules (Fukuda *et al.,* 1996) and CLIQUE (Agarwal *et al.,* 1998), divided the qualitative attributes into many crisps partitions. There were no interactions between the partitions. However, crisp partitions may be unreasonable for some situations. For example, if we tried to partition the range (70, 80 $) of the attribute "COST" for a supplier, into two partitions, then separable point was not different between 75.01 and 74.99$. Hence, interaction of any of the neighborhood partitions can be promised. Moreover, we considered that the fuzzy association rules described by the natural language as well as suited for the thinking of human subjects and will help to increase the flexibility for users in making decisions or designing the fuzzy systems for evaluating agility. Hence, we use fuzzy partition method to find the fuzzy association rules.

Fuzzy partitioning in quantitative attributes

A quantitative attribute can be partitioned into 'L' various linguistic values (L=2, 3, 4….). For example, for the attribute 'cost' (range from 0 to 100), we describe L=2, L=3 in Figures 3 and 4 respectively.

Also, $\psi_{L,\phi_V}{}^{COST}$ can be used to represent a candidate 1-dim fuzzy framework.

Then $\mu_{L,\phi_V}{}^{COST}$ can be represented as follows:

$$\mu_{L,\phi_V}{}^{COST}(y) = Max\left\{1 - \frac{\left|y - \xi_{\phi_V}{}^{L}\right|}{\lambda^{L}}, 0\right\}$$

Where $\xi_{\phi_V}{}^{L} = Min_{AD} + \frac{(Max_{AD} - Min_{AD})(\phi_V - 1)}{(L-1)}$ and $\lambda^{L} = \frac{(Max_{AD} - Min_{AD})}{(L-1)}$.

Min_{AD} and Max_{AD} are the maximum and minimum of the attribute domain.

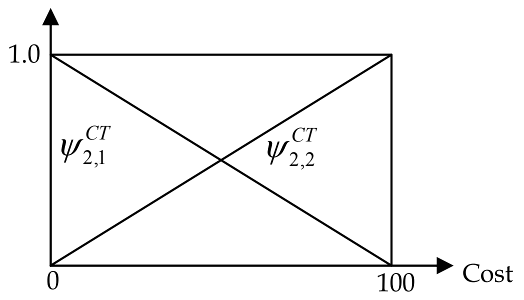


Fig. 3. L=2 for quantitative attribute cost for agility

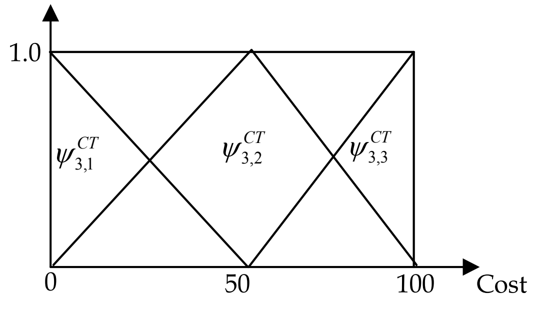


Fig. 4. L=3 for quantitative attribute cost for agility

Fuzzy partitioning in qualitative attributes

Qualitative attributes of a relational database have a finite number of possible values, with no ordering among several values. For example Flexibility (F_X), Profitability (P_T), Quality (Q_L), Innovation (I_V), Pro-activity (P_R), Speed of Response (S_R) and Robustness (R_B)). If the distinct attribute values are η' (η' is finite), then this attribute can only be partitioned by η' linguistic values. In the agility evaluation considered in this second approach, the linguistic sentences of each linguistic value defined by the attributed dependability can be stated as follows: $\psi_{2,1}^{FX} = Low$ and $\psi_{2,2}^{FX} = High$.

Each linguistic value distributed in either quantitative attribute (Cost) or qualitative attributes (Flexibility, Quality, Innovation, etc.) is considered as a potential candidate 1-dim fuzzy framework. The succeeding task is how to use these candidate 1-dim fuzzy frameworks to generate the other large fuzzy frameworks and fuzzy association rules.

Determine large fuzzy frameworks

Once all candidate 1-dim fuzzy frameworks have been generated, we need to determine how to find the other large fuzzy frameworks and fuzzy association rules. Figure 5 describes the proposed model for generating fuzzy association rules.

From figure 5, we can see that large fuzzy frameworks and fuzzy association rules are generated by stages 1 and 2 respectively. To evaluate the agility using fuzzy association rules, the algorithm is given as:

Algorithm

Given by the decision maker, the input comprises of the following specification:

1. A database containing several quantitative and qualitative attributes for evaluating agility.
2. The minimum $F_Z S_P$
3. The minimum $F_Z C_F$

The main algorithm operations comprises of 2 stages:

1. **Stage 1**: Generate large fuzzy frameworks
2. **Stage 2**: Generate effective fuzzy association rules and evaluate the agility

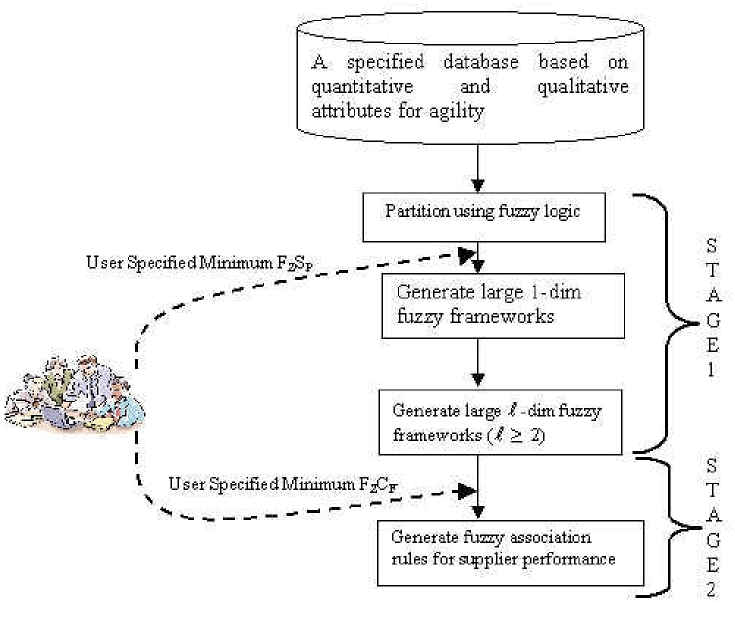


Fig. 5. Two-stage model for generating fuzzy association rules

These two stages are described in detail as following:

Stage 1 (comprises of three different steps)

Begin Step 1:

Step1.1: Generate large fuzzy frameworks

Step1.2: Perform fuzzy partition
Step1.3: Scan the database and construct the table comprising of F_ZF_T, O_PT and F_ZS_P
Step1.4: Generate large 1-dim fuzzy frameworks
Step1.5: Set $\ell = 1$ and eliminate the rows of initials (F_ZF_T, O_PT and F_ZS_P) corresponding to the candidate 1-dim fuzzy frameworks which are not large
Step 1.6: Reconstruct (F_ZF_T, O_PT and F_ZS_P)

Step 2: Generate large ℓ-dim fuzzy frameworks. Set ℓ +1 to ℓ. If there is only one (ℓ -1)-dim fuzzy framework, then go to **Step 3** within the same stage.

For any two unpaired rows F_ZF_T O_PT F_ZS_P [Δ] and F_ZF_T O_PT F_ZS_P [σ], where ($\Delta \neq \sigma$), corresponding to large (ℓ -1)-dim fuzzy frameworks **do**

Step 2.1: **If** any two linguistic values are defined in the same linguistic variable from (F_ZF_T [Δ] OR F_ZF_T [σ]) that corresponds to a candidate ℓ-dim fuzzy framework $\prod$, **then** Discard $\prod$, and skip steps 2.2, 2.3 and 2.4. That is, $\prod$ is not valid.

Step 2.2: **If** F_ZF_T [Δ] and F_ZF_T [σ] do not share ($\ell - 2$) linguistic terms, **then** discard $\prod$ and skip steps 2.3 and 2.4. That is, $\prod$ is invalid.

Step 2.3: **If** there exists integers $1 \leq \text{int}_1 < \text{int}_2 < \dots \text{int}_\ell$ such that (F_ZF_T [Δ] OR F_ZF_T [σ]) (int_1)= (F_ZF_T [Δ] OR F_ZF_T [σ]) (int_2)=…= (F_ZF_T [Δ] OR F_ZF_T [σ]) ($\text{int}_{\ell-1}$) = (F_ZF_T [Δ] OR F_ZF_T [σ]) (int_ℓ)=1, **then** compute [O_PT (int_1). O_PT (int_2)… O_PT (int_ℓ)] and the fuzzy support F_ZS_P of $\prod$.

Step 2.4: Add (F_ZF_T [Δ] OR F_ZF_T [σ]) to table F_ZF_T (O_PT [int_1]. O_PT [int_2]… O_PT [int_ℓ] to O_PT and F_ZS_P when F_ZS_P is $\geq$ Min F_ZS_P, otherwise discard $\prod$.

Step 3: Check whether or not any large ℓ-dim fuzzy framework is generated.

If any large ℓ-dim fuzzy framework is generated,
then go to **Step 2** (of stage 1)
else go to Stage 2.
It is noted that the final F_ZF_T O_PT F_ZS_P only stores large fuzzy frameworks.

End

Stage 2 (comprises of one step)
Begin Step 1:

Step 1.1: Generate effective fuzzy association rules

Step 1.2: **For** two unpaired rows, F_ZF_T [Δ] and F_ZF_T [σ] ($\Delta < \sigma$), corresponding to a large fuzzy frameworks LARΔ and LARσ respectively **do**

Step 1.2.1: Produce the antecedent part of the rule. Let $\hbar$ be the number of nonzero elements in F_ZF_T [Δ] AND F_ZF_T [σ]

Step 1.2.2: **If** the number of nonzero elements in F_ZF_T [Δ] $= \hbar$, **then** LAR$\Delta \subset$ LARσ is hold, and the antecedent part of one rule, say R, is generated as LARσ; **otherwise** skip Steps 1.3 and 1.4

Step 1.3: Generate the consequence of the rule. Use (F_ZF_T [Δ] *XOR* F_ZF_T [σ]) to obtain the consequent part of R_L.

Step 1.4: Check or not whether rule R_L can be generated $F_ZC_P(R_L) \geq \text{Min } F_ZC_P$, then R_L is effective.

End

The efficacy of the presented approach was demonstrated using an illustrative numerical example in (Jain *et al.*, 2008b).

5. Conclusion and perspectives

The ability to build lean and agile supply chains has not developed as rapidly as anticipated, because the development of technologies/techniques/approaches to manage such concepts of lean/agile for integrated supply chains is still under way. Also, due to ill-defined and vague indicators, which exist within leanness/agility assessment, many measures are described subjectively by linguistic terms, which are characterized by vagueness and multi-possibility, and the conventional assessment approaches cannot suitably nor effectively handle such dynamic situations.

In this chapter, firstly, we present a novel approach to model agility and introduce *Dynamic Agility Index* through fuzzy intelligent agents The proposed approach concentrates on the application of linguistic approximating, fuzzy arithmetic and agent technology is developed to address the issue of agility measuring, stressing the multi-possibility and ambiguity of agility capability measurement. Secondly, we discuss a novel approach based on *Fuzzy Association Rule Mining* incorporating fuzzy framework coupled with rules mining algorithm to support the decision makers by enhancing the flexibility in making decisions for evaluating agility with both tangibles and intangibles characteristics. Also, by checking the fuzzy classification rules, the goal of knowledge acquisition can be achieved for users.

As a scope for future work, empirical research is required to study the application of the proposed approaches and to characterize agility in integrated supply chains. Multi-functional workforce and their performance evaluation should also be studied as a scope for further research.

6. References

Agarwal, R., Gehrke, J., Gunopulos, D. and Raghavan, P., 1998. Automatic subspace clustering of high dimensional data for data mining applications, Proceedings of the ACM SIGMOD International Conference on Management of Data, 94–105.

Agarwal, R., Mannila, H., Srikant, R., Toivonen, H. and Verkamo, A.I., 1996. Fast discovery of association rules, in: U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (Eds.), Advances in Knowledge Discovery and Data Mining, AAAI Press, Menlo Park, 307–328.

Berry, M. and Linoff, G., 1997. Data Mining Techniques: For Marketing, Sales, and Customer Support, Wiley, New York.

Bhaskaran, S., 1998. Simulation analysis of a manufacturing supply chain, Decision Sciences, 29(3), 633-657.

Blackwell, R. D. and Blackwell, K., 2001. The Century of The Consumer: Converting Supply Chains Into Demand Chains, Supply Chain Yearbook, McGraw-Hill.

Christian I., Ismail H., Mooney J., Snowden S., Toward M. and Zhang D., 2001. Agile Manufacturing Transition Strategies, In: Hvolby H. (ed), Proceedings of the 4th SMESME International Conference, Department of Production, Aalborg University, Aalborg, 69-77.

Christopher M. and Towill D., 2001. An integrated model for the design of agile supply chains, International Journal of Physical Distribution & Logistics, 31 (4), 234-246.

Christopher M., 1998. Logistics and Supply Chain Management – Strategies for Reducing Cost and Improving Service, Prentice-Hall.

Christopher M., 2004. Supply Chains: A Marketing Perspective, In: New S & Westbrook R (eds) Understanding Supply Chains. Oxford University Press, Oxford, 69-108.

Christopher M., 2000. The Agile Supply Chain: Competing in Volatile Markets, Industrial Marketing.

Ding H., Benyoucef L. and Xie X., 2006. A Simulation-based multi-objective genetic algorithm approach for networked enterprises optimization, EAoAI international journal (Engineering Applications of Artificial Intelligence), 19(6), 609-623.

Evans, G.N., Towill, D. R., and Naim, M.M., 1995. Business process Re-engineering the supply chain, International Journal of Production Planning and Control, 6 (3), 227-237.

Fukuda, T., Morimoto Y., Morishita S. and Tokuyama T., 1996, Mining optimized association rules for numeric attributes, Proceedings of Fifteenth ACM SIGACT–SIGMOD–SIGART Symposium on Principles of Database Systems, 182–191.

Gunasekaran A. and Ngai E.W.T., 2004. Information systems in supply chain integration and management, European Journal of Operational Research, 159, 269–295.

Gunasekaran A., 1999. Agile manufacturing: A framework for research and development, International Journal of Economics, 62, 87–105.

Han, E. H., Karypis, G. and Kumar, V., 2000. Scalable parallel data mining for association rules, IEEE Transactions on Knowledge and Data Engineering, 12(3), 337-352.

Jain, V., Wadhwa, S. and Deshmukh, S. G., 2006. Modeling and Analysis of Supply Chain Dynamics: a High Intelligent Time Petri net based Approach, International Journal of Industrial and Systems Engineering, 1(1/2), 59-86.

Jain, V. Benyoucef, L. and Deshmukh, S. G., 2007a. A new approach to model goal and plan conflicts in a dynamic supply chain, In the proceedings of the IEEE/INFORMS International Conference on Service Operations and Logistics, and Informatics, 487-492, August 27-29, Philadelphia, USA.

Jain, V., Wadhwa, S. and Deshmukh, S. G., 2007b. Supplier Selection using Fuzzy association rules mining, International Journal of Production Research, 45 (6), 1323-1353.

Jain, V., Wadhwa, S. and Deshmukh, S. G., 2007c. A Negotiation-to-Coordinate (N2C) mechanism for Modeling Buyer-Supplier relationship in Dynamic Environment, International Journal of Enterprise Information Systems, 3(2), 1-31.

Jain, V., Benyoucef, L. and Deshmukh, S. G., 2008a. What's the buzz about moving from "Lean" to "Agile" integrated supply chains? A fuzzy intelligent-agents based approach, International Journal of Production Research, (Available Online).

Jain, V., Benyoucef, L. and Deshmukh, S. G., 2008b. A new approach for evaluating agility in supply chains using fuzzy association rules mining, EAoAI international journal (Engineering Applications of Artificial Intelligence), 21, 367-385.

Jharkharia, S. and Shankar R., 2005. IT-enablement of supply chains: understanding the barriers, Journal of Enterprise Information Management, 18(1/2), 11–27.

Kidd P.T., 1994. Agile Manufacturing, Forging New Frontiers, Addison-Wesley, London.

Lambert D.M., Cooper M. and Pagh J.D., 1998. Supply Chain Management: Implementation Issues and Research Opportunities, The International Journal of Logistics Management, 9, 1–19.

Lin C-T, Chiu H. and Chu P-Y, 2006. Agility index in the supply chain, International Journal of Production Economics, 100, 285–299.

Lin, F.-R. and Shaw, M. J., 1998. Reengineering the Order Fulfillment Process in Supply Chain Networks, The International Journal of Flexible Manufacturing Systems, 10, 197-229.

Nakamura, K., 1986. Preference relations on a set of fuzzy utilities as a basis for decision making, Fuzzy Sets and Systems, 20(2), 147-162.

Sharp J.M., Irani Z. and Desai S., 1999. Working towards agile manufacturing in the UK industry, International Journal of Production Economics, 62(1–2), 155–169.

Simchi-Levi D., Kaminsky P. and Simchi-Levi E., 2003. Designing and Managing the Supply Chain: Concepts, Strategies and Case Studies, McGraw-Hill.

Srikant, R. and Agarwal, R., 1996. Mining quantitative association rules in large relational tables, Proceedings of the ACM SIGMOD International Conference on Management of Data, 1–12.

Swafford P. M., Ghosh S. and Murthy N. N., 2006. A framework for assessing value chain agility, International Journal of Operations and Production Management, 26(2), 118-140.

Swaminathan, J. M., Smith, S. F., and Sadeh, N. M., 1998. Modeling Supply Chain Dynamics: A Muliagent Approach, Decision Sciences Journal, 29 (2), 607-632.

Van der Aalst, W. M. P., 1998. The Application of Petri Nets to Workflow Management, Journal of Circuits, Systems, and Computers, 8(1), 21-66

van Hoek R., 2001. Epilogue: Moving forwards with agility, International Journal of Physical Distribution and Logistics Management, 31(4), 290-300.

van Hoek R., 2005. Mitigating the minefield of pitfalls in creating the agile supply chain, In Andersin HE, Niemi E and Hirvonen V (eds) Proceedings of the international conference on agility – ICAM 2005, Helsinki University of Technology, Otaniemi, Finland.

Yusuf Y. Y., Gunasekaran A., Adeleye E. O. and Sivayoganathan K., 2004. Agile supply chain capabilities: Determinants of competitive objectives, European Journal of Operational Research, 159, 379-392.

Yusuf, Y.Y. and Burns, N.D., 1999. Decision Support Method for Agile Enterprise Design (Agile-Enter), EPSRC Project GR/M58085.

Zadeh L. A., 1965. Fuzzy Sets, Information and Control, 8(3), 338-353.

14

Managing and Integrating Demand and Supply Using Web Services and the Service Oriented Architecture

Firat Kart, Louise E. Moser and P. M. Melliar-Smith
Department of Electrical and Computer Engineering
University of California, Santa Barbara
U.S.A.

1. Introduction

The Internet and the World Wide Web have had a significant impact on business management thinking and practice. Globalization has affected how businesses interact with other businesses, and even how divisions within companies interact with each other, and has increased competition and consolidation worldwide (Murch, 2004). The new business models are supported by modern information technology such as the Service Oriented Architecture (Erl, 2005; Moser & Melliar-Smith, 2008; Newcomer and Lomow, 2004; OASIS, 2006) and Web Services (Alonso *et al.*, 2004; Champion *et al.*, 2002; Chatterjee & Webber, 2003). Supply chains (ComputerWorld, 2006; Wilson, 2005) are particularly affected by these developments in business and technology.

The MIDAS (Managing and Integrating Demand and Supply) system that we have developed is an automated supply chain management system based on the Service Oriented Architecture and Web Services. The benefits of MIDAS (and of the Service Oriented Architecture in general) for supply chains are that it increases business flexibility and it enables businesses to adapt more quickly to changing business needs. The MIDAS system provides a loosely coupled distributed environment that allows customers, manufacturers, and suppliers to cooperate over the Internet and the World Wide Web. Generally, supply chain management considers three types of flow:

- **Information flow**, which pertains to placing, transmitting and filling orders, and updating their delivery status
- **Product flow**, which involves movement of goods from a supplier to a customer, as well as customer returns
- **Financial flow**, which relates to credit terms, payments, payment schedules, consignment, and title ownership.

The MIDAS supply chain system described here focuses, in particular, on the management of information flow. MIDAS, which is inspired by the build-to-order business model (Ghiassi & Spera, 2003; Gunasekarana & Ngai, 2005), enables customers to customize their products before they order. At the manufacturer, MIDAS receives orders from the customer, and places orders with the suppliers, automatically and dynamically. The manufacturer can use one of several strategies to aggregate customers' orders before it starts processing them and to accumulate suppliers' quotes before it decides to do business with the suppliers. MIDAS allows manufacturers and suppliers to conduct a business deal online either by accepting a quote as is, or by negotiating.

MIDAS aims to reduce inventory carrying costs and logistics administration costs, yielding a more efficient supply chain, by supporting on-demand, just-in-time manufacturing. MIDAS leverages existing IT infrastructure to enable users to automate their supply chains. It makes it easier for small suppliers to get into business with large manufacturers by facilitating the procurement process. It reduces human intervention on both the customer/manufacturer side and the manufacturer/supplier side. Most importantly, MIDAS aims to meet the needs of the customers on time, and to reduce the costs of the manufacturer by eliminating the need for a large inventory.

In this article we describe the design and implementation of the MIDAS supply chain system. We discuss how MIDAS, and the Service Oriented Architecture and Web Services, support adaptation to change, interoperability, and scalability for supply chains, and how they can substantially improve the efficiency of a supply chain. We also present an evaluation of the MIDAS system, in terms of the customer's satisfaction as measured by the customer's response time, and the manufacturer's gain as measured by the number of orders aggregated or the best price ratio of orders.

2. The MIDAS architecture

As a Service Oriented Architecture (SOA), MIDAS ensures that the IT systems of different enterprises can adapt quickly and easily to support rapidly changing business needs. MIDAS supports horizontal business processes that are distributed across the Internet between, among and within enterprises. Through its use of Web Services, MIDAS provides interoperability between legacy back-end enterprise software systems.

By applying SOA practices to supply chain applications, MIDAS aims to automate the supply chain and, thereby, reduce human intervention, errors, and costs. The MIDAS software is modular, which allows it to be re-used at multiple levels of the supply chain. Moreover, as a SOA based on Web Services, MIDAS allows applications to be composed in a loosely-coupled fashion, and to be modified without disrupting the services provided to the customers.

MIDAS enables customers, manufacturers, and suppliers to cooperate in a dynamic environment, as they have never done before. The MIDAS system in one enterprise interacts dynamically with the MIDAS system in other enterprises, which are accessed over the Internet and the World Wide Web. MIDAS supports communication between the manufacturer and the suppliers, even if the manufacturer did not have any prior business with those suppliers and, thus, it increases the ease of collaboration between them.

MIDAS uses a UDDI Registry to allow a manufacturer to discover and select suppliers dynamically based on price, availability and delivery time of components. If an existing supplier becomes unavailable, it allows the manufacturer to find alternate suppliers on demand and to redirect its requests seamlessly to an alternate supplier. MIDAS uses a Reservation Protocol (Zhao *et al.,* 2008) to improve the performance of business transactions that span multiple enterprises in the supply chain. Use of the Reservation Protocol also decreases the probability of inconsistencies for business transactions between the manufacturer and suppliers.

Typically, a Service Oriented Architecture is fronted by a client user interface that uses the underlying services. Depending on the business application, a customized user interface is provided for customers that use the underlying services. Although the MIDAS client user interface that we have developed is specific to the manufacturing of a computer, the components of which are obtained from different suppliers, the underlying MIDAS system is general and can be used by manufacturers and suppliers of other kinds of products.

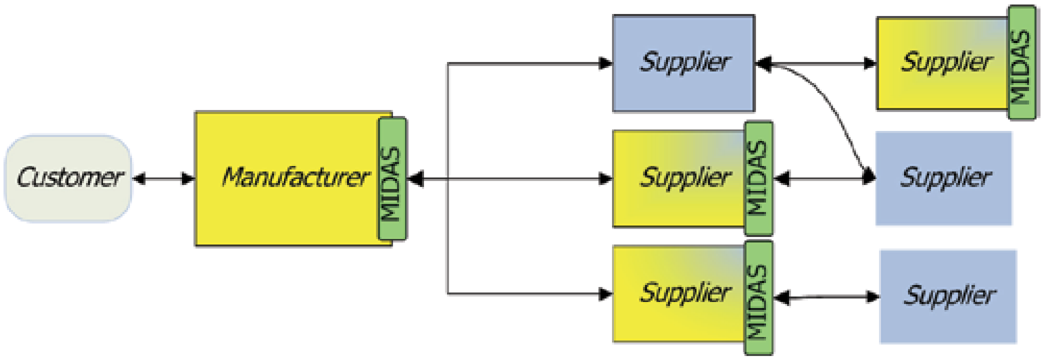


Fig. 1. Use of MIDAS in a supply chain

We consider here a three-level supply chain and a single manufacturer; however, as shown in Figure 1, the MIDAS strategy generalizes to deeper supply chains with N levels, $N \geq 3$, where a manufacturer is a supplier of the products it manufactures and a supplier is a manufacturer of the supplies that it offers. MIDAS is present at the businesses in the supply chain that act as both manufacturer and/or supplier. By considering the entire supply chain, MIDAS captures supply chain needs more effectively and provides faster adaptation to changing supply and demand.

MIDAS makes use of the following concepts in supply chain management. A *material* consists of one or more components. For example, a material for a computer is a particular computer model, such as Dell XPS 1310, Sony Vaio SZ780, *etc*. A *component* is a particular category for which there are one or more supplies. For example, the components of a computer include the processor, memory chip, graphics card, network interface card, *etc*. A *supply* is a product that is produced by a supplier. For example, the supplies associated with a processor might include a 2GHz CPU and a 3GHz CPU. A supply can be obtained from one or more suppliers, and the manufacturer can select the supplier of that supply dynamically. A *supply item* is an instance of a supply. For example, the supply items in a manufacturer's order from a particular supplier might include 100 2GHz CPUs.

MIDAS also uses the concept of *logical inventory* in supply chain management. Logical inventory is data, stored in the computers and databases of the supply chain management system that are related to the customers' needs and the customers' orders. With logical inventory, the manufacturer does not need to maintain a large physical inventory in its warehouses but, rather, can obtain the supplies that it needs on demand.

At the manufacturer, the MIDAS system comprises the following modules: Materials Manager, Orders Manager, Database (DB) Monitor, Communication Manager, and Quotes Manager. These modules are shown in Figure 2.

Customers obtain product information from a catalog provided by a Customer Web Service, which retrieves information from the Materials Manager. The Materials Manager relates a material to its components and a component to its supplies. On receiving orders from the customers, the Materials Manager passes the information to the Orders Manager. The Orders Manager inserts, into the Orders Database, information about the customers and the products that the customers are interested in purchasing, and manages the status of the customers' orders. On receiving an order, the Orders Manager informs the DB Monitor, which scans the orders and triggers a business activity that starts purchasing supplies from suppliers. The DB Monitor checks the Orders Database and decides, depending on the particular strategy chosen, whether to inform the Quotes Manager to initiate a search for suppliers and communicate with them. The Quotes Manager handles Quote requests, and relates Quote replies to Quote requests.

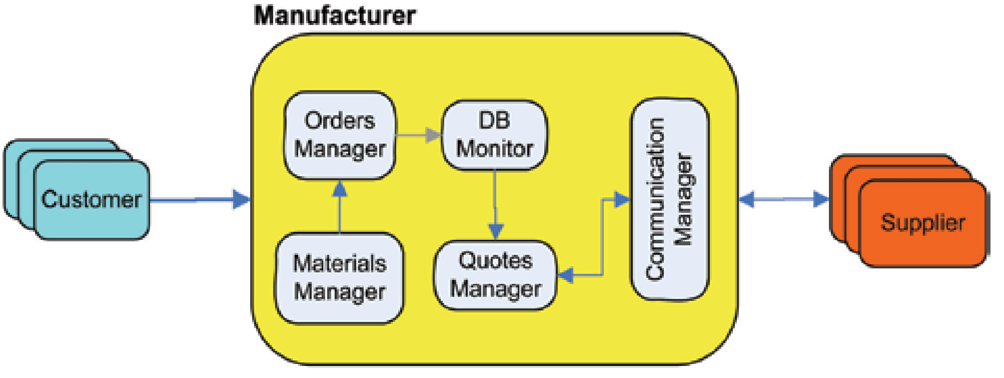


Fig. 2. The components of MIDAS

Each of these modules plays a role in the two phases of the manufacturers' processing an order from a customer. These two phases are:

- **Waiting phase**, which involves the collection of orders from the customers before making Quote requests for aggregated supply items from the suppliers
- **Quotes phase**, which involves the collection of Quote replies from the suppliers before making a decision on which supplier will provide the particular supply.

The strategies used by the DB Monitor in the Waiting phase to decide whether to stop collecting customer orders and to inform the Quotes Manager are discussed in Sections 2.2 and 4.2. The strategies used by the Quotes Manager in the Quotes phase to decide whether to stop collecting Quote replies and to make a decision on a supplier are discussed in Sections 2.2 and 4.3.

The MIDAS architecture has interfaces on the customer side and the manufacturer side. Each interface involves different modules of MIDAS, and some modules of MIDAS serve as a bridge between the two interfaces. The Customer Service and the Manufacturer Service are discussed below.

2.1 The Customer Service

The MIDAS architecture is based on the premise that it is the customer's opinion that counts. Inspired by Dell's build-to-order model, MIDAS enables a customer to customize a product that the customer purchases by choosing the materials that constitute that product. The customer customizes a product order using a Customer Web Service that communicates with other services to obtain the materials necessary to manufacture the product and to arrange shipping and financing of the product.

The customer side of MIDAS deals with customer/manufacturer communication and interactions. The Customer Service provides the following functionality:

- Provides catalog information
- Receives information from the customer
- Displays the status of the customer's orders.

The customer user interface uses the Customer Web Service to interact with MIDAS. Existing Web Services frameworks ease the process of implementing the user interface that interacts with the services. In Section 3 we discuss Web Services frameworks in more detail.

2.2 The Manufacturer Service

The customers send their orders for products to the manufacturer, and the manufacturer processes the orders. At the manufacturer, MIDAS does not contact the suppliers each time it receives an order from a customer. Rather, it uses one of several strategies (discussed below) to aggregate orders from different customers. However, the manufacturer must not take too long to confirm a customer's order with the price, product delivery time, *etc*.
The Manufacturer Service provides the following functionality:

- Monitors orders
- Searches for suppliers
- Contacts relevant suppliers
- Decides on the supplier(s) from which to obtain supplies.

This functionality is discussed below in terms of the Orders Manager, DB Monitor, Quotes Manager, and Registry, shown in Figure 2.

2.2.1 The Orders Manager

MIDAS assumes that, the more orders the manufacturer accumulates, the more gain the manufacturer has. Thousands of supply items have a per item price that is different than the price the price of a single item. The Orders Manager tries to accumulate as many orders for a particular supply as it can.
The Orders Manager collects orders from different customers and aggregates supply items (instances of a particular supply) for those customer orders. The Orders Manager then forwards the orders to the Database Monitor.

2.2.2 The Database Monitor and the Orders Database

The Database (DB) Monitor keeps track of the orders for each supply in the Orders Database and triggers order events, based on one of several strategies discussed below. When the DB Monitor triggers an order event for a particular supply, the processing of the order begins. The DB Monitor then informs the Quotes Manager to find appropriate suppliers for that supply and to communicate with them.
The main problem for manufacturers that depend on suppliers is that, if a supplier is not available, the manufacturer cannot make progress. In a dynamic environment, a manufacturer must be able to find new suppliers as the demand arises and to satisfy the customers' needs in a timely manner.
The DB Monitor uses one of the following strategies to trigger an order event and to initiate the processing of an order for a particular supply by informing the Quotes Manager to place an order for that supply.

- **System user decides when to place orders:**
  The DB Monitor provides information to the system user at the manufacturer about the number of supply items (instances of a particular supply) required for the order of that supply and the wait times for those items. The user triggers manually the Quotes Manager, which initiates communication with the appropriate suppliers.
- **Threshold number of supply items for a particular supply is reached:**
  On receiving orders from the customers, the Orders Manager directs the DB Monitor to scan the database. If the number of supply items is greater than the threshold, the DB Monitor informs the Quotes Manager.
- **Timeout occurs at which an order needs to be placed:**
  The DB Monitor scans the database at defined intervals. For each scan, when a timeout occurs, the DB Monitor informs the Quotes Manager.
- **Hybrid of threshold number of orders and timeout:**
  This strategy combines the threshold strategy and the timeout strategy, such that whichever occurs first triggers the Quotes Manager.
- **Average wait time for a set of orders is reached:**
  The Orders Manager accumulates the number of supply items for the orders from different customers. When a component has an average wait time that is greater than a threshold, the DB Monitor informs the Quotes Manager.

The DB Monitor retrieves information (supply ID, amount to order, average wait time, *etc*) from the Orders Database. According to the strategy used, the DB Monitor decides whether it is time to contact the suppliers of a supply and then informs the Quotes Manager to do so.

2.2.3 The Quotes Manager

The Quotes Manager initiates communication with the suppliers, requesting the aggregated number of supply items. It submits Quote requests to selected suppliers, receives Quote replies with a proposed price, number of items, and proposed delivery time, and then decides on a supplier with which to place an order.

The Quotes Manager retrieves information about the respective suppliers of a supply from the UDDI Registry. Having decided on the suppliers with which it will communicate, the Quotes Manager sends Quote requests to those suppliers initiating the second phase, the Quotes phase. After sending Quote requests, the Quotes Manager uses one of the following strategies to decide on a particular supplier with which to place an order. (It could order items for a particular supply from multiple suppliers but, for simplicity, we consider only a single supplier for that supply.)

- **System user decides on a supplier:**
  The system user has control over the processing of quotes from the suppliers and deciding on a supplier of a particular supply.
- **Threshold percentage of Quote replies:**
  The Quotes Manager continues to aggregate Quote replies until the total number of Quote replies is above a threshold percentage. For example, when the threshold percentage is 100%, the Quotes Manager waits until it receives Quote replies from all of the respective suppliers.
- **Timeout for Quote replies from the suppliers:**
  The Quotes Manager continues to wait until a specific time. If the Quotes Manager receives all of the expected Quote replies before the timeout, the Quotes Manager initiates processing of the Quote replies.

- **Hybrid of threshold number of Quote replies and timeout:**
  This strategy combines the strategies of waiting for a threshold number of Quote replies and waiting for a specific time, whichever occurs first.
- **Average wait time threshold for Quote replies:**
  The Quotes Manager accumulates Quote replies and when the average wait time for the Quote replies is greater than a threshold, the Quotes Manager initiates processing of the Quote replies.

The manufacturer can define the number of items needed, the type of supply, and the delivery time expected. Based on the suppliers' offers, the Quotes Manager then updates its Quote request and sends the updated Quote request to the suppliers again, continuing the negotiation with them.

Figure 3 shows an example sequence diagram for the Quotes phase. In this example, the Quotes Manager could not communicate successfully with Supplier C, perhaps because of a communication failure or unavailability of Supplier C's service. However, the Quotes Manager is aware of the total number of Quote requests that it sent and the number of Quote replies that it must receive. The Quotes Manager decides on the suppliers with which it will do business. Having decided on the status of a quote, the Quotes Manager updates the status of the customer order associated with the quote and informs the Orders Manager. Completion of a quote does not necessarily mean completion of a customer order. The customer order is completed once a decision about all of the different components for that order is made.

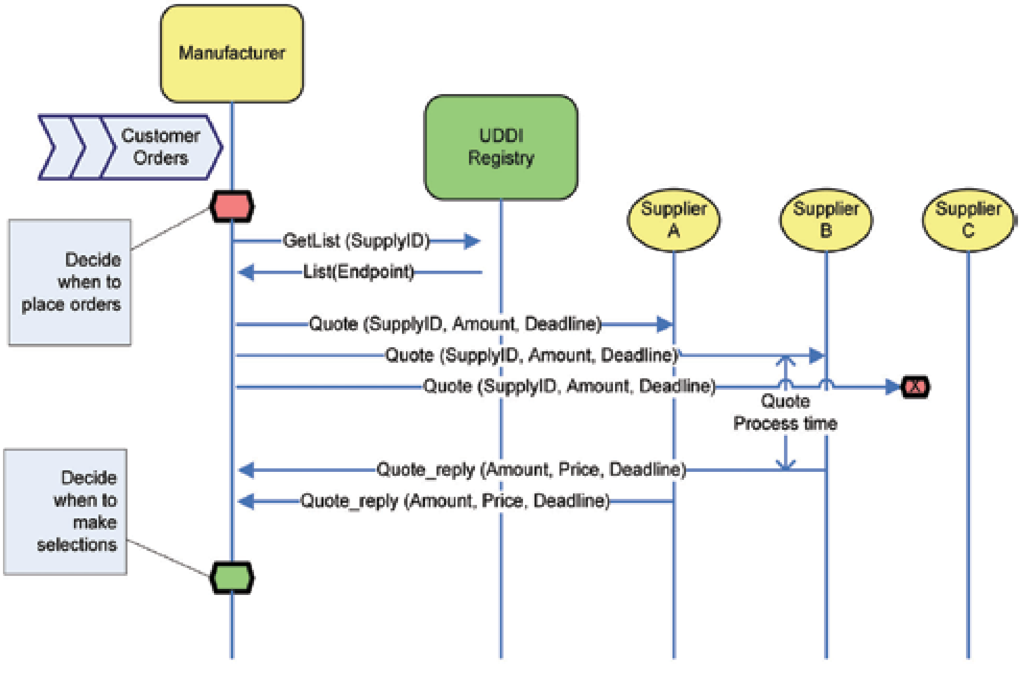


Fig. 3. Quote Request and Quote Reply interactions

As Figure 3 shows, there are two important decision points that affect the time the customer must wait to obtain information about the price and delivery time of the product. First, the Orders Manager accumulates orders for particular supplies, collecting as many as it can, while not increasing the average customer waiting time, before it decides to proceed to the next decision. Next, the Quotes Manager accumulates Quote replies from the suppliers, aiming to make the delay for the customer as short as possible, while not missing better quotes after it decides.

Note that, while waiting for Quote replies, the Quotes Manager is not blocked. All of the messages are sent asynchronously, and no assumption is made about the order of delivery of Quote replies. The time at which the Quotes Manager receives Quote replies from the suppliers varies according to the supplier's quote processing time.

2.2.4 The UDDI Registry

The UDDI Registry enables the manufacturer to find relevant suppliers of the supplies it needs on demand. The Registry Database keeps information about the suppliers, in particular, the contact information of each supplier and the supply IDs of the products that it offers. The manufacturer and the suppliers are assumed to use the same supply IDs, and the supply IDs are assumed to identify, uniquely, the supplies across different manufacturers and different suppliers. The Registry supports a fair business environment for suppliers where they have a chance of doing business along with their competitors.

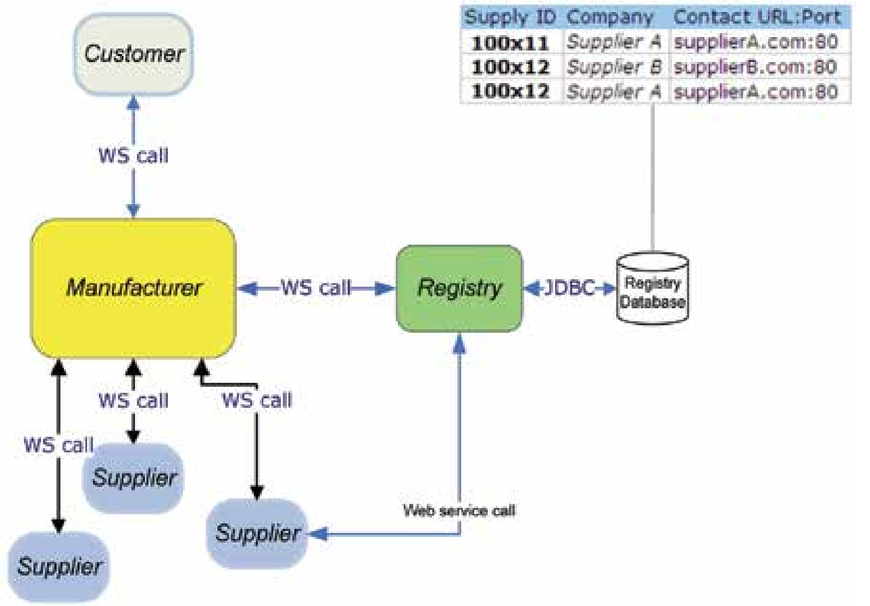


Fig. 4. Communication links and protocols used between the Manufacturer, the Registry, and the Suppliers

Figure 4 shows the Registry and the communication links and protocols, as well as the structure of the Registry. A supplier registers by communicating with the Registry and passing current information about itself to the Registry. The supplier provides its contact information and the products that it offers. If there is a change (such as a change of address, offering of new products, or removing product information), the supplier must update this information in the Registry. After registering, the supplier can receive Quote requests from manufacturers using MIDAS.

There are tradeoffs between the amount of information that is kept in the Registry and the number of Update messages sent by the suppliers. The two extreme strategies for the amount of information maintained by the Registry are:

- **Maintain only contact information for the suppliers:**
  In this case, the manufacturer must contact every supplier listed in the Registry, even though some suppliers do not provide what the manufacturer needs. The suppliers send Update messages (re register) containing only their contact information.
- **Maintain all the information about the suppliers:**
  Here, all the information about the suppliers includes contact information, particular supplies that they offer, maximum number of supply items that they can provide, price per item, projected delivery time, etc. The Registry simply returns a list of suppliers with their contact information sorted according to the criteria that the manufacturer wants to see.

In the first strategy, during the Quotes phase, the manufacturer spends considerable time sending and receiving messages in accumulating Quote replies from the suppliers.

In the second strategy, the suppliers frequently send Update messages to keep their information in the Registry up-to-date. However, during the Quotes phase, there are fewer messages, because the manufacturer needs to contact fewer suppliers; in particular, it can eliminate suppliers that have higher prices and later delivery times. Keeping the price and delivery time information in the Registry not only complicates the running of the Registry but also increases the amount of information that the Registry holds. Moreover, a supplier might not be willing to reveal all of its price information and the kind and number of supply items in the Registry.

Our implementation uses an intermediate strategy between these two extremes. The Registry includes only the contact information of the suppliers and the supply IDs of the supplies that they offer. This choice assumes that a supplier will not frequently update its catalog. If there were thousands of suppliers listed in the Registry, and the Registry contains only contact information, the cost of sending Quote requests to all of them would be quite high, particularly if there were only a few suppliers for a particular supply. Consequently, for the hybrid strategy, the Registry performs better in terms of message overhead than the alternative strategies. Moreover, by not keeping price information in the Registry, MIDAS allows the manufacturer to negotiate the price at which it is willing to buy from a supplier. The supplier might accept the offered price or revise its quote and send a Quote reply with an updated price. Update messages that update the Registry with the suppliers' contact information and supply IDs are still required, but there are fewer of them.

2.3 The Reservation Protocol

MIDAS employs the Reservation Protocol, which is an extended transaction protocol that is designed for business transactions that span multiple enterprises (Zhao *et al.*, 2008). Business activities between the manufacturer and the supplier are executed as two steps. The first step involves an explicit reservation of resources according to the business logic. The second step involves the confirmation or cancellation of the reservation. For example, a manufacturer that is interested in buying a product from a supplier sends Quote requests to the suppliers in the first step. Once the manufacturer receives Quote replies from the suppliers and makes a decision, the manufacturer sends confirmation or cancellation messages to the suppliers.

Alternative transaction methods suffer in terms of response time and throughput when there are concurrent requests. During the transaction of a request, resources are not available to other requests, increasing the response time and decreasing the throughput. Because the Reservation Protocol executes each step as a separate traditional short-running transaction, resources are classified as available and reserved, which differs from blocking the resources until the current request is complete. Use of the Reservation Protocol decreases the wait time of customers, because MIDAS can complete the business transactions between the manufacturer and a supplier much faster, even though there are many concurrent requests arriving at the suppliers.

The Reservation Protocol improves the performance of business transactions that span multiple enterprises in a multi-level supply chain. As shown in Figure 1, an enterprise can deploy both client and/or server middleware depending on business needs. A supplier at level N receiving a Quote request from a manufacturer at level N-1 can contact its suppliers at level N+1 to make a reservation to reply to a Quote request from a manufacturer at level N-1. Receiving Quote replies from the suppliers at level N+1, the supplier at level N can send a Quote reply to a manufacturer at level N-1. Depending on the decision of the manufacturer at level N-1, the supplier at level N sends a decision to its suppliers at level N+1. For traditional transaction protocols, the same scenario can result in longer delays and also inconsistencies in the databases of the different enterprises in the supply chain.

3. Implementation

The MIDAS supply chain system comprises software modules for the customers, the manufacturers, and the suppliers, as described in Section 2. The Web Services for MIDAS are built on the Apache Axis2 Framework (which is the core engine for Web Services built on Apache Axiom) and the Apache Tomcat Server (Apache, 2008).

- An XML client API including WSDL and policy support
- Support for various message exchange patterns
- Synchronous and asynchronous function calls
- WS-Policy driven code generation
- A flexible service lifecycle model
- Support for SOAP, WSDL, WS-Reliable Messaging, WS-Security, WS-Addressing and SAAJ.

In addition, Axis2 provides data bindings that enable application developers to generate SOAP messages without concern for constructing or parsing them.

As pointed out previously, although our implementation is specific to the purchase and manufacturing of a customized computer, the underlying MIDAS system is general and can be used for other kinds of products.

3.1 The customer side

Customers can access the MIDAS system and the manufacturer's catalog of products using their Web browsers. The customers are represented in the manufacturer's database, and authentication is required before the resources for a particular customer can be accessed.

At the manufacturer, servlets use the Customer Web Service interface and make Web Service calls to MIDAS through the API provided. Figure 5 shows an example SOAP reply message for a catalog request.

```
POST /axis2/services/MidasWS HTTP/1.1
Content-Type: application/soap+xml; charset=UTF-8; action="urn:receiveOrder"
User-Agent: Axis2
Host: vanguard.ece.ucsb.edu:9904
Transfer-Encoding: chunked

<?xml version='1.0' encoding='UTF-8'?>
<soapenv:Envelope xmlns:soapenv="http://www.w3.org/2003/05/soap-envelope">
 <soapenv:Body>
  <ns2:receiveOrder xmlns:ns2="http://ws.midas.scc06">
   <ns2:param0>
    <ns1:clientID xmlns:ns1="http://data.midas.scc06/xsd">13783459</ns1:clientID>
    <supplies xmlns="http://data.midas.scc06/xsd">
     <amount>1</amount>
     <brand>Intel</brand>
     <id>2</id>
     <name>Intel Core 2 Duo T5550 (2MB cache/1.83GHz/667Mhz FSB)</name>
     <price>110.0</price>
    </supplies>
    <supplies xmlns="http://data.midas.scc06/xsd">
     <amount>1</amount>
     <brand>Brand</brand>
     <id>8</id>
     <name>4GB Shared Dual Channel DDR2 SDRAM at 667MHz</name>
     <price>600.0</price>
    </supplies>
    <supplies xmlns="http://data.midas.scc06/xsd">
     <amount>1</amount>
     <brand>Brand</brand>
     <id>12</id>
     <name>Item name 1</name>
     <price>200.0</price>
    </supplies>
```


Fig. 5. Example Web Service call to retrieve product catalog information

Once the customer is provided with the product catalog, the customer can create his own customized computer by selecting a particular supply for each component. Figure 6 shows an example SOAP request message for the customer's purchase order to MIDAS.

3.2 The manufacturer side

3.2.1 The Orders Database

The Orders Manager and the Materials Manager at the manufacturer use the Orders Database to insert, query and update the status of customer orders. The Orders Database includes the following tables:

- **Supply Table:** Represents alternative supplies for components
- **Component Table:** Represents the components of materials
- **Materials Table:** Keeps information about the materials that the manufacturer offers to customers
- **SupplyComponent Table:** Relates supplies and components
- **ComponentMaterial Table:** Relates components and materials

```
HTTP/1.1 200 OK
Date: Tue, 01 Apr 2008 02:30:23 GMT
Server: Simple-Server/1.1
Transfer-Encoding: chunked
Content-Type: application/soap+xml; charset=UTF-8; action="urn:getMaterialsResponse"

<?xml version='1.0' encoding='UTF-8'?>
<soapenv:Envelope xmlns:soapenv="http://www.w3.org/2003/05/soap-envelope">
 <soapenv:Body>
  <ns:getMaterialsResponse xmlns:ns="http://ws.midas.scc06"
   xmlns:ax21="http://data.midas.scc06/xsd">
   <ns:return type="scc06.midas.data.Material">
    <ax21:items type="scc06.midas.data.Component">
     <ax21:items type="scc06.midas.data.Supply">
      <ax21:brand>Intel</ax21:brand>
      <ax21:id>1</ax21:id>
      <ax21:name>Intel Core 2 Duo T5450 (2MB cache/1.66GHz/667Mhz FSB)</ax21:name>
      <ax21:price>100.0</ax21:price>
     </ax21:items>
     <ax21:items type="scc06.midas.data.Supply">
      <ax21:brand>Intel</ax21:brand>
      <ax21:id>2</ax21:id>
      <ax21:name>Intel Core 2 Duo T5550 (2MB cache/1.83GHz/667Mhz FSB)</ax21:name>
      <ax21:price>110.0</ax21:price>
     </ax21:items>
     <ax21:items type="scc06.midas.data.Supply">
      <ax21:brand>Intel</ax21:brand>
      <ax21:id>3</ax21:id>
      <ax21:name>Intel Core 2 Duo Processor T7250 (2.0GHz/800Mhz FSB, 2MB</ax21:name>
      <ax21:price>150.0</ax21:price>
     </ax21:items>
```


Fig. 6. Example Web Service call to purchase a customized computer

- **Quotes Request Table:** Keeps information about the Quote requests sent
- **Quotes Reply Table:** Keeps information about the Quote replies received
- **Sales Table:** Keeps information about the customers' order information. For each order item, new sales information is created.

Figure 7 shows the relationships between the tables for supplies, components and materials.

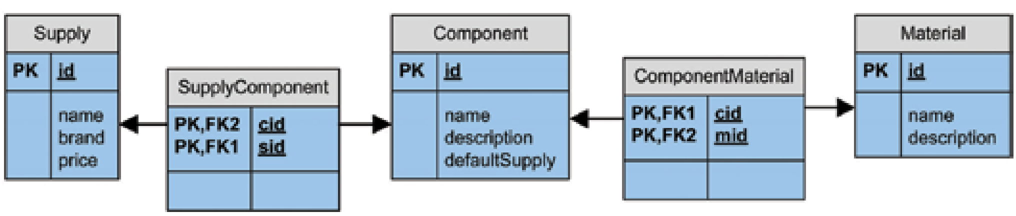


Fig. 7. The relationship between the database tables of MIDAS

3.2.2 The Registry and Registry Database

For the implementation of the UDDI Registry, we have used jUDDI (JUDDI, 2008), which runs with Jakarta Tomcat. The Registry is where the suppliers register their information (using the API provided) and where the manufacturers find that information (using the API provided). The suppliers update their information using authentication provided by jUDDI.

The UDDI Registry is backed with a MySql Server. The Registry accesses the MySql Database using JDBC. The UDDI Registry is deployed on one of the PlanetLab Nodes. The use of PlanetLab is discussed in more detail below.

4. Performance evaluation

In Section 2.2.2 we discussed different strategies that the DB Manager can use in the Waiting phase to aggregate orders from customers before triggering the Quotes Manager to communicate with suppliers. In Section 2.2.3 we discussed different strategies that the Quotes Manager can use in the Quotes phase to accumulate Quote replies from suppliers. The delay for aggregating orders, $delay_{AO}$, affects the delay for accumulating quotes, $delay_{AQ}$, particularly if the supplier processes Quote replies sequentially.

To evaluate the different strategies that the DB Manager and the Quotes Manager can use, we used the PlanetLab global research network (PlanetLab, 2008). Figure 8 shows the deployment of MIDAS on our local network, with suppliers waiting for orders and customers sending their purchase orders from PlanetLab nodes across the United States.

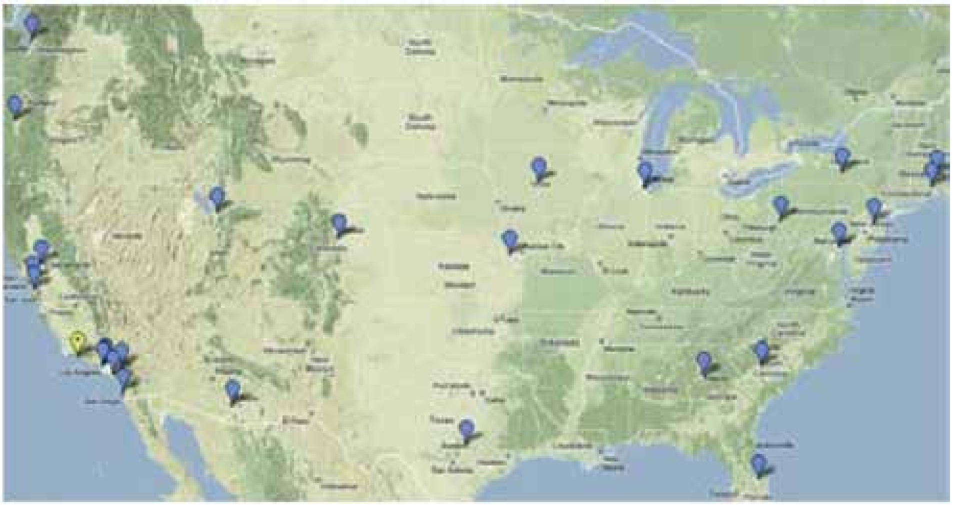


Fig. 8. Deployment of MIDAS on our network and on PlanetLab nodes across the USA

We evaluated the time spent in both the Waiting phase and the Quotes phase of MIDAS, *i.e.*, the time to aggregate orders and the time to aggregate Quote replies, under the different strategies. In addition, we investigated the probability that MIDAS makes a decision for the best quote possible. Finally, we calculated the average time needed to complete a customer order under the different strategies.

4.1 Order process

Before presenting the performance evaluation results, we present customer order timelines for a customized computer in order that the evaluation can be better understood.

Figure 9 depicts the timeline of a customer order submitted to MIDAS in the Waiting phase. The customer retrieves the product catalog for a computer from the manufacturer. Using this catalog, the customer creates the customized computer and submits his order to MIDAS. In this case, the order includes five different supply items (processor, graphics card, memory, hard disk, optical disk drive) for the components of the customized computer. On receiving orders from the customers, MIDAS first aggregates the supply items ordered, and then contacts the suppliers when the total number of items for a particular supply reaches a threshold or, alternatively, when a timeout occurs. Depending on the strategy used, the aggregation time for a supply varies. Frequently chosen supplies are aggregated faster than infrequently chosen supplies. The performance evaluation for the Waiting phase is discussed in Section 4.2.

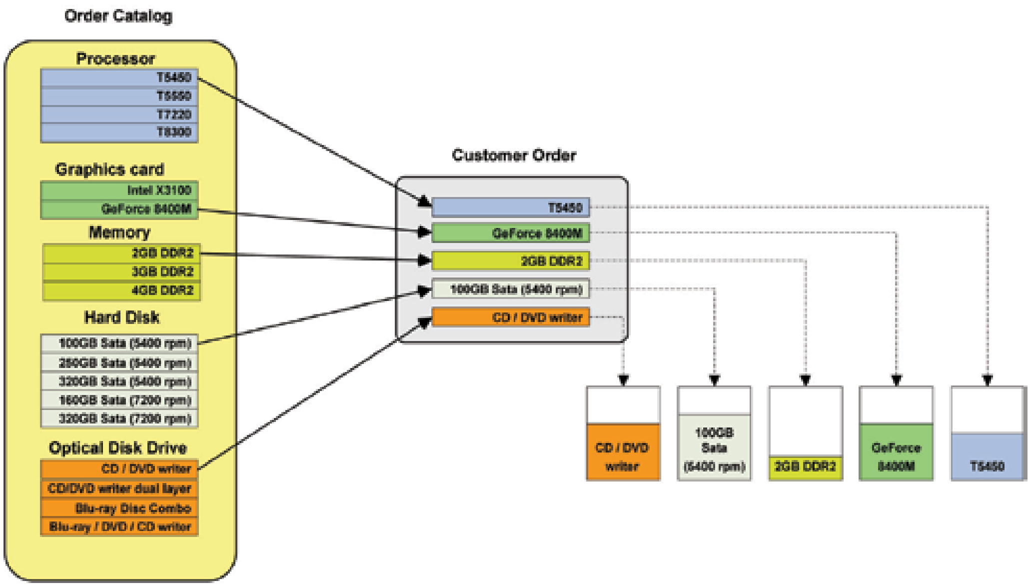


Fig. 9. Waiting Phase. Timeline of a customer order

Figure 10 depicts the timeline of the customer order for the Quotes phase. After deciding to contact the suppliers for a specific supply, MIDAS retrieves information about the respective suppliers for that supply from the UDDI Registry. MIDAS submits a Quote request for that supply to those suppliers. A supplier makes a callback to deliver its Quote reply. Once a supplier has finished the process of making a reservation related to the Quote request, it sends a Quote reply using the callback Web Service. MIDAS collects these Quote replies and makes a decision about the status of a quote, which affects the status of different supply items. The selected supplier is notified with a confirmation related to the purchase. The suppliers not chosen are contacted with a cancellation of the reservation. Once a decision for all order items for a customer order is made, the customer order is complete. The performance evaluation for the Quotes phase is discussed in Section 4.3.

4.2 The Waiting phase

As discussed in Section 2.2.2, the Waiting phase of MIDAS is based on one of several strategies. Here we consider:

- Aggregating a threshold number of orders
- Timeout for a specific amount of time
- Hybrid of threshold number of orders and timeout.

We evaluated these strategies under different customer order arrival rates. In our evaluation, the manufacturer receives a total of 1000 customer orders.

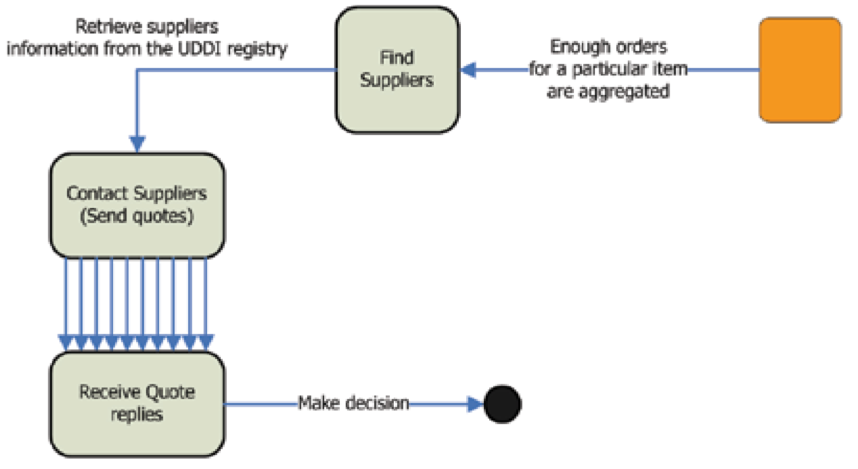


Fig. 10. Quotes Phase: Timeline of a customer order

4.2.1 Aggregating a threshold number of orders

This strategy aggregates orders for a specific number of items for a particular supply before the Quotes Manager issues Quote requests to the suppliers.

We evaluated the time spent to aggregate orders for 5, 10, 15, and 20 supply items with different customer order arrival rates (customer order arrives every 4, 6, 8, 10, and 12 seconds) and for the same number of alternatives (four alternatives) for each of the different components from which the customer can choose. The results are shown in Figure 11.

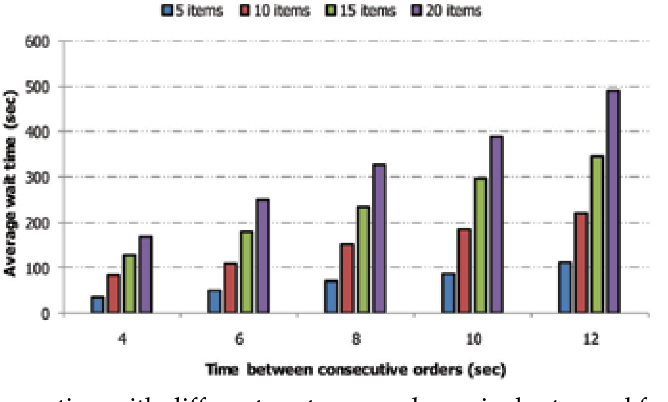


Fig. 11. Order aggregation with different customer order arrival rates and four alternatives for each component

As Figure 11 shows, the average wait time to aggregate enough supply items is higher for low customer order arrival rates. Moreover, the increased threshold delays the time to complete the Waiting phase. The threshold number of orders aggregation strategy keeps the number of supply items at a certain limit; however, a low customer order arrival rate might delay the order of a particular supply if that supply is not popular.

For variable numbers of alternatives for the different components (*e.g.*, 2 choices for the processor, …, 6 choices for the hard disk), the weighted average of the wait time during order aggregation is essentially the same as that for the same number of alternatives. This

behavior is expected because a supply that is aggregated faster has a larger number of occurrences but a smaller wait time.
Customers will be satisfied if their orders are completed earlier. However, contacting suppliers with orders for more supply items is more favorable for the manufacturer than contacting the suppliers with orders for fewer of them. The more supply items purchased from the supplier, the more gain the manufacturer has. Thus, the manufacturer faces a conflict as to whether to be concerned about the number of supply items to aggregate (manufacturer gain) or to limit the wait time for aggregation of supply items (customer satisfaction).

4.2.2 Timeout for a specific amount of time

The timeout strategy keeps the wait time for orders at a certain level. Therefore, we consider the number of orders aggregated as a function of the customer order arrival rate and the timeout value. The results for the timeout strategy are presented in Figure 12.
As Figure 12 shows, an increased timeout value results in the aggregation of a larger number of items for a particular supply. For the same timeout value, the customer order arrival rate affects the number of orders aggregated before the timeout occurs.

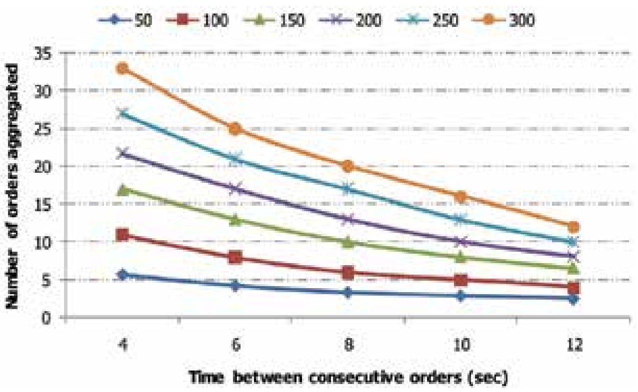


Fig. 12. Order aggregation with different timeout values

4.2.3 Hybrid of threshold number of orders and timeout

Although we have two strategies for the Waiting phase, *i.e.*, aggregating a threshold number of orders and waiting for a timeout to occur , it is better to use both particularly if the customer's request rate is low or is not known a priori. Taking both strategies into account can obtain the threshold number of orders in less than the maximum wait time of the Waiting phase but still bound the wait time to allow the Quotes phase to proceed.

4.3 The Quotes phase

As discussed in Section 2.2.3, the Quotes phase of MIDAS is based on one of several strategies. Here we consider:

- Aggregating a threshold percentage of Quote replies
- Timeout for a specific amount of time
- Hybrid of threshold percentage of Quote replies and timeout.

In the Quotes phase, MIDAS contacts the relevant suppliers and submits a Quote request to each of them. On receiving a Quote request, a supplier incurs some processing time before it replies to MIDAS using the callback Web Service endpoint passed in the Quote request. The manufacturer collects Quote replies concurrently in an asynchronous manner, and the suppliers respond to Quote requests after a processing time that depends on the particular type of supplier.

The processing time for the suppliers, shown in Figure 13, is given by the following formula:

$$Processing\ time \quad = \quad 20 - 5 * \frac{\log\ (random)}{\log\ (factor)}$$

The processing time depends on a random number between 0 and 1, and corresponds to processing times between 20 seconds and ∞. With a high probability, the processing time is close to 20 seconds, and with a very low probability, the processing time is ∞, which represents a supplier that crashes after receiving a Quote request from MIDAS.

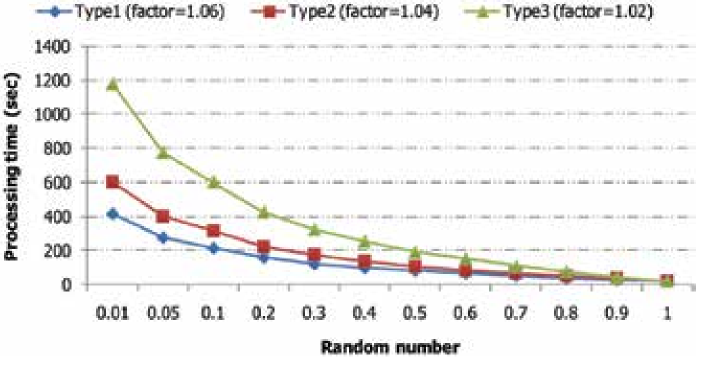


Fig. 13. Modeling of the processing time for the Suppliers

The processing time also depends on a scaling factor to represent different types of suppliers with different processing times. The factors 1.06, 1.04, and 1.02 correspond to processing times in the ranges 20-400, 20-600, and 20-1200. Type 1 suppliers (factor=1.06) respond to Quote requests earlier than Type 3 suppliers (factor=1.02), and Type 2 suppliers respond to Quote requests in between the other two.

The time spent before determination of the status of a quote depends on the strategy and the type of supplier (1, 2, or 3). Depending on the strategy, it is possible to decide the status of a quote before receiving all of the Quote replies. Aggregation of more Quote replies increases the chance of selecting the best quote.

For each of the above strategies in the Quotes phase, we used the threshold strategy in the Waiting phase to evaluate the average wait time and the best quote ratio. The best quote ratio is the probability that MIDAS decides the best possible Quote reply.

4.3.1 Aggregating a threshold percentage of Quote replies

In this strategy, if the number of Quote replies exceeds a certain percentage, the Quotes Manager stops collecting Quote replies from the suppliers and completes the customer orders. Compared to the strategy of collecting all Quote replies, this strategy decreases the customer's response time (increases the customer's satisfaction), at the cost of reducing the best price ratio.

Figure 14 shows the average wait time the threshold percentage strategy in the Quotes phase. Depending on the number of suppliers that the Quotes Manager contacted, the Quotes Manager waits for the number of Quote replies given by:

$$\textit{Expected replies} = \textit{Requests sent} * \textit{Percentage to aggregate}$$

The results show a linear increase and then an exponential increase for the case that MIDAS waits to collect all Quote replies. The average wait time is dominated by the maximum processing time from one of the suppliers for that quote. The average wait time can be kept reasonable, although MIDAS must wait for 90% of the expected Quote replies.

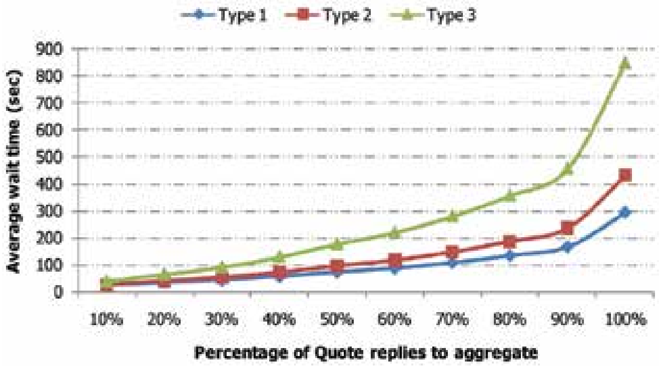


Fig. 14. Average wait time for the threshold percentage strategy

We also evaluated the best quote ratio against the percentage of expected Quote replies for the threshold percentage strategy. The results indicate that the best quote ratio is independent of the processing time at the suppliers, *i.e.*, the type of supplier.

4.3.2 Timeout for a specific amount of time

When the Quotes Manager collects Quote replies from the suppliers, it waits for a specific amount of time, $delay_{AQ}$. When the timeout occurs, the Quotes Manager decides on a supplier for a specific supply for the customer's order and ignores late Quote replies. Because the Quotes Manager doesn't need to collect Quote replies from all of the suppliers, this strategy decreases the customer's response time (increases the customer's satisfaction), compared to the strategy of collecting all Quote replies. On the other hand, the Quotes Manager cannot always obtain the best price from the suppliers, because the Quotes Manager might receive the Quote reply with the best price too late and thus ignore it.

The timeout strategy decides the status of a quote, either when a timeout occurs, or before the timeout when all of the expected Quote replies are received. As expected, our evaluation of the average wait time, *i.e.*, $delay_{AQ}$, shows that the average wait time increases as the time spent aggregating Quote replies increases. Moreover, the average wait time stabilizes around the maximum processing time of the suppliers. After stabilization, increasing the timeout does not affect the time taken to make a decision for the quote.

More interesting is our evaluation of the best quote ratio for the timeout strategy, *i.e.*, the probability that MIDAS catches the best quote given a particular timeout value, as shown in Figure 15. When the Quotes Manager decides the status of a quote early, the probability of catching the best quote is low because a better Quote reply might arrive later. Type 1 suppliers that have a small processing time send their Quote replies early, which increases the number of Quote replies aggregated in the Quotes phase. Once the best quote ratio reaches a maximum value, increasing the timeout does not affect the best quote ratio.

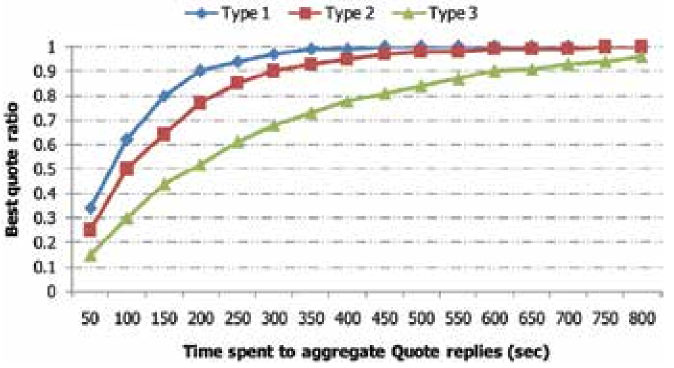


Fig. 15. Best quote ratio for the timeout strategy

4.3.3 Hybrid of threshold percentage of quote replies and timeout

Although we have two strategies for the Quotes phase, *i.e.*, aggregating a threshold percentage of Quote replies and waiting for a timeout to occur, it is better to use both strategies, particularly if the supplier's response rate is low. Taking both strategies into account can obtain the threshold percentage of Quote replies in less than the maximum wait time of the Quotes phase but still bound the wait time of the Quotes phase of MIDAS.

4.4 Customer order analysis

Finally, we provide a customer order analysis in terms of the wait time that elapses before completion of a customer order, *i.e.*, all of the supply items needed for the customer order have been ordered. Thus, the order completion time is the maximum of the completion times for all supply items i needed to fill the customer order, which is given by:

$$Order\ completion\ time \quad = \quad max_{\forall i}\ (Delay_{AO}(i) + Delay_{AQ}(i))$$

The results of the analysis for the timeout strategy are shown in Figure 16 for the case in which MIDAS receives orders every 10 seconds and 5 items are aggregated before MIDAS enters the Quotes phase.

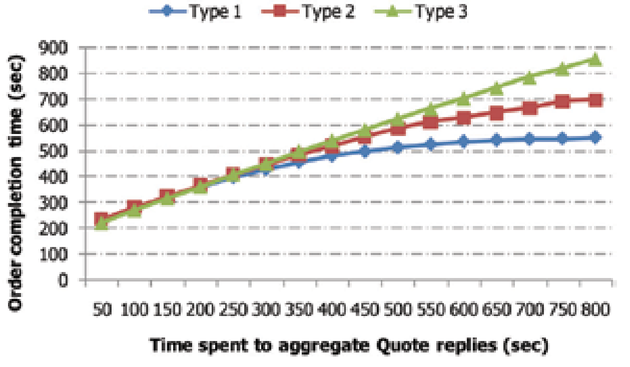


Fig. 16. Order completion time for the timeout strategy

Under the same conditions, we investigated the customer order completion time using the threshold percentage strategy in the Quotes phase. The order completion time looks similar to the behavior seen for the wait time using the threshold percentage strategy in the Quotes phase shown in Figure 14.

From these experimental results, we draw the following conclusions. To enhance the satisfaction of the customers by reducing the customer response time, the system should process orders as soon as possible. On the other hand, aggregating orders can benefit the manufacturer by reducing the manufacturer's costs. Thus, a conflict exists between the customer and the manufacturer, so a balance point must be found. The results also show that the Waiting phase could be adjusted to provide the minimum response time by using a hybrid scheme based on both the threshold and timeout strategies. The analysis of the Quotes phase shows that the manufacturer does not really need to aggregate all possible Quote replies. Above the best quote ratio, increasing the threshold percentage of Quote replies aggregated or the timeout does not affect the manufacturer's gain considerably, but delays the customer's order completion time. The strategies used in the Quotes phase try to reduce the response time by controlling the number of Quote replies received. The best price from the suppliers might not be captured because, again, not all of the Quote replies are taken into account. In the Quotes phase, both the threshold and timeout strategies reduce the customer response time at the cost of decreasing the best price ratio. Again, a hybrid strategy is preferable.

5. Related work

The MIDAS system is an application of the *enterprise-on-demand* model (Stone, 2004), which aims to bridge the gap between business and technology utilizing Business Process Management and the Service Oriented Architecture. Some predict that the enterprise-on-demand model will have a greater impact than the client-server model and will be the killer enterprise application for the Internet.

We have focused in this paper on information flow in supply chains, as realized by MIDAS. Wu *et al. (*2005) have also considered information flow in supply chains. They observe that, in a competitive e-business environment, an enterprise must be involved in managing the supply chain from both the upstream side (the suppliers) and the downstream side (the customers). They propose a novel approach that brings the business processes and services together to support supply chain management.

As noted previously, MIDAS aims to automate the supply chain and, thus, to reduce human intervention, errors and costs, resulting in a more efficient supply chain. Dong and O'Brien (1999) have a similar objective, but they base their business model on the four criteria: profit, lead time, performance, and promptness of delivery. They analyze supply chain performance at two levels: the chain level and the operation level. At the chain level, they set objectives associated with the criteria for each supply chain stage in order to satisfy customer service targets and to select the best supply chain management strategy. At the operation level, they optimize manufacturing and logistics activities for the given targets.

Jinho and Rogers (2005) have investigated the use of UML for building a flexible supply chain business model. They regard a supply chain as five view models with four business domains, where each domain consists of functions, resources, processes, interactions, and business rules. However, they do not describe a system that demonstrates their approach, as we do for MIDAS.

Yang *et al.* (2005) have investigated customization and postponement in supply chains to reduce costs, realize diversity, and improve agility. They consider four kinds of postponement in supply chains: supply, manufacture, delivery, and service. Waller *et al.* (2000) also investigated supply chain customization, and discuss market orientation with supply chain customization. Zhou *et al.* (2003) have also investigated customization in supply chains. They address *mass customization*, which supports customer innovation and which integrates mass production and customized production.

MIDAS realizes the concept of on-demand, just-in-time manufacturing to reduce inventory carrying costs and business logistics costs, by maintaining logical inventory, rather than physical inventory in warehouses. Tanik *et al.* (2001) have proposed a zero-time framework based on what they call the T-strategy, which allows enterprises to adapt in a timely manner to changing market conditions.

In our performance evaluation of MIDAS, we have considered the customer's satisfaction (measured by the average customer response time) and the manufacturer's gain (measured by the number of orders aggregated or the best price ratio). In contrast, Li *et al.* (2005) have considered the average supplier response time in their evaluation of supply chains. Giglio and Minciardi (2003) have investigated the modeling and optimization of supply chains that involve multiple production sites and multiple suppliers, using mathematical programming techniques that aim to minimize costs in the network. Zhao and Jin (2005) have investigated optimized coordination of supply chains based on relationships and dependencies.

Wang *et al.* (2004) observe that appropriate supply chain partners have a large effect on the output value of a supply chain system. They propose an Internet-driven electronic marketplace that can provide an effective platform to select the right partners. In some ways, such a marketplace is like the MIDAS Registry. They present a model of procurement strategies, and discuss factors for the success of such a marketplace.

There exist other papers on build-to-order supply chain management that focus on business and conceptual issues, rather than on the design and implementation of an automated supply chain management system. In particular, Gunasekarana and Ngai (2005) present a review of build-to-order supply chain management and a framework for development. They note that Dell, Compaq and BMW are the best examples of companies that have applied this strategy successfully. Graham and Hardaker (2000) highlight the role of the Internet in building flexible, agile, on-demand supply chains based on the build-to-order strategy. They note that the build-to-order strategy not only addresses diverse customer needs, but also lowers inventory stocks so that parts are pulled from the suppliers as needed, with almost no in-process inventory.

Ghiassi and Spera (2003) discuss industry solutions and best practices for Internet-based supply chains that support mass customization. They emphasize the success of Dell in switching to the build-to-order strategy and gaining a 160% return on its investment. Lancioni *et al.* (2000) discuss the use of the Internet for managing the major aspects of supply chains, particularly for order processing, purchasing/procurement and transportation. They note that enterprises use the build-to-order strategy to achieve lower inventory costs (with somewhat higher production costs), track inventory more accurately, and report the status of orders. Fontanella (2000) observes that Internet technologies make foreign markets more accessible and make it easier to integrate foreign customers, suppliers and intermediate companies into the supply chain, increasing savings and providing innovation.

Forza and Salvador (2002) discuss the difficulties in managing build-to-order systems. They focus particularly on the challenges due to custom customer orders, which must be maintained and processed carefully. Kolish (2000) also discusses the difficulties of the build-to-order model with the production of ships, airplanes and large-scale machine tools. In particular, he focuses on the coordination of fabrication and assembly with respect to scarce capacities.

6. Conclusions and future work

In this article we have presented the design and implementation of the MIDAS system for automated supply chain management. MIDAS allows customers, manufacturers, and suppliers to cooperate over the Internet and the World Wide Web. It aims to meet the needs of the customers on time, and to reduce the costs of the manufacturer by eliminating the need for a large inventory. As a Service Oriented Architecture based on Web Services, MIDAS increases business flexibility and enables businesses to adapt more quickly to changing business needs.

At the manufacturer, MIDAS uses two phases in processing orders from the customers: The Waiting phase in which it aggregates orders from the customers before it makes Quote requests from the suppliers, and the Quotes phase in which it collects Quote replies from the suppliers before deciding on the supplier that will provide the particular supply. The manufacturer can use one of several strategies to decide how long to wait to collect orders from customers or quotes from suppliers. We have presented an evaluation of these strategies based on our implementation.

We have described the services of MIDAS related to information flows. Relationships and dependencies between the components that constitute a material might exist. For example, a particular kind of video card might be usable only with a particular kind of motherboard; therefore, the video card cannot be processed before the motherboard is determined. In the future, we plan to augment the MIDAS system with a workflow component that handles the relationships and dependencies between the components of a material. We also plan to design and implement the services related to product flows and financial flows.

The MIDAS system, presented here, deals with business processes up to the point where the decision to do business with a specific supplier is completed. However, the status of an order might change after the manufacturer has placed an order with a supplier. For example, if the delivery time changes, this information needs to be updated and the customer needs to be informed. MIDAS might use historical data to allow the manufacturer to make decisions at two decision points, the decision to issue Quote requests after collecting orders from the customers and the decision to stop collecting Quote replies and place an order with a particular supplier. Depending on the rates at which the customers place orders, the DB Monitor can adjust its timeout or threshold values accordingly. The Quotes Manager can select better suppliers dynamically by waiting until it receives Quote replies from the suppliers whose Quote replies were previously ignored because they were too late.

7. Acknowledgements

This research was supported in part by the UC Discovery Grant Program and QAD, Inc., grant number COM05-10194.

8. References

G. Alonso, F. Casati, H. Kuno, and V. Machiraju. *Web Services Concepts: Architectures and Applications*. Springer-Verlag, Berlin, 2004

Apache Foundation. Apache Web Server. http://httpd.apache.org/

Apache Foundation. Java implementation of the Universal Description, Discovery, and Integration (UDDI) specification for Web Services. 2008, http://ws.apache.org/juddi/

M. Champion, C. Ferris, E. Newcomer, and D. Orchard. Web Services Architecture. 2002. http://www.w3c.org/TR/2002/WD-ws-arch-20021114/

S. Chatterjee and J. Webber. *Developing enterprise Web Services: An Architect's Guide*. Prentice Hall, Englewood Cliffs, NJ, 2003

L. Dong and C. O'Brien. Integrated decision modeling of supply chain efficiency. *International Journal of Production Economics*, 59(1-3):147-157, March 1999

T. Erl. *Service-Oriented Architecture (SOA): Concepts, Technology, and Design*. Prentice Hall, Englewood Cliffs, NJ, 2005

J. Fontanella. The Web-based supply chain. *Supply Chain Management Review,* 3(4):17-20, 2000

C. Forza and F. Salvador. Managing for variety in the order acquisition and fulfillment process: The contribution of product configuration systems. *International Journal of Production Economics,* 76(1):87-98, 2002

M. Ghiassi and C. Spera. Defining the Internet-based supply chain system for mass customized markets. *Computers and Industrial Engineering,* 45(1):17-41, 2003

D. Giglio and R. Minciardi. Modeling and optimization of multi-site production systems in supply chain networks. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 3, pages 2678-2683, October 2003

G. Graham and G. Hardaker. Supply-chain management across the Internet. *International Journal of Physical Distribution and Logistics Management,* 30(3/4):286-295, 2000

A. Gunasekarana and E. W. T. Ngai. Build-to-order supply chain management: A literature review and framework for development. *Journal of Operations Management,* 23(5):423-451, 2005

K. Jinho and K. Rogers. An object-oriented approach for building a flexible supply chain model. *International Journal of Physical Distribution and Logistics Management*, 35(7):481-502, July 2005

R. Kolisch. Integration of assembly and fabrication for make-to-order production. *International Journal of Production Economics,* 68(3):287-306, 2000

R. A. Lancioni, M. F. Smith, and T. A. Oliva. The role of the Internet in supply chain management: Logistics catches up with strategy. *Industrial Marketing Management,* 29(1):45-56, 2000

Y. S. Li, F. F. Ye, Z. M. Fang, and J. G. Yang. Flexible Supply chain optimization and its SRT analysis. *Industrial Engineering and Management*, 10(1):89-93, 2005

L. E. Moser and P. M. Melliar-Smith. Service Oriented Architecture and Web Services. *Encyclopedia of Computer Science and Engineering*, John Wiley & Sons, New York, NY, 2008

R. Murch. *Autonomic Computing.* On Demand Series, IBM Press, April 2004

E. Newcomer and G. Lomow. *Understanding SOA with Web Services.* Independent Technology Guides, Addison-Wesley Professional Series, December 2004

OASIS. Reference Model for Service Oriented Architecture, 2006. http://www.oasis-open.org/committees/tc home.php?wg abbrev=soa-rm

PlanetLab Global Research Network. http://www.planet-lab.org

A. Stone. Demanding Internet enterprise. *IEEE Internet Computing*, 8(3):13-14, 2004

U. Tanik, M. M. Tanik, and L. Jololian. Internet enterprise engineering. A “zero-time” framework based on “T-strategy". *Proceedings of the IEEE Southeast Conference*, pages 263-270, Clemson, SC, March 2001

M. A. Waller, P. A. Dabholkar, and J. J. Gentry. Postponement, product customization, and market-oriented supply chain management. *Journal of Business Logistics*, 2000

L. P.Wang, F. Y. Qiu, H. R. Dai, and Z. C. Chen. Relationships of electronic marketplace and partner selection of supply chain. *Computer Integrated Manufacturing Systems*, 10(5):550-555, 2004

B.Wu, M. Dewan, L. Li, and Y. Yang. Supply chain protocolling. *Proceedings of the 7th IEEE International Conference on E-Commerce Technology*, pages 314-321, July 2005

J. Yang, S. Zhao, and J. Wang. Study on postponement in customization supply chain. *Industrial Engineering and Management*, 10(4):35-44, 2005

T. Z. Zhao and Y. H. Jin. Optimized coordination of supply chains based on dependency. *Computer Integrated Manufacturing Systems*, 10(8):929-933, 2005

W. Zhao, F. Kart, L. E. Moser, and P. M. Melliar-Smith. A reservation-based extended transaction protocol for coordination of Web Services. *Journal of Web Services Research*, 5(3):64-95, 2008

D. Zhou, B. H. Xiang, and G. S. Zou. The strategy of mass customization and the countermeasure of Chinese enterprises. *Industrial Engineering and Management*, 8(5):12-16, 2003

15

Distributed Supply Chain Planning for Multiple Companies with Limited Local Information

Tatsushi Nishi
Mathematical Science for Social Systems
Graduate School of Engineering Science
Osaka University
Japan

1. Introduction

With rapid progress in market liberalization of a variety of products, intermediates, or electrical power energy, many companies are trying to integrate their enterprises with other organizations by optimizing planning and distribution for multiple companies. Supply chain coordination for Business to Business (B2B) is remarkably increasing for several enterprises at different industry levels with the progress of deregulation of trading products in recent years. Partner companies are reducing distribution costs while keeping high service levels to vendors. For example, in the petroleum industry, several companies at a chemical complex depend on other companies to provide raw materials or to deliver intermediates through a shared pipeline connected to multiple companies at lower delivery costs (Taneda, 2003).

In most companies, strategic instructions are given to each section in the company. Each company is thoroughly capable in regard to its own production planning and scheduling. The supply and demand planning must be coordinated by mutual negotiations across supply chain. Such coordination has been performed by the communications among human operators. However, in recent years, the decision making for each company is becoming increasingly complex with huge number of alternative planning for a number of companies. Conventional planning system has been configured to obtain a near optimal planning with detailed information about multiple companies. Organizations generally have their own private coordination methods, and that accessing others' private information or intruding on their decision-making authority should be avoided. In a practical situation, a plan must be created without sharing such confidential information as production cost, inventory holding costs, or price of products for competitive companies. In this chapter, a framework for distributed supply chain planning system without requiring all of information for multiple companies is proposed. Planning coordination can be efficiently automated by the proposed method.

Various types of supply chain models have been proposed in the literature (Vidal & Goetschalckx, 1997). A midterm planning model involving maximization of total profits and minimization of production, inventory, and transportation costs has been developed within a company (McDonald & Karimi, 1997). Planning coordination problems for production and distribution for multiple organizations are studied by Gaonkar and Viswanadham, 2002, Jackson and Grossmann, 2003 and Jayaraman and Pirkul, 2001. Schedule coordination problem for multiple organizations has been studied in Luh et al. 2003. This work concentrates on a planning model for multiple organizations from different companies.

Typical approach for supply chain planning is to use simulation-based approach combined with discrete optimization methods or to solve integer programming methods. Conventional supply chain planning systems have been configured to obtain near-optimal plans incorporating an information sharing for overall supply chain with detailed data-exchanging with multiple companies. Simulation-based methods for supply chain management (Julka et al. 2002, Tu et al. 2003) require detailed and precise information for all entities to analyze the performance of the entire supply chain. However, in practice, such information as production costs, inventory holding costs, or price of products are considered confidential for competing companies. For such reasons of confidentially, a distributed planning system with partial information sharing is preferable for supply chain planning for multiple companies. It requires the development of a distributed planning system in which each company can generate its own planning with partial information sharing with other companies.

In this chapter, a distributed supply chain planning system for multiple companies with partial information sharing is studied. Supply chain planning problems for multiple companies are formulated as mixed integer programming problems. A Lagrangian relaxation method is applied to decompose the overall problem by relaxing interconnection constraints between suppliers and vendors. Lagrangian relaxation is an optimization method that derives a lower bound by removing complicating constraints from constraint sets and replaces them with a penalty term in the objective function that can be decomposed into multiple solvable subproblems (Fisher, 1973),. Scheduling methods based on Lagrangian relaxation methods have been widely used to improve computation efficiency with near-optimal solution for jobshop problems (Hoitomt et al., 1993). The Lagrangian decomposition and coordination method has been applied to an asynchronous distributed decision making problems (Androulakis & Reklaitis, 1999). In this method, improvements of the Lagrangian multiplier value and generation of a solution of each subproblem are iteratively repeated. For planning problems, the method has been applied to supply chain planning problems in which machine capacity constraints are relaxed (Gupta & Maranus, 1999). Supply chain coordination problems for multiple organizations have been extensively studied by Luh et al. 2003. The problem for overall organizations is decomposed into individual organization-based subproblems relaxing precedence constraints. The performance of a price-based Lagrangian relaxation method is compared with an auction-based method.

The decomposition techniques for supply chain planning problem have been addressed before. For conventional decomposition methods, the solution derived in a distributed manner is coordinated by the heuristic procedure using the entire information for multiple entities. These heuristics are often problem-dependent to generate near-optimal solution. Thus it is difficult to construct good heuristics for general problems.

In this study, an augmented Lagrangian approach with a quadratic penalty function is used to decompose the original problem to eliminate duality gap. The augmented Lagrangian approach has recently extensively studied in short-term hydrothermal coordination (Beltran & Herdia, 1999), unit commitment (Beltran & Herdia, 2002, Georges, 1994) in power systems. However, the approach has not ever been applied to distributed algorithm in supply chain. The main difficulty is that the quadratic penalty term is not decomposable.

Firstly, the original problem is decomposed into several subproblems for each company with an objective function with a weighted penalty function. This method is called multiplier method for continuous convex optimization problems. By applying it to mixed integer programming problems, a feasible solution is expected to be derived by gradually increasing the weighting factor without using the construction procedure of feasible solutions. By adopting the approach, it may be possible to coordinate the entire plan without sharing confidential information among competing companies. The properties and the performance of the proposed approach for supply chain is investigated for a simple delivery/receiving planning problem for multiple companies in a petroleum complex, and a mid-term planning problem for multiple companies for more realistic model.

The rest of the chapter is organized as follows. In Section 2, the supply chain planning problem for multiple companies in a petroleum complex is formulated as a mixed integer linear programming problem. The formulation for distributed optimization using the augmented Lagrangian decomposition and coordination method is introduced in Section 3. Computational results including the discussion of the optimality of the solution by the proposed method, and the comparison of the proposed method with other distributed optimization methods, are shown in Section 4. Section 5 describes our conclusions and future research direction.

2. Supply chain planning for multiple companies

2.1 Problem definition and formulation

In this section, the supply chain planning problem in a petroleum complex is formulated as a mixed integer linear programming problem (MILP). Two layers of supply chain consisting of suppliers and manufacturers are treated in this work. Consider both of supplier companies and vendor companies in a petrochemical complex equipped with a shared pipeline for the delivery of intermediate products. The pipeline has a restricted capacity only for the delivery of one type of product at a time during a time period. Each company has its own demanded supply and demand plan for products. The supply chain planning problem is to determine a supply and demand planning for each supplier and vendor companies satisfying the constraints for the shared pipeline.

The following conditions (i)-(iv) are assumed:

i. Supply/demand quantity is restricted by the capacity of the shared pipeline.
ii. No more than two supplier companies can deliver/receive products into more than two vendor companies during the same time period.
iii. Production delivery/receiving cannot be interrupted for a pre-specified *K* time period once it has started.
iv. Total delivery/receiving quantity for the companies must be equal to the pre-determined quantity of products during total time horizon *H*.

Each company can adopt any optimization model. In this study, the following optimization model is adopted.

Let z_S and z_C denote a set of supplier companies and a set of vendor companies treating a set of products *P*. Supplier company $c \in Z_S$ have a demanded delivery plan $D_{i,t}^c$ $(i \in P; t = 1, \ldots, H)$, and vendor company $c \in Z_C$ have a demanded receiving plan $D_{i,t}^c$ $(i \in P; t = 1, \ldots, H)$, respectively. Each company has its own objective function $f^c\,(c \in Z_S \cup Z_C)$. To simplify the expression, but to formulate general problems, it is assumed for all companies that the objective function for each company f^c consists of the sum of the following costs:

Penalty costs for deviation of delivery/receiving plan $S_{i,t}^c$ from the demanded delivery/receiving quantity $D_{i,t}^c$ that corresponds to inventory holding costs or to due date penalties.

i. For chemical complex or manufacturing systems, it is desirable that the delivery/receiving quantity is almost the same as that of previous periods due to changeover costs for flow rate. Thus, penalty costs are imposed if the delivery/receiving quantity in time period *t* is different from that in time period $(t-1)$.

ii. Transportation costs imposed at each company that correspond to pipeline usage costs.

The objective function f^c for company $c \in Z = \{Z_S \cup Z_C\}$ is given by the following equation.

$$f^c(S_{i,t}^c, X_{i,t}^c, Y_{i,t}^c) = \mu_{i,t}^c \mid D_{i,t}^c - S_{i,t}^c \mid + e_i^c X_{i,t}^c + d_{i,t}^c Y_{i,t}^c \tag{1}$$

The supply chain planning problem for multiple companies is formulated as a mixed integer linear programming problem.

$$(P_0) \ \min \sum_{c \in Z} \sum_{i \in P} \sum_t f^c(S_{i,t}^c, X_{i,t}^c, Y_{i,t}^c) \tag{2}$$

subject to

$$\sum_{c \in Z_S} S_{i,t}^c = \sum_{c \in Z_C} S_{i,t}^c \quad (\forall i \in P; \forall t = 1, \ldots, H) \tag{3}$$

$$\sum_t S_{i,t}^c = m_i^c \quad (\forall c \in Z; \forall i \in P) \tag{4}$$

$$\sum_{i \in P} Y_{i,t}^c \le 1 \quad (\forall c \in Z; \forall t = 1, \ldots, H) \tag{5}$$

$$S_{i,t}^c \le s_i^{\max} Y_{i,t}^c \quad (\forall c \in Z; \forall i \in P; \forall t = 1, \ldots H) \tag{6}$$

$$S_{i,t}^c \ge s_i^{\min} Y_{i,t}^c \quad (\forall c \in Z; \forall i \in P; \forall t = 1, \ldots H) \tag{7}$$

$$Y_{i,t}^c - \sum_{i \in P} Y_{i,t-1}^c \le Y_{i,t+1}^c \quad (\forall c \in Z; \forall i \in P; \forall t = 2, \ldots, H) \tag{8}$$

$$(K+1)(1 - Y_{i,t-1}^c) + K Y_{i,t}^c \ge \sum_{i \in P} \sum_{t'=0}^{K-1} Y_{i,t+t'}^c \quad (\forall c \in Z; \forall i \in P; \forall t = 2, \ldots, H - K + 1) \tag{9}$$

$$X_{i,t}^c = \begin{cases} 1 & (S_{i,t}^c \ne S_{i,t-1}^c) \\ 0 & (S_{i,t}^c = S_{i,t-1}^c) \end{cases} \quad (\forall c \in Z; \forall i \in P; \forall t = 2, \ldots H) \tag{10}$$

$$Y_{i,t}^c = \begin{cases} 1 & (S_{i,t}^c > 0) \\ 0 & (S_{i,t}^c = 0) \end{cases} \quad (\forall c \in Z; \forall i \in P; \forall t = 1, \ldots, H) \tag{11}$$

where $s_i^{\max}, s_i^{\min}$: minimum/maximum quantity of delivery for product *i* in time period *t*,
$D_{i,t}^c$: demanded delivery/receiving quantity of delivery for product *i* in time period *t* for company *c*,
$d_{i,t}^c$: transportation cost of the usage of pipeline for the delivery for product *i* in time period *t* for company *c*,
e_i^c : penalty incurred by the difference of delivery/receiving quantity of product *i* between the previous time period for company *c*,
K : set up time duration,
m_i^c : total delivery/receiving quantity for of product *i* during the total time horizon for company *c*,
$\mu_{i,t}^c$: penalty for deviation from demanded delivery/receiving quantity in time period *t* for company *c*,
$S_{i,t}^c$: delivery/receiving quantity of product *i* for company *c* in time period *t*,
$X_{i,t}^c$: binary variable which takes a value of 1 if delivery/receiving quantity of product *i* in time period *t* is different from that in time period (*t* – 1) for company *c*, and 0 otherwise,
$Y_{i,t}^c$: binary variable which takes a value of 1 if product *i* is delivered/received in time period *t* for company *c*.

The overall objective function given by (1) and (2) consists of sum of the penalty for deviating from the demanded delivery/receiving plan for each company, and the penalty for difference of delivery/receiving quantity, and transportation costs for delivery/receiving. (3) represents that the total delivery/receiving quantity from suppliers is equal to the total quantity of demand for vendor companies at each time period. (4) specifies that the total delivery/receiving quantity must be equal to a pre-determined set point from condition (iv). (5) implies that each company can deliver/receive only one type of product during a time period from condition (ii). (6) and (7) restrict delivery/receiving quantity must be less than $s_i^{\max}$ when $Y_{i,t}^c = 1$ and it is greater than $s_i^{\min}$ in each time period from condition (i). (8) denotes delivery/receiving duration constraints from condition (iii). It indicates that delivery/receiving cannot be interrupted at least two conservative time periods (if $(\sum_{i\in P} Y_{i,t-1}^c = 0) \wedge (Y_{i,t}^c = 1)$, then $(Y_{i,t+1}^c = 1)$ where the operation $\wedge$ stands for conjunction. (9) describes setup time constraints indicating that K conservative time periods are necessary for set up when the delivered/received product type is changed if $(Y_{i,t-1}^c = 1) \wedge (Y_{i,t}^c = 0)$, then $\sum_i \sum_{t'=1}^{K-1} Y_{i,t+t'}^c = 0$. $X_{i,t}^c$ in (10) is a binary variable which takes a value of 1 if the delivery/receiving quantity in time period t is different from that in time period $(t-1)$ and 0 otherwise. It can be realized by (12) and (13).

$$S_{i,t}^c + UX_{i,t}^c \geq S_{i,t-1}^c \quad (\forall c \in Z; \forall i \in P; \forall t = 2,\ldots,H) \tag{12}$$

$$S_{i,t-1}^c + UX_{i,t}^c \geq S_{i,t}^c \quad (\forall c \in Z; \forall i \in P; \forall t = 2,\ldots,H) \tag{13}$$

where U is a sufficiently large constant. (12) ensures $S_{i,t}^c \geq S_{i,t-1}^c$ when $X_{i,t}^c = 0$, on the other hand, (13) indicates $S_{i,t-1}^c \geq S_{i,t}^c$ when $X_{i,t}^c = 0$. If the delivery quantity in time period t is different from that in time period $(t-1)$, then $X_{i,t}^c = 1$, otherwise $X_{i,t}^c = 0$. (11) represents binary variable which takes a value of 1 if product i is delivered in time period t, then $Y_{i,t}^c = 1$, and $Y_{i,t}^c = 0$ otherwise.

3. Distributed optimization using Lagrangian decomposition and coordination method

3.1 Decomposable formulation

The original problem (P_0) can be easily solved by a commercial MILP (Mixed Integer Linear Programming) solver if the problem size is sufficiently small by using all of the information for the companies. However, when competing companies participate in the supply chain, such confidential information as cost data or product prices for other companies cannot be collected. In such situation, it is required to generate a feasible plan in a distributed environment without requiring all of the information. The distributed optimization without using all of information is necessary. From that viewpoint, a distributed optimization method using Lagrangian decomposition and coordination method is explained. The problem (P_0) is decomposed into several subproblems for each company by Lagrangian decomposition and coordination method (LDC method).

The LDC method is considered as a distributed optimization method to derive near-optimal solution efficiently by relaxing several constraints such that the original problem can be decomposed into subproblems. The constraints given by (3) in the original problem (P_0) are relaxed by Lagrangian multiplier $\lambda_{i,t}$, then the Lagrangian relaxation problem (R_0) to minimize Lagrangian function L with fixed Lagrangian multipliers, is described as Lagrangian relaxation problem (R_0).

$$(R_0)\ \min L$$

$$\begin{aligned} L &= \sum_{c\in Z}\sum_{i\in P}\sum_{t} f^c(S_{i,t}^c, X_{i,t}^c, Y_{i,t}^c) - \sum_{i\in P}\sum_{t}\lambda_{i,t}\left(\sum_{c\in Z_S} S_{i,t}^c - \sum_{c\in Z_C} S_{i,t}^c\right) \\ &= \sum_{c\in Z_S}\sum_{i\in P}\sum_{t}\left\{f^c(S_{i,t}^c, X_{i,t}^c, Y_{i,t}^c) - \lambda_{i,t} S_{i,t}^c\right\} + \sum_{c\in Z_C}\sum_{i\in P}\sum_{t}\left\{f^c(S_{i,t}^c, X_{i,t}^c, Y_{i,t}^c) + \lambda_{i,t} S_{i,t}^c\right\} \end{aligned} \tag{14}$$

The Lagrangian function L of (14) is additive for each company. Therefore the original problem (P_0) can be decomposed into subproblems for individual company. The subproblem for supplier $c \in Z_S$ and the subproblem for vendor company $c \in Z_C$ are represented as (15) and (16) when the multipliers are fixed.

For supplier company $c \in Z_S$

$$(SP_{0c}^s)\quad \min \sum_{i\in P}\sum_{t}\left\{f^c(S_{i,t}^c, X_{i,t}^c, Y_{i,t}^c) - \lambda_{i,t} S_{i,t}^c\right\} \tag{15}$$

For vendor company $c \in Z_C$

$$(SP_{0c}^c)\quad \min \sum_{i\in P}\sum_{t}\left\{f^c(S_{i,t}^c, X_{i,t}^c, Y_{i,t}^c) + \lambda_{i,t} S_{i,t}^c\right\} \tag{16}$$

subject to (1), (4)- (11).

3.2 Solving Lagrangian dual problem

The dual problem (D_0) for the original problem is represented as the following equation.

$$(D_0) \quad \max_{\{\lambda_{i,t}\}} q(\{\lambda_{i,t}\}) \quad \text{where } q(\{\lambda_{i,t}\}) = \min L \tag{17}$$

To solve the dual problem, subgradient optimization method is used in most cases. The Lagrangian multipliers are updated according to (18). The steps of the solving subproblems and the updating Lagrangian multipliers are iteratively repeated until dual solution has not been updated.

$$\lambda_{i,t} = \begin{cases} \lambda_{i,t} + \Delta\lambda \ (\sum_{c \in Z_S} \overline{S_{i,t}^c} < \sum_{c \in Z_C} \overline{S_{i,c}^c}) \\ \lambda_{i,t} - \Delta\lambda \ (\sum_{c \in Z_S} \overline{S_{i,t}^c} > \sum_{c \in Z_C} \overline{S_{i,c}^c}) \end{cases} \tag{18}$$

$S_{i,t}^c$ represents tentative delivery quantity derived by solving a subproblem in a previous iteration. $\overline{\lambda_{i,t}}$ represents Lagrangian multiplier in a previous iteration (tentative value of Lagrangian multipliers). If all of the information for other companies is available, a step size of Lagrangian multiplier $\Delta\lambda$ is calculated by (19).

$$\Delta\lambda = \gamma \frac{\overline{L} - \underline{L}}{(\sum_{c \in Z_S} \overline{S_{i,t}^c} - \sum_{c \in Z_C} \overline{S_{i,t}^c})^2} \tag{19}$$

where γ is a positive coefficient satisfying $0 < \gamma < 2$, $\overline{L}$ is upper bound of the original problem, and $\underline{L}$ is lower bound obtained by calculating L for the solution of subproblems.
The algorithm of the LDC method is described in the following steps.
Step 1: Initialization of multipliers.
Step 2: Generation of the solution of subproblem for each company with fixed multipliers. The lower bound $\underline{L}$ is calculated.
Step 3: Generation of a feasible solution using a solution derived at Step 2. The upper bound $\overline{L}$ is calculated.
Step 4: Evaluation of convergence. The condition for convergence is that the duality gap calculated by $\dfrac{\overline{L} - \underline{L}}{\underline{L}}$ has not been updated at a pre-specified number of times.
Step 5: Update of the multipliers by (18), (19) and return to Step 2.
The solution of dual problem is not always feasible for nonconvex optimization problem when the problem includes the setup costs depending on product type in the objective function. In order to obtain a feasible solution at Step 3, the construction of a feasible solution is necessary for LDC method to calculate $\overline{L}$ by modifying the solution of subproblems using a heuristic procedure. Simple priority-based heuristics such like FIFO rules, etc. backward or backward-forward heuristics are often used to generate a feasible solution. The performance of LDC method highly depends on the performance of heuristics to modify the dual solution into a feasible one requiring all of the information corrected to apply these heuristic procedures. However, it is difficult to construct a heuristic to obtain a

good upper bound. Moreover, the solution oscillations often occur if the dual solution is not identical to the primal optimal solution that makes the algorithm to find a feasible solution. To obtain a feasible solution, we applied an augmented Lagrangian approach (Rockafellar, 1974) without using heuristic procedure. The method is called as a multiplier method (Bertsekas, 1976), which is commonly used for continuous optimization problems. The supply chain planning problem for multiple companies is solved by the distributed optimization approach.

3.3 Decomposable reformulation for augmented Lagrangian approach

The main drawback of the augmented Lagrangian approach is that the quadratic penalty term introduced by the augmented Lagrangian is not separable into each subproblem for a company. To make the problem decomposable, a linear approximation technique of the cross penalty terms around a tentative solution has been proposed (Androulakis and Reklaitis, 1999, Cohen and Zhu, 1984, Stephanopoulas and Westerberg, 1975). Let us consider a simple supply chain planning problem for supplier company a and vendor company b treating product $i \in P$. An augmented Lagrangian function with the quadratic penalty term is given by the following equation.

$$L_r = \sum_{i \in P} \sum_t \left\{ f^a + f^b + \lambda_{i,t} (S_{i,t}^b - S_{i,t}^a) \right\} + r \sum_{i \in P} \sum_t (S_{i,t}^b - S_{i,t}^a)^2 \tag{20}$$

The Lagrangian function of (20) cannot be decomposed because the cross-product term $S_{i,t}^a S_{i,t}^b$ is included in the penalty term. To keep the problem decomposable, a first order Taylor series of expansion around the tentative solution $(\overline{S_{i,t}^a}, \overline{S_{i,t}^b})$ is used. The augmented Lagrangian function can be reformulated as:

$$\begin{aligned} L_r &= \sum_i \sum_t (f^a - \lambda_{i,t} S_{i,t}^a) + \sum_i \sum_t (f^b + \lambda_{i,t} S_{i,t}^b) \\ &\quad + r \sum_i \sum_t \left\{ (S_{i,t}^a)^2 + (S_{i,t}^b)^2 - 2 S_{i,t}^a \overline{S_{i,t}^b} - 2 S_{i,t}^b \overline{S_{i,t}^a} + 2 \overline{S_{i,t}^a}\,\overline{S_{i,t}^b} \right\} \\ &= \sum_i \sum_t (f^a - \lambda_{i,t} S_{i,t}^a) + r \sum_i \sum_t \left\{ (S_{i,t}^a)^2 - 2 S_{i,t}^a \overline{S_{i,t}^b} + \overline{S_{i,t}^a}\,\overline{S_{i,t}^b} \right\} \\ &\quad + \sum_i \sum_t (f^b - \lambda_{i,t} S_{i,t}^b) + r \sum_i \sum_t \left\{ (S_{i,t}^b)^2 - 2 S_{i,t}^b \overline{S_{i,t}^a} + \overline{S_{i,t}^a}\,\overline{S_{i,t}^b} \right\} \end{aligned} \tag{21}$$

r is a positive scalar parameter. (21) states that the problem to minimize the Lagrangian function L of (20) with fixed multipliers can be decomposed into subproblems for each company by using the tentative solution at each iteration. The decomposed function consists of sum of the objective function for each company, a multiplier penalty term, and a quadratic penalty term. The minimization problem for each company is a mixed integer quadratic programming (MIQP) problem which is difficult to be solved in reasonable computation time. Therefore, in our study we replaced the quadratic penalty term by a linear penalty term shown in (22). By using this reformulation, we do not have to use nonlinear optimization methods for solving MIQP problem.

The new decomposed function for company a: $L_a^{'}$ can be given by:

$$L_a^{'} = \sum_{i \in P} \sum_t (f^a - \lambda_{i,t} S_{i,t}^a) + r \sum_{i \in P} \sum_t | S_{i,t}^a - \overline{S_{i,t}^b} | \tag{22}$$

The penalty parameter r for a linear penalty term is gradually increased in each iteration. By applying the linear approximation technique around a tentative solution for the proposed method, the solutions derived by solving subproblem for each company cannot provide an exact lower bound of the original problem.

3.4 Coordination of supply chain planning among multiple companies

A sequence of optimization problems E_0^k can be given by (23) where L is given by (14).

$$(E_0^k) \quad \min (L + r_k | \sum_{c \in Z_S} S_{i,t}^c - \sum_{c' \in Z_C} S_{i,t}^{c'} |^2) \tag{23}$$

The decomposed subproblem for each company is reformulated as (24) and (25) by applying the first order Taylor series of expansion around a tentative solution.

For supplier company $c \in Z_S$

$$(EP_{0c}^k) \quad \min \sum_{i \in P} \sum_t \left\{ f^c (S_{i,t}^c, X_{i,t}^c, Y_{i,t}^c) - \lambda_{i,t} S_{i,t}^c \right\} + r_k \sum_{i \in P} \sum_t | S_{i,t}^c + \sum_{c' \in Z_S / \{c\}} S_{i,t}^{c'} - \sum_{c' \in Z_C} S_{i,t}^{c'} | \tag{24}$$

For vendor company $d \in Z_C$

$$(EP_{0c}^k) \quad \min \sum_{i \in P} \sum_t \left\{ f^d (S_{i,t}^d, X_{i,t}^d, Y_{i,t}^d) + \lambda_{i,t} S_{i,t}^d \right\} + r_k \sum_{i \in P} \sum_t | S_{i,t}^d + \sum_{c' \in Z_C / \{d\}} S_{i,t}^{c'} - \sum_{c' \in Z_S} S_{i,t}^{c'} | \tag{25}$$

subject to (1), (4)-(11)

The subproblem for each company is an MILP problem, which can be solved by a commercial solver. r_k represents a weighting factor for penalty function. To derive near-optimal solution for the proposed method, the weighting factor r_k must be gradually increased according to the following equation.

$$r_{k+1} = r_k + \Delta r \tag{26}$$

Δ r is the step size parameter for penalty weighting coefficient which should be determined by preliminary tests. Even though the objective function includes a linear penalty function for each subproblem, a lower bound of the original problem can be obtained by calculating L for the solution of subproblem when r_k is set to zero.

3.5 Scenario of planning coordination for multiple companies

The system generates near-optimal plan in the following steps.

Step 1: Initialization

$k \leftarrow 0$. The multipliers $\lambda_{i,t}$ and the weighting factor r_k are set to an initial value (*e.g.* set to zero).

Step 2: Generation of an initial plan

A manager for each company inputs the demanded delivery/receiving plan at each time period $D_{i,t}^c$ and the total delivery/receiving quantity for each product during time horizon. Each company solves each subproblem and generates a tentative plan with the fixed multipliers.

Step 3: Data exchange of tentative solution

Each company exchanges the data of tentative delivery/receiving quantity of products $\overline{S_{i,t}^{c}}$ derived at each company.

Step 4: Evaluating the convergence

If the plan generated at Step 6 or Step 2 for initial iteration satisfies the following conditions, the algorithm is considered as convergence. Then no more calculation is made and the derived plan is regarded as a final plan.

i. The solution derived at Step 6 is the same as that generated at Step 6 in a previous iteration.
ii. The solution derived at Step 6 satisfies the constraints (3).
iii. The solutions of all other companies also satisfy both of conditions (i) and (ii).

Step 5: Update of the multiplier and the weighting factor

The weighting factors are updated by (26) and the multipliers are updated by (18).

Step 6: Solving subproblems

A company solves its subproblem while the solution of other company is fixed. Then, the tentative solution $S_{i,t}^{c}$ is updated and return to Step 3. If some of the companies derive its solutions concurrently in parallel at Step 6, the same solution is generated cyclically because tentative solution of a previous iteration is used, that makes the convergence of the algorithm more difficult. Skipping heuristic (Nishi et al., 2002) is effective to avoid such situations. Skipping heuristic is a procedure that the Step 6 for each company is randomly skipped. If the proposed method is implemented on a parallel processing system, the procedure must be added to avoid cyclic generation of solutions. Our numerical experiments used a sequential computation that the Step 6 for each company is sequentially executed to avoid the difficulty of convergence without skipping heuristic.

The data exchanged among companies is tentative supply and demand quantity in each time period. This information is not directly concerned with confidential information for each company. The multipliers are updated by (27) without using the information of $\overline{L}-\underline{L}$ for the step size because the upper bound is not calculated for augmented Lagrangian approach.

$$\lambda_{i,t}=\begin{cases}\lambda_{i,t}^{*}+\Delta\lambda & (\overline{S_{i,t}^{c}}<\overline{S_{i,t}^{c'}})(c\in Z_S;\, c'\in Z_C)\\ \lambda_{i,t}^{*}-\Delta\lambda & (\overline{S_{i,t}^{c}}>\overline{S_{i,t}^{c'}})(c\in Z_S;c'\in Z_C)\\ \lambda_{i,t}^{*} & (\overline{S_{i,t}^{c}}=\overline{S_{i,t}^{c'}})(c\in Z_S;c'\in Z_C)\end{cases} \tag{27}$$

$S_{i,t}^{c}$ represents a tentative solution obtained by solving subproblem for company c. $\Delta\lambda$ is the step size given as a scalar parameter, and $\lambda_{i,t}^{*}$ is the value of multipliers at a previous iteration. For the proposed system, $\Delta\lambda$ is considered as a constant step size without generation of a feasible solution for the entire company. All of the information that is exchanged at each iteration during the optimization is the tentative delivery quantity $S_{i,t}^{c'}(c'\in Z_S\cup Z_C\setminus\{c\})$ derived at other companies. Each company has the same value of its own multipliers and updates the value of them for itself. Thus the dual problem can be solved in a distributed environment without exchanging such confidential data as cost information for the proposed method.

4. Computational experiments

4.1 Supply chain planning for 1 supplier and 2 vendor companies

An example of supply chain planning problem for 1 supplier (A) and 2 vendor companies (B, C) treating with 2 types of products is solved. The total time horizon is 30 time periods. The parameters for the problem are generated by random numbers on uniform distribution in the interval shown in Table 1. The demanded delivery/receiving plan which is input data for each company is illustrated in Fig. 1. The result obtained by the proposed method is also shown in Fig. 2. The numbers printed in the figure indicate the delivery and receiving quantity for each company. The program is coded by C++ language. A commercial MILP solver, CPLEX8.0 ILOG(C) is used to solve subproblems. A Pentium IV 2AGHz processor with 512 MB memory was used for computation.

The optimality of solution is minimized when $\Delta r = 0.01$ and $\Delta\lambda = 0.1$ from several preliminary tests. These parameters are used for computation in the following example problems.

	Supplier company $c \in Z_S$	Vendor company $c \in Z_C$
$D_{i,t}^c$	0 – 200	0 – 180
$\mu_{i,t}^c$	1 – 10	1 – 10
$d_{i,t}^c$	1 – 10	1 – 10
e_i^c	10 – 30	10 – 30
$m_{i,t}^c$	1500 – 4000	750 – 2000

Table 1. Parameters for the example problems

Augmented Lagrangian decomposition method (ALDC)

$\Delta\lambda = 0.1$, $\Delta r = 0.01$

Lagrangian decomposition method (LDC)

$\gamma = 0.1$

Penalty method (PM)

$\Delta r = 0.01$ (Case 1, Case 2), $\Delta r = 0.1$ (Case 3)

Table 2. Parameters for the distributed optimization method

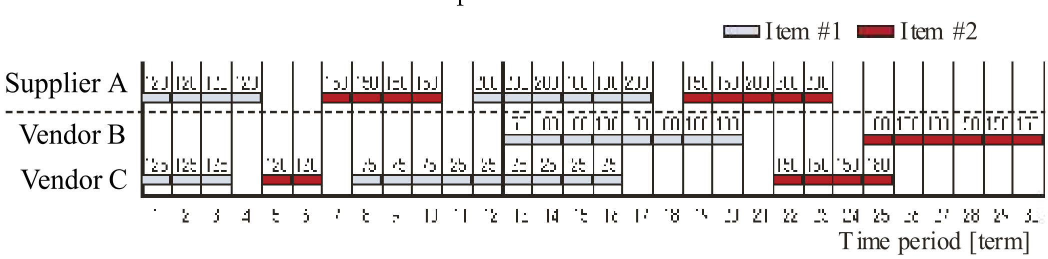


Fig. 1. An initial request for the plan (1 supplier and 2 vendor companies)

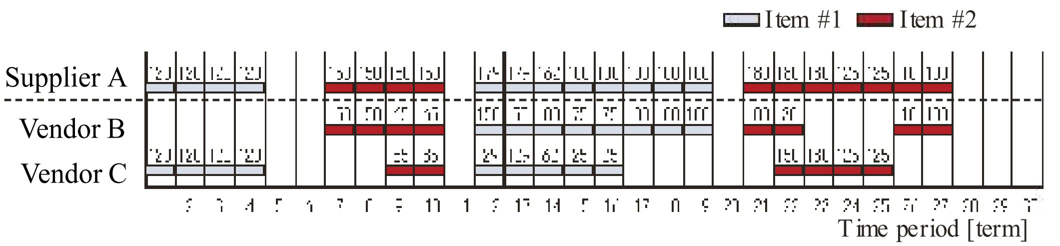


Fig. 2. Result of distributed supply chain planning by the proposed method (after 72 times of data exchanges)

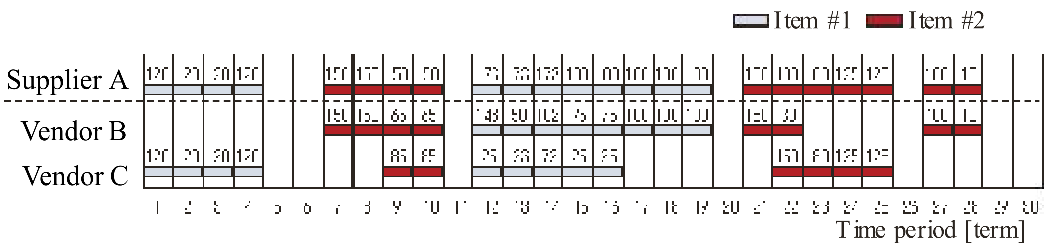


Fig. 3. The optimal solution derived by CPLEX solver

The proposed method generates a feasible solution for the problem after 72 iterations using the parameters shown in Table 2. The result is shown in Fig. 2. An optimal solution derived by commercial solver is also shown in Fig. 3. The result obtained by the proposed method is almost the same as that of an optimal solution. The transition of the value of L_r and the decomposed function $L_r^{'}$ for each company c is shown in Fig. 4. The condition for evaluating convergence is that the difference of the delivery and receiving quantity is less than 0.01 for all products and for all time periods. The optimal value of the objective function of (2) obtained by the proposed method is 9,979. The value for the optimal solution obtained by the commercial MILP solver with all of the information is 9,960. The gap between the derived solution and the optimal solution is 0.18%. It demonstrates that the proposed method can derive near-optimal solution without requiring all of the information for other companies.

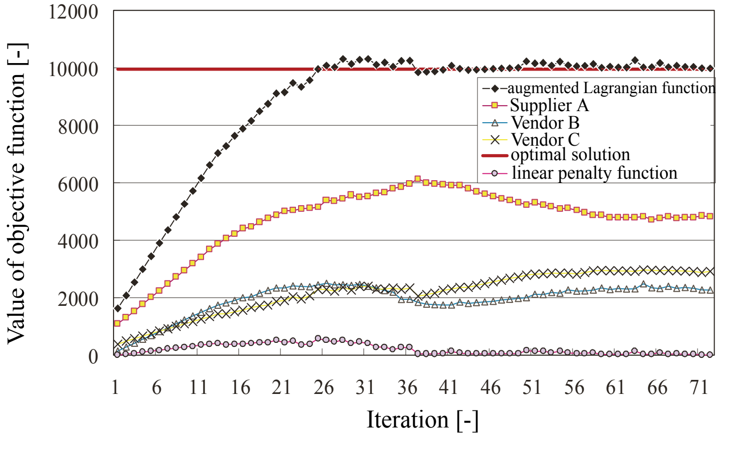


Fig. 4. Transition of the value of objective function for the proposed method

4.2 Comparison with other distributed optimization methods

To investigate the performance of the proposed method, the performance of the proposed method (ALDC method) is compared with other distributed optimization methods: a penalty method (PM method) that the terms of Lagrangian multipliers are removed from (24) and (25), and an ordinary LDC method (LDC method).

For the LDC method, the dual problem D_0 is solved by standard Lagrangian function. The dual solution is modified to generate a feasible solution with the following heuristic procedure at each iteration. The heuristic procedure is constructed so that the constraint violation is checked in forward and the solution is modified to satisfy three types of constraints of (5), (6), (7) and (8), (9) successively satisfying (3).

Step i) Receiving quantity for vendor companies is modified to satisfy the delivery quantity for suppliers. Set

$$S_{i,t}^{c} \leftarrow \overline{S_{i,t}^{c}}\,(\forall c \in Z_S; \forall i \in P; \forall t = 1,\dots,H);$$
$$S_{i,t}^{d} \leftarrow \frac{m_i^d}{\sum_{y \in Z_C} m_i^d} \sum_{c \in Z_S} \overline{S_{i,t}^{c}}\,(\forall d \in Z_C; \forall i \in P; \forall t = 1,\dots,H).$$

Step ii) Find a time period t in forward in which (5) is violated. For a plan in time period t, one type of product is allocated and allocation of other types of products are moved to a neighbour time period e.g. (t -1) or (t +1). If (3) and (5) are not satisfied, then return to step i). Otherwise go to step iii).

Step iii) Find a time period t in forward in which (6) or (7) is violated. For a plan in time period t, the violated delivery/receiving quantity is modified to allocate into a neighbour time period e.g. (t -1) or (t +1). If (3) and (5)-(7) are not satisfied, then return to step i). Otherwise go to step iv).

Step iv) Find a time period t in forward when (8) or (9) is violated. For a plan in time period t, the allocation of delivery/receiving quantity is modified to allocate a neighbour time period e.g. (t -1) or (t +1). If (3) and (5)-(9) are not satisfied, return to step i). Otherwise the heuristic procedure is completed.

Three cases of the supply chain planning problem for 1 supplier and 2 vendor companies are solved by the proposed method, LDC method and PM method. For each case, ten types of problems are generated by using random numbers on uniform distribution with different seeds in the range shown in Table 1. The parameters used for each method are shown in Table 2. The average objective function (Ave. obj. func.), average gap between the solution and an optimal solution (Ave. gap), average number of iterations to converge (Ave. num. iter.), and average computation time (Ave. comp. time) for ten times of calculations for each case are summarized in Table 3. The centralized MILP method uses a branch and bound method to obtain an optimal solution by CPLEX 8.0 using Pentium IV 2GHz processor with 512MB memory.

Computational results of Table 3 show that the ALDC method can generate better solutions than any other distributed optimization methods. The gap between the optimal solutions is within 3% for all cases. This indicates that the proposed method can generate near-optimal solution without using the entire information for each company. The total computation time for ALDC method to derive a feasible solution is shorter than that of MILP method, however, it is larger than that of PM method. The MILP solver cannot derive a solution within 100,000 seconds of computation time for Case 3 (3 types of products). This is why the computational complexity for the problem grows exponentially with number of products. The petroleum complex usually treats multi-products more than 3 types of products. Thus it is very difficult to apply the conventional MILP solver for supply chain planning for multiple companies. The optimality performance of the LDC method is not better than the other methods. This is because the heuristic procedure to generate a feasible solution is not effective for large-sized problems. The LDC method cannot derive a feasible solution by the current heuristic procedure. This is due to the difficulty of finding a feasible solution to satisfy all of such constraints as setup time constraints, and delivery duration constraints. The computation time of penalty method (PM method) is shorter than the proposed method, however, the optimality performance is not better than that of the proposed method. This result implies that the use of Lagrangian multipliers is effective to improve the optimality performance. Even though the proposed method needs a number of iterations to converge to a feasible solution than that of PM method, it is demonstrated that near-optimal solution with less than 3% of gap from the optimal solution can be obtained by the proposed method.

Case 1	Problem for 1 type of product			
Method	MILP	ALDC	LDC	PM
Ave. obj. func. [-]	10,829	10,976	13,297	11,153
Ave. gap [%]	0.00	1.37	23.0	2.90
Ave. num. iter.	-	180	90	53
Ave. comp. time[s]	1783	110	85	27
Case 2	Problem for 2 types of products			
Method	MILP	ALDC	LDC	PM
Ave. obj. func. [-]	48,700	49,975	50,562	51,005
Ave. gap [%]	0.00	2.71	4.06	4.66
Ave. num. iter.	-	137	80	39
Ave. comp. time[s]	16,142	246	149	41
Case 3	Problem for 3 types of products			
Method	MILP	ALDC	LDC	PM
Ave. obj. func. [-]	-	78280	-	79089
Ave. num. iter.	-	827	-	104
Ave. comp. time[s]	-	110	-	30

Table 3. Comparison of the performances of MILP and the distributed optimization methods

5. Conclusion and future works

A distributed supply chain planning system for multiple companies using an augmented Lagrangian relaxation method has been proposed. The original problem is decomposed into several sub-problems. The proposed system can derive a near optimal solution without using the entire information about the companies. By using a new penalty function, the proposed method can obtain a feasible solution without using a heuristic procedure. This is also a predominant characteristic of the proposed algorithm and the improvement of the conventional Lagrangian relaxation methods. It is demonstrated from numerical tests that a near optimal solution within a 3% of gap from an optimal solution can be obtained with a reasonable computation time. The applicability of the augmented Lagrangian function to the various class of supply chain planning problems is one of our future works.

6. References

Androulakis, I.P. & Reklaitis, G.V. (1999). Approaches to Asynchronous Decentralized Decision Making, *Computers and Chemical Engineering*, Vol. 23, pp. 341-355.

Beltan, C. & Herdia, F.J. (1999). Short-Term Hydrothermal Coordination by Augmented Lagrangean Relaxation: a new Multiplier Updating, *Investigaci\'on Operativa*, Vol. 8, pp. 63-76.

Beltran, C. & Heredia, F.J. (2002). Unit Commitment by Augmented Lagrangian Relaxation: Testing Two Decomposition Approaches, *Journal of Optimization Theory and Application*, Vol. 112, No. 2, pp. 295-314.

Bertsekas, D.P. (1976). Multiplier Methods: A Survey, *Automatica*, Vol. 12, pp. 133-145.

Cohen, G. & Zhu D.L. (1984) Decomposition Coordination Methods in Large Scale Optimization Problems: the nondifferentiable case and the use of augmented Lagrangians, *Advances in Large Scales Systems*, vol. 1, pp. 203-266.

Fisher, M.L. (1973). The Lagrangian Relaxation Method for Solving Integer Programming Problems, *Management Science*, vol. 27, no. 1, pp. 1-18.

Gaonkar, R. & Viswanadham, N. (2002). Integrated Planning in Electronic Marketplace Embedded Supply Chains, *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 1119-1124.

Georges, D. (1994). Optimal Unit Commitment in Simulations of Hydrothermal Power Systems: An Augmented Lagrangian Approach, *Simulation Practice and Theory*, vol. 1, pp. 155-172.

Gupta, A. & Maranus, C.D. (1999) Hierarchical Lagrangian Relaxation Procedure for Solving Midterm Planning Problems, *Industrial Engineering and Chemistry Research*, vol. 38, pp. 1937-1947.

Hoitomt, D.J., Luh, P.B. & Pattipati, K.R. (1993) A Practical Approach to Job-Shop Scheduling Problems, *IEEE Transactions on Robotics and Automation*, vol. 9, no. 1, pp. 1-13.

Jackson, J.R. & Grossmann, I.E. (2003) Temporal Decomposition Scheme for Nonlinear Multisite Production Planning and Distribution Models, *Industrial Engineering and Chemistry Research*, vol. 42, pp. 3045-3055.

Jayaraman, V. & Pirkul, H. (2001) Planning and Coordination of Production and Distribution Facilities for Multiple Commodities, *European Journal of Operational Research*, vol. 133, pp. 394-408.

Julka, N., Srinivasan, R. & Karimi, I. (2002) Agent-based Supply Chain Management, *Computers and Chemical Engineering*, vol. 26, pp. 1755-1769.

Luh, P.B. & Hoitomt, D.J. (1993) Scheduling of Manufacturing Systems Using the Lagrangian Relaxation Technique, *IEEE Transactions on Automatic Control*, vol. 38, no. 7, pp. 1066-1079.

Luh, P.B., Ni, M., Chen, H. & Thakur, L.S. (2003) Price-Based Approach for Activity Coordination in a Supply Network, *IEEE Transactions on Robotics and Automation*, vol. 19, No. 2, pp. 335-346.

McDonald, C.M. & Karimi, I.A. (1997) Planning and Scheduling of Parallel Semicontinuous Processes. 1. Production Planning, *Industrial Engineering and Chemistry Research*, vol. 36, pp. 2691-2700.

Nishi, T., Konishi, M., Hasebe, S. & Hashimoto, I. (2002) Machine Oriented Decentralized Scheduling Method using Lagrangian Decomposition and Coordination Technique, *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 4173-4178.

Rockafellar, R.T. (1974) Augmented Lagrange Multiplier Functions and Duality in Nonconvex Programming, *SIAM Journal of Control*, vol. 12, no. 2, pp. 269-285.

Stephanopoulos, G. & Westerberg, A.W. (1975) The Use of Hestenes' Method of Multipliers to Resolve Dual Gaps in Engineering System Optimization, *Journal of Optimization Theory and Application*, vol. 15, no. 3, pp. 285-309.

Taneda, D. (2003) Approach to the Advanced Refining and Petro Chemical Complexes, *Chemical Engineering Journal*, vol. 67, no. 3, pp. 46-183 (in Japanese).

Tu, Y., Luh, P.B., Feng, W., & Narimatsu, K. (2003) Supply Chain Performance Evaluation: A Simulation Study, *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 1749-1755.

Vidal, C. J. & Goetshalckx, M. (1997) Strategic Production-distribution Models: A Critical Review with Emphasis on Global Supply Chain Models, *European Journal of Operational Research*, vol. 98, no. 1, pp. 1-18.

16

Applying Fuzzy Linear Programming to Supply Chain Planning with Demand, Process and Supply Uncertainty

David Peidro, Josefa Mula and Raúl Poler
Research Centre on Production Management and Engineering (CIGIP)
Universidad Politécnica de Valencia,
SPAIN

1. Introduction

A Supply Chain (SC) is a dynamic network of several business entities that involve a high degree of imprecision. This is mainly due to its real-world character where uncertainties in the activities extending from the suppliers to the customers make SC imprecise (Fazel Zarandi et al., 2002).

Several authors have analysed the sources of uncertainty present in a SC, readers are referred to Peidro et al. (2008) for a review. The majority of the authors studied (Childerhouse & Towill, 2002; Davis, 1993; Ho et al., 2005; Lee & Billington, 1993; Mason-Jones & Towill, 1998; Wang & Shu, 2005), classified the sources of uncertainty into three groups: demand, process/manufacturing and supply. Uncertainty in supply is caused by the variability brought about by how the supplier operates because of the faults or delays in the supplier's deliveries. Uncertainty in the process is a result of the poorly reliable production process due to, for example, machine hold-ups. Finally, demand uncertainty, according to Davis (Davis, 1993), is the most important of the three, and is presented as a volatility demand or as inexact forecasting demands.

The coordination and integration of key business activities undertaken by an enterprise, from the procurement of raw materials to the distribution of the end products to the customer, are concerned with the SC planning process (Gupta & Maranas, 2003), one of the most important processes within the SC management concept. However, the complex nature and dynamics of the relationships among the different actors imply an important degree of uncertainty in the planning decisions. In SC planning decision processes, uncertainty is a main factor that can influence the effectiveness of the configuration and coordination of supply chains (Davis, 1993; Jung et al., 2004; Minegishi & Thiel, 2000) and tends to propagate up and down along the SC, affecting its performance appreciably (Bhatnagar & Sohal, 2005).

Most of the SC planning research (Alonso-Ayuso et al., 2003; Guillen et al., 2005; Gupta y Maranas, 2003; Lababidi et al., 2004; Santoso et al., 2005; Sodhi, 2005) models SC uncertainties with probability distributions that are usually predicted from historical data. However, whenever statistical data are unreliable or are even not available, stochastic models may not be the best choice (Wang y Shu, 2005). The fuzzy set theory(Zadeh, 1965) and the possibility theory (Dubois & Prade, 1988; Zadeh, 1978) may provide an alternative simpler and less-data demanding then probability theory to deal with SC uncertainties (Dubois et al., 2003).

Few studies address the SC planning problem on a medium-term basis (tactical level) which integrate procurement, production and distribution planning activities in a fuzzy environment (see Section 2. Literature review). Moreover, models contemplating the different sources of uncertainty in an integrated manner are lacking. Hence in this study, we develop a tactical supply chain model in a fuzzy environment in a multi-echelon, multi-product, multi-level, multi-period supply chain network. In this proposed model, the demand, process and supply uncertainties are contemplated simultaneously.

In the context of fuzzy mathematical programming, two very different issues can be addressed: fuzzy or flexible constraints for fuzziness, and fuzzy coefficients for lack of knowledge or epistemic uncertainty (Dubois et al., 2003). Our proposal jointly considers the possible lack of knowledge in data and existing fuzziness.

The main contributions of this paper can be summarized as follows:

- Introducing a novel tactical SC planning model by integrating procurement, production and distribution planning activities into a multi-echelon, multi-product, multi-level and multi-period SC network.
- Achieving a model which contemplates the different sources of uncertainty affecting SCs in an integrated fashion by jointly considering the possible lack of knowledge in data and existing fuzziness.
- Applying the model to a real-world automobile SC dedicated to the supply of automobile seats.

The rest of this paper is arranged as follows. Section 2 presents a literature review about fuzzy applications in SC planning. Section 3 proposes a new fuzzy mixed-integer linear programming (FMILP) model for the tactical SC planning under uncertainty. Then in Section 4, appropriate strategies for converting the fuzzy model into an equivalent auxiliary crisp mixed-integer linear programming model are applied. In Section 5, the behaviour of the model in a real-world automobile SC has been evaluated and, finally, the conclusions and directions for further research are provided.

2. Literature review

In Peidro et al. (2008) a literature survey on SC planning under uncertainty conditions by adopting quantitative approaches is developed. Here, we present a summary, extracted from this paper, about the applications of fuzzy set theory and the possibility theory to different problems related to SC planning:

<u>SC inventory management</u>: Petrovic et al. (1998; 1999) describe the fuzzy modelling and simulation of a SC in an uncertain environment. Their objective was to determine the stock levels and order quantities for each inventory during a finite time horizon to achieve an acceptable delivery performance at a reasonable total cost for the whole SC. Petrovic (2001) develops a simulation tool, SCSIM, for analyzing SC behaviour and performance in the presence of uncertainty modelled by fuzzy sets. Giannoccaro et al. (2003) develop a methodology to define inventory management policies in a SC, which was based on the echelon stock concept (Clark & Scarf, 1960) and the fuzzy set theory was used to model uncertainty associated with both demand and inventory costs. Carlsson and Fuller (2002) propose a fuzzy logic approach to reduce the bullwhip effect. Wang and Shu (2005) develop

a decentralized decision model based on a genetic algorithm which minimizes the inventory costs of a SC subject to the constraint to be met with a specific task involving the delivery of finished goods. The authors used the fuzzy set theory to represent the uncertainty of customer demands, processing times and reliable deliveries. Xie et al. (2006) present a new bilevel coordination strategy to control and manage inventories in serial supply chains with demand uncertainty. Firstly, the problem associated with the whole SC was divided into subproblems in accordance with the different parts that the SC it was made up of. Secondly, for the purpose of improving the integrated operation of a whole SC, the leader level was defined to be in charge of coordinating inventory control and management by amending the optimisation subproblems. This process was to be repeated until the desired level of operation for the whole SC was reached.
Vendor Selection: Kumar et al. (2004) present a fuzzy goal programming approach which was applied to the problem of selecting vendors in a SC. This problem was posed as a mixed integer and fuzzy goal programming problem with three basic objectives to minimize: the net cost of the vendors network, rejects within the network, and delays in deliveries. With this approach, the authors used triangular membership functions for each fuzzy objective. The solution method was based on the intersection of membership functions of the fuzzy objectives by applying the min-operator. Then, Kumar et al. (2006a) solve the same problem using the multi-objective fuzzy programming approach proposed by (Zimmermann, 1978). Amid et al. (2006) address the problem of adequately selecting suppliers within a SC. For this purpose, they devised a fuzzy-based multi-objective mathematical programming model where each objective may be assigned a different weight. The objectives considered were related to cost cuts, increased quality and to an increased service of the suppliers selected. The imprecise elements considered in this work were to meet both objectives and demand. Kumar et al. (2006b) analyze the uncertainty prevailing in integrated steel manufacturers in relation to the nature of the finished good and the significant demand by customers. They proposed a new hybrid evolutionary algorithm named endosymbioticpsychoclonal (ESPC) to decide what and how much to stock as an intermediate product in inventories. They compare ESPC with genetic algorithms and simulated annealing. They conclude the superiority of the proposed algorithm in terms of both the quality of the solution obtained and the convergence time.
Transport planning: Chanas et al. (1993) consider several assumptions on the supply and demand levels for a given transportation problem in accordance with the kind of information that the decision maker has: crisp values, interval values or fuzzy numbers. For each of these three cases, classical, interval and fuzzy models for the transportation problem are proposed, respectively. The links among them are provided, focusing on the case of the fuzzy transportation problem, for which solution methods are proposed and discussed. Shih (1999) addresses the problem of transporting cement in Taiwan by using fuzzy linear programming models. The author uses three approaches based on the works by Zimmermann (1976). Chanas (1983) and Julien (1994), who contemplate: the capacities of ports, the fulfilling demand, the capacities of the loading and unloading operations, and the constraints associated with traffic control. Liu and Kao (2004) develop a method to obtain the membership function of the total transport cost by considering this as a fuzzy objective value where the shipment costs, supply and demand are fuzzy numbers. The method was based on the extension principle defined by Zadeh (1978) to transform the fuzzy transport problem into a pair of mathematical programming models. Liang (2006) develops an

interactive multi-objective linear programming model for solving fuzzy multi-objective transportation problems with a piecewise linear membership function.

Production-distribution planning: Sakawa et al. (2001) address the real problem of production and transport related to a manufacturer through a deterministic mathematical programming model which minimizes costs in accordance with capacities and demands. Then, the authors develop a mathematical fuzzy programming model. Finally, they present an outline of the distribution of profits and costs based on the game theory. Liang (2007) proposes an interactive fuzzy multi-objective linear programming model for solving an integrated production-transportation planning problem in supply chains. Selim et al. (2007) propose fuzzy goal-based programming approaches applied to planning problems of a collaborative production-distribution type in centralized and decentralized supply chains. The fuzzy elements that the authors consider correspond to the fulfilment of different objectives related to maximizing profits for manufacturers and distribution centers, retailer cost cuts and minimizing delays in demand in retailers. Aliev et al. (2007) develop an integrated multi-period, multi-product fuzzy production and distribution aggregate planning model for supply chains by providing a sound trade-off between the fillrate of the fuzzy market demand and the profit. The model is formulated in terms of fuzzy programming and the solution is provided by genetic optimization.

Procurement-production-distribution planning: Chen and Chang (2006) develop an approach to derive the membership function of the fuzzy minimum total cost of the multi-product, multi-echelon, and multi-period SC model when the unit cost of raw materials supplied by suppliers, the unit transportation cost of products, and the demand quantity of products are fuzzy numbers. Recently, Tarabi and Hassini (2008) propose a new multi-objective possibilistic mixed integer linear programming model for integrating procurement, production and distribution planning by considering various conflicting objectives simultaneously along with the imprecise nature of some critical parameters such as market demands, cost/time coefficients and capacity levels. The proposed model and solution method are validated through numerical tests.

As mentioned before, models contemplating the different sources of uncertainty in an integrated manner are lacking and few studies address the SC planning problem on a medium-term basis which integrate procurement, production and distribution planning activities in a fuzzy environment. Moreover, the majority of the models studied are not applied in supply chains based on real world cases.

3. Problem description

This section outlines the tactical SC planning problem. The overall problem can be stated as follows:

Given:

- A SC topology: number of nodes and type (suppliers, manufacturing plants, warehouses, distribution centers, retailers, etc.)
- Each cost parameter, such as manufacturing, inventory, transportation, demand backlog, etc.
- Manufacture data, processing times, production capacity, overtime capacity, BOM, production run, etc.
- Transportation data, such as lead time, transport capacity, etc.
- Procurement data, procurement capacity, etc.

- Inventory data, such as inventory capacity, etc.
- Forecasted product demands over the entire planning periods.

To determine:

- The production plan of each manufacturing node.
- The distribution transportation plan between nodes.
- The procurement plan of each supplier node.
- The inventory level of each node.
- The sales and demand backlog.

The target is to centralize the multi-node decisions simultaneously in order to achieve the best utilization of the resources available in the SC throughout the time horizon so that customer demands are met at a minimum cost.

3.1 Fuzzy model formulation

The fuzzy mixed integer linear programming (FMILP) model for the tactical SC planning proposed by Peidro et al. (2007) is adopted as the basis of this work. Sets of indices, parameters and decision variables for the FMILP model are defined in the nomenclature (see Table 1). Table 2 shows the uncertain parameters grouped according to the uncertainty sources that may be presented in a SC.

Set of indices	
T:	Set of planning periods (t =1, 2…T).
I:	Set of products (raw materials, intermediate products, finished goods) (i =1, 2…I).
N:	Set of SC nodes (n =1, 2…N).
J:	Set of production resources (j =1, 2…J).
L:	Set of transports (l =1, 2…L).
P:	Set of parent products in the bill of materials (p =1, 2…P).
O:	Set of origin nodes for transports (o =1, 2…O).
D:	Set of destination nodes for transports (d =1, 2…D).
Objective function cost coefficients	
$V\tilde{P}C_{injt}$:	Variable production cost per unit of product i on j at n in t.
$O\tilde{T}C_{njt}$:	Overtime cost of resource j at n in t.
$U\tilde{T}C_{njt}$:	Undertime cost of resource j at n in t.
RMC_{int}:	Price of raw material i at n in t.
$T\tilde{C}_{odlt}$:	Transport cost per unit from o to d by l in t.
$I\tilde{C}_{int}$:	Inventory holding cost per unit of product i at n in t.
$D\tilde{B}C_{int}$:	Demand backlog cost per unit of product i at n in t.
General Data	
B_{pint}:	Quantity of i to produce a unit of p at n in t.
$MP\tilde{R}C_{nt}$:	Maximum procurement capacity from supplier node n in t.
$\tilde{D}_{int}$:	Demand of product i at n in t.
$M\tilde{O}T_{njt}$:	Overtime capacity of resource j at n in t.
$M\tilde{P}C_{njt}$:	Production capacity of resource j at n in t.

$I0_{in}$:	Inventory amount of i at n in period 0.
PR_{injt}:	Production run of i on j at n in t.
MPR_{injt}:	Minimum production run of i on j at n in t.
$DB0_{int}$:	Demand backlog of i at n in period 0.
$SR0_{iodlt}$:	Shipments of i received at d from o by l at the beginning of period 0.
$SIP0_{iodlt}$:	Shipments in progress of i from o to d by l at the beginning of period 0.
$P\widetilde{T}_{injt}$:	Processing time to produce a unit of i on j at n in t.
$T\widetilde{L}T_{odlt}$:	Transport lead time from o to d by l in t.
V_{it}:	Physical volume of product i in t.
$M\widetilde{T}C_{nt}$:	Maximum transport capacity of l in t.
$M\widetilde{I}C_{nt}$:	Maximum inventory capacity at n in t.
χ^{1}_{odlt}:	0-1 function. It takes 1 if $TLT_{odlt} > 0$ and 0 otherwise.
χ^{2}_{odlt}:	0-1 function. It takes 1 if $TLT_{odlt} = 0$ and 0 otherwise.
	Decision Variables
P_{injt}:	Production amount of i on j at n in t / $PT_{injt} > 0$.
k_{injt}:	Number of production runs of i produced on j at n in t.
S_{int}:	Supply of product i from n in t.
DB_{int}:	Demand backlog of i at n in t / $DBC_{int} > 0$.
TQ_{iodlt}:	Transport quantity of i from o to d by l in t / $o <> d$, $TC_{odlt} > 0$, $IC_{i,n=d,t} > 0$.
SR_{iodlt}:	Shipments of i received at d from o by l at the beginning of period t / $o <> d$, $TC_{odlt} > 0$, $IC_{i,n=d,t} > 0$.
SIP_{iodlt}:	Shipments in progress of i from o to d by l at the beginning of period t / $o <> d$, $TC_{odlt} > 0$, $IC_{i,n=d,t} > 0$, $TLT_{odlt} > 0$.
$FTLT_{iodlt}$:	Transport lead time for i from o to d by l in t (only used in the fuzzy model).
I_{int}:	Inventory amount of i at n at the end of period t.
PQ_{int}:	Purchase quantity of i at n in t / $RMC_{int} > 0$.
OT_{njt}:	Overtime for resource j at n in t.
UT_{njt}:	Undertime for resource j at n in t.
YP_{injt}:	Binary variable indicating whether a product i has been produced on j at n in t.

Table 1. Nomenclature (fuzzy parameters are shown with tilde: ~).

FMILP is formulated as follows:
Minimize z =

$$\sum_{i=1}^{I}\sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{t=1}^{T}(V\widetilde{P}C_{injt}\cdot P_{injt})+\sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{t=1}^{T}(O\widetilde{T}C_{njt}\cdot OT_{njt}+U\widetilde{T}C_{njt}\cdot UT_{njt})+ \\ +\sum_{i=1}^{I}\sum_{n=1}^{N}\sum_{t=1}^{T}(RMC_{int}\cdot PQ_{int}+I\widetilde{C}_{int}\cdot I_{int}+D\widetilde{B}C_{int}\cdot DB_{int})+\sum_{i=1}^{I}\sum_{o=1}^{O}\sum_{d=1}^{D}\sum_{l=1}^{L}\sum_{t=1}^{T}(T\widetilde{C}_{odlt}\cdot TQ_{iodlt}) \quad (1)$$

Subject to

$$\sum_{i=1}^{I}(P_{injt}\cdot P\widetilde{T}_{injt})\ \widetilde{\leq}\ M\widetilde{P}C_{njt}+M\widetilde{O}T_{njt} \qquad \forall\, n, j, t \quad (2)$$

Source of uncertainty in supply chains	Fuzzy coefficient	Formulation
Demand	Product demand Demand backlog cost	$\tilde{D}_{int}$ $D\tilde{B}C_{int}$
Process	Processing time Production capacity Production costs Inventory holding cost Maximum inventory capacity	$P\tilde{T}_{injt}$ $M\tilde{P}C_{njt}, M\tilde{O}T_{njt}$ $V\tilde{P}C_{injt}, O\tilde{T}C_{njt}, U\tilde{T}C_{njt}$ $I\tilde{C}_{int}$ $M\tilde{I}C_{nt}$
Supply	Transport lead time Transport cost Maximum transport capacity Maximum procurement capacity	$T\tilde{L}T_{odlt}$ $T\tilde{C}_{odlt}$ $M\tilde{T}C_{nt}$ $MP\tilde{R}C_{nt}$

Table 2. Fuzzy parameters considered in the model.

$$P_{injt} = k_{injt} \cdot PR_{injt} \qquad \forall\, i, n, j, t \tag{3}$$

$$P_{injt} \cdot PT_{injt} \,\tilde{\leq}\, M\tilde{P}C_{njt} \cdot YP_{injt} + M\tilde{O}T_{njt} \cdot YP_{injt} \qquad \forall\, i, n, j, t \tag{4}$$

$$P_{injt} \geq MPR_{injt} \cdot YP_{injt} \qquad \forall\, i, n, j, t \tag{5}$$

$$\begin{aligned} I_{int} = I_{in,t-1} + \sum_{j=1}^{J} P_{injt} + \sum_{o=1}^{O}\sum_{l=1}^{L} SR_{io,d=n,lt} + PQ_{int} - \\ - \sum_{d=1}^{D}\sum_{l=1}^{L} TQ_{i,o=n,dlt} - S_{int} - \sum_{p=1}^{P}(B_{pint} \cdot \sum_{j=1}^{J} P_{i=p,njt}) \end{aligned} \qquad \forall\, i, n, t \tag{6}$$

$$SR_{iodlt} = SR0_{iodlt} + TQ_{iodl,t-T\tilde{L}T} \qquad \forall\, i, o, d, l, t \tag{7}$$

$$SIP_{iodlt} = SIP0_{iodlt} + SIP_{iodl,t-1} + TQ_{iodlt} - SR_{iodlt} \qquad \forall\, i, o, d, l, t \tag{8}$$

$$\sum_{i=1}^{I} I_{int} \cdot V_{it} \,\tilde{\leq}\, M\tilde{I}C_{nt} \qquad \forall\, n, t \tag{9}$$

$$\sum_{i=1}^{I}\sum_{o=1}^{O}\sum_{d=1}^{D} SIP_{iodlt} \cdot V_{it} \cdot \chi^{1}_{odlt} + \sum_{i=1}^{I}\sum_{o=1}^{O}\sum_{d=1}^{D} TQ_{iodlt} \cdot V_{it} \cdot \chi^{2}_{odlt} \,\tilde{\leq}\, M\tilde{T}C_{lt} \qquad \forall\, l, t \tag{10}$$

$$\sum_{i=1}^{I} PQ_{int} \,\tilde{\leq}\, MP\tilde{R}C_{nt} \qquad \forall\, n, t \tag{11}$$

$$DB_{int} \approx DB_{in,t-1} + \widetilde{D}_{int} - S_{int} \qquad \forall\, i, n, t \tag{12}$$

$$OT_{njt} \approx \sum_{i=1}^{I} P_{injt} \cdot P\widetilde{T}_{injt} - M\widetilde{P}C_{njt} + UT_{njt} \qquad \forall\, n, j, t \tag{13}$$

$$\sum_{n=1}^{N}\sum_{t=1}^{T} S_{int} \mathrel{\widetilde{\leq}} \sum_{n=1}^{N}\sum_{t=1}^{T}\left(\widetilde{D}_{int} + DB0_{int}\right) \qquad \forall\, i \tag{14}$$

$$P_{injt},\ k_{injt} \geq 0 \qquad \forall\, i, n, j, t \tag{15}$$

$$S_{int}, DB_{int}, I_{int}, PQ_{int} \geq 0 \qquad \forall\, i, n, t \tag{16}$$

$$SR_{iodlt}, SIP_{iodlt}, TQ_{iodlt} \geq 0 \qquad \forall\, i, o, d, l, t \tag{17}$$

$$OT_{njt}, UT_{njt} \geq 0 \qquad \forall\, n, j, t \tag{18}$$

Eq. (1) shows the total cost to be minimized. The total cost is formed by the production costs with the differentiation between regular and overtime production. The costs corresponding to idleness, raw material acquisition, inventory holding, demand backlog and transport are also considered. Most of these costs cannot be measured easily since they mainly imply human perception for their estimation. Therefore, these costs are considered uncertain data and are modelled by fuzzy numbers. Only the raw material cost is assumed to be known.

The production time per period could never be higher than the available regular time plus the available overtime for a certain production resource of a node (2). Symbol $\widetilde{\leq}$ represents the fuzzy version of ≤ and means "essentially less than or similar to". This constraint shows that the planner wants to make the left-hand side of the constraint, the production time per period, smaller or similar to the right-hand side, the maximum production time available, "if possible". The production time and the production capacity are only known approximately and are represented by fuzzy numbers. On the other hand, the produced quantity of each product in every planning period must always be a multiple of the selected production lot size (3).

Eq. (4) and (5) guarantee a minimum production size for the different productive resources of the nodes in the different periods. These equations guarantee that P_{injt} will be equal to zero if YP_{injt} is zero.

Eq. (6) corresponds to the inventory balance. The inventory of a certain product in a node, at the end of the period, will be equal to the inputs minus the outputs of the product generated in this period. The inputs concern the production, transport receptions from other nodes, purchases (if supplying nodes) and the inventory of the previous period. The outputs are related to shipments to other nodes, supplies to customers and the consumption of other products (raw materials and intermediate products) that are necessary to produce in the node.

Eqs. (7) and (8) control the shipment of products among nodes. The receptions of shipments for a certain product will be equal to the programmed receptions plus the shipments carried out in previous periods. In constraint (7), the transport lead time are considered uncertainty data. On the other hand (8), the shipments in progress will be equal to the initial shipments in progress plus those from the previous period, plus the new shipments initiated in this period minus the new receptions.

Both the transports and inventory levels are limited by the available volume (known approximately). Thus according to Eq. (9), the inventory level for the physical volume of each product must be lower than the available maximum volume for every period (considered uncertainty data). The inventory volume depends on the period to consider the possible increases and decreases of the storage capacity over time. Additionally, the physical volume of the product depends on the time to cope with the possible engineering changes that can occur and affect the dimensions and volume of the different products.

On the other hand, the shipment quantities in progress of each shipment in every period multiplied by the volume of the transported products (if the transport time is higher than 0 periods), plus the initiated shipments by each transport in every period multiplied by the volume of the transported products (if the transport time is equal to 0 periods), can never exceed the maximum transport volume for that period (10). The reason for using a different formulation in terms of the transport time among nodes (TLT_{odlt}) is because the transport in progress will never exist if this value is not higher than zero because all the transport initiated in a period is received in this same period if $TLT_{odlt} = 0$. Finally, the transport volume depends on the period to consider the possible increases and decreases of the transport capacity over time.

Eq. (11) establishes an estimated maximum of purchase for each node and product per period. Eq. (12) contemplates the backlog demand management over time. The backlog demand for a product and node in a certain period will be equal (approximately) to the backlog demand of the previous period plus the difference between supply and demand.

Eq. (13) considers the use of overtime and undertime production for the different productive resources. The overtime production for a productive resource of a certain node in one period is equal (approximately) to the total production time minus the available regular production time plus the idle time. OT_{njt} and UT_{njt} will always be higher or equal to zero if the total production time is higher than the available regular production time, UT_{njt} will be zero as it does not incur in added costs, and OT_{njt} will be positive. On the contrary, if the total production time is lower than the available regular production time, UT_{njt} will be positive and OT_{njt} will be zero.

Conversely, Eq. (14) establishes that the sum of all the supplied products is essentially lower or equal to demand plus the initial backlog demand. At any rate, the problem could easily consider that all demand is served at the end of last planning period by transforming this inequality equation into an equality equation. Finally, Eqs. (15), (16), (17) and (18) guarantee the non negativity of the corresponding decision variables.

4. Solution methodology

In this section, we define an approach to transform the fuzzy mixed-integer linear programming model (FMILP) into an equivalent auxiliary crisp mixed-integer linear programming model for tactical SC planning under supply, process and demand uncertainties. According to Table 2, and in order to address the fuzzy coefficients of the FMILP model, it is necessary to consider the fuzzy mathematical programming approaches that integrally consider the fuzzy coefficients of the objective function and the fuzzy constraints: technological and right-hand side coefficients. In this context, several research works exist in the literature, and readers are referred to them (Buckley, 1989; Cadenas &

Verdegay, 1997; Carlsson & Korhonen, 1986; Gen et al., 1992; Herrera & Verdegay, 1995; Jiménez et al., 2007; Lai & Hwang, 1992; Vasant, 2005). In this paper, we adopt the approach by Cadenas and Verdegay (1997; 2004). The authors propose a general model for fuzzy linear programming that considers fuzzy cost coefficients, fuzzy technological coefficients and fuzzy right-hand side terms in constraints. Fuzziness is also considered in the inequalities that define the constraints. This general fuzzy linear programming model is as follows:

$$\begin{aligned}
&\text{Max } z = \sum_{j=1}^{n} \tilde{c}_j x_j \\
&\text{s.t.} \\
&\sum_{j=1}^{n} \tilde{a}_{ij} x_j \mathrel{\tilde{\lesssim}} \tilde{b}_i \\
&x_j \geq 0, i \in M, j \in N
\end{aligned} \tag{19}$$

where the fuzzy elements are given by:

- For each cost $\exists \mu_j \in F(\Re)$ so that μ_j: $\Re \rightarrow [0,1]$, $j \in N$, which defines the fuzzy costs.
- For each row $\exists \mu_i \in F(\Re)$ so that μ_i: $\Re \rightarrow [0,1]$, $i \in M$, which defines the fuzzy number in the right-hand side of constraints.
- For each $i \in M$ and $j \in N$ $\exists \mu_{ij} \in F(\Re)$ so that μ_{ij}: $\Re \rightarrow [0,1]$, which defines the fuzzy number in the technological matrix.
- For each row $\exists \mu_i \in F[F(\Re)]$ so that μ_i: $F(\Re) \rightarrow [0,1]$, $i \in M$ which provides the accomplishment degree of the fuzzy number for each $x \in \Re^n$

$$\tilde{a}_{i1} x_1 + \tilde{a}_{i2} x_2 + \ldots + \tilde{a}_{in} x_n, \ i \in M$$

with regard to the ith constraint, that is, the adequacy between this fuzzy number and the one $\tilde{b}_i$ in relation to the ith constraint.

Cadenas and Verdegay (1997) define a solution method which consists of substituting (19) by a convex fuzzy set through a ranking function as a comparison mechanism of fuzzy numbers.

Let $A, B \in F(\Re)$; a simple method for ranking fuzzy numbers consists of defining a ranking function mapping each fuzzy number into the real line, $g: F(\Re) \rightarrow \Re$. If this function $g(\cdot)$ is known, then:

$$\begin{aligned}
g(A) < g(B) &\Leftrightarrow \text{A is less than B} \\
g(A) > g(B) &\Leftrightarrow \text{A is greater than B} \\
g(A) = g(B) &\Leftrightarrow \text{A is equal to B}
\end{aligned}$$

Usually, g is called a linear ranking function if:

$$\forall A, B \in F(\Re), g(A + B) = g(A) + g(B)$$

$$\forall r \in \Re, r > 0, \ g(rA) = rg(A), \forall A \in F(\Re)$$

To solve the problem, (19) define: let g be a fuzzy number linear ranking function and given the function, Ψ: $F(\Re) \times F(\Re) \rightarrow F(\Re)$ so that:

$$\psi(\tilde{a}_i x, \tilde{b}_i) = \begin{cases} \tilde{t}_i, & \tilde{a}_i x \leq_g \tilde{b}_i \\ \tilde{t}_i(-)\,\tilde{a}_i x(+)\tilde{b}_i & \tilde{b}_i \leq_g \tilde{a}_i x \leq_g \tilde{b}_i(+)\tilde{t}_i \\ 0, & \tilde{a}_i x \leq_g \tilde{b}(+)\tilde{t}_i \end{cases}$$

Where $\tilde{t}_i \in F(\Re)$ is a fuzzy number in such a way that its support is included in $\Re^+$, and $\leq_g$ is a relationship that measures that $A \leq_g B$, $\forall A, B \in F(\Re)$, and (–) and (+) are the usual operations among fuzzy numbers.

According to Cadenas and Verdegay (2004), the membership function associated with the fuzzy constraint $\tilde{a}_i\, x \tilde{\leq} \tilde{b}_i$, with $\tilde{t}_i$ a fuzzy number giving the maximum violation of the ith constraint is:

$$\mu^i : F(\Re) \to [0,1] / \mu^i(\tilde{a}_i x, \tilde{b}_i) = \frac{g(\psi(\tilde{a}_i x, \tilde{b}))}{g(\tilde{t}_i)} \tag{20}$$

where g is a linear ranking function.

Given the problem (19), $\tilde{\leq}$ with the membership function (20) and using the Decomposition Theorem (Cadenas, 1993; Negoita & Ralescu, 1975) for fuzzy sets, the following is obtained:

$$\mu^i(\tilde{a}_i x, \tilde{b}_i) \geq \alpha \Leftrightarrow \frac{g(\psi(\tilde{a}_i x, \tilde{b}))}{g(\tilde{t}_i)} \geq \alpha \Leftrightarrow \frac{g(\tilde{t}_i(-)\tilde{a}_i x(+)\tilde{b}))}{g(\tilde{t}_i)} \geq \alpha \Leftrightarrow$$
$$g(\tilde{t}_i) - g(\tilde{a}_i x) + g(\tilde{b}_i) \geq g(\tilde{t}_i)\alpha \Leftrightarrow g(\tilde{a}_i x) \leq g(\tilde{b}_i(+)\tilde{t}_i(1-\alpha)) \Leftrightarrow$$
$$\tilde{a}_i x \leq_g \tilde{b}_i + \tilde{t}_i(1-\alpha)$$

where $\leq_g$ is the relationship corresponding to g.

Therefore, an equivalent model to solve (19) is the following:

$$\begin{aligned} &\text{Max } z = \sum_{j=1}^{n} \tilde{c}_j x_j \\ &\text{s.t.} \\ &\sum_{j=1}^{n} \tilde{a}_{ij} x_j \leq_g \tilde{b}_i + \tilde{t}_i(1-\alpha), \\ &x_j \geq 0, i \in M, j \in N, \alpha \in [0,1] \end{aligned} \tag{21}$$

To solve (21), the different fuzzy numbers ranking methods can be used in both the constraints and the objective function, or ranking methods can be used in the constraints and α-cuts in the objective, which will lead us to obtain different traditional models, which allows to obtain a fuzzy solution (Cadenas & Verdegay, 2004).

Specifically in this paper and for illustration effects of the method, we apply a linear ranking function for the constraints (the first index of Yager (1979; 1981)) and β-cuts in the objective, although the approach could be easily adapted to the use of any other index.

Thus, if we effect β-cuts in the coefficients of the objective and we apply the first index of Yager as a linear ranking function to the constraint set, we obtain the following α, β-parametric auxiliary problem.

$$\begin{aligned} &\text{Max } z = \sum_{j=1}^{n} \left[\left(c_j - \beta \cdot d'_{c_j} \right), \left(c_j + \beta \cdot d_{c_j} \right) \right] x_j \\ &\text{s.t.} \\ &\sum_{j=1}^{n} \left(a_{ij} + \frac{d_{a_{ij}} - d'_{a_{ij}}}{3} \right) x_j \le \left(b_i + \frac{d_{b_i} - d'_{b_i}}{3} \right) + \left(t_i + \frac{d_{t_i} - d'_{t_i}}{3} \right)(1-\alpha), \\ &x_j \ge 0, i \in M, j \in N, \alpha, \beta \in [0,1] \end{aligned} \tag{22}$$

where, for instance, d_{cj} and d'_{cj} are the lateral margins (right and left, respectively) of the triangular fuzzy number central point c_j (see Fig. 1).

Solving Eq. (22) by weighting objectives (w_1, w_2 / $w_1 + w_2 = 1$) the FLP problem defined in Eq. (21) is transformed into the crisp equivalent linear programming problem defined in Eq. (23) (Cadenas and Verdegay, 1997) .

$$\begin{aligned} &\text{Max } z = w_1 \cdot \sum_{j=1}^{n} \left[\left(c_j - \beta \cdot d'_{c_j} \right) x_j \right] + w_2 \cdot \sum_{j=1}^{n} \left[\left(c_j + \beta \cdot d_{c_j} \right) x_j \right] \\ &\text{s.t.} \\ &\sum_{j=1}^{n} \left(a_{ij} + \frac{d_{a_{ij}} - d'_{a_{ij}}}{3} \right) x_j \le \left(b_i + \frac{d_{b_i} - d'_{b_i}}{3} \right) + \left(t_i + \frac{d_{t_i} - d'_{t_i}}{3} \right)(1-\alpha), \\ &x_j \ge 0, i \in M, j \in N, \alpha, \beta \in [0,1], w_1 + w_2 = 1 \end{aligned} \tag{23}$$

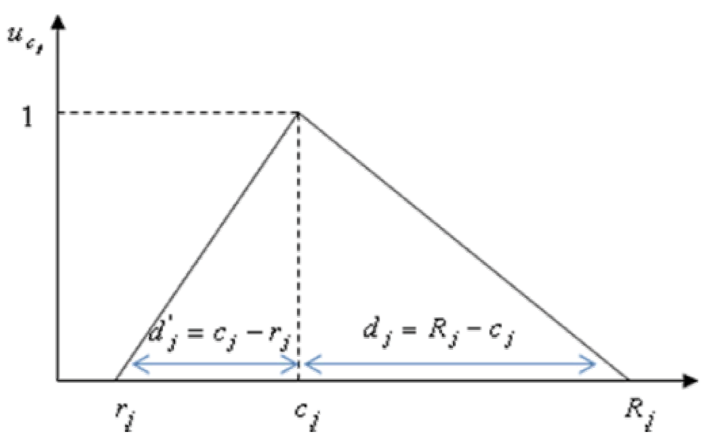


Fig. 1. Triangular fuzzy number

Consequently, by applying this approach to the previously defined FMILP model, we would obtain an auxiliary crisp mixed-integer linear programming model (MILP) as follows:

Minimize z =

$$
w_1 \cdot \left[\begin{array}{l}
\sum_{i=1}^{I}\sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{t=1}^{T}\left[\left(VPC_{injt} - \beta \cdot d'_{VPC}\right) \cdot P_{injt}\right] + \\
\sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{t=1}^{T}\left[\left(OTC_{njt} - \beta \cdot d'_{OTC}\right) \cdot OT_{njt} + \left(UTC_{njt} - \beta \cdot d'_{UTC}\right) \cdot UT_{njt}\right] + \\
\sum_{i=1}^{I}\sum_{n=1}^{N}\sum_{t=1}^{T}\left[RMC_{int} \cdot PQ_{int} + \left(IC_{int} - \beta \cdot d'_{IC}\right) \cdot I_{int} + \left(DBC_{int} - \beta \cdot d'_{DBC}\right) \cdot DB_{int}\right] + \\
\sum_{i=1}^{I}\sum_{o=1}^{O}\sum_{d=1}^{D}\sum_{l=1}^{L}\sum_{t=1}^{T}\left[\left(TC_{odlt} - \beta \cdot d'_{TC}\right) \cdot TQ_{iodlt}\right]
\end{array} \right] +
$$

$$
w_2 \cdot \left[\begin{array}{l}
\sum_{i=1}^{I}\sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{t=1}^{T}\left[\left(VPC_{injt} - \beta \cdot d_{VPC}\right) \cdot P_{injt}\right] + \\
\sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{t=1}^{T}\left[\left(OTC_{njt} - \beta \cdot d_{OTC}\right) \cdot OT_{njt} + \left(UTC_{njt} - \beta \cdot d_{UTC}\right) \cdot UT_{njt}\right] + \\
\sum_{i=1}^{I}\sum_{n=1}^{N}\sum_{t=1}^{T}\left[RMC_{int} \cdot PQ_{int} + \left(IC_{int} - \beta \cdot d_{IC}\right) \cdot I_{int} + \left(DBC_{int} - \beta \cdot d_{DBC}\right) \cdot DB_{int}\right] + \\
\sum_{i=1}^{I}\sum_{o=1}^{O}\sum_{d=1}^{D}\sum_{l=1}^{L}\sum_{t=1}^{T}\left[\left(TC_{odlt} - \beta \cdot d_{TC}\right) \cdot TQ_{iodlt}\right]
\end{array} \right] \tag{23}
$$

Subject to

$$
\begin{array}{l}
\sum_{i=1}^{I}\left[P_{injt} \cdot \left(PT_{injt} + \frac{d_{PT} - d'_{PT}}{3}\right)\right] \le MPC_{njt} + \frac{d_{MPC} - d'_{MPC}}{3} + MOT_{njt} + \frac{d_{MOT} - d'_{MOT}}{3} + \\
\left(t_1 + \frac{d_{t_1} - d'_{t_1}}{3}\right)(1-\alpha)
\end{array} \qquad \forall\, n, j, t \tag{24}
$$

$$
\begin{array}{l}
\left(PT_{injt} + \frac{d_{PT} - d'_{PT}}{3}\right) \cdot PT_{injt} \le \left(MPC_{njt} + \frac{d_{MPC} - d'_{MPC}}{3}\right) \cdot YP_{injt} + \\
\left(MOT_{njt} + \frac{d_{MOT} - d'_{MOT}}{3}\right) \cdot YP_{injt} + \left(t_3 + \frac{d_{t_3} - d'_{t_3}}{3}\right)(1-\alpha)
\end{array} \qquad \forall\, i, n, j, t \tag{25}
$$

$$
SR_{iodlt} = SR0_{iodlt} + TQ_{iodl,t-FTLT} \qquad \forall\, i, o, d, l, t \tag{26}
$$

$$
\sum_{i=1}^{I} I_{int} \cdot V_{it} \le \left(MIC_{nt} + \frac{d_{MIC} - d'_{MIC}}{3}\right) + \left(t_8 + \frac{d_{t_8} - d'_{t_8}}{3}\right)(1-\alpha) \qquad \forall\, n, t \tag{27}
$$

$$
\begin{array}{l}
\sum_{i=1}^{I}\sum_{o=1}^{O}\sum_{d=1}^{D} SIP_{iodlt} \cdot V_{it} \cdot \chi^1_{odlt} + \sum_{i=1}^{I}\sum_{o=1}^{O}\sum_{d=1}^{D} TQ_{iodlt} \cdot V_{it} \cdot \chi^2_{odlt} \le \\
\left(MTC_{lt} + \frac{d_{MTC} - d'_{MTC}}{3}\right) + \left(t_9 + \frac{d_{t_9} - d'_{t_9}}{3}\right)(1-\alpha)
\end{array} \qquad \forall\, l, t \tag{28}
$$

$$\sum_{i=1}^{I} PQ_{int} \leq \left(MPRC_{nt} + \frac{d_{MPRC} - d'_{MPRC}}{3} \right) + \left(t_{10} + \frac{d_{t_{10}} - d'_{t_{10}}}{3} \right)(1-\alpha) \quad \forall\, n,t \tag{29}$$

$$DB_{int} - DB_{in,t-1} + S_{int} \leq \left(D_{\text{int}} + \frac{d_D - d'_D}{3} \right) + \left(t_{11} + \frac{d_{t_{11}} - d'_{t_{11}}}{3} \right)(1-\alpha) \quad \forall\, i,n,t \tag{30}$$

$$DB_{int} - DB_{in,t-1} + S_{int} \geq \left(D_{\text{int}} + \frac{d_D - d'_D}{3} \right) - \left(t_{12} + \frac{d_{t_{12}} - d'_{t_{12}}}{3} \right)(1-\alpha) \quad \forall\, i,n,t \tag{31}$$

$$\sum_{i=1}^{I} P_{injt} \cdot PT_{injt} + UT_{njt} - OT_{njt} \leq \left(MPC_{njt} + \frac{d_{MPC} - d'_{MPC}}{3} \right) + \left(t_{13} + \frac{d_{t_{13}} - d'_{t_{13}}}{3} \right)(1-\alpha) \quad \forall\, n,j,t \tag{32}$$

$$\sum_{i=1}^{I} P_{injt} \cdot PT_{injt} + UT_{njt} - OT_{njt} \geq \left(MPC_{njt} + \frac{d_{MPC} - d'_{MPC}}{3} \right) - \left(t_{14} + \frac{d_{t_{14}} - d'_{t_{14}}}{3} \right)(1-\alpha) \quad \forall\, n,j,t \tag{33}$$

$$\sum_{n=1}^{N}\sum_{t=1}^{T} S_{int} \leq \sum_{n=1}^{N}\sum_{t=1}^{T} \left(D_{\text{int}} + \frac{d_D - d'_D}{3} + DB0_{int} \right) + \left(t_{15} + \frac{d_{t_{15}} - d'_{t_{15}}}{3} \right)(1-\alpha) \quad \forall\, i \tag{34}$$

$$FTLT_{odlt} \leq \left(TLT_{odlt} + \frac{d_{TLT} - d'_{TLT}}{3} \right) + \left(t_{16} + \frac{d_{t_{16}} - d'_{t_{16}}}{3} \right)(1-\alpha) \quad \forall\, o,d,l,t \tag{35}$$

$$FTLT_{odlt} \geq \left(TLT_{odlt} + \frac{d_{TLT} - d'_{TLT}}{3} \right) - \left(t_{17} + \frac{d_{t_{17}} - d'_{t_{17}}}{3} \right)(1-\alpha) \quad \forall\, o,d,l,t \tag{36}$$

$$FTLT_{odlt} \geq 0 \quad \forall\, o,d,l,t \tag{37}$$

The non fuzzy constraints (3), (5), (6), (8), (15), (16), (17) and (18) are also included in the model in a similar way.

In order to solve the problem and according to Eq. (22) α, β is settled parametrically to obtain the value of the objective function for each of these α, $\beta \in [0, 1]$. The result is a fuzzy set and the SC planner has to decide which pair (α, β, z) is more adequate to obtain a crisp solution. Although the descomposition theorem could be applied in different scales to the objective and to the constraint set (the decision maker's aspirations on the objective could be different from his/her satisfaction degree on the accomplishment of the constraints), in this work, the auxiliary crisp mixed-integer linear programming model (MILP) presented before is solved by using the same values for the parameters α and β.

5. Application to an automobile supply chain

The proposed model has been evaluated by using data from an automobile SC which comprises a total of 47 companies (see Figure 2). In fact, these companies constitute a segment of the automobile SC. Specifically, this SC segment supplies a seat model to an automobile assembly plant. The nodes that form the SC are a seat assembly company, its first tier suppliers, a manufacturing company of foams for seats and a second tier supplier that supplies chemical components for foam manufacturing. The automobile assembly plant weekly transmits the demand information (automobile seats) with a planning horizon for six months. However, these demand forecasts are rarely precise (Mula et al., 2005). This section validates whether the proposed fuzzy model for SC planning can be a useful tool for improving the decision-making process in an uncertain decision environment.

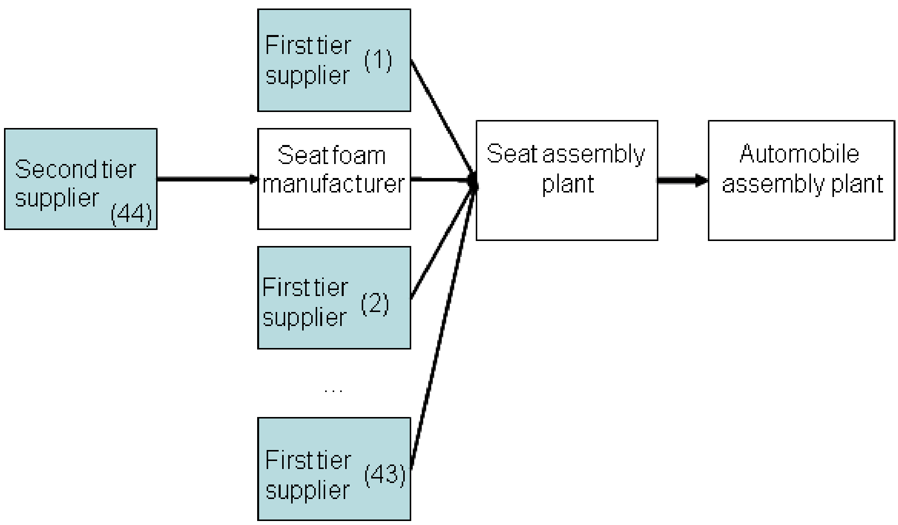


Fig. 2. Supply chain

5.1 Implementation and resolution

The model has been developed with the modelling language MPL (Maximal Software Incorporation, 2004) and solved by the CPLEX 9.0 solver (ILOG Incorporation, 2003). The input and output data are managed through a MS SQL Server database.

The model has been executed for a rolling horizon over a total of 17 weekly periods. These periods correspond to 17 different demand forecast programs, which are transmitted weekly by the automobile assembly plant. The total set of planning periods considered by the demand forecast programs is 42 weeks. Figure 3 depicts the execution of the models based on the rolling horizon technique. Each model calculation in the different planning horizon periods updates the data for the period being considered, and the results of the decision variables for the remaining periods are ruled out. Some of the stored decision variables are used as input data to solve the model in the following periods. These data include: demand backlog, shipments received, shipments in progress and inventory. This process is repeated for all the rolling horizon planning periods. The results of the model are evaluated from the data of the decision variables stored in each model execution. The experiments were run in an Intel Xeon PC, at 2.8 Ghz and with 1GB of RAM memory.

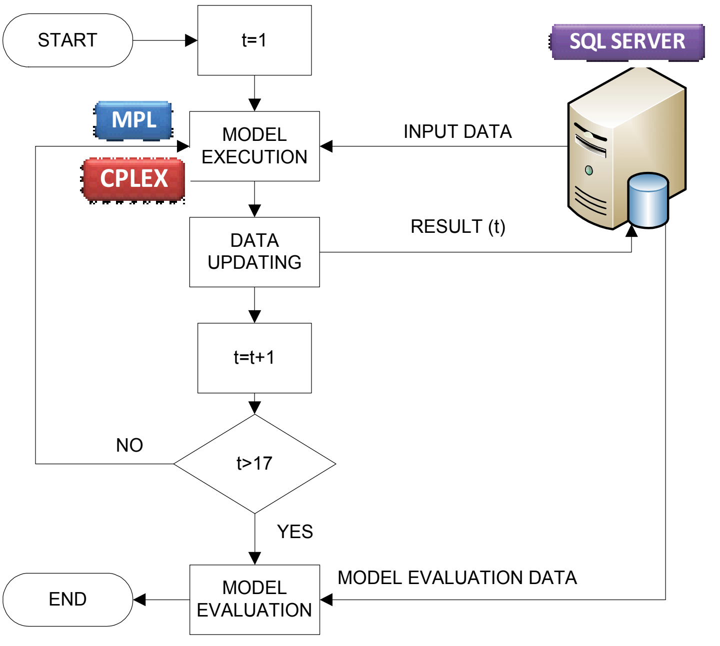


Fig. 3. Computational experiment diagram

5.2 Assumptions

The main characteristics and assumptions used in the experiment are presented below:

- The study considers a representative single finished good, i.e. a specific seat which can be considered to be a standard seat. The bill of materials of the standard seat is composed of 53 elements arranged in a three-level structure.
- The decision variables S_{int}, k_{injt}, DB_{int}, TQ_{iodlt}, SR_{iodlt}, SIP_{iodlt}, PQ_{int}, and P_{injt} are considered integer. Therefore, a mixed integer linear programming model is required to be solved.
- Only the finished good has external demand.
- The demand backlog for the finished good is considered but with a high penalization cost since the service level required by a sequenced and synchronized automobile seat supplier is 100%.
- A single productive resource restricts the capacity of the productions nodes (i.e. by focusing on the bottleneck resource).
- Triangular fuzzy numbers were defined by the decision makers involved in the planning process from the deviation percentages on the crisp value. These percentages range from an average 5% to 30%, depending on the parameter to be evaluated.
- A maximum violation of 5% is contemplated on the right-hand side of fuzzy constraints.
- The demand values for the first period of each model run, according to the rolling planning horizon, are considered to be firm. This means that the fuzzy intervals of the demand for this period will be the equivalent to a crisp number. The same happens for all the demand values of the last program. Thus, all the models will have the same net requirements to fulfill.
- A maximum calculation time of 100 CPU seconds is set.

5.3 Evaluation of the results

Here, we compare the behaviour of the proposed fuzzy model with its deterministic version. The aim is to determine the possible improvements that can provide the fuzzy model, which incorporates the uncertainties that may be presented in a SC.

Table 3 shows the computational efficiency of the deterministic model and the fuzzy SC planning model proposed. The data are related to the iterations, number of constraints, variables, integers, non zero elements, calculation time and the average density of the array of constraints for the set of the 17 planned executions of the models. Although the fuzzy model obtains higher values for these parameters, the CPU time has not markedly increased.

	Deterministic	*Fuzzy*
Iterations	636,128	717,377
Constraints	4,475,429	4,759,207
Variables	4,840,812	5,853,474
Integers	5,823,864	6,836,526
Non zero elements	15,793,161	23,251,766
Array density (%)	12.35 %	16.25 %
CPU time (seconds)	1,298.50	1,494.73

Table 3. Efficiency of the computational experiments.

Table 4 summarizes the evaluation results with the different α and β values, according to a group of parameters defined in (Mula et al., 2006): (i) the average service level, (ii) the inventory levels, (iii) planning nervousness in relation to the planned period and planned quantity and (iv) the total costs.

i. The average service level for the finished good is calculated as follows:

$$Average\ service\ level\ (\%) = \sum_{t=1}^{T} \frac{\left(1 - \dfrac{DB_{int}}{\sum_{t'=1}^{t} D_{int'}} \right) \times 100}{T} \qquad \forall i,n \tag{38}$$

ii. The inventory level is calculated as the sum of the total quantity of inventory of the finished good and parts at the end of each planning period T= (1,…,42). Then the following rules are applied to determine which model presents, on average, the minimum and maximum inventory levels:
   - If for each model the minimum inventory level is presented, it is assigned the value of 1, while a null value is assigned to the rest. The model which obtains the highest number will have the minimum levels of inventory. The maximum inventory levels can be determined in a similar way but by assigning the value of 1 to the maximum inventory level for item and model.

iii. Planning nervousness with regard to the planned period. "Nervous" or unstable planning refers to a plan which undergoes significant variations when incorporating the demand changes between what is foreseen and what is observed in successive plans, as defined by Sridharan et al. (1987). Planning nervousness can be measured according to the demand changes in relation to the planned period or to the planned quantity. The demand changes in the planned period measure the number of times that a planned

order is rescheduled, irrespectively of the planned quantity (Heisig, 1998). The next rule proposed by Donselaar et al. (2000) is summarized as follows:
At time t we check for each period $t + x$ ($x = 0, 1, 2, \ldots, T\text{-}1$):

- If there is a planned order in t + x, and this order is not planned in the next planning run, we increase the number of reschedules by 1.
- If there was no planned order in t + x, and there is one in the next planning run, we increase the number of reschedules by 1.

Planning nervousness with regard to the planned period measures the demand changes in the planned quantity as the number of times that the quantity of a planned order is modified (De Kok and Inderfurth 1998). The rule is described as follows:
In the period $t = 1, \ldots, T$, where T is the number of periods that forms the planning horizon, it is checked for every period $t + x$ ($x = 0, 1, 2, \ldots, T\text{-}1$):

- If a planned order exists in the period t + x, then if the quantity of the planned order is not the same as in the next planning run, we increase the number of reschedules by 1.

In the computation of planning nervousness, we measure the number of changes. Another way to compute it would be to take into account the rate of the changes.

iv. Total costs are the sum of all the costs that are generated in every period of the considered planning horizon, and derived from the procurement, production and distribution plans provided by the model.

$\alpha = \beta$	Service level (%)	Number of min/max inventory levels	Planning nervousness (period)	Planning nervousness (quantity)	Total cost (€)
0	98.32%	10/11	1.31	20,69	4,528,053.0
0.1	98.31%	12/12	1.31	20,56	4,565,027.9
0.2	98.28%	8/10	1.31	20,56	4,601,774.4
0.3	98.28%	9/6	1.31	20,56	4,624,121.9
0.4	98.28%	9/12	1.31	20,56	4,640,859.9
0.5	98.28%	7/9	1.31	20,63	4,655,773.8
0.6	98.24%	11/11	1.31	20,56	4,698,218.3
0.7	98.24%	13/8	1.31	20,56	4,711,902.7
0.8	98.24%	15/8	1.31	20,56	4,737,356.4
0.9	98.21%	12/9	1.31	20,56	4,769,438.2
1	98.21%	12/9	1.31	20,56	4,769,438.2
Deterministic model	98.21%	8/7	1.31	20.56	4,768,579.1

Table 4. Evaluation of results

As seen in Table 4, all the fuzzy models, in general, obtain better results than the deterministic model. Only those models whose α values come close to 1 obtain similar results to the deterministic model. This situation is logical because the closer the α value comes to 1, the more similar the triangular fuzzy number model will be to a deterministic model. As seen in Figure 4, the fuzzy models obtain service levels that are better than or equal to the deterministic model, and these fuzzy models have better adapted to the existing uncertainties in the demand forecasts considered in this work because these demand forecasts in this sector are rarely precise (Mula et al., 2005), as previously mentioned.

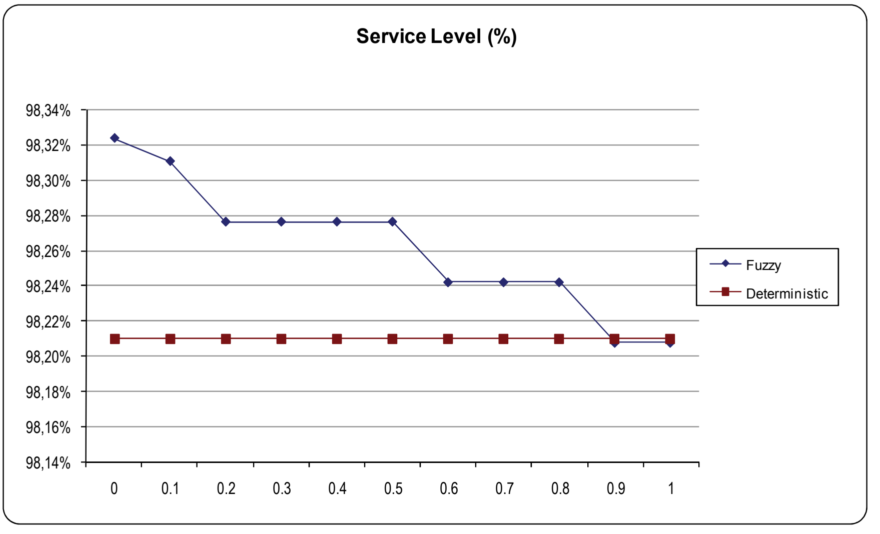


Fig. 4. Service Level (%)

With regard to inventory levels (see Fig. 5), the fuzzy model obtains better results for the number of minimum inventory levels for almost all the α values. It is important to highlight that for α > 0.3, the fuzzy model generates better results for minimum and maximum inventory levels. Besides, the levels of nervousness of fuzzy model are similar to those of the deterministic model. Finally, all the fuzzy models (see Fig. 6) obtain lower or similar costs than the deterministic model. This is because the demand backlog in this work is very heavily penalized, which means that those models with higher service levels achieve lower costs.

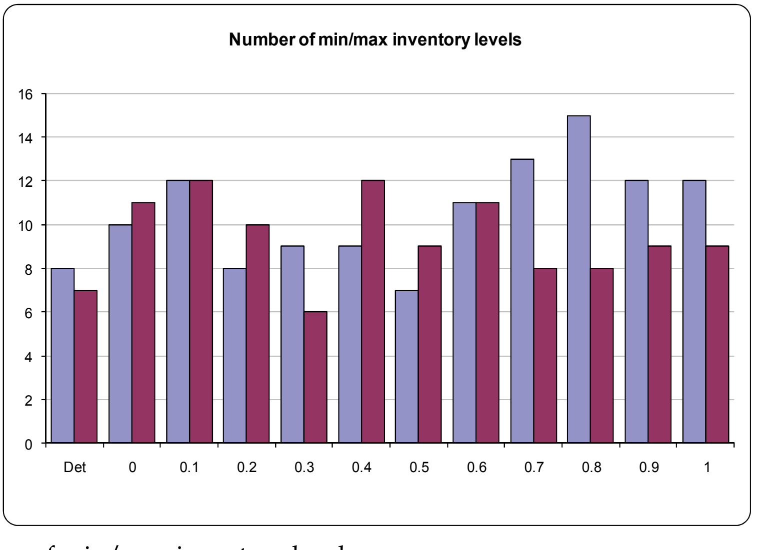


Fig. 5. Number of min/max inventory levels

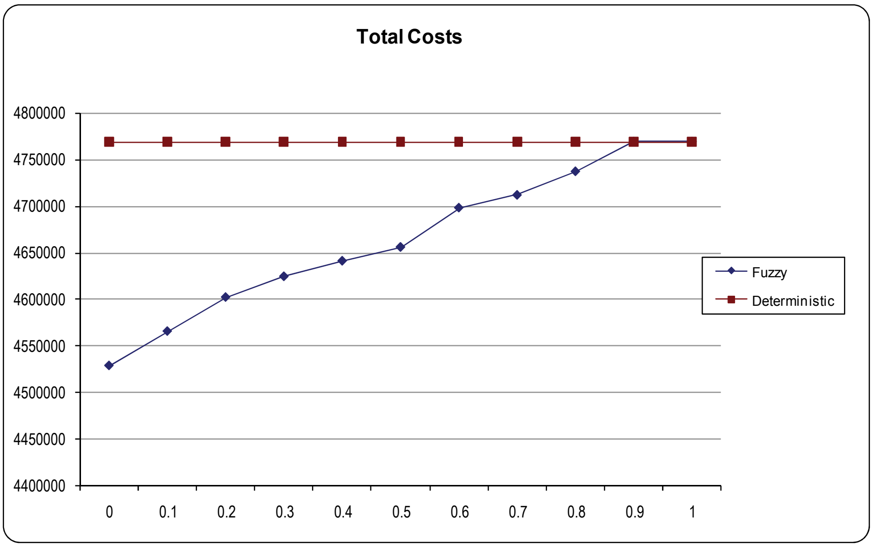


Fig. 6. Total Costs

6. Conclusions

This paper has proposed a novel fuzzy mixed integer linear programming (FMILP) model for the tactical SC planning, by integrating procurement, production and distribution planning activities into a multi-echelon, multi-product, multi-level and multi-period SC network. The fuzzy model integrally handles all the epistemic uncertainty sources identified in SC tactical planning problems given lack of knowledge (demand, process and supply uncertainties). This model has been tested by using data from a real-world automobile SC applying the rolling horizon technique over a total of 17 weekly periods. The evaluation of the results has demonstrated the effectiveness of a fuzzy linear programming approach for SC planning under uncertainty. The proposed fuzzy formulation is more effective than the deterministic methods for handling the real situations where precise or certain information is not available for SC planning. Additionally, the fuzzy model behaviour has been clearly superior to the deterministic model, as previously shown. Furthermore, the fuzzy model has not generated an excessive increment of the computational efficiency.

Finally, further research will consider: (1) other fuzzy mathematical programming-based approaches; (2) to design an expert system to solve the problem in which each decision maker, according to its aspirations, experiences and business, could have that index for ranking fuzzy numbers that better is adapted to its requirements; (3) the use of evolutionary computation in order to solve the fuzzy multi-objective, non linear SC planning problems; and (4) the application of hybrid models based on the integration of analytical and simulation models as an interesting option to integrate the best capacities of both types of models for SC planning problems.

7. Acknowledgment

This work has been funded by the Spanish Ministry of Science and Technology project: 'Simulation and evolutionary computation and fuzzy optimization models of transportation and production planning processes in a supply chain. Proposal of collaborative planning supported by multi-agent systems. Integration in a decision system. Applications' (EVOLUTION) (Ref. DPI2007-65501).

8. References

Aliev, R. A.; Fazlollahi, B.; Guirimov, B. G. & Aliev, R. R. (2007). Fuzzy-genetic approach to aggregate production-distribution planning in supply chain management. *Information Sciences*, 177(20): 4241-4255.

Alonso-Ayuso, A.; Escudero, L. F.; Garin, A.; Ortuno, M. T. & Perez, G. (2003). An approach for strategic supply chain planning under uncertainty based on stochastic 0-1 programming. *Journal of Global Optimization*, 26(1): 97-124.

Amid, A.; Ghodsypour, S. H. & O'Brien, C. (2006). Fuzzy multiobjective linear model for supplier selection in a supply chain. *International Journal of Production Economics*, 104(2): 394-407.

Bhatnagar, R. & Sohal, A. S. (2005). Supply chain competitiveness: measuring the impact of location factors, uncertainty and manufacturing practices. *Technovation*, 25(5): 443-456.

Buckley, J. J. (1989). Solving Possibilistic Linear-Programming Problems. *Fuzzy Sets and Systems*, 31(3): 329-341.

Cadenas, J. M. (1993). Design and implementation of an interactive system to solve fuzzy optimizations problems. *Ph.D.Dissertation*, Univ. Granada, Spain.

Cadenas, J. M. & Verdegay, J. L. (1997). Using fuzzy numbers in linear programming. *Ieee Transactions on Systems Man and Cybernetics Part B-Cybernetics*, 27(6): 1016-1022.

Cadenas, J. M. & Verdegay, J. L. (2004). Métodos y Modelos de Programación Lineal Borrosa. *in: Optimización Bajo Incertidumbre (Alonso-Ayuso, A., Cerdá, E., Escudero, L.F.y Sala, R.)*, Ed. Tirant Lo Blanch(Monografías Nº 2): 77-98.

Carlsson, C. & Fuller, R. (2002). A fuzzy approach to taming the bullwhip effect. *ADVANCES IN COMPUTATIONAL INTELLIGENCE AND LEARNING: METHODS AND APPLICATIONSInternational Series in Intelligent Technologies*, 18: 247-262.

Carlsson, C. & Korhonen, P. (1986). A Parametric Approach to Fuzzy Linear-Programming. *Fuzzy Sets and Systems*, 20(1): 17-30.

Chanas, S. (1983). The Use of Parametric Programming in Fuzzy Linear-Programming. *Fuzzy Sets and Systems*, 11(3): 243-251.

Chanas, S.; Delgado, M.; Verdegay, J. L. & Vila, M. A. (1993). Interval and fuzzy extensions of classical transportation problems. *Transportation Planning and Technology*, 17(2): 203-218.

Chen, S. P. & Chang, P. C. (2006). A mathematical programming approach to supply chain models with fuzzy parameters. *Engineering Optimization*, 38(6): 647-669.

Childerhouse, P. & Towill, D. R. (2002). Analysis of the factors affecting real-world value stream performance. *International Journal of Production Research*, 40(15): 3499-3518.

Clark, T. H. & Scarf, H. (1960). Optimal polices for a multi-echelon inventory problem. *Management Science*, 6: 475-490.

Davis, T. (1993). Effective Supply Chain Management. *Sloan Management Review*, 34(4): 35-46.

Donselaar, K. V.; Nieuwenhof, J. V. & Visschers, J. (2000). The impact of material coordination concepts on planning stability in supply chains. *International Journal of Production Economics*, 68: 169-176.

Dubois, D. & Prade, H. (1988). Possibility theory. *New York, London*.

Dubois, D.; Fargier, H. & Fortemps, P. (2003). Fuzzy scheduling: Modelling flexible constraints vs. coping with incomplete knowledge. *European Journal of Operational Research*, 147(2): 231-252.

Fazel Zarandi; Turksen, I. B. & Saghiri, S. (2002). Supply chain: crisp and fuzzy aspects. *International Journal of Applied Mathematical Computer Sciences*, 12(3): 423-435.

Gen, M.; Tsujimura, Y. & Ida, K. (1992). Method for solving multiobjective aggregate production planning problem with fuzzy parameters. *Computers & Industrial Engineering*, 23(1-4): 117-120.

Giannoccaro, I.; Pontrandolfo, P. & Scozzi, B. (2003). A fuzzy echelon approach for inventory management in supply chains. *European Journal of Operational Research*, 149(1): 185-196.

Guillen, G.; Mele, E.; Bagajewicz, M. J.; Espuna, A. & Puigjaner, L. (2005). Multiobjective supply chain design under uncertainty. *Chemical Engineering Science*, 60(6): 1535-1553.

Gupta, A. & Maranas, C. D. (2003). Managing demand uncertainty in supply chain planning. *Computers & Chemical Engineering*, 27(8-9): 1219-1227.

Heisig, G. (1998). Planning stability under (s,S) inventory control rules. *OR Spectrum*, 20(4): 215-228.

Herrera, F. & Verdegay, J. L. (1995). Three models of fuzzy integer linear programming. *European Journal of Operational Research*, 83(3): 581-593.

Ho, C. F.; Chi, Y. P. & Tai, Y. M. (2005). A structural approach to measuring uncertainty in supply chains. *International Journal of Electronic Commerce*, 9(3): 91-114.

ILOG Incorporation (2003). CPLEX 9.0. *USA*.

Jiménez, M.; Arenas, M.; Bilbao, A. & guez, M. V. (2007). Linear programming with fuzzy parameters: An interactive method resolution. *European Journal of Operational Research*, 177(3): 1599-1609.

Julien, B. (1994). An Extension to Possibilistic Linear-Programming. *Fuzzy Sets and Systems*, 64(2): 195-206.

Jung, J. Y.; Blau, G.; Pekny, J. F.; Reklaitis, G. V. & Eversdyk, D. (2004). A simulation based optimization approach to supply chain management under demand uncertainty. *Computers & Chemical Engineering*, 28(10): 2087-2106.

Kumar, M.; Vrat, P. & Shankar, R. (2004). A fuzzy goal programming approach for vendor selection problem in a supply chain. *Computers & Industrial Engineering*, 46(1): 69-85.

Kumar, M.; Vrat, P. & Shankar, R. (2006a). A fuzzy programming approach for vendor selection problem in a supply chain. *International Journal of Production Economics*, 101(2): 273-285.

Kumar, V.; Prakash; Tiwari, M. K. & Chan, F. T. S. (2006b). Stochastic make-to-stock inventory deployment problem: an endosymbiotic psychoclonal algorithm based approach. *International Journal of Production Research*, 44(11): 2245-2263.

Lababidi, H. M. S.; Ahmed, M. A.; Alatiqi, I. M. & Al Enzi, A. F. (2004). Optimizing the supply chain of a petrochemical company under uncertain operating and economic conditions. *Industrial & Engineering Chemistry Research*, 43(1): 63-73.

Lai, Y. J. & Hwang, Ch. L. (1992). Fuzzy Mathematical Programming: Methods and Applications. *Springer-Verlag, Heidelberg*..

Lee, H. L. & Billington, C. (1993). Material Management in Decentralized Supply Chains. *Operations Research*, 41(5): 835-847.

Liang, T. F. (2007). Applying fuzzy goal programming to production/transportation planning decisions in a supply chain

SO. *INTERNATIONAL JOURNAL OF SYSTEMS SCIENCE*, 38(4): 293-304.

Liang, T. F. (2006). Distribution planning decisions using interactive fuzzy multi-objective linear programming. *Fuzzy Sets and Systems*, 157(10): 1303-1316.

Liu, S. T. & Kao, C. (2004). Solving fuzzy transportation problems based on extension principle. *European Journal of Operational Research*, 153(3): 661-674.

Mason-Jones, R. & Towill, D. R. (1998). Shrinking the supply chain uncertainty circle. *IOM Control*, 17-22.

Maximal Software Incorporation (2004). MPL modeling system. Release 4.2e. *USA*.

Minegishi, S. & Thiel, D. (2000). System dynamics modeling and simulation of a particular food supply chain. *Simulation Practice and Theory*, 8(5): 321-339.

Mula, J.; Poler, R. & García J.P. (2006). MRP with flexible constraints: A fuzzy mathematical programming approach. *Fuzzy Sets and Systems*, 157(1): 74-97.

Mula, J.; Poler, R.; García J.P. & Ortiz, A. (2005). Demand Uncertainty Effects of First Tier Suppliers of an Automobile Industry Supply Chain. *The ICFAI Journal of Supply Chain Managament*, 2(3): 19-39.

Negoita, C. V. & Ralescu, D. A. (1975). Application of fuzzy sets to system analysis. *Basel*, Stuttgart.

Peidro, D.; Mula, J. & Poler, R. (2007). Supply chain planning under uncertainty: a fuzzy linear programming approach. *Fuzzy Systems Conference, 2007.FUZZ-IEEE 2007.IEEE International*, 1-6.

Peidro, D.; Mula, J.; Poler, R. l. & Lario, F. C. (2008). Quantitative models for supply chain planning under uncertainty: a review. *The International Journal of Advanced Manufacturing Technology*, Available online (http://dx.doi.org/10.1007/s00170-008-1715-y).

Petrovic, D. (2001). Simulation of supply chain behaviour and performance in an uncertain environment. *International Journal of Production Economics*, 71(1-3): 429-438.

Petrovic, D.; Roy, R. & Petrovic, R. (1998). Modelling and simulation of a supply chain in an uncertain environment. *European Journal of Operational Research*, 109(2): 299-309.

Petrovic, D.; Roy, R. & Petrovic, R. (1999). Supply chain modelling using fuzzy sets. *International Journal of Production Economics*, 59(1-3): 443-453.

Sakawa, M.; Nishizaki, I. & Uemura, Y. (2001). Fuzzy programming and profit and cost allocation for a production and transportation problem. *European Journal of Operational Research*, 131(1): 1-15.

Santoso, T.; Ahmed, S.; Goetschalckx, M. & Shapiro, A. (2005). A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1): 96-115.

Selim, H.; Araz, C. & Ozkarahan, I. (2007). Collaborative production-distribution planning in supply chain: A fuzzy goal programming approach. *Transportation Research Part E: Logistics and Transportation Review*, In Press, Corrected Proof.

Shih, L. H. (1999). Cement transportation planning via fuzzy linear programming. *International Journal of Production Economics*, 58(3): 277-287.

Sodhi, M. S. (2005). Managing demand risk in tactical supply chain planning for a global consumer electronics company. *Production and Operations Management*, 14(1): 69-79.

Sridharan, V.; Berry, W. & Udayabhanu, V. (1987). Freezing the master production schedule stability under rolling planning horizons. *Management Science*, 33(9): 1137-1149.

Torabi, S. A. & Hassini, E. (2008). An interactive possibilistic programming approach for multiple objective supply chain master planning. *Fuzzy Sets and Systems*, 159(2): 193-214.

Vasant, P. M. (2005). Fuzzy linear programming for decision making and planning under uncertainty. *International Journal of Information Technology & Decision Making*, 4(4): 647-662.

Wang, J. T. & Shu, Y. F. (2005). Fuzzy decision modeling for supply chain management. *Fuzzy Sets and Systems*, 150(1): 107-127.

Xie, Y.; Petrovic, D. & Burnham, K. (2006). A heuristic procedure for the two-level control of serial supply chains under fuzzy customer demand. *International Journal of Production Economics*, 102(1): 37-50.

Yager, R. (1979). Ranking fuzzy subsets over the unit interval. *In: Proceedings of 17th IEEE International Conference on Decision and Control*, San Diego, CA: 1435-1437.

Yager, R. (1981). A procedure for ordering fuzzy subsets of the unit interval. *Information Sciences*, 24: 143-161.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, 8(3): 338-&.

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1: 3-28.

Zimmermann, H. J. (1976). Description and Optimization of Fuzzy Systems. *International Journal of General Systems*, 2(4): 209-215.

Zimmermann, H. J. (1978). Fuzzy programming and linear programming with several objective functions. *Fuzzy Sets and Systems*, 1: 45-55.

17

Research Issues on Collaborative Product Design and Development

Jiun-Yan Shiau
Department of Logistics Management
National Kaohsiung First University of Science and Technology
Taiwan

1. Introduction

1.1 What is collaborative product design and development

Collaborative product design and development (CPD) is also known as collaborative product definition management (cPDM). It is about business strategy, workflow and collection of software applications that facilitates different vendors to work together on development/design of a product. The early participation of vendors in the design process is considered critical in order to improve the product quality and reduce the development cycle time. CPD is becoming more valuable because of the increasing coordination and management complexity of organizational information, responsibilities, schedules, deliverables, product information, and business process. As outsourcing and globalization increase the number of design chain participants, a CPD speeds up the decision-making of trusted partners, employees, suppliers, and customers in design chains. Design chain is a subset of supply chain. The major collaborative activities between suppliers and manufacturers are design activities. Therefore, how to manage the design flow in a design chain is as important as how to manage the material flow in a supply chain.

1.2 What are the main phases of product design and development

Before discuss the collaboration issues for product design and development, we need to briefly review the phases of product design and development. The major phases of product design and development are normally defined as conceptual design, preliminary design, and detail design and development (Blanchard, 2004). During the conceptual design phase, the concepts, which are also called scheme, are built in order to completely and efficiently design the transceiver. The concepts may include such as product operational requirements and maintenance, current product problem (or deficiency), functional analysis for the product, applicable technical performance measures (TPMs), and specific performance measures and design-to criteria. Preliminary design phase begins with a "functional baseline" product configuration described in the product specification prepared during conceptual design phase. The functional baseline is translated into detailed qualitative and quantitative design requirements for allocating applicable elements of the product. An "allocated baseline" configuration in the form of development, product, and process specifications is established during this phase. At the beginning of detail design and development phase, a rough product configuration has been defined, a functional analysis has been accomplished, and the requirements for detail design have been included in the appropriate specifications. The information above must be converted into the proper mix of hardware, software, people, data, and specific items of support during the detail design and development.

1.3 What is the core technology of collaborative product design and development

The core technology comes for CPD does vary depending on who you ask. However, it usually consists of the product lifecycle management (PLM), product data management (PDM), product visualization, team collaboration and conferencing tools, supplier sourcing software, and data translation technology. It is generally not including CAD geometry authoring tools. In this chapter, we will concentrate on PLM and PDM.

1.4 What is PLM/PDM

PLM, which is known as PDM formerly, is the process of managing the entire lifecycle of a product from its conception, through design and manufacture, to service and disposal. PDM systems first appeared in the 1980s. The early PDM systems were effective in the engineering domain, but failed to encompass non-engineering activities, such as sales, marketing, and supply and customer management. With development of newer information technologies, web-based PDM systems were introduced and better accessibilities to suppliers and customers were provided. PDM, however, was still confined to engineering information management (Ameri & Dutta, 2005). Around 2003, PDM was expected to focus on product lifecycle stages in general; an improved support of engineering collaboration functionality, the name of PLM was thus given.

1.5 How to apply PLM/PDM to collaborative product design and development

The purpose of this chapter is to discuss research issues on how PLM/PDM is applied to CPD. From collaborative environment perspective, we categorize CPD into (1) single firm or multiple firms, (2) centralized managed or distributed managed, and (3) localized or global. From strategy perspective, Krishnan and Ulrich (2001) categorized product design and development as (1) marking, (2) organizations, (3) engineering design, and (4) operations management. From project management perspective, product design and development can be categorized as (1) conceptual design, (2) design chain, (3) detail design, and (4) production ramp-up. Based on these three perspectives (i.e., collaborative environment, strategy and project management), this chapter surveys related literatures on PLM/PDM and proposes a research issue cube (as shown in Figure 1) for CPD.

2. Research issues on applying PLM/PDM to collaborative product design and development

2.1 Configuration management theory

Before further discussion on how to apply PLM/PDM to each cell inside the research issue cube (as shown in Figure 1), the preliminary background about the theory behind PLM/PDM is required. Configuration management (CM) is the theory behind a PDM system. Some researchers considered PDM as an implementation of CM principles (Lyon, 2002). There are also some researchers developed distributed CM principles or web-based PDM for distributed collaborative environment.

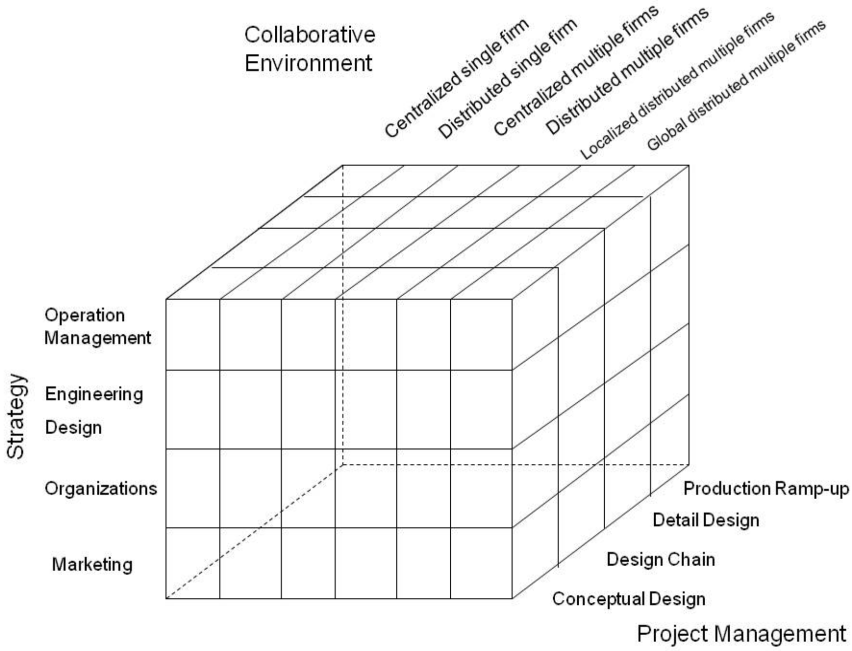


Fig. 1. Research Issue Cube on Collaborative Product Design and Development

Configuration management was first introduced by US Department of Defense in 1992 (Lyon, 2002). It is a discipline applying technical and administrative direction, and a surveillance over the life cycle of configuration items (CI's) to:

- Identify and document the functional and physical characteristics of CI's.
- Control change to CI's and their related documentation.
- Record and report information needed to manage CI's effectively, including the status of proposed and approved changes.
- Audit CI's to verify conformance to documented requirements

Some forms such as ECR (enterprise change request) and ECN (enterprise change notice) are commonly used in the configuration management. The forms, used in the configuration management process, serve two purposes.

- To provide authorization to do work.
- To provide a historical record plus proof of conformance.

Also configuration management is a theory proposed for tracing and maintaining the integrity among valuable outputs during the lifecycle of product development. According to IEEE standard for software configuration management plans, a configuration includes configuration items and their structures at each project control point. Configuration items of software could be physical and functional characteristics of the code, specifications, design, data elements, outputs of the development process, and elements of the support environment. Structures mean the way of combinations among configuration items. Shiau et al. (2008) proposed a formulization of configuration management. Let's denote a structure among configuration items as a matrix S. The equation below represents the concept of a configuration.

$$Configuration = (CI_i, S) \quad where\ i = 1, 2, \ldots, n\ and\ n\ is\ a\ constant\ number$$

Normally structure (*S*) utilized in a configuration management plan should be static and unchangeable during the development cycle (Gruhn et. al., 2003). When different versions of CI's with static structure are created, approved and finally released, they form a new version of configuration. The equation below represents the concept of a version of configuration.

$$\begin{aligned} &Version(Configuration) \\ &= Version(CI_i, S) \\ &= (Version(CI_1), Version(CI_2), \ldots, Version(CI_n), S) \end{aligned}$$

When different versions of configurations are based on a static structure, they are defined as a version-aware configuration and the changing history of the configurations are then traceable. However, if two versions of configurations have different structures, they are defined as non-version-aware configurations. For example below, $Version(Configuration_1)$ and $Version(Configuration_2)$ have structures, S_1 and S_2, respectively. The changing history of the two configurations may be unable linked and therefore untraceable.

$$Version(Configuration_1) = (Version(CI_1), Version(CI_2), Version(CI_3), S_1)$$

$$Version(Configuration_2) = (Version(CI_2), Version(CI_4), Version(CI_5), S_2)$$

A version-aware configuration at a specific project control point is called a baseline. The concept of a baseline concentrates on the status of configurations. The equation below represents the concept of a baseline of configuration.

$$\begin{aligned} &Baseline(Configuration) \\ &= Status(CI_i, S) \\ &= (Status(CI_1), Status(CI_2), \ldots, Status(CI_n), Status(S)) \end{aligned}$$

When two baselines are established at two specific project control points, the status of configurations is recoverable. For example below, $Baseline(Configuration_1)$ and $Baseline (Configuration_2)$ are two configurations with different structures at two project control points. It is able to recover the status back to either $configuration_1$ or $configuration_2$ once the two baselines are established.

$$Baseline(Configuration_1) = (Status(CI_1), Status(CI_2), Status(CI_3), Status(S_1))$$

$$Baseline(Configuration_2) = (Status(CI_2), Status(CI_4), Status(CI_5), Status(S_2))$$

Based on these two concepts (i.e., version control and baseline management) plus a set of automated computer modules (for example, a workflow management module, an authorization module, status accounting module, configuration auditing module, and so on), configuration management can help an enterprise to maintain the consistency among CI status while changes occur.

2.2 Issues in each dimension

2.2.1 Research issues on configuration items identification related to collaborative environment

A CI identification principle expressed in EIA/IS-649 is that each CI, which is usually represented in electronic document format, must have a unique identifier so that it can be associated correctly with the configuration of the physical item to which it relates. The US Department of Defense and all military components use the following three elements to assure the unique identity of any document: CAGE code, document type and document identifier. A configuration items identification activity guide is provided in Table 1.

Preferred Identifier Element	Definition
Document Identifier	
CAGE Code or NSCM (NATO Supply Code for Mfg.)	CAGE (and NSCM) Codes identify the source of the document. The codes are provided in **Defense Logistic Agency (DLA) Cataloging Handbook H4/H8 Series.** The codes are affixed to all CIs, and their *replaceable* subordinate parts and assemblies. They are also part of the identification marking of each item of configuration documentation, software media and software product.
Document Identifier	The document Identifier distinguishes one document produced by the organization referenced by the CAGE code from another. Each document and each revision thereto, requires the document identifier. There are as many schemes for identifying documents as there are organizations producing them, so there is no standard format for all documents. There are however, a few common sense constraints on the numbering content for some specifications, and engineering drawings, as defined in applicable standards
Revision/Version identifier	
Revision Identifier	Revision Identifier clearly establishes which issue of a particular document is current or applicable.
Version Identifier	Conceptually the same as revision, version is the term typically used for files
Date	Date is an additional discriminator. It is good common sense business practice to date every document and every revision
Restrictive Markings: ***These requirements apply to digital data files and digital media as well as to paper documents and are all intended to limit the access to such data to those entitled to access them.*** **[Ref: DoD FAR Supplements 252.227-7015, 7018, 7052 and -7057]**	
Security Markings	Security markings are required to be clearly marked on all classified data and special handling requirements apply. Each contract contains classification guidance and direction, which must be strictly adhered to.
Distribution Statements	Specific distribution statements and export restrictions must be marked on information subject to secondary distribution limitations as prescribed by law and as indicated by the contract. The purpose of these markings is to inform the secondary distributor, such as a Government repository whether they can legally provide the subject information to third parties, and if the data are allowed to be exported to foreign countries.
Data Rights	Documents which contain data for which the Government or other parties do not have unlimited rights, must be appropriately labeled to indicate the data rights limitations, so that proprietary information disclosed to the Government for specific purposes is protected.

Table 1. Document Identification Activity Guide (MIL-HDBK-61A, 2008)

Research issues on configuration items identification for a single firm are related to the activities shown in Table 1 plus a systematically process to generate product configurations. Normally, a company will not generate any product configuration (such as functional baseline, allocated baseline, etc.) from scratch. Therefore, a systematical normalization process is required. With such process, one can normalize all the collected data/documents into several hierarchical styles of configurations (such as functional baseline, allocated baseline, E-BOM, M-BOM etc.) systematically. This is similar to normalization processes of relational database. Database engineers can decompose all collected persistent attributes into several relational tables in order to fulfill criteria of 1st, 2nd, or 3rd normalization forms. The aspect oriented configuration identification model presented in (Shiau et al., 2008) is one example in configuration items identification area. Jiao and Zhang (2005) presented an association rule mining approach for product portfolio identification is another example. Wang and Lin (2003) proposed a fuzzy multicriteria group approach for selecting configuration items is also an example in this area.

In addition to the issues above, research issues on configuration items identification for multiple firms or distributed single firm include zoning and partition procedures for further categorizing those hierarchical styles of configurations. The inter-company configuration and intra-company configuration presented in (Shaiu & Wee, 2008) is an example of the outcomes of such procedures.

Data exchange and format conversion are critical research issues for localized and global firms. The differences among international currencies, metrologies and regulations cause the needs of data exchange and/or format conversion among configuration items. The data exchange issues will be even more complex if structures of a configuration are diverted due to globalization

2.2.2 Research issues on change control workflow related to collaborative environment

A change control workflow is a logistic procedure to control and coordinate changes among configuration items for ensuring the consistency of a configuration. The Institute of Configuration Management (2002) proposed a closed-loop change control workflow (see Figure 2) within the CMII principles as a reference model for managing changes. In addition to CM principles, CMII shifts the emphasis of CM to (1) accommodate change, (2) accommodate the reuse of standards and best practices, (3) assure that all requirements remain clear, concise and valid, (4) communicate (1), (2) and (3) to each user promptly and precisely and (5) assure conformance in each case. As shown in Figure 2, an enterprise change request (ECR) is provided and passed to Change Administration I, when a document in the baseline is intended to be changed (i.e, an engineering change is requested).

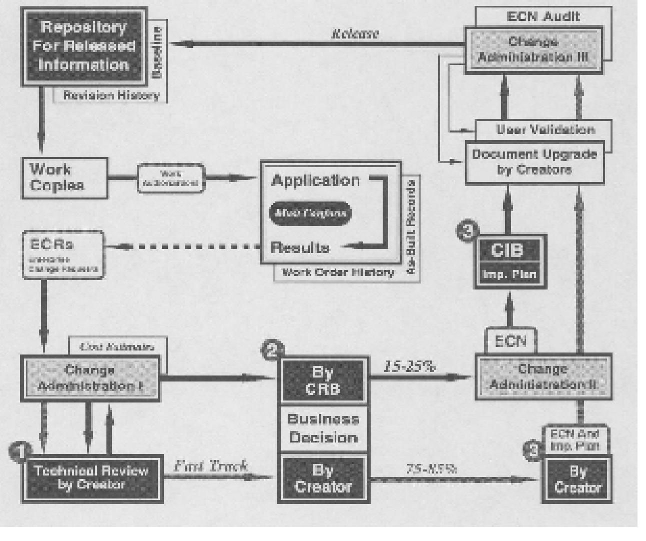


Fig. 2. Closed-Loop Change Control Workflow (Institute of Configuration Management, 2002)

ECR is a kind of document that records what to change, the reason to change and the priority of changes. Change Administration I accepts or denies the ECRs based on the results consulted from professionals in charge with each configuration item (CI). Accepted ECRs are then passed to change review board (CRB) or original creators for approval and then for making business decision based on further discussion in CRB meetings. Approved ECRs are organized as enterprise change notices (ECNs) by Change Administration II. ECN is a document recording how to change and when to change. Change implementation board (CIB) is held by Change Administration II for planning the detail of ECN implementations. Finally, Change Administration III audits the consistencies of ECNs, releases the revised documents, and updates new states to the baseline. CRB and CIB together are so called change control board (CCB).
Research issues on change control workflow for a single firm are about minimal disruption to services, reduction in back-out activities, and economic utilization of resources involved in implementing change. Techniques of workflow management are helpful while building such workflow. A closed-loop design change control workflow (see Figure 3) during conceptual design phase proposed by (Shiau and Li, 2007) relates to this area.

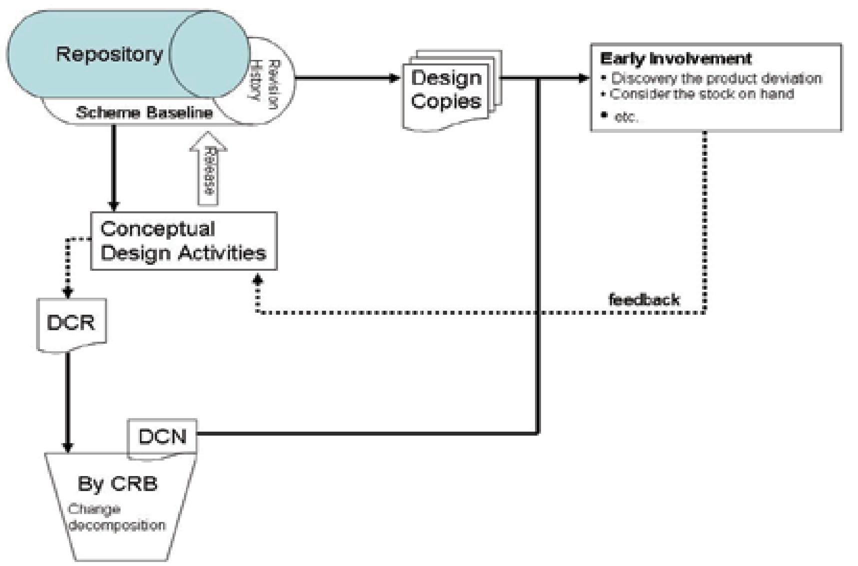


Fig. 3. A Closed-loop Design Change Control Workflow

Research issues on change control workflow for multiple firms either reside in local area or global areas are more depending on distributed workflow management and technology. The distributed change control workflow demonstrated in (Shiau and Wee, 2008) is probably the first one in this area. The distributed change control workflow, which is also a kind of distributed algorithm, is illustrated as below:

Send: Every t days or when collaborative design table, Tl, changes, send Tl to each related companies
Receive: whenever Tr is received from another company via interface n:
For all rows Rr in Tr {
    Determine if (Rr.C_{ij} <> Rl.C_{ji}) {

```
        // calculate new utility of Cji
        if (Rr.company is not in Tl) {add Rr to Tl }
        else for all rows Rl .Cji.utility in Tl {
                if (Rr.company = Rl.company and Rr.Cij.utility > Rl.Cji.utility) {
                        Rl = Rr
                }
        }
}
}
```

Every t days or when collaborative design table (*Tl*) changes, *Tl* is send to each outgoing interface. When a collaborative design table (*Tr*) is received on interface n, it compares all rows (*Rr*) in *Tr* with it's own collaborative design table (*Tl*). If an integration checking table, C_{ij}, in any *Rr* is not equal to the integration checking table, C_{ji} in *Rl*, calculate the utility of *Rr* in the integration checking table and set the interface of *Rr* to n. If any company in *Rr* is not in *Tl*, add *Rr* to *Tl*. Compare all *Rl* in *Tl*, if the company in *Rr* is equal to the company in *Rl* and the utility of collaborative design table in *Rr* is less than the utility of collaborative design table in *Rl*, then replace *Rl* with *Rr*. A collaborative design table records the information of how a company determines which assembly interface to deal with when a design change occurs. The recorded information includes the company names, the assembly interfaces of end-products, and the integration checking tables. An integration checking table contains a list of preference values, called utilities, to every assembly interfaces for a company.

2.2.3 Research issues on configuration status accounting related to collaborative environment

Configuration status accounting is a task of CM concerned with recording the state of a CI at any point in time, past, present or future. There are three distinct but overlapping areas in configuration status accounting activities. They are configuration status accounting data capture, configuration status accounting data processing, and configuration status accounting data reporting (Lyon, 2002). The aim of configuration status accounting is to ensure that not only the physical configuration item, but also the configuration documentation describing that physical configuration item, is always at a known state commensurate with the grade of CM being applied.

Research issues on configuration status accounting for single firm are about the implementation of system modules to track the location and actual build state of individual CI or whole configurations. Burgess et. al. (2003) explored how the European aerospace industry views and practices 'configuration status accounting' is an example in this area.

2.2.4 Research issues on configuration auditing related to collaborative environment

An audit is an independent evaluation of a configuration to ascertain compliance with specifications, standards, contractual agreements or other authorized criteria. As such, audits are a quality assurance function, and all configuration auditing processes must be integrated with existing quality assurance/management procedures. There are three categories of configuration audits. They are functional configuration audits, physical configuration audits, and configuration verification audits (Lyon, 2002). Configuration audits should not be viewed simply as a test for compliance; they should be considered as method for determining the level of compliance achieved, with the aim being to identify any areas requiring additional effort. Significant pre-audit checks and consultation should be carried out prior to official configuration audit activities. There is lack of research and case study in this area today.

2.2.5 Research issues on configuration management related to strategy of product design and development

As shown in Table 2, Krishnan and Ulrich (2001) did a comprehensive survey of papers in design and development research according to the four common perspectives, which are marketing, organizations, engineering design, and operations management. Research issue on this area is to develop new strategies for collaboration among venders in terms of marketing, organizations, engineering design, and operations management. For example, the concept of vendor managed inventory (VMI) could be borrowed for CPD among joint development manufacturers (JDMs). VMI is a kind of business model in which the buyer of a product provides certain information to a supplier of that product and the supplier takes full responsibility for maintaining an agreed inventory level of that product, and sometimes even responsibility for maintaining agreed inventory levels of related products, called product family. Borrow such strategy, a JDM might consider to provide vendor managed change control service. Besides providing services of design and manufacturing, a JDM may now provide design logistics management for its client. Any change among components of a product from other JDMs could cause the integrity issue and require design logistics management. A modification of today's PDM/PLM for such collaborative strategy could be an interesting topic. A prototype of conceptual design information system proposed in (Shiau et. al., 2004; Lin et. al., 2004; Huang et. al., 2006) is partial of such system.

	Marketing	Organizations	Engineering Design	Operations Management
Perspective on Product	A product is a bundle of attributes	A product is an artifact resulting from an organizational process	A product is a complex assembly of interacting components	A product is a sequence of development and/or production process steps
Typical Performance Metrics	"Fit with market" Market Share Consumer utility (Sometimes profits)	"Project success"	"Form and function" Technical performance Innovativeness (Sometimes direct cost)	"Efficiency" Total cost Service level Lead time Capacity utilization
Dominant Representational Paradigm	Customer utility as a function of product attributes.	No dominant paradigm. Organizational network sometimes used.	Geometric models. Parametric models of technical performance.	Process flow diagram Parametric models of process performance.
Example Decision Variables	Product attribute levels, price	Product development team structure, incentives	Product size, shape, configuration, function, dimensions	Development process sequence and schedule Point of differentiation in production process
Critical Success Factors	Product positioning and pricing Collecting and meeting customer needs	Organizational alignment Team characteristics	Creative concept and configuration Performance optimization	Supplier and material selection Design of production sequence Project Management

Table 2. Comparison of Perspectives of the Academic Communities in Marketing, Organizations, Engineering Design, and Operations Management (Krishnan & Ulrich, 2001)

2.2.6 Research issues on configuration items identification related to project management

Project management is the discipline of planning, organizing, and managing resources under constraints such as scope, quality, time, and budget to bring about the successful completion of specific project goals and objectives. Different phases of product design and development usually own different types of projects. Research issues on configuration items identification in this area are about establishing As-Planned and As-Released baselines during each phase of CPD. Baseline is the compilation and accumulation of all documentation plus digital design files that represent a product at a specific point in time. Gruhn et. al. (2003) presented a case study about how to apply software CM while developing a software application is an example in this domain.

Configuration items related to project management are also categorized as resource oriented or activity oriented. Depend on types of projects (i.e., conceptual design project, detail design project, production ramp-up project, or design chain project), the configuration items are very diverse. For example, CI's for production ramp-up project might be more resource oriented. Materials in E-BOM and materials in M-BOM are examples of CI's in production ramp-up project. Oppositely, CI's for conceptual design project might be more activity oriented since the resources (or outputs) of conceptual design is unable to define during the early stage of conceptual design.

Research issues in this domain are about when to identify project resources as CI's and when to identify project activities as CI's. Sometimes, if it is hard or impossible to identify project resources, one might needs to identify them in the dual plane of original configuration. The aspect oriented configuration identification model presented in (Shiau et al., 2008) is an example of identifying CI's in dual plane. Similar to the Fourier analysis in mathematics, if a line in x and y plane (see Figure 4) is:

$$y = px + q$$

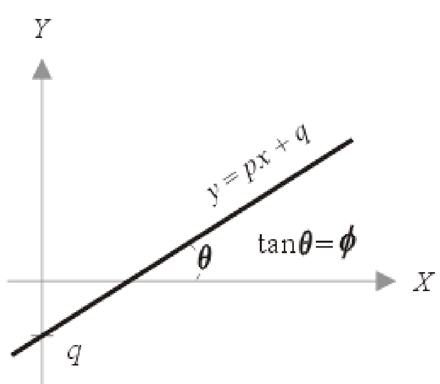


Fig. 4. XY Plane

One can solve the equation above by solving its dual plane (see Figure 5) below:

$$q = -px + y$$

Let's formulate component configuration for a system as:

$$\textit{configuration} = (\textit{vertical viewpoint})\ \textit{CI's} + \textit{static structure}$$

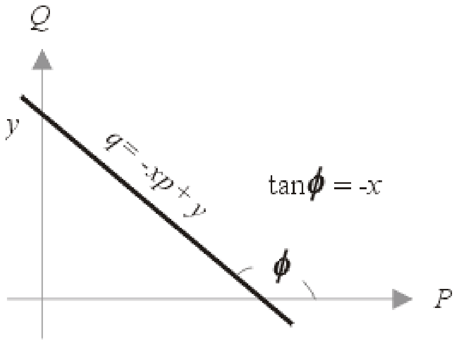


Fig. 5. PQ Plane

If we could not find a static component configuration from the vertical viewpoint (for example, class perspective in software applications), we can try to find one from the horizontal viewpoint (for example, crosscut-classes perspective in software applications). The dual plane is as below:

static structure
= -(vertical viewpoint) CI's + configuration
= (horizontal viewpoint) CI's + configuration

2.2.7 Research issues on change control workflow related to project management

There are two types of change control workflow for product design and development. One is called engineering change control; the other is called design change control. Normally, the change control process is launched while the first edition of manufacturing bill of materials (M-BOM) is generated and recorded in the repository. Due to continuous changes in production engineering in nature, and inevitable errors and changes in products, ECR is not avoidable during the whole product lifecycle. Engineering change control is designed for in-time feedback from production phase to product design phase. A common misunderstanding of this change control model is to launch engineering change management during the conceptual design phase.

Design change request (DCR) and design change notice (DCN) are the forms designed to give in-time feedback from conceptual design phase to production phase and service phase. These are opposite to the directions of feedback of engineering change control. Although DCR and DCN forms are similar to ECR and ECN forms, the change control workflow for engineering change is no longer suitable for design changes. Therefore there is a need to develop aother change control workflow for phases before detail design. A design change control workflow (as shown in Figure 3) is different from an engineering change control workflow in two essential aspects:

- The "feedback" here is a forward notification mechanism that goes from the upstream conceptual design to downstream activities including detailed design, production planning, etc. Such changes are mostly useful when they occur at the early development stages of a product, i.e. when most efforts are spent in conceptual design.
- The working forms or messages and their functions are, largely due to the above difference, very different from the ECR and ECN. In this sense these forms should function more as active triggers rather than passive responses. This requires special

attention to the control and synchronization of the lifecycle, authorization and structure workflows in terms of project, document and scheme configuration management.

A closed-loop design change control workflow (Shiau & Li, 2007) was developed for the purpose of continuous improvement of conceptual design. It is also a workflow adapting concept of early involvement of concurrent engineering.

Depend on performance indexes (i.e., quality, time, or budget) of a project, the analysis and evaluation approaches in CRB (see Figure 2), which is part of change control workflow, are very diverse and also time-consuming. The decisions of the CRB are about when a change is to be made (effectivity) and what should be done to the existing inventory of the old configuration assemblies and components. In practice, CRB members have to fill in a planned effectivity in each ECR during analyzing and evaluation phase of an engineering change. In the survey of (Huang et. al., 2003), shop floor workshop, design office, and quality control department are the major representatives in the CRB. Habhouba et. al. (2006) reported that most PDMs do not offer any intelligence or decision-making assistance during change control. Shiau (2007) presented an effectivity date maintenance model for PDM is an example in this domain.

2.2.8 Research issues on configuration status accounting and configuration auditing related to project management

Research issues on configuration status accounting and auditing in this area are similar to issues for single firm describted before. Burgess et. al. (2003) explored how the European aerospace industry views and practices 'configuration status accounting' is an example. Procedures of configuration status accounting and auditing proposed in (Lyon, 2002) probably is one of the most compreshesive approaches in this area.

3. Discussion and further research

Table 3 shows the overview of research issues on CPD in terms of collaborative environment, strategy, and project management. To solve those research issues, it requires the wide spectrum of CM or PDM/PLM knowledge. The major problem in apply CM to CPD is lack of standard procedures for identifying configuration, controlling changes, accounting configuration status, and auditing configurations (Schuh, 2008). Most researches available today are about the principles and guidelines in applying CM. In this chapter, we reviewed several specific procedures for identifying configuration, controlling changes, accounting configuration status, and auditing configurations from previous researches.

There are still some undiscovered cells inside our proposed research issue cube. There is also no generic procedure discovered till the date this chapter is written. Despite these shortages, we believe the generic procedures will be inducted in the near future with more and more specific implementations of CM to industries. For example, the idea of apply PLM to maintenance services in aerospace industry (Lee et. al., 2008).

4. Conclusion

The trend of component manufacturing has been changed from EMS (electronic manufacturing services) provider to JDM (joint development manufacturer). EMS is an

Dimension	*CM area*	*Research Issues*	*References*
Collaborative Environment	Configuration Identification	systematically normalization process; zoning and partition procedures; data exchange and format conversion	(Jiao and Zhang, 2005); (Shaiu & Wee, 2008); (Shiau et al., 2008); (Shaiu & Wee, 2008); (Wang & Lin, 2003)
	Change Control	centralized and distributed workflow design and management	(Institute of Configuration Management, 2002); (Shiau and Li, 2007); (Shaiu & Wee, 2008)
	Configuration Status Accounting	data capture, data processing, and data reporting	(Burgess et. al., 2003); (Lyon, 2002)
	Configuration Audits	quality assurance function	(Lyon, 2002)
Strategy	Configuration Management	New strategies on marketing, organizations, engineering design, and operations management	(Huang et. al., 2006); (Lin et. al., 2004); (Shiau et. al., 2004)
Project Management	Configuration Identification	baseline establish, resource oriented and activity oriented configuration items	(Gruhn et. al., 2003); (Shiau et al., 2008); (Wang & Lin, 2003)
	Change Control	forward and backward oriented notifications, decision making models	(Huang et. al., 2003); (Shiau, 2007); (Shiau and Li, 2007)
	Configuration Status Accounting	data capture, data processing, and data reporting	(Burgess et. al., 2003); (Lyon, 2002)
	Configuration Audits	quality assurance function	(Lyon, 2002)

Table 3. Overview of Research Issues on Collaborative Product Design and Development

industry based on providing contract design, manufacturing and product support services on behalf of OEMs (original equipment manufacturers). However, all intellectual property of the new product belongs to the OEM. JDM is a company that helps design parts of a product for OEM customers. Unlike EMS, JDM may own the copyright of its design and provide joint design services to its OEM customer. The core competitions of JDM are joint design and productivity of manufacturing. Once basis requirement is hand-off from customer, JDM takes the job of completing the design, performs design verification, assembles and tests prototypes, assembles and tests qualification units, assembles and test proof manufacturing units, and finally produces the production units (Kaylor, 2004). Lots of stakeholders have recognized that the earlier the manufacturer becomes involved in the design process, the better the product. When a product is outsourcing to OEM and OEM outsource it to down tiers of JDMs, a design chain is formed. In order to design a product, a JDM has to joint design with its OEM customer and also collaborate design with other JDMs in a design chain environment. With the trend above, CPD is becoming more valuable.

In this chapter, the research issues of collaborative product design and development based on CM principles are explored. The four major CM areas, which are configuration identification, configuration change control, configuration status accounting, and configuration audits, were introduced under three dimensions (see Figure 1). Although this chapter may not provide the exhaustive review for all research issues of applying CM/PDM/PLM to CPD, we think that our work has laid the cornerstone of CM/PDM/PLM and CPD studies.

5. Acknowledgements

This work is supported in part by the National Science Council (NSC) of Republic of China under the grant NSC 95-2221-E-033-045-MY3 and in part by the College of Electrical Engineering and Computer Science at Chung Yuan Christian University of Taiwan, Republic of China.

6. References

Ameri, F. and D. Dutta (2005). Product Lifecycle Management: Closing the Knowledge Loops., *Computer-Aided Design & Applications*, Volume: 2, Issue: 5, pp.577-590.

Blanchard, Benjamin S. (2004). Logistics Engineering and Management, *Prentice Hall*, ISBN 978-8120327634, USA.

Burgess, T.F.; Byrne, K.; Kidd, C. (2003). Making project status visible in complex aerospace projects, *International Journal of Project Management,* Volume: 21, Issue: 4, May, 2003, pp. 251-259

Gruhn, V., Ijioui, R., Peters, D., Schäfer, C. (2003). Configuration management for Lyee software, *Knowledge-Based Systems*, Volume: 16, Issue: 7-8, November, 2003, pp. 441-447.

Habhouba, Dounia, Alain Desrochers, and Soumaya Cherkaoui (2006). Engineering Change Management and Decision-Making Assistance Using Software Agent, *IEEE Canadian Conference on Electrical and Computer Engineering*, May 2006 , pp. 1694-1697, Ottawa.

Huang, Cheng-Wei, Jiun-Yan Shiau, and Hui Ming Wee (2006) “A Conceptual Design Formation Information System for Joint Design Manufactures,”*The 36th International conference on Computers and Industrial Engineering*, June 20-23, 2006, Taipei, Taiwan.

Huang, G.Q., W.Y. Yee, and K.L. Mak (2003). Current Practice of Engineering Change Management in Hong Kong Manufacturing Industries, *Journal of Materials Processing Technology*, 2003, Vol. 139, pp. 481-487.

Institute of Configuration Management (2002). CMII for Business Process Infrastructure, *Holly Publishing Company*, ISBN 978-0972058209, USA.

Jiao, Jianxin; Zhang, Yiyang (2005). Product portfolio identification based on association rule mining, *Computer-Aided Design*, Volume: 37, Issue: 2, February, 2005, pp. 149-172.

Kaylor Jeff (2004). Defense & Aerospace Electronics Manufacturing Outsourcing – Realities & Trends, *the Proceedings of the SMTA International Conference*, 2004, pp.27-30, Chicago, Illinois, USA.

Krishnan, V. and Karl T. Ulrich (2001). Product Development Decisions: A Review of the Literature, *Management Science*, Vol. 47, No. 1, January 2001, pp. 1-21.

Lee, S.G., Y.-S. Ma, G.L. Thimm, J. Verstraeten (2008). Product lifecycle management in aviation maintenance, repair and overhaul, *Computers in Industry*, March 2008, Vol. 59, Issue 2-3, pp. 296-303.

Lin, Yu-Jen, Cheng-Wei Huang, Ju-Ching Tseng, Jiun-Yan Shiau (2004) "Issue Resolution for Conceptual Design Using AHP," *Intelligent systems in Design and Manufacturing V, Proceedings of SPIE vol. 5605*, October 25-26, 2004, Philadelphia, Pennsylvania, USA.

Lyon, D. Douglas (2002). Practical CM 4th Edition, *Raven Publishing Company*, ISBN 978-0966124842, USA.

MIL-HDBK-61A: Configuration Identification (2008). < http://www.product-lifecycle-management.com/mil-hdbk-61a-5-6.htm>, Sep 2008.

Schuh, Günther, Henrique Rozenfeld, Dirk Assmus, Eduardo Zancul (2008). Process oriented framework to support PLM implementation, *Computers in Industry*, March 2008, Vol. 59, Issue 2-3, pp. 210-218.

Shiau, Jiun-Yan (2007). Effectivity Date Maintenance for a Product Data Management System, *16th International Conference on Software Engineering and Data Engineering*, July 9-11, 2007 , Las Vegas, Nevada USA.

Shiau, Jiun-Yan, James C. Chen, Yu-Jen Lin, and Chih-Ming Chang (2004) "Architecture of Distributed Conceptual Design System for EMS Providers," *The 34th International conference on Computers and Industrial Engineering*, November 14-16, 2004, San Francisco, CA, USA.

Shiau, Jiun-Yan and Hui Ming Wee (2008). A Distributed Change Control Workflow for Collaborative Design Network, *Computers in Industry*, March 2008, Vol. 59, Issue 2-3, pp. 119-127.

Shiau, Jiun-Yan and Xiangyang Li (2007). A Closed-Loop Design Change Control Workflow in Product Data Management for Conceptual Design Phase, *International Journal of Intelligent Control and Systems*, March 2007, Vol. 12, Number 1, pp. 24-36

Shiau, Jiun-Yan, Xiangyang Li, and Pei-Hua Tseng (2008). An Aspect Oriented Configuration Identification Model for Incremental Development, *IEEE International Conference on Networking, Sensing and Control*, April 6-8, 2008, China.

Wang, Juite and Yung-I Lin (2003). A fuzzy multicriteria group decision making approach to select configuration items for software development, *Fuzzy Sets and Systems*, Volume: 134, Issue: 3, March 16, 2003, pp. 343-363.

18

Improvement of Supply Chain Performances Using RFID Technology

Cornel Turcu, Cristina Turcu and Adrian Graur
Stefan cel Mare University of Suceava
Romania

1. Introduction

As markets become more global and competition intensifies, firms are beginning to realize that competition is not exclusively a firm versus firm domain, but a supply chain against supply chain phenomenon (***a, 2008). Under these circumstances, an increasing strategic importance to any organization independent of size or of sector, is to deliver information, goods and services in full, on time and error-free to customers.

From demand forecasting, to the sourcing of raw materials, right through to manufacture and dispatch- visibility in the supply chain is becoming an important facet of any modern operation (Coltman et al., 2008). But at this moment, the interconnectivity between various links in the supply chain is incomplete and inaccurate, every link in the chain being an individualistic entity with different processes. This leads to poor product visibility and stock transparency across the supply chain. For companies looking at multiple markets, the lack of visibility in their supply chain can lead to tremendous loss of revenue.

But even if information technology is used within a supply chain to share information on end-customer demand and inventory levels, there is still often a discrepancy between this information and the real physical flow of products. This discrepancy frequently derives from the missing real-time or near real-time data in concordance with the physical flow of goods. The result is inaccurate inventory information. Reasons why information system inventory records are inaccurate include external and internal theft, unsaleables (e.g. damaged, out-of-date, discontinued, promotional, or seasonal items that cannot be sold any longer), incorrect incoming and outgoing deliveries (Raman et al., 2001; Fleisch & Tellkamp, 2003), as well as misplaced items (Raman et al., 2001). Thus, even when inventory records are accurate, misplaced items mean that they were not out of stock, but rather misplaced in storage areas or in the wrong location within the store.

The phenomenon of inventory inaccuracy is well-known. As Raman et al. (Raman et al., 2001) show in their case study, most retailers cannot precisely identify the number of units of a given item available at a store; thus for more than 65% of stock keeping units (SKUs) in retail stores, information on inventory in the inventory management system was inaccurate (i.e. the information system inventory differed from physical inventory). The difference was on average 35% of the target inventory. In a second case study, the authors found that a median of 3.4% of SKUs were not found on the sales floor although inventory was available

in the store. In the first case, inventory inaccuracy reduced profits by 10 %, while in the second case, misplaced items reduced profits by 25%.
Inventory record inaccuracy and misplaced items can lead to a substantial decrease in profits due to lost sales, additional labor costs, and higher inventory carrying costs. All these problems may also have a long-term negative impact on firm image.
RFID technology can be a solution to these problems by tracking and tracing products at any point across the supply chain. Thus, RFID will have a significant impact on every facet of supply chain management—from the mundane, such as moving goods through loading docks, to the complex, such as managing terabytes of data as information about goods on hand is collected in real time (Caton, 2004).
For the perishable goods industry, demand management is crucial. In the United States, up to 20 per cent of foods are discarded due to spoilage in the supply chain (Rangarajan et al., 2005). Monitoring and control of time-sensitive products can be facilitated by the application of RFID technology.

2. RFID and supply chain

2.1 RFID technology overview

RFID technology is classified as a wireless Automatic Identification and Data Capture (AIDC) technology that can be applied to the identification and tracking of entities. An RFID device called RFID tag or transponders can be attached to a product as a means of identification. This tag contains an integrated circuit for storing information (including serial number, configuration instructions, activity history, etc.), modulating and demodulating a (RF) signal, and other specialized facilities. The circuit is attached to a miniature antenna within a set upon a label to permit attaching the tag to the desired physical object. The RFID tag transmits their data in response to an interrogation received from a read-write device called RFID reader or interrogators. This device decodes the tag signal and transfers the data to a computer through a cable or wireless connection. The tags and readers are designed with a specific operating frequency. Given the wireless communication between the RFID chip and the RFID reader, all data may be read from a distance. The reading range varies in accordance with the operating frequency, the size of the reader antenna, the orientation of the RFID tag towards the antenna, the tag position with respect to the antenna core, as well as with the tag type.
RFID tags come in a large variety of designs; they can be classified in many different ways and multiple criteria could be used. Thus, RFID tags can be categorized in accordance with the following criteria:

- power source
- operating frequency
- data storage
- memory size

Each of them is briefly presented below.
Tags use a variety of power sources:

- *active tags* - contain their own power source (a battery) that is used to run the microchip's circuitry and to broadcast a signal to a reader when prompted;
- *passive tags* – with no internal power source. Instead, they draw power from the reader;
- *semi-passive tags* - which use a battery to run the chip's circuitry, but communicate by drawing power from the reader.

Because the active and semi-passive tags contain more hardware than passive RFID tags, they are more expensive. Active and semi-passive tags are reserved for costly items that are read over greater distances. Yet, this flexibility does have a cost; active tags require more maintenance and have a limited life span due to onboard power supplies (5-10 years). Passive RFID tags have lower production costs, meaning that they can be applied to less expensive items. In fact, improved passive tag technology is responsible for the current wave of RFID adoption, as costs are reduced and operating ranges increase.

In some cases, active tags and tags with sensors can be used to monitor product quality. Thus tags can record temperature, humidity, pressure, shock/vibration, leakage and other data that could help determine the physical condition of the items monitored. For example, companies handling fresh produce such as vegetables can ensure product freshness by ensuring first expiry first out (FEFO) instead of the regular first in first out (FIFO).

A factor that also influences the cost of RFID tags is *data storage*. There are three storage types: read-write, read-only and WORM (write once, read many) (Gibson & Bonsor, 2005). A read-write tag's data can be added to or overwritten. Read-only tags are programmed with a serial number or other unalterable data when they were made and cannot be added to or overwritten. WORM tags can have a user-defined secure read-only area that may contain additional data (like another serial number) added once, but they cannot be overwritten.

Another tag classification criterion is *memory size*. Generally speaking, tag memory size can vary from 1 bit to 32 kbits and up. Active tags are able to retain more memory than passive tags. But more data on the tag leads to increased data reading time. One of the most challenging RFID implementation issues is the choice of the right memory capacity to support specified requirements.

Frequency is the leading factor that determines RFID range, resistance to interference and other performance attributes. RFID systems are available in a wide range of frequencies to suit various performance needs and they can be classified based on the band in which they operate. For the moment, there is no global frequency standard for RFID communication, bandwidth availability being regulated by telecommunications authorities in each country. RFID uses a range from 125 kilohertz (low frequency) to 5.3 gigahertz (microwave), generally divided in four distinct categories: Low Frequency (LF), High Frequency (HF), Ultra-High Frequency (UHF) and Microwave systems. Most commercial RFID systems operate at either the UHF band, between 859 and 960 MHz, or HF, at 13.56 MHz. Not all frequencies are available for use throughout the world and this is an important point to consider when planning supply chain applications. Most RFID technology used in warehousing and distribution operates at 13.56 MHz (HF), 860-930MHz (UHF) or the 2.45GHz (microwave) band. For material handling, logistics and supply chain applications RFID systems are concentrated in the UHF band and 13.56MHz.

The *reading range* of RFID systems is given by the maximum distance between the tag and the reader antenna that allows the reading of the information stored on the tag chip. The reading range varies from a few centimeters to tens of meters, depending on the frequency used, the power output, immediate physical environment and the directional sensitivity of the antenna. For read/write tags, the reading range typically exceeds the write range. HF range is limited to the near field only. Thus HF technology is used for short-range applications and can be read from up to about three meters; this means it cannot be used on cases and pallets where warehouses and distribution center logistics require longer range RFID operations. UHF technology provides a reading range of 20 meters or more. The

detection range of active tags is relatively large (up to 300 feet), whereas passive tags only operate at smaller distances (a few inches up to 30 feet).
The material composition of the tagged item and the contents of the items to be tagged can have a serious impact on the reading performance. Tag performance generally decreases with size, so it's advisable to use the largest size possible that fits the object. Longer ranges require larger tags, and it's a reality of physics that with longer ranges, the read rates are slower, and more reader power or more sensitive tags are needed. Extra range may be required if the application calls for reading a large number of tags moving very quickly past the antenna.
Given current tags costs, Byrne indicates that only medium to high value products should be tagged (Byrne, 2004). Industry is hoping that tag manufacturers can hit 5 cents per unit, and that is being regarded as a breakthrough level, and Gaughan sets the item/product cost delineation at least $15 (Gaughan, 2005).
RFID technology is emerging as a powerful and proven tool for streamlining production at manufacturing facilities of all sizes. As RFID is integral to the future of supply chain management and items tracking, it is important to examine RFID in detail and to compare its capabilities to an existing industry standard, the barcode.

2.2 RFID vs. barcode

RFID is similar to another AIDC technology, barcode technology. Conceptually, bar coding and RFID are quite similar. In fact, an RFID tag can be attached to a product as a means of identification, in much the same way as a barcode label. The two technologies differ in terms of the technology employed: barcode uses optical technology, while RFID uses radio technology. However, RFID tags have numerous advantages over barcodes.
The major advantage is that RFID has the capacity to store larger amount of information. Barcode is based on WORM (write once read many) technology, which means that once printed, a barcode cannot be modified. But an RFID tag can be read and written with a reader for thousands of times, acting as a portable database. In fact, RFID-enabled supply-chains can generate 10 to 100 times more information than traditional barcode technology.
Another advantage of RFID technology is that information gathering is faster than in the case of barcodes, while the read/write operations can be performed through different materials such us paper, plastic or wood, with the exception of metals.
RFID also allows easy, uninterrupted and upon-request access to the tag data. Unlike the barcode where identification is limited by line-of-sight, RFID technology requires neither a line of sight for identification, nor a straight-line alignment between the tags and readers. This means that packaging never needs to be opened to read a product tag. RFID tags are also sturdier than barcodes, allowing for use in adverse conditions (including exposure to dirt, outdoors, etc.), and tags can be affixed or embedded on the product packaging or inside the item. Barcodes are scanned one at a time, requiring much more time and effort to scan than RFID tags, when a large number of items are to be counted or tracked. The barcode is generally used to identify a product family, not the single item. The RFID tags can track items more precisely than traditional barcodes, and they can be read faster with less human intervention, thus allowing for more rapid product movement. Furthermore, by anti-collision mechanisms, several RFID tags in the field of a writer/reader can be addressed at the same time. For example, if a large amount of pallets are being unloaded into a warehouse, they can simply be crossed through docking doors attached with RFID readers instead of being unpacked and scanned manually.

Barcode presents some privacy and security issues. Although the data encoded on the barcode could be encrypted, there is no protection to prevent the barcode data from being copied and decrypted using commercial tools. Thus barcodes may be duplicated and attached to products. RFID tags allow more sophisticated forms of data protection and encryption than barcode. Each RFID tag has its own unique identity code or serial number from the manufacturer embedded on the tag. This number may never be modified making the tags counterfeit proof.

Barcodes are cheaper than RFID tags; a barcode label costs fractions of a penny instead of RFID tags cost that vary from 20 cents to a couple of dollars (for specialized tags). But, as time passes, it is estimated that RFID tags costs will decrease due to an increase in demands and lower costs from suppliers. On the other hand, barcode costs would are likely to remain the same because companies have already invested enough in the technology and its corresponding equipment.

Table 1 summarizes these aspects and provides a brief comparison between RFID and barcodes technologies (Vempati, 2004; ***b, 2007).

Characteristic	RFID	Barcode
Reads Per Second	40-200	1-2
Read Range	Up to 25 feet for passive RFID and up to 100's of feet or more for active RFID	Several inches
Read/Write	Yes	No
Anti-collision capabilities (simultaneously read capabilities)	Yes	No
Cost	More (>$.20)	Less (pennies)
Reusability	More	Less
Human Intervention	Less	More
Line of Site Required	No	Yes
Read Speed	Milliseconds	> second
Dirt Influence	No effect	Very high
Security	More	Less
Reader Interoperability	Limited, but growing	Yes

Table 1. RFID versus Barcodes

Speaking in enterprise terms, it is evident that the usage of RFID tags in a supply chain system enjoys considerable benefits (***c, 2007): high efficiency in collecting, managing, distributing and storing information on inventory, business processes, and security controls; increased productivity; products are processed at high speeds, so the time allotted to product scanning is considerably reduced; the time involved in product handling is reduced; inventory activities are simplified and data accuracy increases. Thus, various studies have proved that all inventory procedures may be performed faster than those involving barcodes (Davis & Luehlfing, 2004). Moreover, if one user gets near the products holding a mobile reading system, the handheld device will immediately collect and store data; product management is improved thanks to the re-programmable memory which also allows instant product location; customer services are considerably improved; RFID will allow receiving authorities to verify the security and authentication of shipped items.

RFID technology is not likely to replace barcodes in the near future. In fact, since barcode and RFID technology exchange data in different ways, nowadays the two technologies complete each other in real applications. They are both valuable in different situations, and can often be used together effectively for many purposes. In such a hybrid solution, a tag may be linked with a preprinted barcode.
But the differences in data exchange between the RFID and barcodes can help the user to decide where each technology can be most effective. The implementation of RFID technology will focus initially on pallets and crates containing products. Only when passive RFID tag prices are sufficiently low and adoption is more widespread, will the barcode be under threat in the retail industry. However, in the coming years, RFID tags and barcodes will still coexist.

2.3 ISO standards

The International Organization for Standardization (ISO) has developed RFID standards for automatic identification and item management that tried to solve the compatibility problems. This standard, known as the ISO 18000 series, deals with the air interface protocol (the way tags and readers communicate) for systems likely to be used to track goods in the supply chain. They cover the major frequencies used in RFID systems around the world. There are seven parts:
18000–1: Generic parameters for air interfaces for globally accepted frequencies
18000 - Part 2: Parameters for Air Interface Communications below 135 KHz (ISO standard for Low Frequency)
18000 - Part 3: Parameters for Air Interface Communications at 13.56 MHz (ISO standard for High Frequency)
18000 - Part 4: Parameters for Air Interface Communications at 2.45 GHz (ISO standard for Microwave Frequency)
18000 - Part 5: Parameters for Air Interface Communications at 5.8 GHz
18000 - Part 6: Parameters for Air Interface Communications at 860 – 930 MHz (ISO standard for UHF Frequency)
18000 – Part 7: Parameters for Air Interface Communications at 433.92 MHz.
ISO has also created standards that define how data is structured on the tag. For example, ISO 11784 and 11785 describe the structure and the information content of the codes stored in the tag for RF identification of animals.
There are also standards that deal with supply chain applications (i.e. how standards are used in different domains):

- ISO 17358 - Application Requirements, including Hierarchical Data Mapping
- ISO 17363 - Freight Containers
- ISO 17364 - Returnable Transport Items
- ISO 17365 - Transport Units
- ISO 17366 - Product Packaging
- ISO 17367 - Product Tagging (DoD)
- ISO 10374.2 - RFID Freight Container Identification

The usage of RFID to track items in open supply chains is relatively new and fewer standards have been finalized. For example, ISO has proposed standards for tracking 40-foot shipping containers, pallets, transport units, cases and unique items. These are at various stages in the approval process (***a, 2008).

2.4 RFID privacy & security

RFID data must be used in compliance with clear regulations concerning IT security as well as consumer and data protection (Heintz, 2005). A primary RFID security concern is the illicit tracking of RFID tags. Unauthorized readout of the RFID tag memory content has raised privacy concerns from both retailers and consumers. The issue of consumer privacy in RFID applications has received a great deal of attention from consumer groups and has garnered high visibility through the media. Therefore, it is necessary to provide counter measures which enhance consumer privacy and eliminate the concerns when consumer-sensitive data like pharmaceuticals are involved. In fact, RFID technology, when combined with a secure tag and data infrastructure, can assure both package authenticity and pedigree while creating new revenue opportunities.

A method of defense against unauthorized readers uses cryptography to prevent tag cloning. Thus, some tags use a form of "rolling code" scheme, wherein the tag identifier information changes after each scan, thus reducing the usefulness of observed responses. Nevertheless, cryptographically-enabled tags typically have dramatically higher cost and power requirements than simpler equivalents, and as a result, deployment of these tags is much more limited (***d, ****).

2.5 RFID applications

The RFID technology has been available for decades, but given the current significant lowering of tag costs, it is expected that their usage will be considerably increased. RFID allows the identification, location, tracking and monitoring of individual physical entities such as people, individual products or palleted goods. RFID may be viewed as a means of explicitly labeling objects to facilitate their "perception" by computing devices; thus, real-time information about these objects can be easily obtained from the factory, through shipping and warehousing, to the retail location (Finkenzeller, 2003). In fact, the RFID term is often used to describe the entire system of supply chain management using RFID, from the physical tags to the processing of information on electronic databases.

Almost all industries have used automatic identification (Auto-ID) in many applications: access and security systems, item tracking systems, inventory management and simplified checkout at retail stores. For example, automatic identification technology offers the potential to achieve inventory accuracy and thus reduce supply chain costs as well as the out-of-stock level. The relatively new technology, RFID upgrades the Auto-ID capabilities and enhances implementation in various industries with significantly hard and soft savings. Employed in a wide range of applications, RFID technology has become an indispensable asset.

RFID technology will benefit lots of industries and applications are constantly being developed and refined as the technology advances. The potential applications of RFID technology in supply chain are vast and refer to any organisation engaged in the production, movement or sale of physical goods. This includes retailers, distributors, logistics service providers, manufacturers and their entire supplier base, hospitals and pharmaceuticals companies, and the entire food chain. For example, the logistical tracking of goods will increase efficiency and will make available accessible supply chain transport and route information to everyone involved from the producers to the consumers. RFID tags in car sub-assemblies will make safety checks and recalls faster and easier. Tags in sub-sea structures like oil and gas pipelines will make maintenance and repair simpler. Hospitals will be able to maximise their return on assets by tracking the whereabouts of expensive and life-saving equipment at all times. The pharmaceutical industry will be able to reduce or even eliminate counterfeiting by giving each unit of dosage a unique EPC number. This will allow pharmaceutical data to be properly recorded. In fact the location of certain drugs will be made accessible to all supply chain partners; they will know the exact location of any drug and historical locations, the time spent for to transport it from one place to another, as well as the environmental storage conditions from its production to its usage.

Perhaps the most significant sign of transition to RFID was Wal-Mart's announcement in June 2003 of its intention to have top suppliers begin using RFID tags on pallets and cases by January 2005. In USA, the Department of Defense, Target, Best Buy, Albertson's, and others followed with their own RFID initiatives.

In retail industry, it is imperative that perishable products remain within a fixed temperature range across the entire supply chain. Temperature levels can be monitored in real-time by a temperature sensor connected as an additional device on an RFID tag attached, for example, to a shipping container, an individual product or a vehicle. In the not too distant future RFID tags will offer seamless product temperature records from point of manufacture to the time of purchase (Smith, 2005).

RFID, with its expected advantages, has currently been a major trend in many industries. Ranging from commercial to military uses, RFID technology is a modern resource which has not exhausted yet its applicability potential.

2.6 Integrating RFID technology into supply chain

An important application of RFID technology is supply chain management, where RFID helps close information gaps by enabling real-time supply chain visibility. By placing RFID tag on a product, users can track the product throughout the supply chain- from the manufacturer all the way to the customer.

In most cases the RFID tag can be written and rewritten so that the information in the tag doesn't remain static. For instance, at first, the tag may only contain manufacturing information; later on, additional information from the distributor may be added. RFID can enable the vision of real-time, multidimensional coordination for all the players in the supply chain (Grackin, 2004).

In fact, RFID is considered the most intelligent technology for managing and collecting product data or tracking it as it moves through the supply chain.

Today, companies looking to adopt RFID have to deal with three key challenges:

1. *RFID Hardware* - Selecting tags, readers, and antennas; placing RFID tags on the products; placing and configuring readers and antennas in the stores, warehouses, and other locations.
2. *Software Infrastructure* - Capturing and managing data from the RFID readers, integrating the data into different levels of enterprise information systems, and sharing data with trading partners for business collaboration.
3. *Evolving business processes* - Supporting finer granularity, more real-time product data, automating supply chain execution, and developing new business processes for exploiting RFID technology.

The non-line of sight capability of RFID makes it a perfect supply chain technology (Gibson & Bonsor, 2005). Passive tags operating under the ultra high frequency (UHF) band are common to supply chain applications because tag costs are low and the read range and rate is adequate (Tajima, 2007). In the case of supply chain management (SCM), RFID is not just about the identification of an individual pallet, case or item but about the relationships between objects, between organizations, between space and time. RFID is about process level change that can streamline business-to-business (B2B) operations and bring about major changes to organizational policies, culture, performance and structure (Lefebvre et al., 2007).

2.7 Electronic Product Code (EPC)

The use of RFID in supply-chain application is based on EPC. The EPC was conceived as a means to uniquely identify all physical entities. In fact EPC is a numbering scheme that provides unique identification for physical objects, assemblies, components and systems. Information is not stored in the code, but serves only as a reference to on-line – or Internet-based – information (Brock, 2007).

EPC ID numbers assigned to an entity are used with RFID tags in the same way that UPC (Universal Product Code) numbers are used with barcodes. In fact, the EPC is considered the electronic equivalent of the UPC barcode and a possible successor to the barcode. An RFID tag stores a single EPC number in its memory, just as a barcode holds a UPC number. Barcodes have been in use for over 30 years and have become an integral part of product identification. Product identification is performed differently by RFID tags and conventional barcodes. A UPC refers to an object class or generic category of products. For example, when a barcode gets scanned at a store checkout counter, it will return product information such as product name and price; the same information is valid for every product in the same category. An EPC refers to a specific instance of product, allowing the unique identification of any tagged item. Thus, an EPC makes it possible to automatically track individual items, for example, products from the manufacturing line that reach the shelves.

The current version of the specification comprises encodings for EPCs of 96 to 202 bits length. But EPC codes are typically 96 bits in length (24 four-bit characters) divided into four fixed length components (header, Domain Manager Number, Object Class Number and Serial Number), each part containing specific information.

The first section consists of an 8-bit *header* indicating the number, type and length of subsequent data sections. Practically this header provides a real extensibility for future, unanticipated data requirements.

The second section of EPC code identifies the company or entity responsible for maintaining the subsequent codes. This entity is known as the *EPC manager*. Its responsibility is to maintain both object type codes and serial numbers in their domain. The EPC manager section covers a 28-bit section, encoding a maximum of 228 = 268,435,456 companies.

The next section of the EPC code, called the *object class*, occupies the next 24-bits. The object class number is used to identify a class of product, meaning a group of products sharing similar characteristics. When applied to retail products, the object class is often considered the skew or stock keeping unit (SKU), lot number or any other object-grouping scheme considered by the EPC manager. For each organization is allowed more than 16 million object types, so this section could encode all the current UPC SKUs and many other object classes. This allows expanding beyond retail applications into general supply chain.

The final section of the EPC code encodes a unique object identification number that serves to identify a particular item belonging to the specified object class. It is the managing entity responsibility to assign unique serial number for every instance within each object class. For

all objects of a similar type, the *EPC serial number* provides 36-bits, or 236 = 68,719,476,736, unique identifiers. Together with the product code, this provides $1.1x10^{18}$ unique item numbers for each company – currently beyond the range of all manufactured products.
In January 2008, 1347 companies from different industries were EPCglobal subscribers (Schmitt & Michahelles, 2008).

2.8 Benefits

The RFID technology will bring benefits to a wide range of industries, but one of the main domains of RFID adoption has been the supply chain for retail sector. RFID technology can help improve efficiency and visibility; it will cut management costs, influence considerably the production of higher quality goods and enhance the utilization of products; it will also reduce shrinkage and counterfeiting, and increase sales by reducing out-of-stocks.
For example, RFID technology has the potential to:

- reduce the time taken to re-order shipments;
- minimize warehouse discrepancies by validating the accuracy of deliveries and shipments;
- reduce product shrinkage and theft;
- improved tracking of pallets, cases and individual products;
- provide better planning and optimization of inventory and reusable products;
- allow more efficient use of labor by automation data handling and reading multiple products;
- monitor expiration dates of an organization's complete inventory list;
- automate supplier receiving and billing procedures;
- reduce manual entry errors (e.g. data typing mistakes);
- allow more efficient transport and distribution;
- allow information sharing to better collaboration between partners;
- increase visibility and lead to better decision making capabilities.

The main benefit of RFID integrating in supply chain process is to allow the constantly monitoring and improving the whole system by using all the available data.
But to maximize competitive advantage in a supply chain context, RFID needs to be used by multiple companies to do all sorts of things, creating widespread advantages for all supply chain participants (Coltman et al., 2008). RFID is expected to be worth billions of dollars in new investments. According to IDTechEx, a leading market research and advisory firm, the RFID market will increase from US$4.96 billion in 2007 to US$26.88 billion in 2017 (Das & Haropp, 2007).

2.9 Obstacles

The path to RFID technology integration in supply chain is not without some obstacles and they can be enumerated briefly:

- tags, infrastructure and implementation costs are still high;
- unclear cost/benefit sharing models;
- the integration in existing systems;
- readers can't always read all the cases on a pallet;
- technology incompatibilities: inability for a single reader to read tags from multiple frequencies;
- standards are in a state of flux;

- some of RFID standards are not globally accepted; different standards coexist in parallel and there are a lot of them incompatible one with other;
- tag reliability can be impacted by humidity, liquid or metals;
- radio interference can upset even the best-laid plans: in close proximity to one another, tags may seriously affect performance. Antenna placement and orientation can also negatively impact the RFID solution;
- end-users lack real RFID knowledge.

3. RFID@B2B

3.1 Business to Business (B2B)

Twenty-first century technology has revolutionized the way companies do business with each other and with their customers. The Internet has connected companies around the world and changed the global economy as a whole. Indeed, the Internet has emerged as a most cost-effective means of driving supply chain integration. A new concept has been devised: e-business - as the marriage between the Internet and supply chain integration (Johnson & Whang, 2002). Following today’s economic globalization, e-business has become a necessity for companies to remain competitive. As one major component of e-business, business to business (B2B or B-to-B) includes all applications intended to enable or improve relationships within firms and between two or more companies (***e, ****). In fact, B2B is commonly used to describe any electronic business transaction occurring between two separate business entities. This includes the exchange of both products and service. Examples of exchanged products and services might include the selling of raw material inputs from one firm to another, the sale of capital equipment, the purchasing of commercial insurance or the contracting of one firm with another for the procurement of accounting services. B2B is all about product and materials procurement and the supply chain is the vehicle through which business-to-business is ultimately achieved. In fact, a B2B infrastructure links buyers, suppliers and logistics service providers into a global trading network. One of the new concepts that will further define traditional B2B Internet commerce is RFID (Gerhards, 2006).

3.2 Integrating RFID technologies in B2B applications for enterprise supply chain

In the case of B2B, RFID is supposed to benefit not only the identification of individual pallets, cases or items, but also the relationships between/among objects, between/among organizations, between space and time. RFID is about process level change that can streamline business-to-business operations and bring about major changes to organizational policies, culture, performance and structure (Lefebvre et al., 2006). In fact, according to S.F. Wamba et al. (Wamba et al., 2007), RFID technology and the EPC are enablers of intelligent B2B e-commerce supply chain management.

Our research team has developed an RFID_B2B integrated system which combines the advantages of B2B with those of RFID technology and which presents itself as a viable solution for the problems raised by globalization. The software system deals with business relations between corporations, big companies and groups of companies, in order to optimize the flow of materials among them and the supply chain management inside every company. To identify both parts and finite products, our system uses passive 13.56 MHz tags. Unique IDs are used to control and trace every part of a finite product. The RFID_B2B system could be tailored to the diverse needs of the companies and the different roles of employees in each company. If this system is embraced by the entire supply chain management, final consumers will be able to follow the entire production chain of a finite product. And this is possible if the traceability information is memorized on each tag attached to some part of the final product.

The RFID_B2B system architecture is flexible and easily extensible. The research team chooses to design a layered architecture arranged in such a way that the lower layers support and enable the upper layers. This architecture has some advantages: divide the complex system into several more manageable components, allow different groups to work on different layers concurrently etc. The RFID_B2B system is structured on three levels: the corporation level, the local level and data collection level at the material control departments (Figure 1).

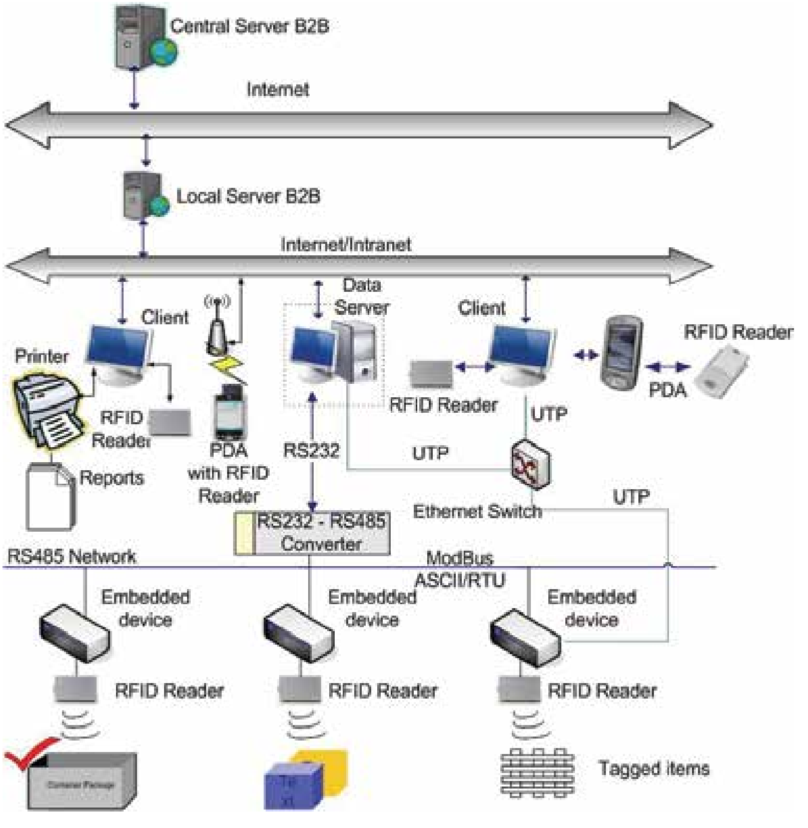


Fig. 1. The RFID_B2B system architecture

At the platform's *corporation level,* the following achievements have been made:

- services are offered to ensure the support for concluding contracts, along with the agreements, the conventions at the board level of the corporations, the firm groups or the representatives of the firm groups;
- informational management of the group/corporation enterprises, ensuring supervision of the material flows (with effects on the establishing of supply and sales strategies), as well as of the good functioning of the internal network of the group/ corporation;

- supply of reports regarding the current activities within the group/ corporation, including alarm in case of generation of specified events.

The following are provided at the *local level* or at the enterprise level:

- proper administration of the received, sent, defective, repaired, returned entities at the enterprise level;
- access to the company servers network, as well as communication management along the supply- sales main chain, providing the opportunity to manage and access the information referring to the route followed by materials, assemblies and finite products;
- coordination of the materials/assemblies flow in order to ensure adequate distribution to corresponding departments, as well as to deliver the order to the gates in departments;
- documents delivery for controlling the production, materials, finite products, assemblies, including those in the service department.

Different applications of RFID are implemented in the *data collection level* in order to write and read the data from the tags attached to the materials, assemblies and finite products. At this level, the communication is wired or wireless.

The integrated system comprises the following elements:

- an IBM-PC compatible computer which runs an OPC (OLE for Process Control) server with two main components: communication and data acquisition;
- an IBM-PC compatible computer which runs an OPC dedicated client; in fact one and the same computer may be used to run both the server and the dedicated client;
- a network comprising several low-resource embedded devices that have attached low-cost RFID readers; these device processes the local data and are capable of connecting and controlling other devices;
- PDA devices with attached RFID readers;
- an IBM-PC compatible computer which runs the local B2B server (Giza & Cerlinca, 2007);
- an IBM-PC compatible computer which runs the central B2B server (Giza & Cerlinca, 2007).

3.3 Improving supply chain management with RFID mobility solution

Enterprise supply chain systems can offer access to desired tasks (for example, inventory management, demand projections, production planning) from handheld computers, such as Pocket PCs.

Inefficient manual data-handling process in supply-chain management is a common problem with serious repercussions that affected the whole supply-chain, product traceability etc. There are needed solutions that could automate data handling.

Integrating a mobile solution in an RFID-based supply chain system is the answer (Cerlinca et al., 2008). Handheld computers can perform many of the tasks that are commonly executed on a desktop computer or standard laptop. In fact, these devices can be used to extend the capability and reach of an existing information infrastructure by enabling workers to collect, access and analyze desired data at any time from anywhere. Thus, the RFID labeled product can be read and tracked through the entire supply chain with handheld mobile devices (for example, PDA-Personal Digital Assistant) endowed with RFID readers. The collected data is stored on the mobile device using a mobile database software technology. This data is transferred between the mobile device database and PC database whenever it's necessary. For security reasons, the data stored in the handheld device memory can be transferred to a PC, for instance, when the handheld device is placed in a docking station. But, for integration in our complex enterprise application, encrypted data can be securely transferred across any kind of Internet-connected network. The users can also set up a virtual private network. These solutions are less expansive than mobile communications infrastructure.

The visual space on the handheld screen is far too small; the low display resolution and small display screen have inhibited information to be displayed completely and clearly. But the windows of RFID-based mobile application can be adapted to display a plurality of RFID tag information.

With an efficient RFID-based supply chain solution with integrated mobile support, the companies can reduce errors and cut costs. Implementing a mobile solution helps companies improve efficiency, extending the power of enterprise computing to new processes, people and places. Employees are more productive and businesses are more competitive.

3.4 RFID_B2B system benefits

The presented system offers a high degree of flexibility and helps companies of all sizes enable their customers to do business on demand — when they want, where they want and how they want. Other system benefits are:

- Assures realtime inventories so the users can always receive accurate, up-to-date inventory information;
- Offers the possibilities to share meaningful data with supply chain partners;
- Permits strengthening customer and partner relationships with collaboration;
- Speeds and simplifies the deployment and management of e-commerce sites;
- Maximizes performance, scalability and adaptability of partners systems;
- Provides rich, ready capabilities for products catalog and content management;
- Permits a greater visibility through realtime product updates, availability and pricing information;
- Offers personalization capabilities.

3.5 Future improvements

With the growing number of B2B sites available through Internet, a useful addition to the RFID_B2B system would be an intelligent software agent for information gathering. The agent will be able to perform semantic query optimization and to offer data mining facilities. It will dynamically plan for alternative information source when a source or a B2B site goes down. This agent will organize the results and display them in an easily interpreted manner to the user. To face the new global market and to provide an effective collaborative relationship between trading partners, an environment to support the semantic integration could help. Another useful feature would be a special section that enhances the management of production planning to ensure good deliveries and productive efficiencies.

Transition to B4B (Business for Business) – next evolution in B2B communication (Jones, 2007) is the following aspect that might be taken into account as future direction for system development.

The Internet, electronic business and RFID technology are changing the history of supply chains, and modifying the way that consumers select, purchase, and use products and services of partners. The result will be the surfacing of new business-to-business supply chains that are consumer-focused rather than product-focused.

4. Conclusion

This chapter helps to improve readers understanding of the RFID and EPC potential for business processes. RFID technology is classified as a wireless AIDC technology that uses digital data encoded into a radio tag embedding a microchip with an antenna. The data stored on the tag is collected by a reader using radio waves. There are a large variety of RFID tags designs; they also have many different functional characteristics such as power source, carrier frequency, read range, data storage capacity, operational life, cost etc.

RFID has immediate benefits over barcodes. Thus RFID tags are an improvement over barcodes because the tags have read and write capabilities. The data stored on RFID tags can be changed, updated and locked. RFID technology offers a better way to track items with minimal human intervention, for stocking and marketing purposes. Benefits come in the form of inventory, shrinkage and labor reduction on the one hand, and sales increasing due to reducing out-of-stock and getting real-time demand information on the other hand.

RFID technology represents one of a number of possible solutions to enhance supply chain. It is therefore important to do a cost-benefit analysis to evaluate each alternative solution. The majority of the costs of integrating RFID in supply chain application come from IT, tags, hardware and services. But due to the actual relative high cost of integrating RFID technology, each company needs to evaluate its own business processes to determine where and if RFID can be applied (incorporated) to provide substantial business benefits. Thus, all RFID solutions have to evaluate different performance and cost factors, including the operating environment, on-tag memory storage, and signal transmission restrictions. Each of these issues has significant cost impacts on both tags and readers. The costs of RFID readers have already fallen to a considerable extent. The cost of tags is expected to decrease over time and as quantities increase. Passive tags are undoubtedly less expensive than active tags and most companies are focusing on passive tags. The different studies proves that at a lower quantity the barcode is the cheapest alternative for supply chain, but, as quantity of product increase the optimum choice is RFID. In some applications, RFID and barcodes system will still coexist and this redundancy cost must be considered. However, to realize maximum return on investment (ROI) for RFID integration, the enterprises need to leverage their information architecture strategically.

Any industrial domain may benefit from RFID technology, and the number of applications is on the rise. Thus, RFID technology is applied in a vast area of industrial, commercial and military domains, including manufacturing and logistics, retail, animal tracking, etc. This chapter focuses on how RFID technology can be used to solve problems faced by supply chain. In fact, RFID has the potential to radically change the entire supply chain by improving inventory management, asset visibility, and interoperability in an end-to-end integrated environment. The ability to track, at item level, material flows among partners until they reach the consumer, while maintaining the data accuracy advantages of various types of automatic identification technology (AIT), is the perfect solution to the many issues of enterprises in the past. RFID technology permits the unique identification of each container, pallet, case and item to be manufactured, shipped and sold, thus allowing an increased visibility throughout the supply chain. Thus the RFID has the potential of helping retailers provide the right product at the right place at the right time, allowing maximizing sales and profits.

The EPC represents a low-cost method of tracking products using RFID technology. The EPC is a short, simple and extensible code designed for the unique identification of individual physical objects such as spare parts and whole products; the identification process may be extended to cover further information related to container, packages, shipments or manufacturers. The EPC can provide up to 268,435,456 companies identifiers, more than 16 million object types and 1.1 x 10^{18} unique item numbers for each company.

In a global market where change is continuous, companies require tools that allow them to respond quickly to new opportunities. The presented RFID_B2B system can be considered as a viable solution for potential problems raised by globalization process, contributing to a significantly more efficient business process. Thus, the presented system helps small, medium and enterprise organizations to improve productivity and provide better service to their customers by providing a flexible solution for all of a company's B2B needs. Many mobile systems already employed in supply chain management have proved their importance through significant return on investment. Not only can they extend corporate data outwards to mobile devices for viewing and querying, but users can use any mobile device endowed with an RFID reader for data collection. In this way, manual entry data has been eliminated. Moreover, users can read the tags wherever the items are placed, which enables a more flexible storage environment and an efficiency increase of supply chains. The RFID_B2B system is so adaptable that many types of businesses can use it and allows enabling new business opportunities and growth. Using the developed system may help customers sharpen data accuracy, process supply chain transactions faster, and improve supply chain and inventory management. Given slim profit margins, companies are looking for ways to save on costs while remaining globally competitive. RFID@B2B may be their answer.

5. References

Brock, D. (2003), The Electronic Product Code (EPC) as a Meta Code, *Massachusetts Institute of Technology*, Available at: http://www.autoidlabs.org/uploads/media/MIT-AUTOID-WH020.pdf

Byrne, P. M. (2004), Supply Chain Rfid: Lessons from a Leader, *Logistics Management*, 43(5), 2004, pp. 31-32

Caton, M. (2004), RFID Reshapes Supply Chain Management, http://www.eweek.com/c/a/Mobile-and-Wireless/RFID-Reshapes-Supply-Chain-Management/

Cerlinca, T.; Turcu, C.E. & Cerlinca, M. (2008), Integrating Mobile Applications into RFID Based B2B Systems, *22nd International Conference on Advanced Information Networking and Applications*, 25-28 March, 2008, Gino-wan, Okinawa, Japan, pp. 1341-1345, ISSB/ISBN: 978-0-7695-3096-3

Coltman, T.; Gadh, R. & Michael, K. (2008), RFID and Supply Chain Management: Introduction to the Special Issue, *Journal of Theoretical and Applied Electronic Commerce Research*, (3)1, pp. iii-vi, Available at: :http://works.bepress.com/kmichael/27

Das, R. & Harrop, P. (2007), Complete RFID analysis and forecasts 2007-*2017, RFID Forecasts, Players & Opportunities 2007*-2017, IDTechEx Available at: www.idtechex.com/products/en/view.asp?productcategoryid=119

Davis, H. & Luehlfing, M. (2004), Radio Frequency Identification: The Wave of the Future, *Journal of Accountancy*, Online Issues, Available at: http://www.aicpa.org/PUBS/JOFA/nov2004/davis.htm

Finkenzeller, K. (2003), RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, Second Edition, *John Wiley & Sons*

Fleisch, E. & Tellkamp, C. (2003), The Impact of Inventory Inaccuracy on Retail Supply Chain Performance: A Simulation Study, *Auto-ID Center, University of St. Gallen*, Institute of Technology Management, Switzerland, Available at http://www.autoidlabs.org/single-view/dir/article/6/206/page.html

Gaughan, D. (2005), RFID Hype Collides with Reality, *Supply Chain Management Review*, 9(2), 2005, pp. 14-15

Gerhards, E. (2006), Latest developments in the field of B2B e-markets, *B2B group Meeting*, Available at: http://ec.europa.eu/enterprise/ict/studies/b2b_2006_05_10_dev.pdf

Gibson, C. & Bonsor, K. (2007), How RFID Works, Available at: http://electronics.howstuffworks.com/ rfid3.htm

Giza, F., Cerlinca, T., A N-Tier Application Architecture for B2B Systems, *Distributed Systems*, Vol: 5, September, 2007, ISSN/ISBN: 1842-6808

Grackin, A. (2004), RFID for Consumer/Retail Supply Chains, *CHAINLINK Research*, http://prescient-news.com/0905/RFID_for_Consumer-Retail_Supply_Chains.pdf

Heintz, B. (2005), RFID White Paper Technology, Systems, and Applications, *Bitom*, Germany, Available at: http://www.rfidconsultation.eu/docs/ficheiros/White_Paper_RFID_english_12_12_2005_final.pdf

Johnson, M. & Whang, S. (2002), e-Business and Supply Chain Management: An Overview and Framework, *Production and Operations Management*, Vol. 11, No. 4, 2002, Available at: http://mba.tuck.dartmouth.edu/ digital/ Research/ Academic Publications/POMSArticle.pdf

Jones, M. (2007), Business 4 Business – ASN and Data Accuracy Drive Higher Order Functionality and Redefinition of Retail B2B Communications, *VCF & GXS*, November 2007, Available at: http://www.gxs.com/forms/0711_B4B_wp_VCF.htm

Lefebvre, L.; Lefebvre, E.; Bendavid, Y.; Fosso, S. & Boeck, H. (2006), RFID as an enabler of B-to-B e-Commerce and its impact on business processes: a pilot study of a supply chain in the retail industry, *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, Hawaii, 2006, pp. 1-10

Raman, A.; DeHoratius, N. & Ton, Z. (2001), Execution: The Missing Link in Retail Operations, *California Management Review*, 43, 2001, pp. 136–52.

Rangarajan, T.S.; Vijaykumar, A. & Subramaniam, S. (2005), Stop Getting Strangled by your Supply Chain. Enhancing SCM using RFID, *Tata Consultancy Services Limited*, 2005

Schmitt, P. & Michahelles, F. (2008), Economic Impact of RFID Report, *Building Radio frequency Identification for the Global Environment*, April 2008

Smith, J.N. (2005), Specialized Logistics for a Longer Perishable Supply Chain, *World Trade Magazine*, Available at: www.pressroom.ups.com/staticfiles/articles/456.pdf

Tajima, M. (2007) Strategic Value of RFID in Supply Chain Management, *Journal of Purchasing and Supply Management*, vol 13, 2007, pp. 261-273

Vempati, S. (2004), RFID Architecture Strategy, *Infosys Technologies Limited*, http://www.infosys.com/rfid/Infosys_White_Paper_on_RFID_Architecture_Strategy.pdf

Wamba, F.; Lefebvre, L. & Lefebvre, E. (2007), Integrating RFID Technology and EPC Network into a B2B Retail Supply Chain: a Step Toward Intelligent Business Processes, *Journal of Technology, Management & Innovation*, , Vol. 2, No. 2, 2007, pp. 114-124, ISSN 0718-2724

***a (2008), A Summary of RFID Standards, *RFID Journal*, Available at: http://www.rfidjournal.com/article/view/1335/1/129

***b, (2007), RFID Is A Hot Topic, *Nutech Systems,* Available at : http://www.nutechsystems.com/news_rfid.html

***c, (2007), Advantages of RFID, *ActiveWave Inc.*, http://www.activewaveinc.com/technology_rfid_advantage.html

***d, (****), Radio-frequency identification, *Wikipedia*, Available at: http://en.wikipedia.org/wiki/RFID

***e, (****), E-business Definitions, http://www.cheshirehenbury.com/ebusiness/ebdefinitions.html

19

RFID Technology, Security Vulnerabilities, and Countermeasures

Qinghan Xiao[1], Thomas Gibbons[2] and Hervé Lebrun[2]
[1]*Defence Research and Development Canada – Ottawa*
[2]*Canadian Operational Support Command*
Canada

1. Introduction

Radio frequency identification (RFID) is a means of automatic identification that uses radio waves to detect, track, identify, and thus manage a variety of objects. The purpose of an RFID system is transmitting data from a portable device, called a tag, to an RFID reader to execute a particular application based on the tag provided identification or location information (Graafstra, 2006; O' Brien, 2006).

RFID technology has been around for about 60 years and was originally developed for improving warfare technologies. The first application was developed by Britain as the Identify Friend or Foe (IFF) system, which was used to distinguish friendly aircraft from enemy aircraft during World War II (Landt, 2001). The second era of RFID technology began with the commercial activities in the 1960s. An early application was the development of electronic article surveillance (EAS) equipment to counter theft in retail stores. EAS as an early forerunner to RFID uses a '1-bit' signal to represent the presence or absence of a tag (Landt, 2005). The third era of RFID technology started in 1999, when the Auto-ID Centre was established at MIT to investigate new ways for improving bar code technology and implementing RFID technology in various commercial applications. The 1990's were a significant decade for RFID because of increased commercialization of RFID systems and the standardization activities on RFID technologies. Electronic toll collection systems were widely deployed in the United States; RFID tags were affixed to over 3 million rail cars in North America; and the International Organization for Standardization (ISO) developed several standards in the RFID field, including, for example, the ISO 18000 series of standards that define the air interfaces, collision detection mechanisms, and communication protocols for different frequency bands (Knospe & Pohl, 2004). In the 21st century, with the development of RFID standards, decreasing prices, and mandates from large organizations such as Wal-Mart and the U.S. Department of Defense (DoD), RFID has become "the first important technology of the twenty-first century" (Garfinkel & Rosenberg, 2005).

An RFID system has three key components: the tag, the reader, and the backend system. RFID tags, also known as transponders, are identification devices that are attached to objects. Each tag typically consists of an antenna constructed from a small coil of wires; a microchip used to store information electronically about the object (e.g. a vehicle or a container); and encapsulating material to enclose the chip and the coil. Like there are various types of barcode, RFID tags are available with different memory sizes and encoding options. However, an RFID tag offers the capability to store a unique serial number and product information for each item, not just the class of the items. The tag can also incorporate sensors to record temperature, shock, or humidity, for example, providing the ability to track and report on an object's environmental characteristics dynamically.

An RFID reader, also called an interrogator or scanner, is the device used to communicate with the RFID tag. It emits RF signals to, and receives radio waves from, the tag via an antenna or antennas. The reader converts the received radio waves into digital information that is usually passed to a backend system. Readers, either as stationary or handheld devices, consist of a transmitter, receiver, antenna, microprocessor, controller, memory and power source.

A backend system, sometimes referred to as an online database, is needed to collect, filter, process, and manage the RFID data. The backend stores complete records of product information, tracking logs, and key management information associated with the RFID tags. It is critical for an RFID application to perform data collection, data management, and data analysis accurately and efficiently.

There are various areas in which RFID technology has been implemented, including the following significant applications.

- RFID electronic toll collection systems identify vehicles mounted with RFID transponders and automatically deduct toll fees electronically without impeding traffic flow.
- Animal RFID implant tags have been used to identify and track animals and obtain information about their owners. Combined with GPS, it is possible to perform around-the-clock surveillance of individual animals and fish in the wild.
- RFID book tracking is a hot topic in the library community for use in managing extensive collections of books, manuscripts, and rare items, as well as offering self checkout and protection against theft.
- Healthcare providers are considering the use of RFID technology to improve the ability to accurately identify and track patients, hospital staff, medical equipment, and blood products. A number of case studies have demonstrated that not only does RFID technology make treatment safer and more efficient, but it also has the side benefits of preventing identity theft and reducing paper work, both of which cut costs. However, a newly published research study revealed that RFID systems in hospitals might cause critical care medical equipment to malfunction (van der Togt et al., 2008).
- The pharmaceutical industry deploys RFID technology to track drugs, reduce inventory cost, and prevent counterfeiting and theft. The administrator for the U.S. Centers for Medicare and Medicaid Services, Mark McClellan, called RFID "the most promising technology" for dealing with drug-counterfeiting problems (Whiting 2004).
- Access control has been among the most common applications of RFID technology because RFID badges provide many advantages over traditional access control badges, including fast access, durability without mechanical wear, and a superior ability to protect data on the card. In addition to traditional applications such as building access, RFID access cards have been used in less traditional applications such as ski passes, metro passes, and toll gates.
- The e-passport is the next generation of passport, which is equipped with an embedded RFID chip to store digital information and biometric data of the passport holder. The objective is to provide a trusted document to reduce fraud, make immigration control faster, and enhance the level of security.

- Supply chains are the biggest beneficiary of the RFID technology. The use of RFID in supply chains makes it possible to provide instant inventory management, increase asset visibility, track shipments, trace recalled products, and prevent theft.

Although RFID technology has been around for more than half a century, only recently have RFID security and privacy issues begun to attract attention from both academic and corporate research communities. In a research survey (Juels, 2005a), Juels provides an excellent overview of various RFID security and privacy concerns. In particular, when dealing with passive RFID tags, the author suggests that there is a need to divide the read range of a tag into four different ranges as follows:

- Nominal read range − the standard operating range of a tag under normal intended use
- Rogue scanning range − the read range of a tag when using a sensitive reader equipped with a powerful antenna or an antenna array and/or a higher signal transmission power
- Tag-to-reader eavesdropping range − the range that another reader can eavesdrop tag emissions without powering the tag itself, which can even be larger than the rogue scanning range
- Reader-to-tag eavesdropping range − the range that another reader can capture the signal sent by the reader to the tag, which is larger than any of the above ranges

It is necessary to point out that in addition to the above ranges there exists another range — detection range. This is the range from which the presence of a tag or a reader can be detected without the need to be able to send or capture information. We are carrying out a proof-of-concept study to examine whether an RFID system can be attacked by detecting the presence of the tag and reader communication.

From an information security point of view, Knospe and Pohl considered the RFID communication model to be similar in nature to the TCP/IP networking model used for computer networks. Their model consists of an application layer, a data link layer, and a physical layer for both the RFID reader and tag. They define RFID security from information security principles of confidentiality, integrity, availability, authenticity, and anonymity perspectives (Knospe & Pohl, 2004). In addition, RFID systems have their own vulnerabilities and security threats that are separate from the network model (Xiao et al., 2007). The security and privacy issues of RFID systems have been reviewed and evaluated in several studies (Ranasinghe & Cole, 2006; Rieback et al., 2006a; Aragones-Vilella et al., 2007; Rotter, 2008). The threats can be categorized based on their point of attack: the tag, the reader, or the air interface between the tag and the reader (Lieshout et al., 2007).

In recognizing the potential risks when deploying RFID technology, government agencies have played important roles in closely collaborating with industry groups and academia. In Germany, for example, the Federal Office for Information Security conducted a study about RFID security aspects to help German companies to understand security and privacy threats, such as eavesdropping, unauthorized reading of data, cloning, and tracking of people, along with possible protection strategies (Oertel et al., 2005). To meet requirements of the Federal Information and Security Management Act of 2002, the US National Institute of Standards and Technology (NIST) has published guidelines and a set of best practices for the use of radio frequency technology by federal agencies and private organizations. The guidelines focus specifically on the use of RFID technologies for asset management, tracking, matching, and process and supply chain control. NIST recommends the use of firewalls between RFID databases and an organization's IT systems. It also advises the use of encryption of RFID signals, authentication to identify approved users, and shielding of RFID tags to prevent unauthorized skimming of information (Karygiannis et al., 2007). The U.S. Department of Defense has initiated the DoD RFID security taxonomy through the Office of the Assistant Secretary of Defense (OASD) for Networks and Information Integration. Three areas of concern are being addressed: network-based risks, mission assurance risks, and order of battle risks. Since risks generally increase with system complexity, eleven high-level RFID security vulnerabilities have been identified, which include common threats, such as unauthorized reading of tag data and leaking tags' electronic information, and special threats specific to military applications, such as using RFID tags as trigger devices for explosives and using RFID readers as platforms for attack (Norton, 2006).

There are numerous publications focusing on RFID privacy issues. One of the major applications of RFID technology is for tracking and tracing of objects. However, the technology becomes a major privacy threat if it is used to track people (Thornton et al., 2006). Consumers are most afraid of being tracked without their consent or knowledge by RFID tags that are ubiquitously hidden in clothing and other consumer items (Ayoade, 2007). The privacy threats in RFID systems can be categorized into two classes: data privacy and location privacy (Oertel et al., 2005; Kim et al., 2006). The threats to data privacy involve discovering personal information stored on the tag and/or in the associated database, while the threats to location privacy comprise the information about a person's current location and past movement through a tag ID associated with that person (Langheinrich, 2007a). Both of these types of threats need to be addressed because the information could be used to profile the victims' preferences, movement, and/or social network. A survey has been conducted to review up-to-date RFID privacy approaches and their attributes (Langheinrich, 2007b). Governments have paid attention to RFID privacy challenges and may regulate RFID use to address privacy concerns. For example, the Ontario Commissioner for Information and Privacy released a set of guidelines to address privacy issues regarding use of item-level RFID technology in the retail/commercial sector. The guidelines were aimed at promoting RFID technology by addressing concerns about the potential threat to privacy (Cavoukian, 2006).

The Canadian Forces (CF) needs a lot of materiel from consumable logistics items to tanks in military operations. To deal with challenges of the visibility, tracking and traceability of its logistics assets, the Canadian Operational Support Command (CANOSCOM) has implemented RFID technology to provide effective and efficient support to CF operations at home and abroad. Defence R&D Canada – Ottawa (DRDC Ottawa), sponsored and funded by CANOSCOM, has been working on RFID security issues, such as the analysis of security threats and the identification of appropriate countermeasures. This chapter is based primarily on the results of our research, previous publications, and the currently available literature. We first present an overview of RFID technology with detailed exposition of the basic components and the essentials of RFID systems. Next, we analyze underlying vulnerabilities and security threats that exist in the RFID system. Then, we propose possible countermeasures to defeat the discussed attacks. Some case studies are presented to illustrate the real attacks. Finally, we conclude the chapter with a discussion of possible future research directions. Through identifying the common vulnerabilities and threats to RFID systems and providing possible countermeasures to resolve the security issues, the objective of this chapter is to provide information and defence techniques against the potential attacks.

2. Brief introduction to RFID technology

RFID is an emerging technology that uses radio waves as the means to identify items or objects. In order to analyze security and privacy issues, it is necessary to give a brief introduction to the basic components of RFID systems. As shown in Figure 1, a typical RFID system contains one or more RFID tags, a reader, and a backend system.

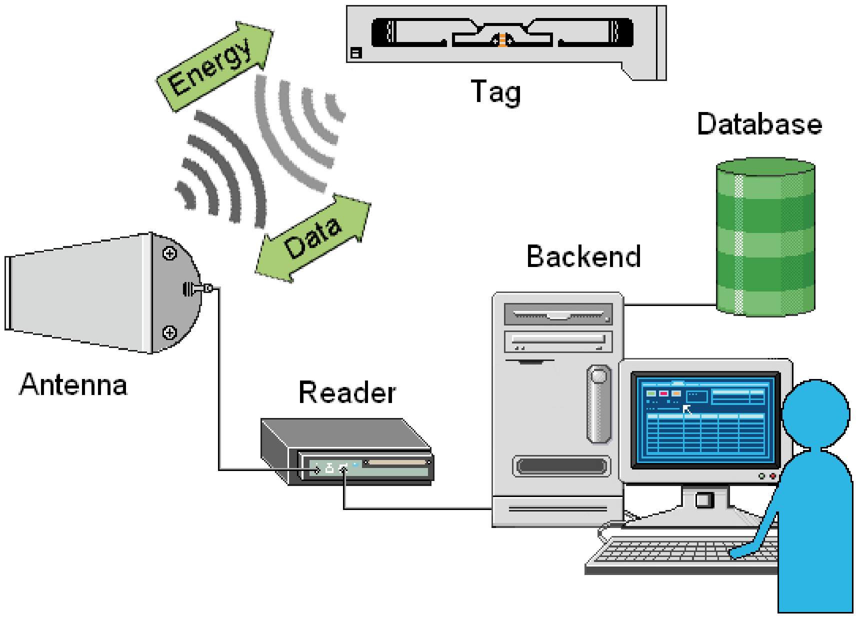


Fig. 1. A generic RFID system

2.1 System components

RFID tags consist of a microchip with an antenna. They come in a wide variety of sizes, from pencil lead thin tags used for animal tracking and credit-card sized ones for access control to heavy duty transponders used for tracking shipping containers, vehicles, and railroad cars. RFID tags can be categorized as either 'active', 'passive', or battery-assisted (semi-active/semi-passive), depending on how they are powered. Active tags are powered by a long-life internal battery and usually have both read and write capabilities. Passive tags are powered by the electromagnetic field generated from the reader and are usually read-only with shorter read ranges. Battery-assisted tags contain a battery that enables them to monitor, process, store, and transmit data over extended ranges. Based on the memory type, RFID tags can be further classified as read-only, write once read many (WORM), or read/write.

An RFID reader is a powered device that wirelessly communicates with the RFID tags and facilitates data transfer between the tag and the backend system. A typical reader consists of a radio frequency module, a control unit, and a coupling element to interrogate electronic tags via radio frequency (RF) communications. The basic functions of the reader include activating tags by sending querying signals, supplying power to passive tags, encoding the data signals going to the tag, and decoding the data received from the tag. RFID readers differ considerably in complexity, depending on the type of tags being supported and the functions being performed, such as sophisticated signal conditioning, parity error checking, and correction. They can either be portable handheld units or fixed devices.

RFID systems also rely on software. The software can be divided into three groups: front-end tag reading algorithms, middleware, and backend system interface. The front-end algorithms carry out the signal processing tasks. RFID middleware connects readers to the backend server and database. It also filters the data acquired from the reader and handles various user interfaces. The real power of RFID comes in integrating RF technology with a backend system to perform functions such as matching digital information received from the reader against the backend database and routing the retrieved information to the correct application.

2.2 RFID tag categories

RFID tags are at the heart of an RFID system, and can be categorized as active, semi-active/semi-passive and passive in relation to power, as well as read/write and read only in terms of their memory (Thornton et al., 2006; Wyld, 2006).

Passive tags do not have an internal power source and need to draw power from an RFID interrogator. The interrogator emits electromagnetic waves that induce a current in the tag's antenna and power the chip on the tag. When the power to the tag's chip passes the minimum voltage threshold, the circuit turns on and the tag transmits its information to the reader. Because of the absence of a battery, passive tags have a relatively short reading range of only a few meters.

Active tags contain their own battery that is used for both powering the chip and boosting the return signal. The battery gives the tags the ability to continuously monitor high-value goods or a container's seal status. Compared to passive tags, active tags have wider read ranges (tens of meters and even hundreds of meters), larger memory capacities, and faster processing times. However, battery life limits the life of the tag.

There are three terms that are used to describe passive tags that contain batteries: semi-active, semi-passive, and battery-assisted. Some think the terms are interchangeable, while others think that "a semi-active tag is an active tag that remains dormant until it receives a signal from the reader to wake up. The tag can then use its battery to communicate with the reader" … "semi-passive tags are similar to active tags in that they have their own power source, but the battery only powers the microchip and does not broadcast a signal. The RF energy is reflected back to the reader like a passive tag" (Karygiannis et al., 2007). A problem in the above definitions is that some active tags can be put into sleep mode and later reactivated upon receiving a wake up signal from a reader. Therefore, it is more accurate to use the term "battery-assisted tag". Battery-assisted tags have a power source that can keep the chip on the tag constantly powered. However, they can be programmed to preserve battery power by only signaling if an alert condition is detected, or only at predetermined time intervals. In addition, battery-assisted tags may incorporate one or more sensors, enabling them to monitor environmental conditions, such as temperature, humidity, shock, or vibration.

Depending on the memory type, the tags can be further classified as read-only, write once read many (WORM), or read/write. Read-only tags are typically passive and are similar to bar codes because they only carry a serial number. Data stored on the tag cannot be modified or appended unless the microchip is reprogrammed electronically. Read-only tags are available in many versions, varying in range, data bits, and operating temperature.

WORM memory allows users to encode tags one time during production or distribution. After the initial encoding, the tag's data becomes locked and cannot be changed.

Read/write tags function like computer disks because the data stored can be edited, added to, or completely rewritten an unlimited number of times. These tags are often implemented on reusable containers or other assets in logistic applications. When the contents of the container are changed, new information can be updated on the tag.

In this chapter, we will use RFID as a generic term to describe any automated tagging and reading technology, including active, passive, and battery-assisted RFID technologies, and various formats and applications.

2.3 Frequency bands

In addition to the types of tags used, RFID systems can also be distinguished by their radio frequency. The four primary radio frequency bands, ranging from 30 KHz to 5.8 GHz, are low frequency (LF), high frequency (HF), ultra-high frequency (UHF), and microwave frequency (MW) (Wyld, 2006). The choice of frequency is dependent on the application, the size of the tag, and the read range required. In general, the higher the frequency, the faster the data transfer or throughput rates, but the more expensive the system.

LF ranges from 30 KHz to 300 KHz. In this band, RFID systems commonly operate in a long waveband of 125 to 135 KHz. LF RFID systems generally use passive tags with short read ranges (up to 20 inches) and have lower system costs. The LF tags perform very well in most manufacturing environments. They work well around metal and are resistant to rain. The application areas include security access control, animal identification, and asset tracking.

HF ranges from 3 MHz to 30 MHz, but most HF RFID systems operate at 13.56 MHz. A typical HF RFID system uses passive tags that have a maximum read range of up to 3 feet and a faster data transfer rate than LF tags. This wavelength is robust against water, dust, and other environmental factors. Not only have HF systems been widely used in libraries, pharmaceutical manufacturing sites, and logistics, they have also been adopted for smart identification such as the e-passport.

The next frequency range is UHF that ranges from 300 MHz to 1000 MHz. The passive UHF tags operate around 865-868 MHz in Europe and 902-928 MHz in the United States, while active UHF RFID systems operate at 315 MHz and 433 MHz. Since UHF tags can be read at longer distances with a faster communication speed than LF and HF tags (from 3-6 meters for passive tags and more than 30 meters for active tags), this frequency band is emerging as the preferred band for supply-chain applications.

A typical MW RFID system operates either at 2.45 GHz or 5.8 GHz. The former frequency is traditionally used in long-range access control applications and has a read range of up to 1 meter as a passive tag or longer as an active tag. In Europe, the 5.8 GHz frequency band has been allocated for road traffic and road-tolling systems.

An overview of the characteristics of each RFID frequency band is presented in Table 1, which includes read ranges, data transfer rates, application areas, and corresponding ISO standards. From a standards perspective, the ISO 18000 standard covers the air interface protocol—the way RFID tags and readers communicate—for major frequencies used in RFID systems.

	LF	HF	UHF	MW
Frequency	30 – 300 KHz	3 – 30 MHz	300 – 1000 MHz	2 – 30 GHz
Typical Frequencies	125-134 KHz	13.56 MHz	433 MHz (Active) 865 – 956 MHz	2.45 GHz 5.8 GHz
Read Range	Up to 1m with long-range fixed reader	Up to 1.5m	433 MHz: Up to 100m 865-956 MHz: 0.5 – 5m	Passive ≈ 3 m Active ≈ 15m
Data Transfer Rate	Less than 1 kilobit per second (kbit/s)	≈ 25 kbit/s	≈ 30 kbit/s	Up to 100 kbit/s
Common Applications	Access control, Animal identification, Inventory control, Vehicle immobilizers	Smart cards, Contact-less access and security, Item level tracking, Library books, Airline baggage	Logistics case/pallet tracking, Baggage handling	Railroad car monitoring, Automated toll collection
Pros and Cons	LF signal penetrates water. It is the only technology that can work around metal. LF tags have a short read range and low data transfer rate, and are more expensive than HF and UHF tags because of their longer copper antennas.	HF system is able to read tags that are placed in a very close proximity to each other. HF signal penetrates water but not metal. HF tags are less expensive and offer higher read rate than LF tags.	Active RFID has a very long read range with high price tags. Since using a battery, tags have a finite lifespan (typically 5 years). UHF system is capable of reading multiple tags quickly. However, UHF tags are highly affected by water or metals.	Microwave transmission is highly directional, and enables precise targeting. MW tags provide the fastest data transfer rate. However, they cannot penetrate water or metal.
ISO Standards	11784/85, 14223	14443, 15693, 18000	15693, 18000	18000

Table 1. RFID frequency bands and standards

3. Security aspects

Like other information systems, RFID systems are vulnerable to attack and can be compromised at various stages of their use. Attacks against an RFID system can be categorized generally into four major groups: attacks on authenticity, attacks on integrity, attacks on confidentiality, and attacks on availability. Besides being vulnerable to common attacks such as eavesdropping, man-in-the-middle, and denial of service, RFID technology is, in particular, susceptible to spoofing and power attacks (Figure 2). This section illustrates different kinds of attacks and provides countermeasures against these attacks.

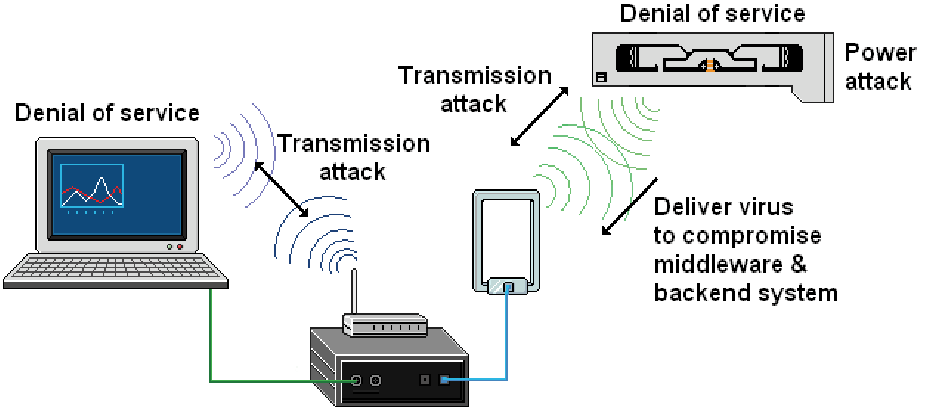


Fig. 2. Attack points

3.1 Reverse engineering

Reverse engineering is the process of taking something apart to discover how it works. Figure 3 shows an example of RFID physical elements (MacGillivray & Sheehan, 2006). Considering privacy issues related to the biometric e-passport, it may be possible for an attacker to gain access to the chip and read its memory contents optically to retrieve the PIN, biometric data, personal information, etc. The technical ability and equipment needed to reverse engineer an integrated circuit can be rated at three different levels from a knowledgeable individual using low cost and easily available tools to a highly skilled team, using equipment not commonly available in the commercial market (Actel, 2002). Unfortunately, the methods of attacking ASIC technology are not a secret and can be easily accessed (Blythe et al., 1993).

Countermeasures

A FIPS standard refers to chip coatings as an anti-reverse engineering method to prevent attacks. Various tamper proof techniques have been developed to defend against reverse engineering attacks. For instance, by adding a tamper-release layer to RFID tags, operations personnel can be alerted if a tag has been tampered with.

3.2 Power analysis

Power analysis is a form of side-channel attack that is intended to retrieve information by analyzing changes in the power consumption of a device. It has been proven that the power emission patterns are different when the card received correct and incorrect password bits or cryptographic keys. It is possible to breach smart card security by monitoring power consumption signals. Professor Adi Shamir demonstrated the ability to use a password to kill an RFID tag during the RSA Conference 2006. He also predicted that a power analysis attack on a RFID tag could be performed using a very common device such as a cell phone

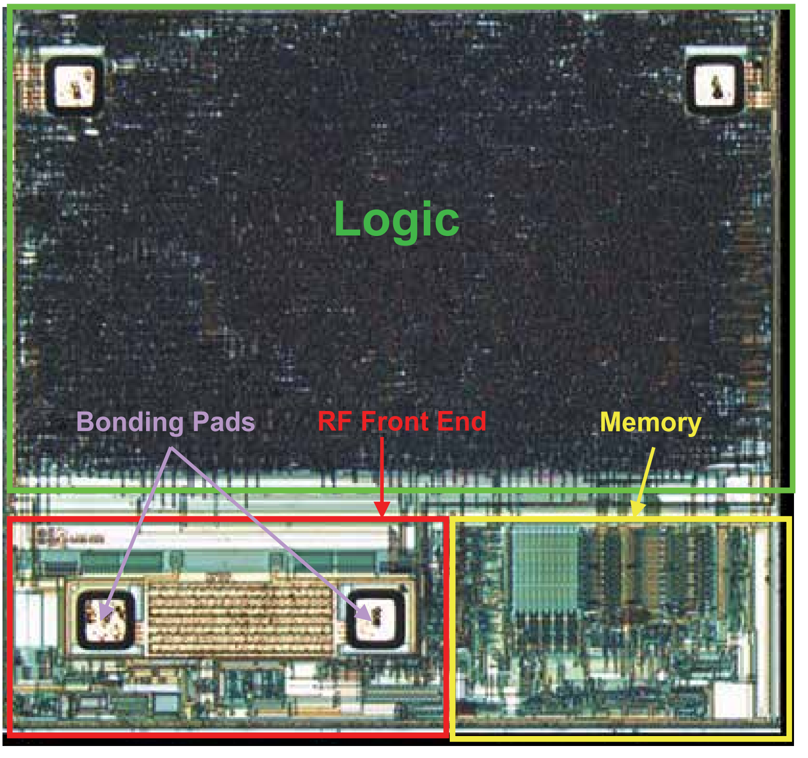


Fig. 3. Reverse engineering

(Merritt, 2006). Two methods—either masking the spikes in power consumption or improving the hash algorithm—can be used to protect the password for being cracked with power analysis attack. Figure 4 shows an example of using Hamming weight data to break the Data Encryption Standard (DES) through analyzing the power consumption (Messerges et al., 2002).

Countermeasures

The common methods used to defeat power analysis attacks are filtering or adding an element of randomness. Filtering power signals or delaying the computation randomly can increase the difficulty for the attacker to identify the power consumption patterns. Another method implemented in some smart card designs is adding an element that simply consumes a random amount of power. Unfortunately, this approach may cause a problem for RFID systems where minimizing power consumption is a priority.

3.3 Eavesdropping

Since an RFID tag is a wireless device that emits data, usually a unique identifier, when interrogated by an RFID reader, there exists a risk that the communication between tag and reader can be eavesdropped. Eavesdropping occurs when an attacker intercepts data with a compliant reader—one for the correct tag family and frequency—while a tag is being read by an authorized RFID reader. Since most RFID systems use clear text communication, due to tag memory capacity or cost, eavesdropping is a simple but efficient means for the attacker to obtain information on the collected tag data. The information picked up during the attack can have serious implications—it can be used in subsequent attacks against the RFID system. It is necessary to point out that in passive RFID systems readers have

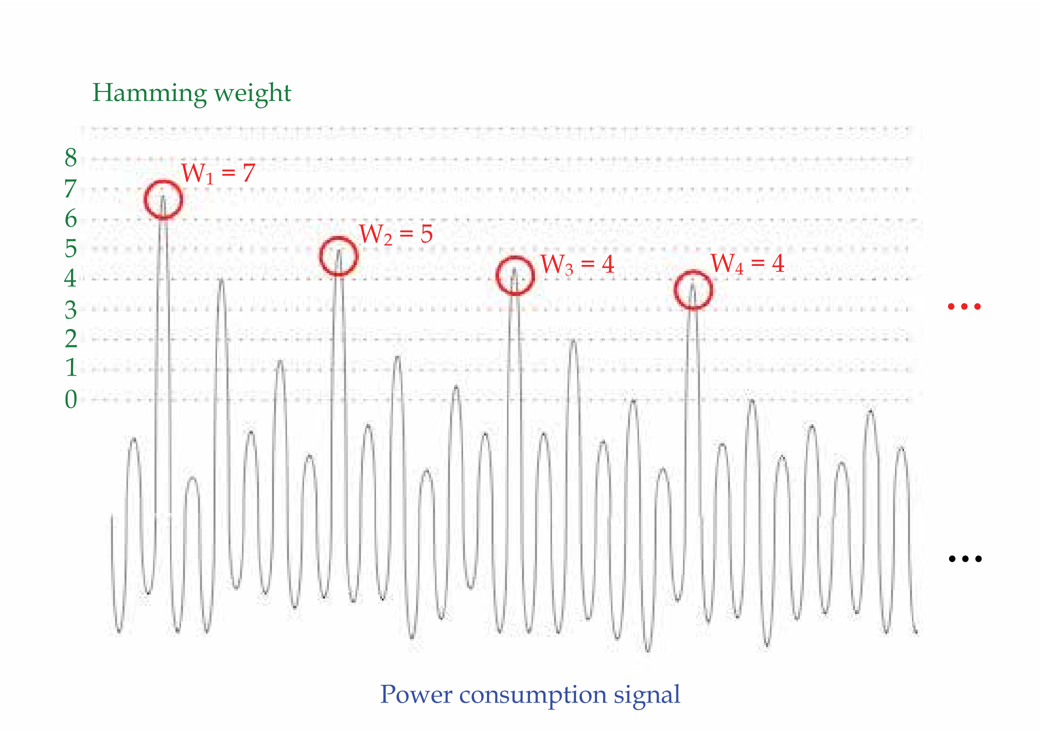


Fig. 4. An example of power analysis

significantly longer transmission ranges than tags. When passive tags modulate and backscatter the signal from the reader to communicate, they have only a fraction of the transmission power of the reader. Therefore, passive tags have a more limited transmission range and are less susceptible to eavesdropping (Karygiannis et al., 2007). However, it is necessary to keep in mind that even if the eavesdropper is out of the range of the tag signal, he or she may still be able to listen to the commands sent out from the reader (Figure 5).

Countermeasures

Countermeasures against eavesdropping include establishing a secure channel and/or encrypting the communication between tag and reader. Another approach is to only write the tag with enough information to identify the object. The identity is used to look up relevant information about the object in a back end database, thus requiring the attacker to have access to both the tag and the database to succeed in the attack.

3.4 Man-in-the-middle attack

Depending on the system configuration, a man-in-the-middle (MITM) attack is possible while the data is in transit from one component to another. An attacker can interrupt the communication path and manipulate the information back and forth between RFID components (Figure 6). This is a real-time threat. The attack reveals the information before the intended device receives it and can change the information en route (Welch & Lathrop, 2003). Even if it received some invalid data, the system being attacked might assume the problem was caused by network errors and would not recognize that an attack occurred. An RFID system is particularly vulnerable to MITM attacks because the tags are small in size and low in price, all of which means that there is generally a lack of sophisticated protection circuitry.

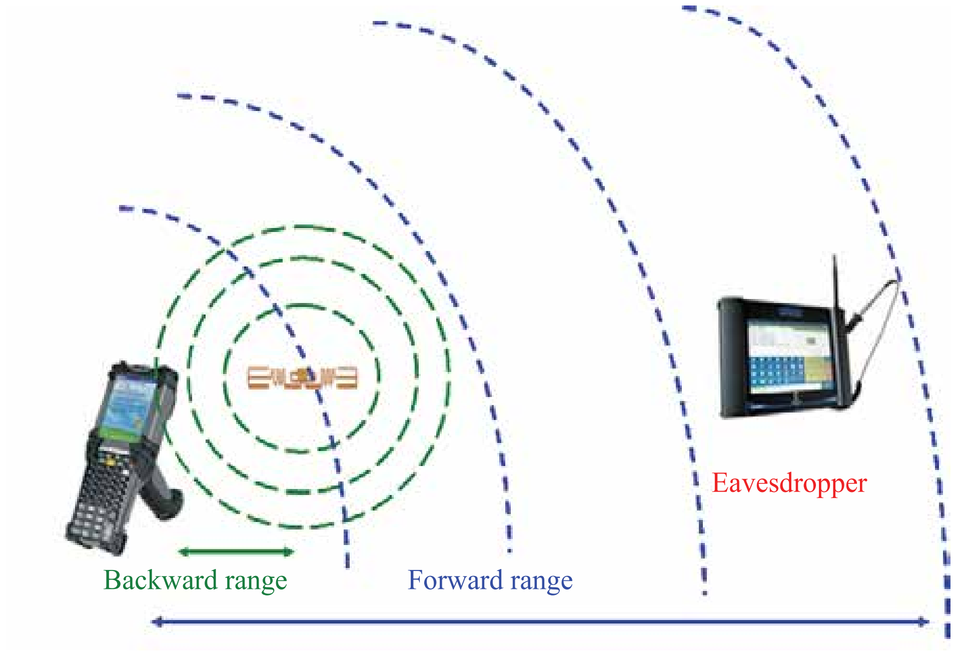


Fig. 5. Eavesdropping on reader-tag communication

Fig. 6. Man-in-the-middle attack

Countermeasures

Several technologies can be implemented to reduce MITM threats, such as encrypting communications, sending information through a secure channel, and providing an authentication protocol.

3.5 Denial of Service (DoS)

DoS attacks can take different forms by attacking the RFID tag, the network, or the backend. The purpose is not to steal or modify information, but to disable the RFID system so that it cannot be used. When talking about DoS attacks on wireless networks, the first concern is physical layer attacks, such as jamming and interference. Jamming using noise in the RFID system's frequency range can reduce the throughput of the network and ruin network connectivity resulting in overall supply chain failure (Egli, 2006). Jamming happens when a device that actively broadcasts radio signals can block and disrupt the operation of any and all nearby RFID readers. Interference with other radio transmitters can also launch a DoS attack to obscure the communications between the tags and reader. Another form of DoS is to destroy or disable RFID tags by removing them from the items, washing out their contents completely, or wrapping them with metal foil.

Countermeasures

In general, it is easier to detect DoS attacks than prevent them from happening. However, once detected, the attacks can generally be stopped before they do too much harm. For example, countermeasures against jamming can use passive listening to detect the tags whose transmission exceeds a predefined volume, and then use block functions to thwart them. Countermeasures against detaching the tags from the targeted items could be either through enhancing the mechanical connection between the tags and items, or adding an alarm function to active tags.

3.6 Spoofing

In relation to RFID technology, spoofing occurs when a forged tag masquerades as a valid tag and thereby gains an illegitimate advantage. Tag cloning is a spoofing attack where the attacker captures the data from a valid tag and creates a copy of the captured sample on a blank tag. Another example is an attacker reading a tag's data from a cheap item in a store and then uploading the data onto another tag attached to a similar but more expensive item. Mr. Lukas Grunwald, a German security expert, said "I was at a hotel that used smartcards, so I copied one and put the data into my computer, … Then I used RF Dump to upload the room key card data to the price chip on a box of cream cheese from the Future Store. And I opened my hotel room with the cream cheese!" (Newitz, 2006)

Countermeasures

A common way to defeat a spoofing attack is to implement an RFID authentication protocol and data encryption, which increases the cost and technology complexity needed for a successful attack.

3.7 Cloning

Tag cloning is a process that first captures the data from a legitimate tag and then creates an unauthorized copy of the captured sample on a new chip. Researchers from Johns Hopkins University and RSA Labs published experimental results of cloning a cryptographically protected Texas Instruments digital signature transponder (DST) that was used to buy gasoline and activate a car's ignition (Rieback et al., 2006a).

Countermeasures

In order to defeat physical cloning attacks, Tuyls and Batina proposed to use Physical Unclonable Functions (PUFs) as secure memory for the storing of the secret key on an RFID tag (Tuyls & Batina, 2006). The authors claimed that "both the physical cloning attack as well as the cloning attack based on (actively or passively) attacking the protocol between the tag and the reader can be prevented."

3.8 Replay

In a replay attack, an attacker intercepts communication between a reader and a tag to capture a valid RFID signal. At a later time, the recorded signal is re-played into the system when the attacker receives a query from the reader. Since the data appears valid, it will be accepted by the system.

Countermeasures

The most popular solution is the use of a challenge and response mechanism to prevent replay attacks. Time-based and counter-based schemes can also be used as countermeasures against replay attacks.

3.9 Viruses

Since most of the passive RFID tags currently only have a small storage capacity of 128 bits, viruses are probably not a credible threat to RFID systems. However, the situation may be changing since three computer researchers released a paper in March 2006, which reported that RFID tags could be used as a medium to transmit a computer virus. It also explained how the RFID virus works in a supply chain. If a container arrived in a distribution center and the container's RFID tag had been infected with a computer virus, this particular RFID virus could use SQL injection to attack the backend servers and eventually bring an entire RFID system down (Rieback et al., 2006b).

Countermeasures

The virus attacks which have been demonstrated on RFID-based systems are the common attacks against information systems, such as buffer overflow attacks, code or SQL injection attacks, etc. Well-developed middleware can be used to avoid virus attacks by blocking anomalous bits from the tag.

3.10 Tracking

Unlike the previously discussed RFID attacks, tracking is a threat directed against an individual. Within the next few years, manufacturers may put item-level RFID tags into many more household products. There is a privacy concern because instead of tracking books and consumer products such as clothing, RFID systems can be used to track people's movements and even create a precise profile of their purchases.

Countermeasures

An easy method to disable tracking is to deactivate the RFID tags, which is known as "killing" the tag that will be introduced in the following section.

3.11 "Killing tag" approach

Typically killing an RFID tag is done to prevent it from communicating thus making it impossible to be read anymore. For example, a kill command is defined in standard Electronic Product Code (EPC) format, which is used to permanently disable the tags for purposes of privacy. Since it is necessary to make sure that RFID tags are not killed by an unauthorized party, the kill command is secured by a password called the kill password. However, the kill command also brings some drawbacks:

- Although the kill command was introduced to protect consumer privacy, consumers cannot easily detect whether a tag has been deactivated. Furthermore consumers cannot kill the tag by themselves. Because in order to do this, not only would they need an interrogator, but also a valid kill password.
- The kill feature also brings up a new threat to an RFID system. If an enemy deactivates RFID tags in a supply chain, the supported application will not function properly because the item identification numbers cannot be read anymore. Furthermore, once killed, a tag can never be re-activated for any further application, such as item recalls, product returns, etc.
- If the kill password is weak (for example, EPC Class-1 Generation-1's 8-bit kill password has only 256 possibilities, while EPC Class-1 Generation-2 has improved significantly with a 32-bit password—more than 4,000,000,000 possibilities), unauthorized parties can kill the tag very easily.
- Although the killed tags cannot emit radio frequency anymore, data are still stored in the tags' memory.

Countermeasures

In many applications, it is important to protect the tag from a kill command that permanently disables the tag's functionality. In order to do so, tag memory and reading can be password protected, and a command of permanent lock can make password and/or the tag data permanently unchangeable.

3.12 Block tag

Another method to protect against unwanted scanning of RFID tags attached to items that people are carrying or wearing is to block the tags. Blocking the tags can be accomplished with different approaches, such as Faraday Cage, active jamming, or "blocker tags". A Faraday cage is a metal or foil-lined container, which is impenetrable to radio-frequency waves. Petty thieves are already known to use foil-lined bags in retail shops to defeat shoplifting detection mechanisms. Since active jamming violates the regulations of most governments, a device called a "blocker" tag has been proposed to protect against inappropriate scanning (Juels et al., 2003). The blocker tag obstructs the RFID scanning process by simulating that all the possible IDs are present. Let us take the tree-walking protocol, which is often used to avoid collisions while reading, as an example. Because in tree-walking protocol the space of k-bit identifiers is viewed as the leaves in a tree of depth k, a reader traversing the tree needs to figure out whether the next bit is a "0" or "1". When a blocker tag is queried, it always responds with both "0"and "1" and causes a reader to stall. A blocker can be made from a cheap passive RFID tag. Therefore it is possible to embed a blocker into a portable device to actively prevent inappropriate scanning (Juels, 2005b).

3.13 Summary of security aspects

As mentioned above, there are many ways to attack various parts of RFID systems. Many efforts have been taken to study the countermeasures needed to defend against these threats. As a summary of this section, a threat-countermeasure map is provided for visualizing the relationships between security threats and countermeasures. In Figure 7, each attack is mapped onto as many countermeasures as possible to show that one threat can be defeated with a specific countermeasure or several countermeasures. In this way, decision makers can easily determine the most efficient strategy to protect their RFID system.

From the threat-countermeasure map, it is clear that authentication and encryption are the most important security techniques for the protection of RFID systems. We can use them to address a wide variety of security threats. The purpose of authentication is to confirm that an entity is what it claims to be. In an RFID system, authentication is performed by tags to verify an authorized reader, making sure that an RFID reader cannot communicate with the tag unless being successfully authenticated. With different designs, the authentication can be either one way or two ways. Encryption is another major countermeasure, which is a process of scrambling data to make it difficult to unscramble or decipher. To heighten the security in an RFID application, both the data stored on a tag and the data communicated between a reader and the tags need to be encrypted. In real-life applications, encryption and authentication protocols are often used in a combination to enhance the security of an RFID system. However, such a solution will certainly increase the cost of the implementation.

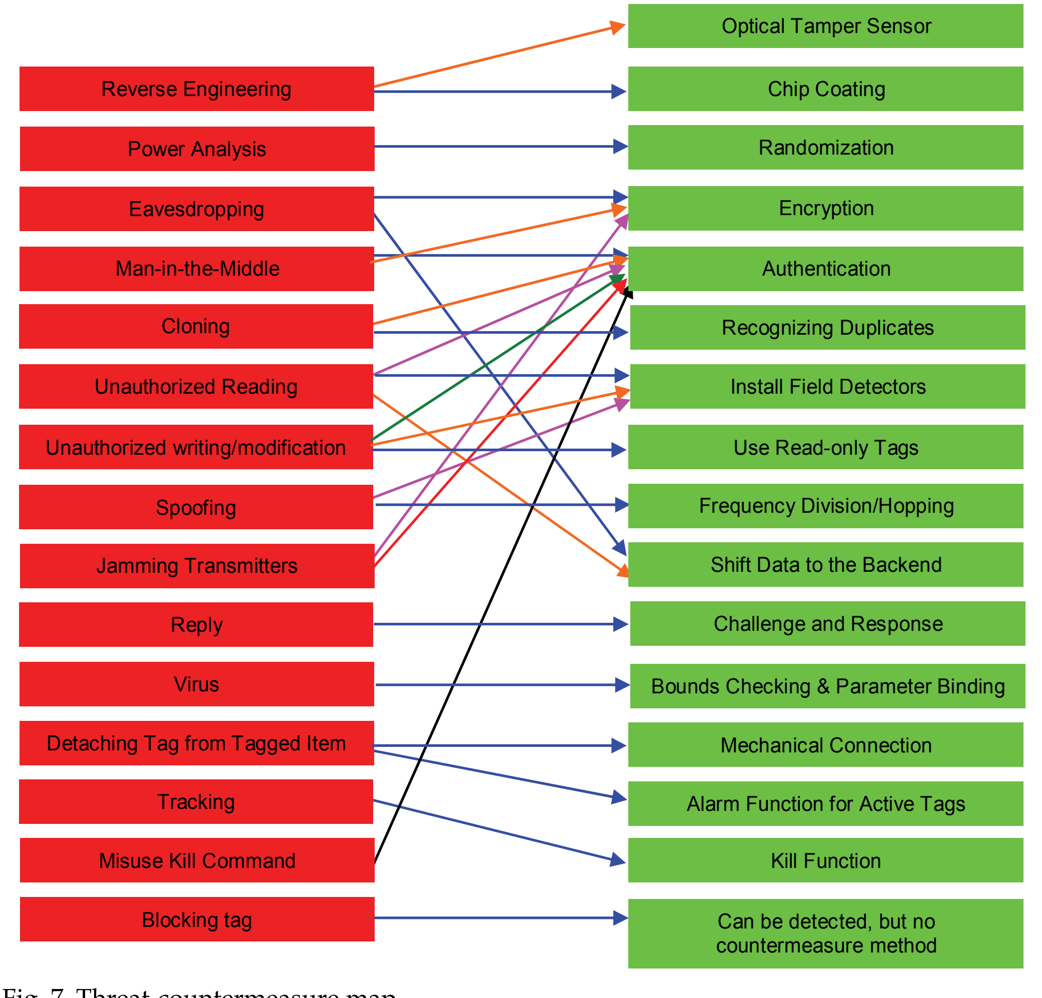


Fig. 7. Threat-countermeasure map

4. Authentication

From the above threat-countermeasure map, we can see that authentication is a primary method for ensuring RFID security. Authentication is a process of confirming the identity claimed by an entity. In the context of a tamper-resistant authentication protocol for an RFID system, the tag and reader establish a trusted relationship and agree on a common, secret, session key to secure the communication between them. It is not difficult to develop a trusted authentication protocol to make high-cost RFID tags directly authenticate RFID readers. However, the majority of RFID applications use low-cost and high-volume passive tags. Under this circumstance, developing a secure authentication protocol is a challenge because tags, compared with readers and back-end servers, are highly resource limited and cannot perform strong encryption. In order to solve this problem, various lightweight authentication algorithms and protocols have been proposed, debated, and tested.

4.1 Pseudonym

The pseudonym technique has been proposed to deal with this problem in low-cost RFID systems: each tag stores a list of pseudonyms that can only be understood by authorized verifiers (Juels, 2003). When the tag is queried, it emits the next pseudonym from the list. Since the protocol uses only the XOR operation and does not require the tag to perform any cryptographic operations, it fits with the restrictions of the low-cost RFID tags very well. A problem is that the tag can only store a small list of pseudonyms because of its small data capacity. One method used to solve this problem is renewing the list each time when the tag is queried. However, to allow the list to be renewed, a mutual authentication protocol is required between the tag and the reader to prevent an attacker from updating the list.

4.2 YA-TRAP, O-TRAP and YA-TRAP+

Yet Another Trivial RFID Authentication Protocol (YA-TRAP) presents a novel idea for RFID authentication. It uses monotonically increasing timestamps and a keyed hash to distinguish anonymous (adversary) tags from legitimate tags (Tsudik, 2006). In the beginning, a reader sends the current system time to a tag. The tag decides if the time is valid by checking if it is in the interval between the stored timestamp and a maximum system allowable timestamp. If the received time is valid, the tag will use it to update the stored timestamp and send the key-hashed timestamp to the reader. Otherwise, the tag will send a pseudo-random number to the reader. The information is forwarded to the backend system that maintains a hash lookup table and is able to quickly compare the values to validate the tag. However, YA-TRAP is susceptible to a trivial DoS attack when the adversary sends a wildly inaccurate timestamp to the tag.

In order to overcome the weakness of YA-TRAP, modified authentication protocols, such as O-TRAP and YA-TRAP+, have been proposed (Burmester et al., 2006). The protocol O-TRAP stands for "Optimistic" Trivial RFID Authentication Protocol, i.e., the security overhead is minimal when the parties are honest. O-TRAP is a revision of YA-TRAP with added one-pass anonymity for authenticated transponders and solves some vulnerabilities of YA-TRAP. YA-TRAP+ was proposed to deal with large scale DoS attacks by introducing an extra optional pass in which a server authenticates the timestamp. A major drawback for both O-TRAP and YA-TRAP+ is that the server workload is increased so that more computational resources are required on a per-tag basis for authentication.

4.3 HB, HB+ and HB++

Hopper and Blum proposed a human-to-computer authentication protocol, named the HB protocol (Hopper & Blum, 2001). Its extremely low computational cost makes the protocol well suited for resource-constrained devices like RFID tags. Unlike other classical symmetric key cryptography solutions, the security of the HB protocol is based on the hardness of the Learning Parity with Noise (LPN) problem. A random k-bit binary vector is generated by the reader and transmitted to the tag for challenge. The tag computes the inner dot product of the k-bit vector and a shared key, and XORs the value with a noise bit (=1 with probability $\eta \in [0, 1/2]$). The calculated value is sent back to the reader for checking to result in a pass or fail. This is one round of HB authentication with the same process being repeated several times. However, the HB protocol is only secure against passive attacks and not against active attacks. For example, an adversary can transmit a fixed k-bit vector to the tag several times and potentially deduce the key.

Addressing this problem, HB+ was proposed to secure against both passive and active attacks with some additions (Juels & Weis 2005). The first is an additional shared key so that the tag and reader share two independent keys (instead of using one shared key in the HB protocol). The other is a random "blinding vector" that is generated by the tag at the beginning of the process and is used in calculations later on. In HB+, a basic authentication step consists of three rounds. First, the tag sends a random "blinding factor" to the reader. Then the reader replies with a random challenge in the same way as HB protocol. Finally, the tag calculates a return value that is the inner dot product of the newly introduced key and blinding vector XORs with the HB return signal as before, and replies with it to the reader. However, it has been shown that HB+ is vulnerable to a simple man-in-the-middle attack that was not considered in HB protocol (Gilbert et al., 2005). As an improvement, a further modified HB protocol, HB++, was proposed to overcome the weakness of HB+ protocol (Bringer et al., 2005). However, it has been discovered that the HB++ is not immune to attacks from an adversary that pretends to be an authentic reader (Piramuthu 2006).

5. Case studies

In the previous sections, we analyzed the security threats and provided the corresponding countermeasures. "Unfortunately, businesses and governments are not the only ones interested in RFID. Civil liberties groups, hackers and criminals are also keenly interested in this new development, albeit for very different reasons" (Rieback et al., 2006c). Following are four case studies that illustrate how the security of some RFID systems could be compromised. The purpose is not to teach people how to attack an RFID system, but to help people become aware of the kinds of threats that need to be taken into account when designing a secure RFID application.

5.1 Cracking crypto-enabled RFID products

The Texas Instruments Registration and Identification System (TIRIS) is an RFID system that uses a 3.6x29mm cylindrical tag with a reading range of up to 40 inches. The TIRIS DST tags have been adopted by different companies to make millions of SpeedPass payment transponders and automobile ignition keys. In 2005, researchers from Johns Hopkins University and RSA Laboratories demonstrated a way to crack the encryption of Exxon Mobil's SpeedPass. The RFID tag they compromised was a DST-40 tag that consists of a small microchip and an antenna coil that uses a secret 40-bit cryptographic key.

The Mifare Classic RFID smartcard is a wireless card protected by an encryption algorithm, which has been used by transit operators in London, Boston and the Netherlands, as well as public and private organizations to control access to restricted areas. In 2008, two research groups managed to crack the encryption and reported the security flaw that allowed them to do so. They revealed that the method to retrieve cryptographic keys is relatively easy and does not rely on expensive equipment (Nohl et al., 2008 & Schreur et al., 2008).

The procedures to crack a crypto-enabled RFID tag, including collecting data, revealing encryption key, and creating a clone tag, are as follows:

1. Reverse engineering: The encryption algorithm can be reverse engineered through flawed authentication attempts. The method involves sending RFID devices carefully chosen electronic queries and recording the responses of the devices. The response information gives clues as to what is happening inside the microchip, and therefore makes it possible to reconstruct the encryption algorithm.
2. Key cracking: Once the algorithm is known, the keys can be figured out by brute force attack, i.e. simply trying all possible keys. Since the DST-40 tag uses a proprietary 40-bit and Mifare Classic uses a 48-bit encryption algorithm, it will take 9 to 10 hours to try all possible keys for both devices on advanced equipment.
3. Simulation: After obtaining the key (and serial number), it is possible to create a clone tag.

Lessons to learn

The impact of compromising tag encryption on supply chain RFID systems has not been determined. However, one lesson that can be learned from the details of these cases is that the cryptographic algorithm needs to be built into the RFID system correctly.

- It is better to use a longer key length, such as industry-standard 128-bit Triple DES encryption or AES encryption.
- It is necessary to use standard cryptographic algorithms that have been through peer reviews, instead of a proprietary cryptographic algorithm.
- It is better to use public key (asymmetric) encryption rather than secret key (symmetric) encryption.

5.2 RFID-Zapper

There are several ways to deactivate passive RFID tags. The RFID-Zapper is an easy-to-build electronic device that can permanently deactivate passive RFID-Tags. The device was developed by two German students in 2006. Their motivation was a privacy concern over the potential use of RFID tags on individual items purchased by consumers (Juels, 2005b). The technique is so simple that everyone can build his/her own RFID-Zapper. The concept of RFID Zapper was presented at the 22nd annual Chaos Communication Congress (22C3) as follows (MiniMe & Mahajivana, 2005).

> *Basically it copies the microwave-oven-method, but in a much smaller scale. It generates a strong electromagnetic field with a coil, which should be placed as near to the target-RFID-Tag as possible. The RFID-Tag then will receive a strong shock of energy comparable with an EMP and some part of it will blow, thus deactivating the chip forever.*

A prototype was built by modifying the circuit board of a single-use camera with a flash. The voltage of flash capacitor needs at least 100 V to supply enough electrical current. The flash bulb is replaced with a coil of 5 windings of 1 mm diameter copper wire. Disconnected from the flash, the capacitor is re-connected to the coil with an added switch to turn the device on or off. The flash indicator light is re-connected so that it glows when the capacitor is fully charged and can be clearly observed. Since a large amount of energy can be emitted into the environment in a very short time, the magnetic field of the Zapper is sufficient to destroy an RFID forever at close range.

The prototype device was tested on the passive 13.56 MHz RFID tags successfully. Currently, the RFID-Zapper has been tested only on 13.56 MHz tags; however, the inventors hope to be able to try their device on other tags running at different frequencies soon. Another threat is that a German privacy advocacy group FoeBuD plans to manufacture and sell such a device that consumer could use it to disable RFID tags permanently (Collins, 2006).

Lessons to learn

The demonstration was performed with the capacitor loaded about 100 V. The RFID-Zapper was able to destroy the RFID tags placed right next to it.

- It is necessary to test the working range of the RFID-Zapper with a capacitor that can supply more power.
- Even though RFID technology offers several advantages over optical bar codes, it is a good practice to use both technologies in some critical applications because RFID tags could be deactivated or killed.

5.3 Trigger a bomb

The increasing threat of identity fraud has produced worldwide efforts on strengthening security features in identity documents. The International Civil Aviation Organization (ICAO) has been working on the new e-passports fitted with RFID tags for wireless processing when people pass through Customs. However, at Black Hat 2006, a group of security experts from Flexilis demonstrated that the proposed American RFID passports might be used by terrorists as potential bomb triggers (Coverson, 2006). Kevin Mahaffey, Director of Software Development at Flexilis, and his colleagues used a mockup e-passport, equipped with an RFID chip, and set up a small explosive charge nearby. With the passport opened about 1/2 inch, Mahaffey demonstrated how the explosive could be set off when a passport was detected by a nearby inquiring RFID reader.

In a report "RFID Passport Shield Failure Demonstration", they explained the mechanism behind their proof of concept experiment (Flexilis, 2006):

> *When present in a reading field, a passport RFID tag will wirelessly draw power from the reader in order to operate. The change in antenna current is detectable by the RFID-reading hardware; therefore, even if a tag is not directly sending data, it intrinsically discloses its proximity to the reader by its presence in the reading field.*

A potentially dangerous security breach is that the terrorists could potentially use the RFID tag to trigger a bomb.

Lessons to learn

At the moment, Flexilis's proof of concept experiment only demonstrated the ability to read passive RFID. However, the demonstration alerts us to a real world security threat.

- It is necessary to study whether it is possible to use active RFID tags as bomb trigger.
- It is necessary to research authentication methods for passive RFID tags so that they cannot be activated with an unauthorized reader.

5.4 Snooping attack

One of the advantages of RFID technology over barcodes is that RFID does not require a line of sight between tag and reader. However, it brings a new security threat to supply chains if RFID-tagged items can be read even if they are inside a truck. Sniffing the truck's payload was selected as "one of the five coolest hacks of 2007" (Higgins, 2007a). Researchers from PacketFocus Security Solutions and Atlas RFID Solutions have demonstrated reading EPC codes from tagged products on 18-whheeler tractor-trailers with standard tag readers and antennas (Higgins, 2007b). The test explored a vulnerability of RFID tags used in supply chain applications — business competitors or enemy forces could use the sniffed information for intelligence purposes. The detailed logistics information is vital to business and military success.

To evaluate the risk of such an attack, we carried out an experiment with four different passive RFID tags: Alien 9554, Avery AD222, Symbol Four T, and StongTech. First, we measured their readability in an indoor open area that is about 11 meters long and 6 meters wide. The test result in Figure 8 shows that all the tested tags can be reliably read within 4.5 meters when the tag and reader are parallel to each other.

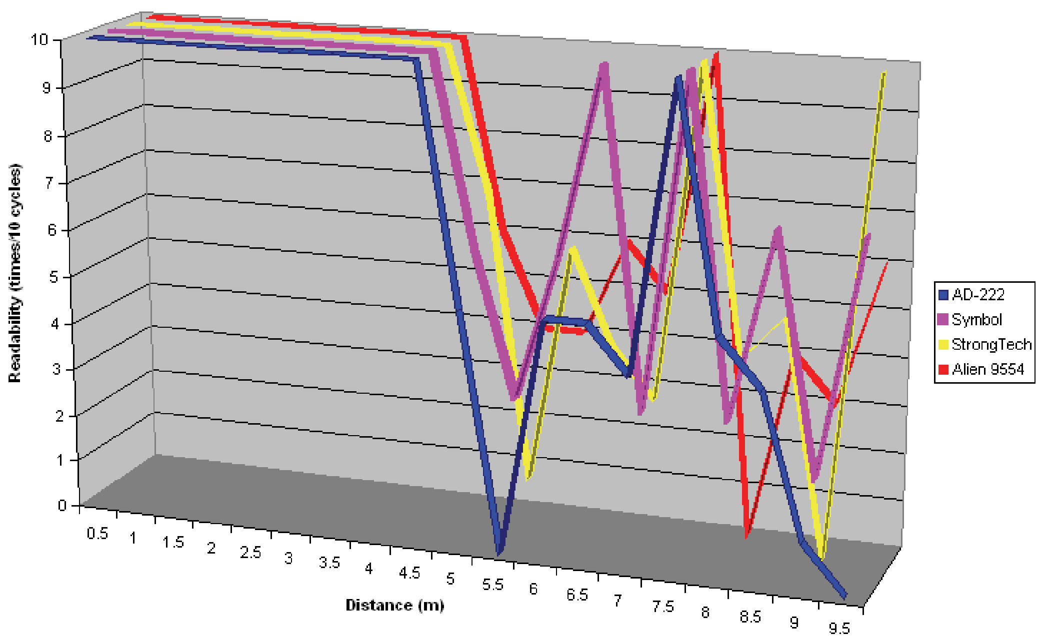


Fig. 8. Free space tag performance test

Then, we attached the tags to paper boxes and put them into a mini van (Figure 9 (a)). The reader, antenna and router were set up on the top of a wood frame on the side of the road. The router communicated with the backend computer wirelessly. We performed the tests in two situations: 1) when the vehicle was stopped; 2) when the vehicle was being driven at slow speeds, 5km/h to 30 km/h (Figure 9 (b)). The experimental results showed that the tag could only be read at an angle, through the door's edge, but not through metal, when the vehicle was stationary and the reader was within a short distance (less than 30 cm).

Lessons to learn
With an off-the-shelf reader and antenna, it is possible to scan and hack EPC labels on products being loaded on a car or truck. Several methods could be useful in protecting against sniffing information from the RFID tags:

- Limit the amount of sensitive information on the tag
- Use a masking technique that masks the structured tag information with a row of zeros
- Encrypt the EPC tag data

(a) Tagged boxes

(b) Reading test

Fig. 9. Sniffing test

6. Conclusion

In this chapter, we focused on the issues and potential solutions for a range of security vulnerabilities of RFID systems. Recent advances in the uses of RFID technology have generated significant interest in society, not only because they have brought change to the industry and business sectors, but also because they will begin to influence our daily life more and more. As mentioned above, the use of RFID has grown exponentially across a variety of core industries, such as logistics, manufacturing, retail and healthcare. Although each application has its own special requirements, security vulnerabilities will be always a major concern when deploying RFID applications. Like the Internet or mobile telephony, RFID is a wireless networking technology. While the non-contact and non-line-of-sight properties of RFID systems increase the convenience and efficiency of their applications, these properties also increase the system's vulnerability. In this chapter, we have analyzed the underlying vulnerabilities that exist in RFID systems, illustrated the threats of possible attacks, and provided corresponding countermeasures. Case studies have been presented and discussed to examine four specific security threats. The objective of the chapter is to try to make life for an attacker very difficult, if not impossible.
The directions for further study are suggested in there major areas: technology standards, authentication protocols, and operational policies/guidelines. Security has not historically been the focus of technology standards for RFID systems and their components. With the increasing usages of RFID, such as passports, personal ID cards and consumer products, potential security threats and compliance risks in the future are enormous. It is necessary to pay attention to standardization of RFID systems. There are many different RFID standards

at the moment. The technology standards typically describe the physical and the link layers, covering aspects such as air interface, anti-collision mechanisms, communication protocols, host interface, data syntax, etc. However, there are some standards that are more important than others — ISO and EPC Global are main contributors in defining RFID standards. EPC tags were originally designed for supply chain and logistical applications. The people who established the EPC standards aimed on low-cost tags with high potential reading rates. Therefore, security was not a high priority issue resulting in the first generation EPC tags lacking the computational resources for cryptographic authentication.

It is important to put more effort into developing authentication protocols for passive RFID. The reasons, as mentioned above, are that 1) current proposed lightweight authentication protocols can still be compromised by attacks; 2) it is difficult or impossible to use computationally intensive cryptographic algorithms for low-cost RFID tags.

A well-designed RFID policy can reduce the risk of attacks. When dealing with security and risk management, policy decisions also play an important role in the security of an RFID system. An RFID security policy is a document that states how an organization plans to protect its physical RFID devices and information data assets. Since, sooner or later, new threats will appear, an RFID security policy should be considered a "living" document that needs to be continuously updated as the RFID technology and implementation requirements change. The policy also needs to take into account how the users will be trained in the proper use of RFID, and explain how security measures will be carried out and enforced.

7. References

Actel Corporation (2002). Design security in nonvolatile flash and antifuse FPGAs security backgrounder, Sunnyvale, California, 2002, from http://www.actel.com/documents/DesignSecurity_WP.pdf

Aragones-Vilella, J.; Martínez-Ballesté, A. & Solanas, A. (2007). A brief survey on RFID privacy and security, *Proc. of the World Congress on Engineering 2007*, pp. 1488-1493, ISBN 978-988-98671-5-7, July 2-4, 2007, London, UK.

Ayoade, J. (2007). Roadmap to solving security and privacy concerns in RFID systems, *Computer Law & Security Report*, Vol. 23, No. 6, pp. 555-561, ISSN: 0267-3649.

Blythe, S.; Fraboni, B.; Lall, S.; Ahmed, H; & de Riu, U. (1993). Layout reconstruction of complex silicon chips, *IEEE Journal of Solid-State Circuits*, Vol. 28 No. 2, pp. 138-145, ISSN: 0018-9200.

Bringer, J.; Chabanne, H. & Dottax, E. (2005). HB++: A lightweight authentication protocol secure against some attacks, *Proc. of the Second International Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing* (SecPerU'06), pp. 28-33, ISBN: 0-7695-2549-0, 29 June 2006, Lyon, France.

Burmester, M.; van Le, T. & de Medeiros, B. (2006). Provably secure ubiquitous systems: Universally composable RFID authentication protocols, *Proc. of the 2nd IEEE/CreateNet International Conference on Security and Privacy in Communication Networks* (SECURECOMM 2006), pp. 1-9, IEEE Press, ISBN: 1-4244-0423-1, 28 Aug.–1 Sep. 2006, MD, USA.

Cavoukian, A. (2006). *Privacy Guidelines for RFID Information Systems (RFID Privacy Guidelines)*, Information and Privacy Commissioner/Ontario, Toronto, ON, June 2006, from http://www.ipc.on.ca/images/Resources/up-rfidgdlines.pdf

Collins, J. (2006). RFID-Zapper shoots to kill, *RFID Journal*, Hauppauge, NY, 23 January 23 2006, from http://www.rfidjournal.com/article/articleview/2098/1/1/

Coverson, L. (2006). Could the new passport trigger a bomb? *ABC News*, New York, NY, 16 August 2006, from abcnews.go.com/Technology/story?id=2319734&page=1

Flexilis (2006). RFID passport shield failure demonstration, *FLX[2006-0605]*, from www.flexilis.com/download/RFIDPassportShieldFailureDemonstration.pdf

Garfinkel, S. & Rosenberg, B. (2005). *RFID: Applications, Security, and Privacy*, Addison-Wesley Professional, ISBN: 0321290968, Boston, MA.

Gilbert, H.; Robshaw M. & Sibert, H. (2005). An active attack against HB+ — A provably secure lightweight authentication protocol, *Cryptology ePrint Archive*: Report 2005/237, from http://eprint.iacr.org/2005/237.pdf

Graafstra, A. (2006). *RFID Toys: 11 Cool Projects for Home, Office and Entertainment*, John Wiley & Sons, ISBN: 0471771961, New York.

Higgins, K. J. (2007a). The five coolest hacks of 2007, *Dark Reading*, Manhasset, NY, 31 December 2007, from
http://www.darkreading.com/document.asp?doc_id=142127&page_number=4

Higgins, K. J. (2007b). Hacking truckers, *Forbes.com*, New York, NY, 25 June 2007, from http://www.forbes.com/2007/06/25/cx_0625darkreading.html?partner=alerts

Hopper, N. J. & Blum M. (2001). Secure human identification protocols, In C. Boyd (ed.) *Advances in Cryptology - ASIACRYPT 2001*, Vol. 2248 of LNCS, pp. 52-66, Springer-Verlag, ISBN 3-540-42987-5, Gold Coast, Australia.

Juels, A. (2003). Privacy and authentication in low-cost RFID tags, *RSA Laboratories*, from http://www.rsasecurity.com/rsalabs/staff/bios/ajuels/

Juels, A. (2005a). RFID security and privacy: A research survey. *IEEE Journal on Selected Areas in Communication*, Vol. 24, No. 2, pp. 381-394, ISSN: 0733-8716.

Juels, A. (2005b). RFID privacy: A technical primer for the non-technical reader, In K. Strandburg & D. S. Raicu (Eds.), *Privacy And Technologies of Identity: A Cross-disciplinary Conversation*, pp 57-74, Springer-Verlag, ISBN: 0387260501, New York, NJ, USA.

Juels, A.; Rivest, R. L. & Szydlo, M. (2003). The blocker tag: Selective blocking of RFID tags for consumer privacy. *Proc. of the 10th ACM Conference on Computer and Communications Security*, pp. 103-111, ACM Press, ISBN: 1-58113-738-9, 27-30 October 2003, Washington D.C., USA.

Juels, A. & Weis, S. A. (2005). Authenticating pervasive devices with human protocols, In V. Shoup (Ed.): *Advances in Cryptology – Crypto 2005*, Vol. 3621 of LNCS, pp. 293–308, Springer-Verlag, ISBN: 3-540-28114-2, Santa Barbara, California, USA.

Karygiannis, T.; Eydt, B.; Barber, G.; Bunn, L. & Phillips, T. (2007). *Guidelines for Securing Radio Frequency Identification (RFID) Systems*, Special Publication 800-98, National Institute of Standards and Technology (NIST), April 2007, Gaithersburg, MD.

Kim, H. W.; Lim, S.Y. & Lee, H. J. (2006). Symmetric encryption in RFID authentication protocol for strong location privacy and forward-security, *Proc. of the 2006 International Conference on Hybrid Information Technology*, pp. 718-723, ISBN 0-7695-2674-8, Nov. 9-11, 2006, Cheju Island, South Korea.

Knospe, H. & Pohl, H. (2004). RFID security, *Information Security Technical Report*, Vol. 9, No. 4, pp. 39-50. ISSN: 1363-4127.

Landt, J. (2001). Shrouds of time: The history of RFID, *An AIM Publication*, Pittsburg, PA, Oct. 2001, from www.rfidconsultation.eu/docs/ficheiros/shrouds_of_time.pdf

Landt, J. (2005). The history of RFID, *IEEE Potentials*, Vol. 24, Issue 4, Oct.-Nov., pp. 8-11, ISSN: 0278-6648.

Langheinrich, M. (2007a) RFID and privacy, In M. Petkovic & W. Jonker (Eds.): *Security, Privacy, and Trust in Modern Data Management*, Springer, ISBN 978-3-540-69860-9, pp. 433-450, Berlin Heidelberg New York, July 2007.

Langheinrich, M. (2007b). A survey of RFID privacy approaches, *The Fifth Workshop on Ubicomp Privacy - Technologies, Users, Policy*, September 16, 2007, Innsbruck, Austria.

Lieshout, M. van; Grossi, L.; Spinelli, G.; Helmus, S.; Kool, L.; Pennings, L.; Stap, R.; Veugen, T.; Waaij, B. van der & Borean, C. (2007). RFID technologies: Emerging issues, challenges and policy options, *Technical Report EN 22770*, European Commission Joint Research Centre Institute for Prospective Technological Studies, ISBN 978-92-79-05695-6, Printed in Spain.

MacGillivray, G. & Sheehan, C. (2006). RFID security, *Semiconductor Insights*, RFID Security Issues Briefing to CANOSCOM, July 27, 2006.

Merritt, R. (2006). Cellphone could crack RFID tags, says cryptographer, *EE Times Online*, Manhasset, NY, February 24, 2006, from http://www.eetimes.com/news/semi/showArticle.jhtml?articleID=180201688

Messerges, T. S.; Dabbish, E. A. & Sloan, R. H. (2002). Examining smart-card security under the threat of power analysis attacks, *IEEE Trans. on Computers*, Vol. 51, Issue 5, pp. 541–552, ISSN: 0018-9340.

MiniMe & Mahajivana (2005). RFID-Zapper (EN), *the 22nd annual Chaos Communication Congress (22C3)*, Berlin, Germany, 27-30 December 2005, from http://events.ccc.de/congress/2005/wiki/RFID-Zapper(EN)

Newitz, A. (2006). The RFID hacking underground, *Wired*, May 14, 2006, from http://www.etdot.com/2006/05/09/the-rfid-hacking-underground/

Ngai, E. W. T.; Moon, K. K. L.; Riggins, F. J. & Yi, C. Y. (2008). RFID research: An academic literature review (1995-2005) and future research directions, *International Journal of Production Economics*, Vol. 112, No. 2, pp. 510-520, ISSN: 0925-5273.

Nohl, K.; Evans, D.; Starbug & Plotz, H. (2008). Reverse-engineering a cryptographic RFID tag. *USENIX Security Symposium*, San Jose, CA, 31 July 2008, from http://www.cs.virginia.edu/~evans/pubs/usenix08/usenix08.pdf

Norton, M. (2006). RFID security issues, *Presentation at the Wireless and RFID 2006 Conference*, Feb. 27-Mar. 1, 2006, Ronald Reagan Building, Washington, DC.

O' Brien, D. (2006). RFID - Introduction and security considerations, *Presentation at the ISS World*, Washington, DC., Dec. 4-6, 2006.

Oertel, B.; Wolk, M; Hilty, L.; Kohler, A.; Kelter, H.; Ullmann, M. & Wittmann, S. (2005). Security aspects and prospective applications of RFID systems, *Federal Office for Information Security*, Bonn, January 2005, from http://www.bsi.bund.de/fachthem/rfid/RIKCHA_englisch.pdf

Piramuthu, S. (2006). HB and related lightweight authentication protocols for secure RFID tag/reader authentication, *Proc. of CollECTeR Europe*, Basel, Switzerland 9-10 June 2006, from http://www.collecter.org/archives/2006_June/23.pdf

Ranasinghe, D. C. & Cole, P. H. (2006). Confronting security and privacy threats in modern RFID systems, *Proc. of the Fortieth Asilomar Conference on Signals, Systems and*

Computers, pp. 2058-2064, ISBN: 1-4244-0785-0, Oct. 29 –Nov.1, 2006, Pacific Grove, California.

Rieback, M. R.; Crispo, B. & Tanenbaum, A. S. (2006a). The evolution of RFID security, *IEEE Pervasive Computing*, Vol. 5, No. 1, pp. 62-69, ISSN: 1536-1268.

Rieback, M. R.; Crispo, B. & Tanenbaum, A. S. (2006b). Is your cat infected with a computer virus? *Proc. of the 4th Annual IEEE International Conference on Pervasive Computing and Communication*, pp. 169-179, ISBN: 0-7695-2518-0, 13-17 March 2006, Pisa, Italy.

Rieback, M. R.; Simpson, P. N. D.; Crispo, B. & Tanenbaum, A. S. (2006c). RFID viruses and worms, *Vrije Universiteit Amsterdam*, 02 March 2006, from http://www.rfidvirus.org/

Rotter, P. (2008). A framework for assessing RFID system security and privacy risks, *IEEE Pervasive Computing*, Vol. 7, Issue. 2, pp. 70-77, ISSN: 1536-1268.

Schreur, R. W.; van Rossum, P.; Garcia, F.; Teepe, W.; Hoepman, J. H.; Jacobs, B.; de Koning Gans, G.; Verdult, R.; Muijrers, R.; Kali, R.; Kali, V. (2008). Security flaw in Mifare Classic, *Press release, Digital Security Group*, Radboud University Nijmegen, 12 March 2008, from http://www.sos.cs.ru.nl/applications/rfid/pressrelease.en.html

Thornton, F.; Haines, B.; Das, A. M.; Bhargava, H. & Campbell, A. (2006). *RFID Security*. Syngress Publishing, Inc., ISBN: 1597490474, Rockland, MA, USA.

Tsudik, G. (2006). YA-TRAP: Yet another trivial RFID authentication protocol, *Proc. of the 4th IEEE International Conference on Pervasive Computing and Communication Workshops*, pp. 640-643, IEEE Computer Society Press, ISBN: 0-7695-2520-2, 13-17 March 2006, Pisa, Italy.

Tuyls, P. & Batina, L. (2006). RFID-tags for anti-counterfeiting, In D. Pointcheval (Ed.), *Topics in Cryptology – CT-RSA 2006*, Vol. 3860 of LNCS, pp.115-131, Springer-Verlag, ISSN: 0302-9743, San Jose, CA, USA.

van der Togt, R.; van Lieshout E. J.; Hensbroek R.; Beinat E.; Binnekade, J. M. & Bakker, P. J. M. (2008). Electromagnetic interference from radio frequency identification inducing potentially hazardous incidents in critical care medical equipment, *Journal of the American Medical Association (JAMA)*, Vol. 299, No. 24, pp. 2884-2890, ISSN: 0098-7484.

Welch, D. & Lathrop, S. (2003). Wireless security threat taxonomy, *Proc. of the 2003 Workshop on Information Assurance*, pp. 76 – 83, ISBN: 0-7803-7808-3, 18-20 June 2003, West Point, NY.

Whiting, R. (2004). Focusing on E-payments at Medicare, Medicaid, *InformationWeek*, April 5, 2004, from http://www.informationweek.com/ story/showArticle.jhtml?articleID=18900188

Wyld, D. C. (2006). RFID 101: The next big thing for management, *Management Research News*, Vol. 29, No. 4, pp. 154-173, ISSN: 0140-9174.

Xiao, Q.; Boulet, C. & Gibbons, T. (2007). RFID security issues in military supply chains, *Proc. of the Second International Conference on Availability, Reliability and Security*, pp. 599-605, ISBN: 0-7695-2775-2, 10-13 April 2007, Vienna, Austria

20

Ontology and Its Application in Supply Chain Information Management

Zetian Fu[1], Jun Yue[1, 2] and Zhenbo Li[3]
[1]*China Agricultural University*
[2]*LuDong University*
[3]*Telecom & Management SudParis*
[1,2]*China*
[3]*France*

1. Introduction

Since ontology was put forward by Neches and Fikes in 1991, the research on ontology is becoming increasingly matured after more than 10 years development. Although usages on concepts and terms related with ontology do not keep completely identical, the agreement of actual use has emerged. Most researches are concentrated on theoretic in current ontology research, while practical application of ontology is still less.

Ontology takes the role as a philosophical concept originally, which regarded as an explanation or description to an objective existing system from the philosophy category, and concerns the abstract nature of the objective reality. Later, it is given a new definition by the artificial intelligence category with the development of artificial intelligence. Neches and Fikes (1991) define ontology as the "explicit formal specifications of the terms in the domain and relations among them". The terms and relations are abstract description of phenomenon happened in the word. Gruber (1993) gives a most popular definition to ontology:"ontology is the clear specification on a conceptual model". Based on this definition, another new concept of ontology is given by Borst, i.e. (1997) as "ontology is the clear specification on a sharing conceptual model". Studer (1998) thinks that ontology consists of four layers meaning after an in-depth studying on the two above definitions, which are "conceptualization", "explicit", "formal" and "share". Conceptualization refers to models obtained by abstracting some related phenomenon in the objective world, which performs independent of the specific state of the environment. Explicit refers to explicit definitions to used concepts and their restriction. Formal refers to the ontology model can be processed by computer. Share refers to ontology aims at group consensus instead of individual consensus. The ontology goal is to capture the knowledge of relevant domain, provide a common understanding to the knowledge, confirm vocabularies of the domain and define the relationships between the vocabularies clearly.

Some scholars regard ontology as an approach to construct knowledge base. There are different ontology category modes according to different ontology attributes. Gruber (1993) divides ontology into four classes as its level of detail and level of dependence: top level ontology, domain ontology, task ontology and application ontology. Ontology is divided into meta-ontology, general ontology, domain ontology and application ontology as the different applied situations. Mizoguchi etc. (1995) divides ontology into domain ontology, general ontology and task ontology.

Perez etc. (1999) think that ontology can be organized by taxonomy. Ontology consists of five basic modelling primitive, which are “class”, “relations”, “function”, “axioms” and “instance”. In general terms, class is called concept as well. Concept (class) is with very wide meaning, which can be considered as anything. For example, work description, function, action, strategy and reasoning processes etc.. Relationship represents the interacting effect between two concepts in the domain, which is defined as a subset of n-dimensions Cartesian set: $R : C_1 \times C_2 \times \cdots \times C_n$. Function shows a kind of specific relationship, the element *n* can be decided by other *n-1* elements. The formalized definition is as: $F : C_1 \times C_2 \times \cdots \times C_{n-1} \rightarrow C_n$. Axioms represent tautology. Instance refers to elements. Analyzing it form the view of semantic, instance describe objects, while concept mean the set of objects, relations correspond to the set of object element group. The definitions to concept adopt frame structure, which consists of the names of concepts, relations to other concepts and descriptions to concepts in natural language. There are four relations of ontology: “part-of”, “kind-of”, “instance-of” and “attribute-of”. Part-of describes the relationship of part and whole. Kind-of describes the inheritance between concepts, similar to the relationship of parent and child of object-oriented. Instance-of describes the instance of concept and concept itself, similar to the relationship of object and class of object-oriented. Attribute-of describes one concept taking a role as attribute of another concept. In actual application, it is unnecessary to apply the meta-language above exactly to construct ontology. And meantime, relationships between concepts are not limited to the four types listed above. The corresponding relationships can be defined according to the status of specific domain to satisfy the requirement of application.

There are five widely applied ontology as below: WordNet(2006), Framenet(2008), GUM (Bateman *et. al.,* 2001), SENSUS(Knight, K. & S. Luk.,1994), Mikrokmos(1996). WordNet is an English dictionary based on psychological language rules, taking synsets as a unit to organize information. Framenet is another English dictionary, adopting description frame called Frame Semantics and providing the semantic analyzing ability. GUM adopts natural-language-oriented processing and supports multi-language processing. SENSUS consisting of more than 70000 concepts adopts natural-language-oriented processing and provides concept-structure to computer understanding. Mikrokmos adopts interlingua TMR to express knowledge, orients natural-language processing and supports multi-language processing.

It is widely recognized that the participation of experts is necessary in domain ontology construction. The process is various proceeding form different ontology projects. Since there is not a standard approach of ontology construction, some standards benefiting to ontology construction are put forward from practice. The most effective is pointed out by Gruber (1995), which consists of five rules: explicit and objective, perfectibility, consistency, expansibility of maximum monotonicity and minimum stability. Explicit and objective refers to that ontology should give out explicit and objective semantic definition using natural language. Perfectibility refers to that the definition should be perfect, and the meaning of a term can be perfectly expressed. Consistency refers to that the deduction from a term should be consistent with the term itself. Expansibility of maximum monotonicity refers to that it is unnecessary to modify the previous ontology when some new general or specific terms are added to it and minimum stability refers to that modelling object will be given as few as restricts.

With the emergence and application of ontology in various domains, some kinds of ontology construction approaches appeared. The main methods relevant to well-known ontology projects and its modelling processes includes frame methodology (T. R. Gruber, 1993), TOVE ontology and Gruninger and Fox's methodology(M. Gruninger & M. Fox, 1995), KACTUS and Bernaras methodology(A. Th. Schreiber *et. al.* 1995), MHONTOLOGY methodology (Fernández-López. M. *et. al.* 1997, 1999), SENSUS ontology and methodology (Knight, K. & S. Luk, 1994), IDEF5 methodology(Benjamin, P. C. et. al. 1994) and seven steps methodology (Corcho. O. & Gómez-Pérez. A., 2000). Seven steps methodology was put forward by School of Medicine of Stanford University. The seven steps are confirming the domain scope, investigating the possibility to use the previous ontology, listing important terms of ontology, defining classes and their grading system, defining attributes of classes, defining distribution of attributes and establishing instances.

2. Formal representation of ontology models

The formalized ontology language provides a possibility for users to describe concepts of domain model explicitly and formally. Therefore, it should meet the following requirements: a well-defined syntax, a well-defined semantic, efficient reasoning support, sufficient expressive power, convenience of expression.

Till now, there appear some kinds of ontology formalized languages, such as RDF and RDF(S), OIL, DAML, OWL, KIF, SHOE, XOL, OCML, Ontolingua, CycL, Loom. RDF(S) and OWL are wildly used formalized languages.

2.1 RDF(S) language

RDF, RDF(S) (The Resource Description Framework) is a language for representing information about resources in the World Wide Web, which is a recommended standard based on XML by W3C. A simple model is put forward from RDF to express any types of data and the data type consists of marked joint arcs between nodes, which display sources on Web. The arcs display attributes of sources. Therefore, the data model can describe objects and their relationships. RDF data model is an expression of a duality-relationship. As any complex relationships can be decomposed into several simple duality-relationships, so RDF data model can be taken as a basic model for any complex model.

RDF(S) is complementary with XML. First, RDF intends to standardize XML by standardizing and inter-operating modes. XML documents can be cited by RDF in a simple way. Next, since RDF expresses data semantics in a model constructing way, RDF can not be restricted by specific grammar. However, RDF still needs a suitable grammar format to realize its application on Web. Serializing RDF to XML expression can make RDF to obtain a better processing, and make RDF data can be used, transferred and stored as easy as XML. Combination of XML and RDF benefits to data retrieval and discovery of relevant knowledge. It not only realizes the description of data based on semantics, but also fully displays merits of XML and RDF(S).

Similar to tags of XML, property set of RDF(S) have no restricts. In another word, there are thesaurus and polysemy phenomena. RDF(S) can not solve the problems. Although RDF(S) can provide the glossary for it properties and types, data semantic description has semantic

conflicts. In order to clear up the semantic conflicts, a further restrict to semantic description result has been done by citing relevant ontology technology. Fortunately, on the time of providing semantic model understandable to computer, RDF also provides modeling base to ontology language (OIL, OWL) for a certain domain. Thus the application based on RDF(S) can be conveniently combined with the ontology expressed by those ontology languages. The characteristic of RDF(S) makes semantic describing results to have ability to interact with some more domain knowledge, also makes Web data description based on XML and RDF to have a perfect vitality.
In a short, one RDF document includes various source descriptions, while one source description consists of several RDF sentences which are triples consisting of source, property type and property value to express a property of source. The sentences in source description correspond to those sentences of natural, the source corresponds to the subject of natural, the property type corresponds to predicate, and property value corresponds to object. As sentences of natural language can be passivity, so the simple correspondence above is just an analogy of concepts.

2.2 OWL language

OWL (Web Ontology Language) (Grigoris Antoniou & Frank van Harmelen, 2004) is a standard of ontology description language in semantic internet recommended by W3C, which is developed from a combined description language (DAML+OIL) in some research institution in Europe and America. DAML is from an American overture DAML-ONT, while OIL is a kind of ontology description language from Europe. OWL is on the top layer of the ontology language stack put forward by W3C, which is shown in fig.1.

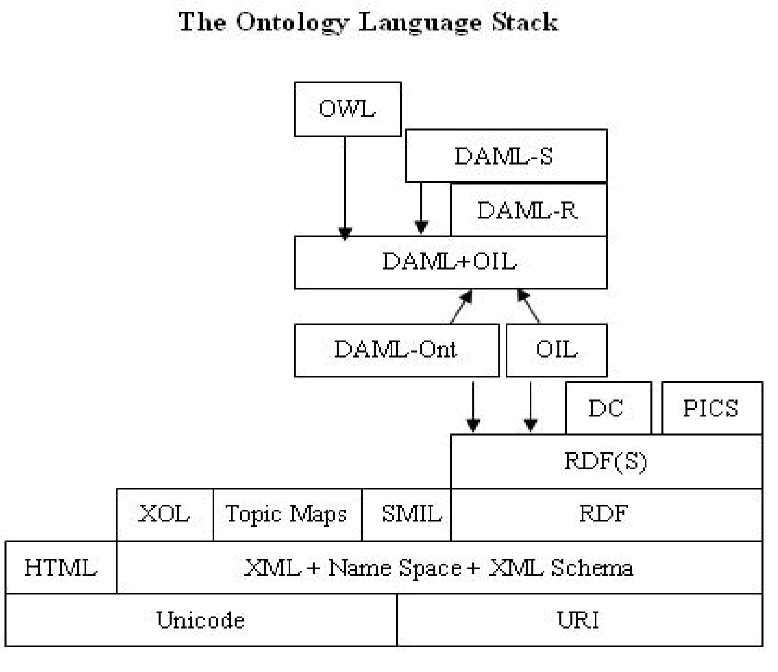


Fig. 1. Ontology language structure

Aiming at various requirements, there are three sublanguages as shown in table 1.

Sublanguage	Description	Instance
OWL Lite	Provided to users who need only one classified layer and simple restrict of property.	Support cardinality and permit it can only be 0 or 1.
OWL DL	Support those who need to express to the maximum extent in the reasoning system, which can ensure the computational completeness and decidability. It contains all restricts of OWL, but it can only be put into certain restricts.	When a class can be a sub-class of multi-classes, it can not be an instance of another class.
OWL Full	Support those who need to express to the maximum extent under free-grammar RDF without calculating guarantee. It permits vocabulary to be added to the pre-defined glossary of an ontology.	A class can be expressed as a set of individuals, and also an individual of a set.

Table 1. Three sublanguages of OWL

The relationships between these three sublanguages are:

- Each suitable OWL Lite is a suitable OWL DL and each suitable OWL DL is a suitable OWL Full.
- Each effective OWL Lite conclusion is an effective OWL DL conclusion.
- Each effective OWL DL conclusion is an effective OWL Full conclusion.

The considerations in choosing which language to use are:

- Choosing OWL Lite or OWL DL mainly depends on the expressiveness degrees of restriction.
- Choosing OWL Lite or OWL Full mainly depends on the requirement degrees of RDF meta-model mechanism.
- While using OWL Full rather than OWL DL, it is unpredictable to support reasoning as there is no complete realization of OWL Full.

The relationships between these three sub-language and RDF(S) are:

- OWL Full can be taken as the expandation of RDF(S).
- Both OWL Lite and OWL Full can be taken as a restricted expandation.
- All OWL documents (Lite, DL, Full) are RDF documents.
- All RDF documents are OWL Full documents.
- Only a part of RDF documents are suitable OWL Lite and OWL DL documents.

3. Design and implementation of supply chain information management system based on ontology

3.1 System architecture

The knowledge management system is to gather relevant knowledge, process the gathered knowledge in semantic level, and provide service.

The system is divided into three parts, Knowledge Gathering, Knowledge Processing and Knowledge Service. Which are shown in fig.2.

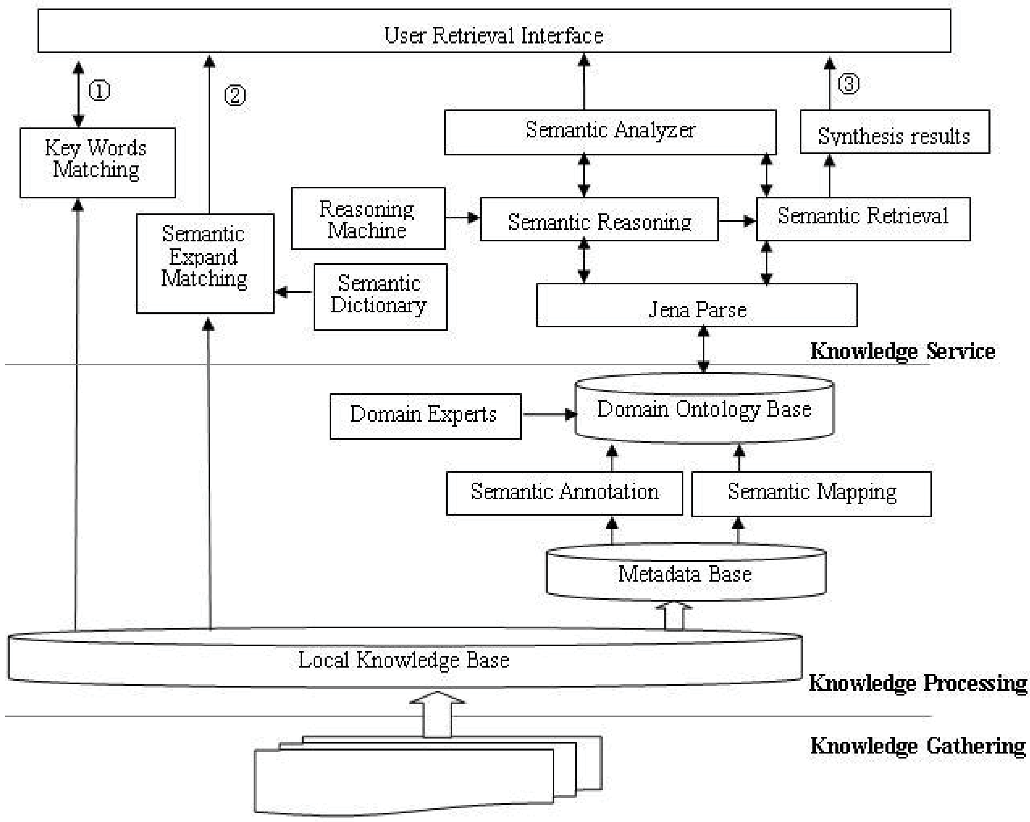


Fig. 2. System architecture

3.2 System module design

The system provides three knowledge acquisition modes which are keywords retrieval, semantic extension retrieval and ontology based retrieval.

3.2.1 Keywords retrieval

Keywords retrieval is a traditional information retrieval mode. At present, many well-known search engines such as Google (www.google.com) and Baidu (www.baidu.com) etc. query information by keywords submitted by users.

Keywords retrieval uses a group of representative keywords (indexing term) to describe each database item. The keywords index can be established in order to improve the efficiency of retrieval generally. Keywords are selected according to the content of database items. Moreover, each keyword can be given a weight to describe its importance.

Keywords retrieval has the same obviously merits and weaknesses. Merits are simple, with faster retrieval speed. The main shortcomings are:

It is hard to express retrieval intent. In general terms, the retrieval intent can be expressed by one or several simple keywords hardly, which lead to the retrieval effect not high.

As the polysemy and synonymy phenomena of language, the problem of thesaurus retrieval can not be solved easily. Such as "Apple" can be understood as either apple or the famous

brand of Apple Inc. In addition, as the differences in cultural and educational background etc., different keywords may be chosen for the same information retrieval.
Another problem is “Information Island”. Keywords can only express the description of original data items instead of the real content of data items, which leads to that the relevant information of a concept can not reflect the intrinsically links of concepts. Therefore, the information relevant to a data item can not be obtained only by the data item itself in retrieval. Moreover, in keywords retrieval, the too much pursuit to the retrieval recall leads a large number retrieval results. This makes users hard to analyze and use the results.

3.2.2 Semantic extension retrieval

In order to solve the shortcomings of keywords retrieval, semantic extension retrieval is introduced to the system. Consulting the modus operandi of WordNet and analyzing domain concepts semantically the semantic vocabulary dictionary is established to describe the relationship of domain concepts.
When semantic retrieval is going on, a query condition sets which contains a group of relevant concepts is obtained for users’ initial query conditions by semantic vocabulary dictionary semantic expand, semantic intension, semantic extension and semantic association etc.. In the process of semantics explanation, the semantic reasoning can be done according to descriptions to relationships between concepts.

3.2.3 Ontology based semantic retrieval

Ontology based supply chain knowledge semantic retrieval is the focus module of systems development. For this, supply chain ontology model has to be established firstly.

3.2.3.1 Construct domain ontology

RDF (Resource Description Framework) has been taken as the best choice to express and process semi-structured data, and it has become the W3C recommended standard of both XML and SOAP.
The core of expressing domain ontology by RDF is to construct triple description. In another word, describe complex things as a series of triples. Each statement of RDF contains three parts-main body, predication and object, and the core description of RDF is relationships between things. In RDF model, the definition to relationship is different with it in other system, such as Object-Oriented system. It considers that relationship exists forever in the world and the source object consist of various kinds of complex relationships, while Object-Oriented considers that the world consists of many sources and relationships exist depending on the specific source. Two different points of view make them applied in different environments, and the RDF model shows more powerful description ability. Therefore, RDF model is chosen as the reference in ontology model semantic coding. Another reason for storing ontology model as RDF format is that the domain ontology model can be parsed in Java program by Jena tool kit.

3.2.3.2 Parse domain ontology by Jena

In order to realize the semantic retrieval based on ontology, Jena is adopted to parse and use the constructed ontology, which is the precondition of semantic reasoning based on ontology model.
At semantic retrieval based on ontology, retrieval requires submitted by users are transferred into RDF source objects, by which the more proper ontology model will be obtained.

At the time of semantic processing to ontology model, relevant metadata are processed according to domain ontology and reasoning rules to obtain the connotative information, which services the following operations. Reasoning rules can be reinforced according to the actual requirements, and axiom reasoning and theorem reasoning will be required in this processing. However, axiom reasoning should be used as much as possible and reduce the percentage of theorem reasoning.

4. A case-ontology based vegetable supply chain information system

The vegetable supply chain knowledge management system is to gather vegetable supply chain relevant knowledge, process the gathered knowledge in semantic level, and provide service.

4.1 System ontology model

The system ontology model consists of vegetable supply chain ontology model, vegetable supply chain knowledge ontology model and vegetable supply chain knowledge user ontology model.

4.1.1 Vegetable supply chain knowledge ontology model

There are three sub-sets which are nominal class subset, individual and organization class in vegetable supply chain and verbal class subset. The vegetable supply chain is marked Veg.SCM in the system.

4.1.1.1 Nominal concept subset (Norminal.SC)

Individual and Organization in Vegetable Supply Chain
    Instance: Mr.Zhang (producer), Import&Export Company,…
Vegetable
    Frozen Vegetable
        Instance: frozen mushroom, frozen green soy bean,…
    Fresh Vegetable
        Putrescible Vegetable
            Instance: tomato, cucumber,…
        Disputrescible
            Instance: potato, onion, cabbage,…
Contract
    Exchange Contract
    Transportation Contract
    Process Contract
Time
    Instance: Jan.10th, 2005…
Address
    Instance: Huangshan Road, Shouguang, China…
Person
    Instance: Tongxin Zhang…
Organization
    Instance: China Agricultural University…
Vehicle
    Instance: Huanghe truck…
…….

4.1.1.2 Individual and organization class in vegetable supply chain (Individual-Organization.SC)

The individual and organization class in vegetable supply chain subset is shown in fig.3.

Class	1st 2nd 3rd 4th Sub-class	Instance
Individual / organization	Vegetable producer	
	Non-contract-individual	Mr.Zhang
	Cooperative producer	Mr.Li, some famers…
	Contract producer	Some vegetable produce base…
	Merchant	
	purchase merchant	
	Non-contract	Mr.Wang, one company…
	With-contract	Mr.Qian, one company…
	sale in China	
	Whole sale	Mr.Zhao, one company…
	Retail	
	Individual farmer	Mr.Tian…
	Saler in wet market	Mr.Gao…
	Super market	Wal-mart supermarket…
	Exportant	
	User	
	Restaurant	…
	Family	…
	Others	…
	Process Individual / organization	
	Briefly	…
	Process factory	…
	Transportation	
	Individual	…
	Company	…
	Storage Individual / organization	…

Fig. 3. Individual and organization model

4.1.1.3 Verbal class subset (Verbal.SCM)

Concepts in verbal class subset have the character of verb, but they can be used as noun. They are also called duality concepts. They all have a common property-object, but not are noted very clearly. The values of property are class or instance of individual and organization class in vegetable supply chain. The verbal class subset is as below.

Process
Transportation
Storage
Exchange
…….

4.1.2 Vegetable supply chain knowledge ontology model

Knowledge organization and expression are crucial to the knowledge management system. In order to provide users the convenient and speedy knowledge retrieval, the vegetable supply chain knowledge is organized together as the following mode.

The vegetable supply chain knowledge is divided into four sub-classes: Basic Knowledge, Academic Knowledge, Practical Knowledge and Case Study. Each class has the properties of their own. It is shown in the following figure.

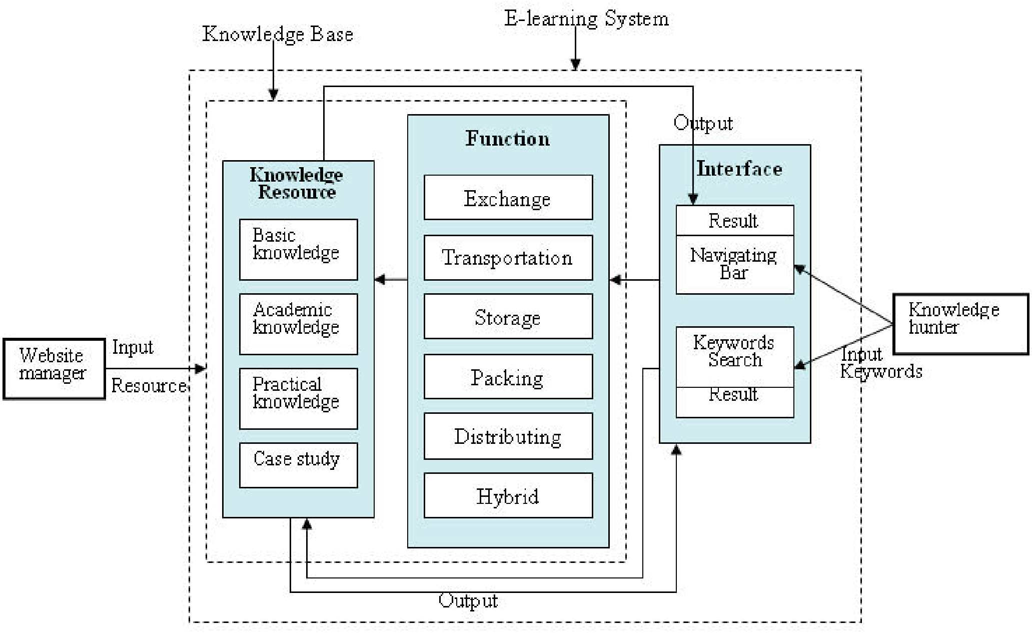


Fig. 4. Framework of Veg. SC knowledge

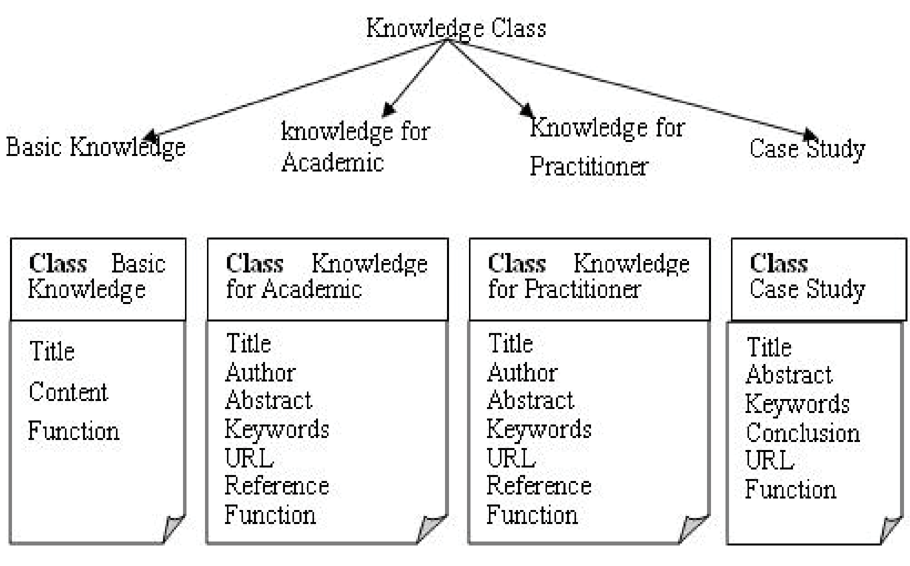


Fig. 5. Knowledge classes and their properties

The vegetable supply chain is divided into five classes considering its functions, exchange, transportation, storage, packing, and hybrid, which are the values or instances of individual and organization class of vegetable supply chain ontology. It is shown in fig.6.

Combining four knowledge classes and five functions, the knowledge retrieval process is shown in fig.7 when the requiring intent is submitted.

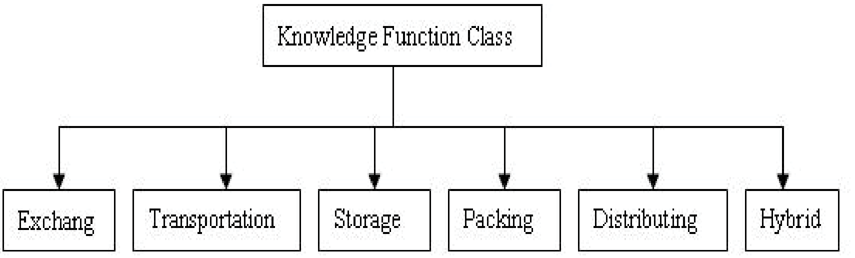


Fig. 6. Function category

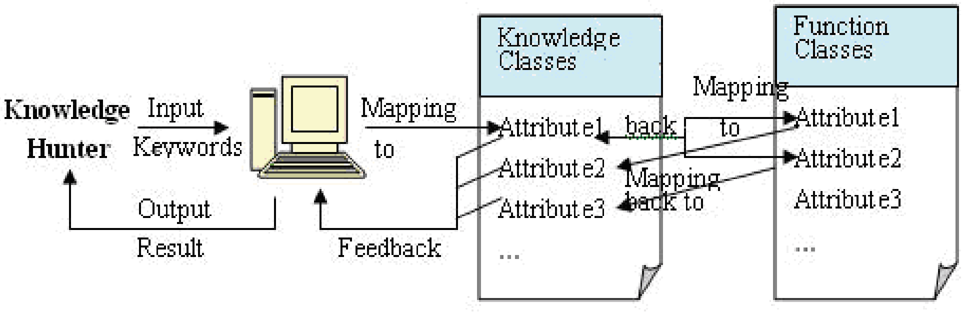


Fig. 7. Retrieval process

4.1.3 Vegetable supply chain knowledge user ontology model

The final user is divided into two classes, academic user and practical user. The academic user pays more attention to theoretical knowledge and case study, while practical user cares more about basic knowledge, application knowledge and case study, which is shown in fig.8.

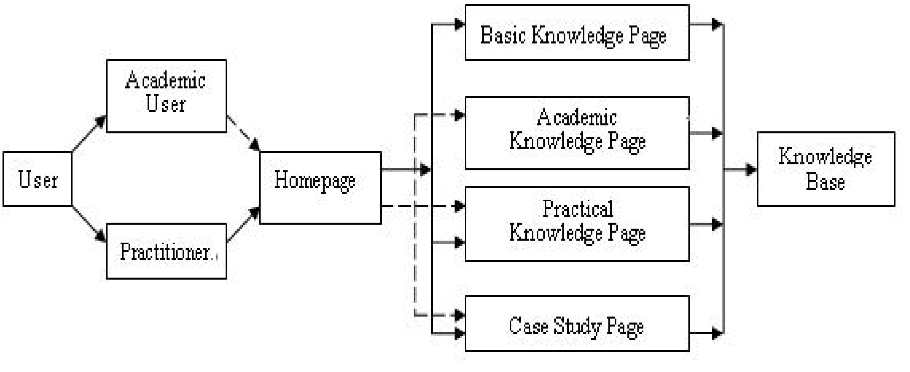


Fig. 8. User classes and their considering knowledge scopes

4.1.4 Integration of ontology models

There are four relationships between two classes which are: “part_of”, “kind_of”, “attribute_of” and “instance_of”. By these four relationships, we integrate all ontology models as a whole which are shown in fig.9.

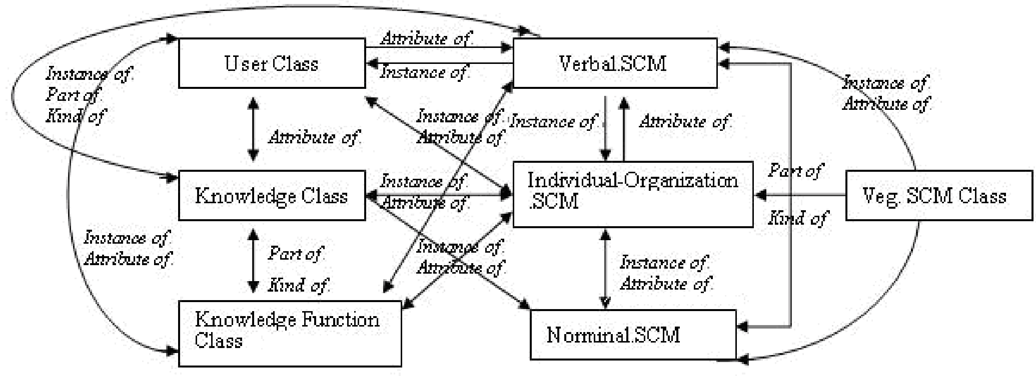


Fig. 9. System ontology model

4.2 Formalization of system ontology

After constructing vegetable supply chain knowledge ontology model, vegetable supply chain knowledge user ontology model and vegetable supply chain ontology model, these ontology models are formalized in order to realize semantic reasoning. RDF(S) and OWL exported by Protégé are adopted here, of which OWL formalization expresses the structure formalized by RDF(S).

4.2.1 RDF(S) formalization

To illustrate the usage of Veg. supply chain knowledge management system, we use our project web site (http://icb.cau.edu.cn/vegnet) as a data-intensive website example. As a website instance, the Veg. supply chain knowledge searching system website contains an index page and a list of resources to present information. Fig.9 shows a fragment of RDF statements describing our website. (The namespace prefix 'vegnet' refers to the namespace of Veg. Supply Chain Knowledge System site ontologies:
xmlns:vegnet=”http://icb.cau.edu.cn/vegnet/e-learning/siteontology/”).

```
<rdf:Description rdf:about="http://icb.cau.edu.cn/vegnet/e-learning">
<rdf:type rdf:resource="http://icb.cau.edu.cn/vegnet/e-learning/siteontology#Site"/>
<vegnet:IndexResource rdf:resource="http://icb.cau.edu.cn/vegnet"/>
<vegnet:Resource>
<rdf:Bag>
<rdf:li rdf:resource=" http://icb.cau.edu.cn/vegnet/e-learning/bsicknowledge"/>
<rdf:li rdf:resource=" http://icb.cau.edu.cn/vegnet/objective"/academicknowledge>
<rdf:li rdf:resource=" http://icb.cau.edu.cn/vegnet/e-learning"/practicalknowledge>
<rdf:li rdf:resource=" http://icb.cau.edu.cn/vegnet/evens"/casestudy>
...
</rdf:Bag>
...
</vegnet:Resource>
</rdf:Description>
```

Fig. 10. Description of the website

We describe the knowledgebase more detail in fig.10. Knowledge class is divided into six sub-classes: Computer Basic Knowledge, Computer Application Knowledge, Vegetable Supply Chain Basic Knowledge, Vegetable Supply Chain Academic Knowledge, Vegetable Supply Chain Practical Knowledge and Vegetable Supply Chain Case Study. Function is a 2nd sub-class of Basic Knowledge and its value is a sub-class of Knowledge Function Class.

```
<?xml version="1.0"?>
<rdf:RDF
xmlns rdf=http://www.w3c.org/1999/02/22-rdf-syntax-ns#
xmlns rdfs="http://www.w3c.org/2000/01/rdf-schema#">

<rdf:Description rdf:ID="KnowledgeClass">
<rdf:type rdf:resource= "http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>

<rdf:Description rdf:ID="Computer BasicKnowledge_ KnowledgeClass">
<rdf:type rdf:resource= "http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="#KnowledgeClass"/>
</rdf:Description>

<rdf:Description rdf:ID="Computer Applicationknowledge_ KnowledgeClass">
<rdf:type rdf:resource= "http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="#KnowledgeClass"/>
</rdf:Description>

<rdf:Description rdf:ID="VegetableSupplyChain BasicKnowledge_ KnowledgeClass">
<rdf:type rdf:resource= "http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="#KnowledgeClass"/>
</rdf: Description>

<rdf:Description rdf:ID=" VegetableSupplyChain AcademicKnowledge_ KnowledgeClass">
<rdf:type rdf:resource= "http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="#KnowledgeClass"/>
</rdf:Description>

<rdf:Description rdf:ID=" VegetableSupplyChainPracticalKnowledge_ KnowledgeClass">
<rdf:type rdf:resource= "http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="#KnowledgeClass"/>
</rdf:Description>

<rdf:Description rdf:ID=" VegetableSupplyChain CaseStudy_ KnowledgeClass">
<rdf:type rdf:resource= "http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="#KnowledgeClass"/>
</rdf:Description>

<rdf:Description rdf:ID="KnowledgeFunctionClass">
<rdf:type rdf:resource= "http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>

<rdf:Description rdf:ID="Function">
<rdf:type rdf:resource= "http://www.w3c.org/1999/02/22-rdf-syntaxns#
Property"/>
<rdfs:domain rdf:resource= "# BasicKnowledge_ KnowledgeClass"/>
<rdfs:range rdf:resource="# KnowledgeFunctionClass"/>
</rdf:Description>

</rdf:RDF>
```

Fig. 11. Description of the website knowledge

4.2.2 Formalize ontology by protégé and OWL

Protégé software (2008) can be used to construct vegetable supply chain domain ontology. Classes are organized as hierarchy in Protégé, each class has it own sub-class, and the sub-class has it own property. The property in Protégé is described by slot, and the current domain is decided by domain option. Part class structure is defined as fig.12 shown.

Fig. 12. Class structure represented by protégé

Properties of class are described by slot in Protege. The process of constructing slot is similar to that of constructing class. Thereinto, the choice of defaults can be used to setup and inherit the property class and its default value. And the domain option is used to confirm the domain of present slot. Fig.13 shows the property slot of Contract Producer sub-class.

OWL format is chosen as the document storage type when Protégé stores ontology. OWL can be used to express items of glossary and relationship between these items, which is called ontology. When the stored ontology document needs to be cited and understood by a computer, OWL will be introduced.

At the time of storing the constructed ontology in OWL format, an initial standard module of OWL is contained in a series of namespace statements of rdf:RDF tag. These statements are adopted to exactly explain identifiers of the document, by which the other parts of ontology can also be understood. A typical namespace statement is shown as below.

```
<rdf:RDF
   xmlns="http://www.example.org/wine#"
   xmlns:vin ="http://www.example.org/wine#"
   xmlns:food="http://www.example.org/food#"
```

```
   xmlns:owl ="http://www.w3.org/2002/07/owl#"
   xmlns:rdf ="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
   xmlns:xsd ="http://www.w3.org/2000/10/XMLSchema#"
xmlns:dte ="http://www.example.org/wine-dt#" >
```

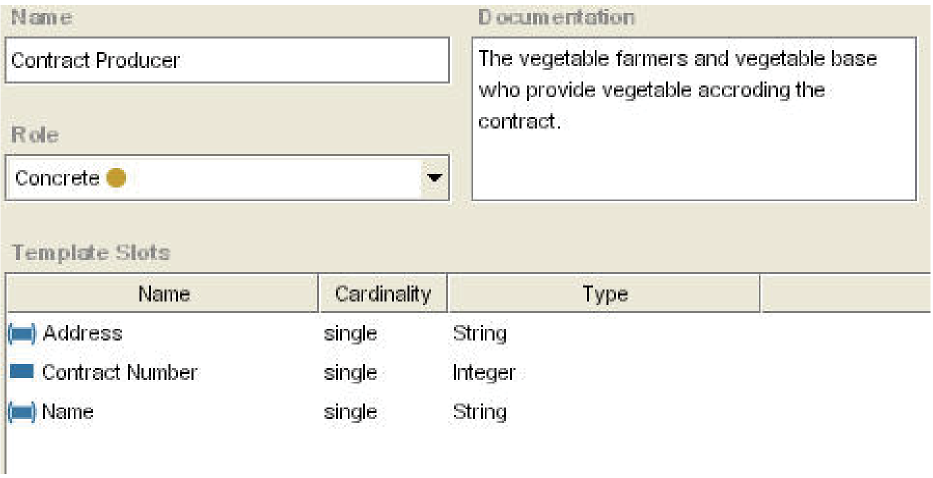


Fig. 13. Property slot of contract producer class

4.3 System development

4.3.1 System development environment and tools

System environment： Windows XP (SP2)
JAVA environment：JDK1.5
Implement tool： Eclipse3.1
Server：Tomcat5.0
Semantic network tools：Jena 2.3, Protégé 3.1

4.3.2 Parse domain ontology by Jena

Jena is adopted to parse vegetable supply chain ontology stored in OWL format in order to realize the semantic retrieval based on ontology. The kits used in parsing process are as below.

```
java.lang.*
java.lang.String.*
java.util.*
java.io.*
com.hp.hpl.jena.rdf.model.*
com.hp.hpl.jena.util.*
com.hp.hpl.jena.rdf.*
com.hp.hpl.jena.ontology.*
com.hp.hpl.jena.reasoner.*
com.hp.hpl.jena.vocabulary.*
com.hp.hpl.jena.reasoner.rulesys.*
```

The first step that Jena parses vegetable supply chain ontology is to read ontology model into memory, before which the createDefaultMode( ) in ModelFactory Class is adopted to create an empty model base on memory storage. Jena also contains other realization modes of Model interface. For example, the Model interface can be created by ModelFactory for modes using relational database.

```
Model model = ModelFactory.createDefaultModel();
```

After the empty model created, the function read of Model interface is adopted to read the domain ontology constructed by Protégé into memory.

```
model.read(new InputStreamReader(vegetable supply chain ontology model document), "");
```

Then a source will be created. The most obvious character of intelligent information retrieval is that source is introduced in the retrieval process. Source can be considered as whatever will be understood, and marked by Uniform Resource Identifier (URI).

```
Resource  myresource=model.createResource();
```

Source holds its property, and each property has its own value. The name of the property is an URI as well.

When vegetable supply chain information retrieval is going on, the retrieval requirements of users will be transferred into RDF source, by which contacts the constructed vegetable supply chain ontology model. Then all the properties and their value of the ontology model are listed by listSubjectsWithProperty method as a source for a given retrieval value, and ResIterator type values will be returned. The final sources can be obtained by hasNext method, which is coded as below.

```
ResIterator  iter=model.listSubjectsWithProperty(searchProperty,searchValue);

while(iter.hasNext() ){
        Resource  r=iter.nextResource();
}
```

For the words and their relations submitted by retrieval users, all the retrieval items will be listed by listObjectsOfProperty method and hasNext method.

```
NodeIterator result=model.listObjectsOfProperty(r,searchProperty);

while(result.hasNext()){
        temp=result.next();
}
```

In the reasoning system, axiom is usually described by standardization terms such as sub-class, sub-property, property definition domain, property value range, base restriction, disjoint etc. The two reasoning rules of the axiom used in the system are as below.

[equating relationship: (*?a* equate to *?c*), (*?b* equate to *?c*),notEqual(*?a*, *?b*)->(*?a* equate to *?b*)]
[similar relationship:(*?a* similar to *?b*),(*?a* similar to *?c*),notEqual(*?b*, *?c*)->(*?b* similar to *?c*)]

The supplement is necessary for reasoning rules.

4.3.3 Vegetable supply chain knowledge acquisition system

The interface of the system consults the design of www.google.com and www.baidu.com etc., which are with concise retrieval engine. The main interface of the system is shown in fig.14.

Supply chain is taken as retrieval item, which is retrieved in keywords based mode. 166 items are returned, which is shown as fig.15.

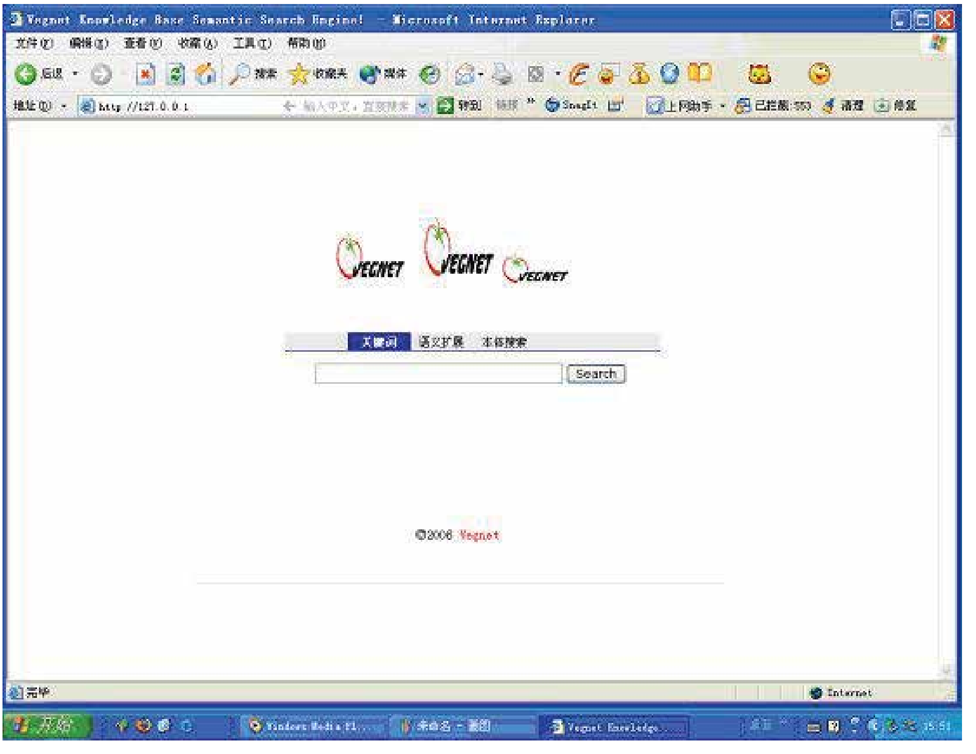


Fig. 14. System interface

Retrieve supply chain in semantic extension retrieval module, 239 items are returned. It is shown in fig.16.

The return items which are retrieved in ontology based is shown in fig.17. The first retrieval results list the semantic relationship of concepts according to ontology model. As it is hard for users to describe the retrieval requires properly in the first retrieval, the second retrieval will be done for the reasonable results, which can lead users to obtain the needed literatures.

4.4 Retrieval results analysis

As ontology model and semantic extension module is realized offline, the retrieval time is occupied by read-in ontology model time and reasoning time. The system in response to user retrieval requests can be done on a real-time response for the mount of current vegetable supply chain knowledge.

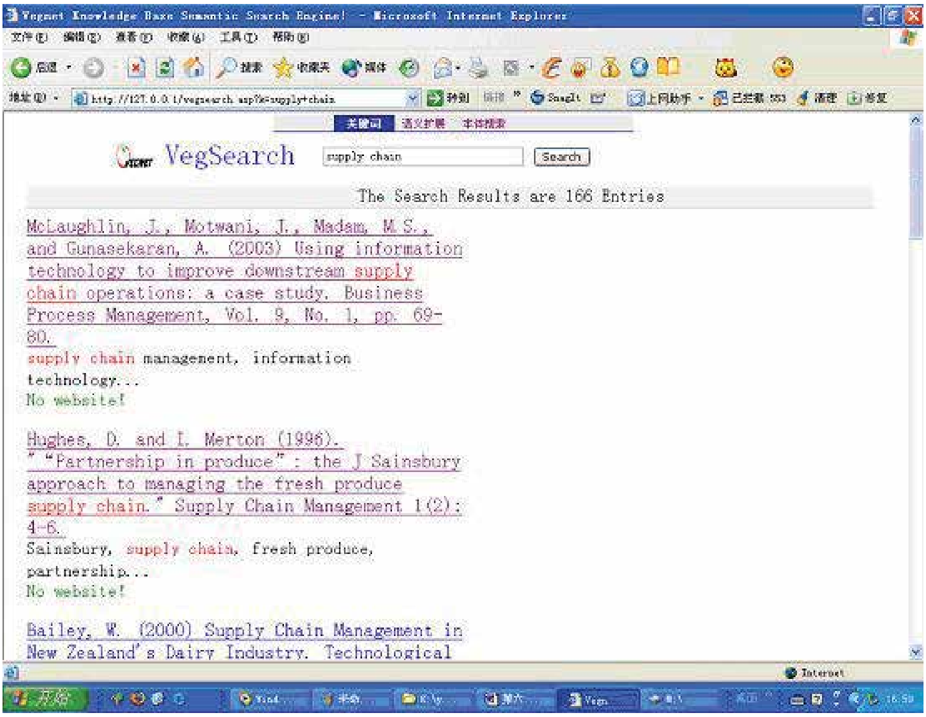


Fig. 15. Keywords based retrieval module

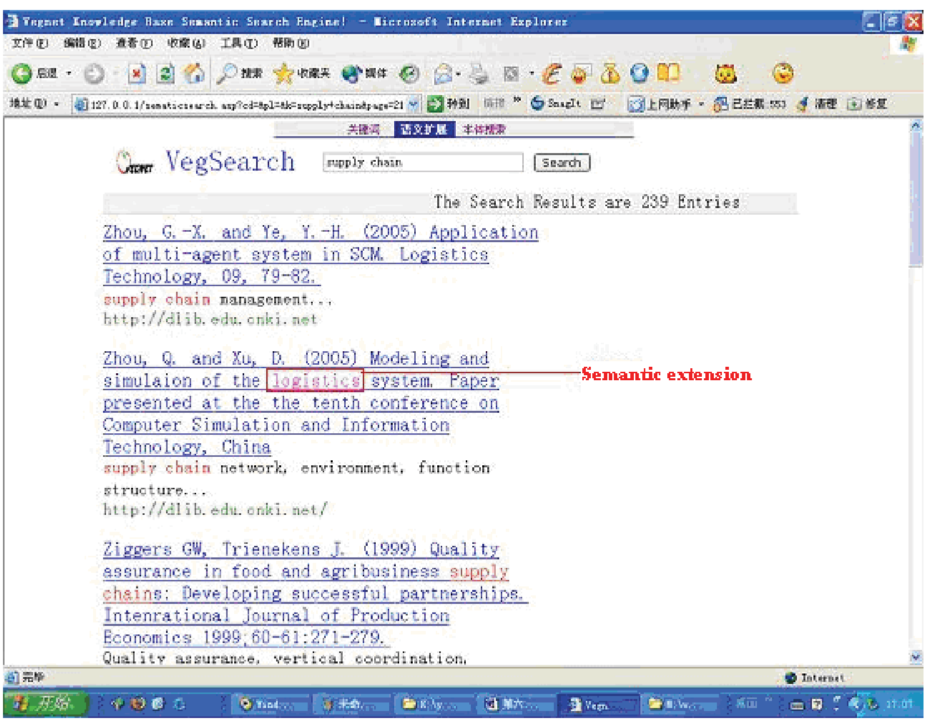


Fig. 16. Semantic extension retrieval module

Fig. 17. Ontology based retrieval module

From the perspective of users, more attention will be paid to function assessment, that is to say that user will pay more attention to whether the retrieval system meet user's retrieval requests. Precision Ratio and Recall Ratio are two basic indicators.

Five concepts (agri_product, fruit, inventory, logistics, transportation) are selected to test the Precision Ratio and Recall Ratio. Due to the system is completed aiming at vegetable supply chain domain. The data is analyzed and processed in the collection processing. Therefore the Precision Ratio is high for each retrieval mode and the Recall Ratio of different retrieval strategy will be inspected. The retrieval results are shown in table 2.

	agri_product	fruit	inventory	logistics	transportation	Average Recall
Keywords	56	16	18	149	48	0.268
Semantic Extension	56	66	18	228	48	0.4412
Ontology Based	91	91	286	350	316	1

Table 2. Retrieval results

Due to the relevant concepts and their relationships in our database are defined in the ontology model, the recall is considered as 1. The average recall of keywords retrieval and semantic expansion retrieval are calculated above this. It can be seen that vegetable supply chain knowledge is semantic tagged in the ontology model. The recall of ontology based retrieval will be the highest recall rate, and the recall of semantic extension retrieval gets the higher recall rate than recall of keywords base retrieval. It is also can be seen that the effect of semantic extension retrieval depends on the definition of dictionary. The average recall of it will be further improved by the expansion of semantic dictionary.

5. Conclusions and future works

5.1 Conclusions

In this chapter, aiming at the problems in traditional knowledge retrieval, ontology is introduced. An approach is put forward to supply chain knowledge management construction, which consists of construction of domain ontology, formalization of ontology model, and development of supply chain knowledge management system based on ontology. Finally, taking vegetable supply chain as a case, complete the vegetable supply chain knowledge management system based on ontology. The better recall and precision shows that the approach is effective. Conclusions are as follows:

- Concepts of supply chain domain and their relationships can be expressed clearly by ontology model. In order to implement an ontology based supply chain knowledge management system, supply chain ontology model, supply chain knowledge ontology model and supply chain knowledge user ontology model are constructed firstly. Each ontology model consists of classes and sub-classes, and there are four kinds of relationships between them, inheritance, part-whole, instance and properties, by which all concepts of supply chain domain are integrated. These ontology models are the base of semantic retrieval;
- Ontology models are formalized in order to get the understanding of computer, RDF(S) is taken to formalize ontology models. And the semantic reasoning is realized by setting reasoning rules;
- The ontology based supply chain knowledge management system can be implemented by Protégé, Java and Jena, by which the developing cost and difficulty will be reduced. Meantime, fasten the developing speedy. And the system can be transferred and used in other operating system convenient.

5.2 Future works

The researches in this chapter will be done further in the followings:

- The expanding of ontology model and the setting of reasoning rules need to be improved. Ontology models and reasoning rules fitting for the present situation have been established, while they need to be expanded and developed in accordance with the changing application situation.
- The storing mode of data and the ordering of retrieval results need to be improved. It will influenced the system capability seriously when the knowledge base and ontology model enlarge to a certain degree.

6. References

A. Th. Schreiber, B. J. Wielinga, and W. H. J. Jansweijer. (1995). The KACTUS view on the 'O' word. *In IJCAI Workshop on Basic Ontological Issues in Knowledge Sharing*.

Benjamin, P. C., Menzel, C. P., Mayer, R. J., Fillion, F., Futrell, M. T., deWitte, P. S., and Lingineni, M. (1994). IDEF5 Method Report. *Knowledge Based Systems*, Inc., September 21, 1994.

Bateman, J. A., Henschel, R. and Rinaldi, F. (2001). The generalized upper model 2.0. *http://www.darmstadt.gmd.de/publish/komet/gen-um/newUM.html*

Borst W N. (1997). Construction of Engineering Ontologies for Knowledge Sharing and Reuse. *PHD thesis,* University of Twente, Enschede.

Corcho, O., Gómez-Pérez, A.. (2000). A Roadmap to Ontology Specification Languages, material for *EKAW00 12th International Conference on Knowledge Engineering and Knowledge Management.* *http://babage.dia.fi.upm.es/ontoweb/wp1/OntoRoadMap/index.html.*

Fernández-López M, Gómez-Pérez A, Juristo N. (1997). METHONTOLOGY: From Ontological Art Towards Ontological Engineering. *Spring Symposium on Ontological Engineering of AAAI*. Stanford University, California, pp 33–40.

Fernández-López M, Gómez-Pérez A, Pazos A, Pazos J. (1999). Building a Chemical Ontology Using Methontology and the Ontology Design Environment. *IEEE Intelligent Systems & their applications* 4(1):37–46.

Framenet. (2008). *Http://www.icsi.berkeley.edu/~framenet/.*

Grigoris Antoniou, Frank van Harmelen. (2004). Web Ontology Language: OWL. *Handbook on Ontologies* .pp: 67-92.

Gruber T R. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition.* pp: 199-220.

Gruber T R. (1995). Towards Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies*, 43,pp:907-928.

Knight, K. and S. Luk.. (1994). Building a Large Knowledge Base for Machine Translation". *In Proceedings of the American Association of Artificial Intelligence Conference (AAAI-94)*, July 31, August 3. Seattle, WA. http://www.isi.edu/natural-language/resources/sensus.html#pubs.

M. Gruninger and M. Fox. (1995). Methodology for the design and evaluation of ontologies. *In Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing held in conjunction with IJCAI-95*, Montreal, Canada.

Mikrokmos. (1996). *Http://crl.nmsu.edu/Research/Projects/mikro/.*

Mizoguchi, R, M. Ikeda, K. Seta, et al. (1995). Ontology for modeling the world from problem solvingperspectives, *Proc. of IJCAI Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal.*

Neches R ,Fikes R E ,Gruber T R ,et al. (1991). Enabling Technology for Knowledge Sharing. *AI Magazine*, 12(3).pp:36-56.

Protégé. (2008). *Http://protege.stanford.edu/.*

RDF. (2008). *Http://www.w3.org/RDF/.*

Studer R, Benjamins V R, Fensel D. (1998). Knowledge Engineering, Principles and Methods. *Data and Knowledge Engineering*, 25(1-2), pp: 161-197.

T. R. Gruber. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5:199–220.

WordNet. (2006). *Http://www.cogsci.princeton.edu/~wn/*.

W3C. (2008). *Http://www.w3.org/*.

Edited by Yanfang Huo and Fu Jia

With the ever-increasing levels of volatility in demand and more and more turbulent market conditions, there is a growing acceptance that individual businesses can no longer compete as stand-alone entities but rather as supply chains. Supply chain management (SCM) has been both an emergent field of practice and an academic domain to help firms satisfy customer needs more responsively with improved quality, reduction cost and higher flexibility. This book discusses some of the latest development and findings addressing a number of key areas of aspect of supply chain management, including the application and development ICT and the RFID technique in SCM, SCM modeling and control, and number of emerging trends and issues.

ISBN 978-953-51-6406-7

Photo by raigna / iStock

IntechOpen