J Intell Manuf (2006) 17:307–319
DOI 10.1007/s10845-005-0003-z

Optimization of order fulfillment in distribution network problems

Felix T.S. Chan · S.H. Chung · K.L. Choy

Received: October 2004 / Accepted: September 2005

Abstract This paper focuses on optimization of order due date fulfillment reliability in multi-echelon distribution network problems with uncertainties present in the production lead time, transportation lead time, and due date of orders. Reliability regarding order due date fulfillment is critical in customer service, and customer retention. However, this reliability can be seriously influenced by supply chain uncertainties, which may induce tardiness in various stages throughout the supply chain. Supply chain uncertainty is inevitable, since most input values are predicted from historical data, and unexpected events may happen. Hence, a multi-criterion genetic integrative optimization methodology is developed for solving this problem. The proposed algorithm integrates genetic algorithms with analytic hierarchy process to enable multi-criterion optimization, and probabilistic analysis to capture uncertainties. The optimization involves determination of demand allocations in the network, transportation modes between facilities, and production scheduling in manufacturing plants. A hypothetical three-echelon distribution network is studied, and the computation results demonstrated the reliability of the proposed algorithms.

Keywords Multi–criterion decision making · Genetic algorithms · Analytic hierarchy process · Uncertainties · Distribution network · Production scheduling

Glossary

The following notation is used in the distribution network model:

AR	average reliability.
AT_k	mean arrival time (total lead time) of demand k.
CMD_{ij}	delivery unit cost for each unit of product transported from warehouse j to customer k.
CWD_{jk}	delivery unit cost for each unit of product transported from manufacturing plant i to warehouse j.
$\mathrm{D}(x_\mathrm{d})$	normal probability distribution of the required due date.
DC_k	demand quantity from customer k.
DC_{jk}	a share of DC_k allocated to warehouse j.
DCR_m	delivery unit cost ratio for each unit of product transported for delivery mode m.
D_k	customer k.
dL	lower boundaries of order due date of $D(x_\mathrm{d})$.
DMM_{ijm}	transportation mode selected to delivery product from manufacturing plant i to warehouse j.
dU	upper boundaries of order due date of $D(x_\mathrm{d})$.
DU_k	due date of demand k.
DW_j	demand quantity from warehouse j.

F.T.S. Chan (✉) · S.H. Chung
Department of Industrial and Manufacturing Systems Engineering,
The University of Kong Hong,
Pokfulam Road,
Hong Kong
e-mail: ftschan@hkucc.hku.hk

K.L. Choy
Department of Industrial and Systems Engineering,
The Hong Kong Polytechnic University,
Hung Hom, Kowloon,
Hong Kong

DW_{ij} a share of DW_j allocated to manufacturing plant i.
DWM_{jkm} transportation mode selected to delivery product from warehouse j to customer k.
LM_{ijm} delivery lead-time for each unit of product transported from manufacturing plant i to warehouse j for delivery mode m.
LW_{jkm} delivery lead-time for each unit of product delivered from warehouse j to customer k for delivery mode m.
m transportation mode.
MCA_i maximum production capacity of manufacturing plant i.
M_i manufacturing plant i.
MLT_i production lead-time per unit product of manufacturing plant i.
MPC_i production unit cost of manufacturing plant i.
MS_{ik} storage unit time of demand k spent in manufacturing plant i.
MSC_i storage unit cost per unit time of manufacturing plant i.
MT_{ik} mean production lead time of demand k in manufacturing plant i.
PC penalty cost per unit of product incurred from tardiness.
PR_k probability of demand k being supplied before due date.
$S(x_s)$ normal probability distribution of the supply date.
SCH_{ik} ranking number of demand k in production scheduling in manufacturing plant i.
SL lower boundaries of supply date of $S(x_s)$.
STD R standard deviation of reliability.
SU upper boundaries of supply date of $S(x_s)$.
TC total system cost.
TD total deviation of tardy supplies.
T_k time of tardiness for demand k.
TN total number of tardy supplies.
TT total duration of tardiness times.
UM_i capacity utilization of manufacturing plant i.
UW_j capacity utilization of warehouse j.
WCA_j maximum inventory storage capacity of warehouse j.
WHC_j inventory handling unit cost of warehouse j.
W_j warehouse j.
WLT_j order-processing lead time per order of warehouse j.
WPR_k probability of demand k being supplied before due date with weighting assigned.
WR weighted reliability.
WSC_j storage unit cost per unit time of warehouse j.
WS_{jk} storage unit time of demand k spent in warehouse j.
x_d a set of possible required order due date.
x_s a set of possible supply arrival date.
$\mu(*)$ mean value of $*$.
$\sigma(*)$ standard deviation of $*$.

Introduction

Distribution problems are concerned about with the allocation of a number of demand points to a number of source points, such as material suppliers, manufacturing plants, warehouses, distribution centers, and customers, connected by with various transportation facilities to form a supply network. Efficient integration of these facilities has been recognized to be an important strategy to increase and maintain competitive strength (Alshawi, 2001; Marcel & Evers, 1996; Yrjölä, 2001). A good distribution network can improve the operational costs, traveling distance, and the customer satisfaction level, by determining the optimal links between each pair of points, the product flow routine, transportation policy, and scheduling of the carrier loading/unloading points. Meanwhile, it can increase the flexibility, responsiveness, and reliability (Lumsden, Dallari, & Ruggeri, 1999; Milgate, 2001; Stank & Goldsby, 2000).

Accurate due date fulfillment is critical in winning customer orders, and maintaining customer retention. It relies on the reliability of various supply chain activities including material supply, production, and transportation. Earliness or tardiness happened in any activity may affect the completion time. Managing them as a whole has been recognized as an important issue (Thomas & Griffin, 1996), including buyer and vendor, production planning and inventory control, distribution and logistics, production-distribution, and inventory and distribution, etc. (Beamon, 1999; Cohen & Lee, 1988; Lee, Kim, & Moon, 2002). Failure of integration may result in inefficient utilization of resources, over-loaded/idle capacity, long production lead time, high in-transit inventory level, larger buffer stock, and unreliable due date assignment (earliness and tardiness).

This paper develops an optimization algorithm to maximize the reliability of customer order due date fulfillment in a demand driven multi-echelon distribution network. The proposed algorithm simultaneously optimizes demand allocations, transportation, and production scheduling and is capable of handling uncertain input in the customer due date, production lead time, and transportation lead time. It adopts the optimization concept of GA, and multi-criterion decision-making technique (MCDM) of Analytic Hierarchy Process (AHP) (Chan & Chung, 2004). The content of this paper is organized as follows. Next section gives a literature review.

Third section discusses the order fulfillment reliability with the probability approach. Section The Problem structure defines the distribution network problem. section Optimization Methodology presents the proposed optimization methodology. Final section analyzes the computation results, and lastly the paper will be concluded with a conclusion and directions of future research.

Literature review

Concurrently planning the production capability, production capacity, production costing, rough production scheduling, and demand due date, can help in providing products and services to customers at a lower cost and with a higher customer service level (Wu, Fuh, & Nee, 2002). With a set of demands, the inter-relationship between demand allocations, inventory management, location of facilities, and determination of transportation policy has been studied (González & Fernández, 2000; Wilson, 1995). From a macro view, the completion time of an order mainly depends on the production, warehousing, and transportation deployed between manufacturing plant and warehouse, and warehouse and customer. Among them, production scheduling influences the most since production capability is limited. In this connection, during demand allocations, the optimization methodology should also determine the production priority of each demand allocated.

Supply chain uncertainties are inevitable and they will affect the scheduling(s) (Schmidt & Wilhelm, 2000). Chen and Paulraj (2004) summarized supply chain uncertainties into three areas, which are supplier, manufacturing, and customer. However, there is a lack of paper studying in concurrent optimization of demand allocations, transportation, and rough production scheduling, especially with the consideration of supply chain uncertainties.

Optimization methodologies

To solve distribution network problems, some researchers applied linear programming, non-linear programming, mixed integer programming, fractional programming, and multi-objective linear fractional programming (Chakraborty & Gupta, 2002; Mármol, Puerto, & Fernández, 2002; Wilson, 1995). These optimization approaches can generate optimal solutions. However, they are usually time consuming in computation, and complicated in model construction.

In recent years, near optimal solutions obtained from heuristic approaches such as genetic algorithms (GAs), tabu search, simulated annealing, neural networks, etc. are more preferable because of their effective computational time (Berry, Murtagh, McMahon, Sugden, & Welling, 1998; Chang & Lo, 2001; Leão & Matos, 1999). Among them, GAs have been widely adopted by many researchers. Abdinnour-Helm (1999) testified the reliability and robustness of GA, by deploying five different distribution networks with the geographical layouts of Continental USA, Regional USA, Canada, and Western Europe. The results obtained were compared with the one obtained from mixed integer programming and it was proven that GA is reliable and robust. Vignaux and Michalewica (1991) adopted GA to minimize the operating cost for a linear transportation problem. Some researchers combined GAs with other heuristic methods, such as Tabu Search to enhance the optimization results (Abdinnour-Helm, 1998, 1999; Glover, 1986, 1989). GAs can also obtain good results in Job-shop scheduling problems (Cheung, Gen, & Tsujimura, 1996, 1999; Sakawa, 2002).

To capture supply chain uncertainties, many researchers applied probabilistic methods, and fuzzy logic theory (Cheung et al., 1996, 1999; González & Fernández, 2000; Miznuma & Watada, 1995). Sakawa (2002) deployed GA to maximize the fulfillment reliability of job due dates in a job-shop scheduling problem, in which the uncertainties of production lead time, and the due dates are represented by fuzzy sets. Some researchers combined GA with fuzzy theory to overcome uncertainties, such as demand fluctuation, which always exists in supply chain, (González & Fernández, 2000; Miznuma & Watada, 1995).

Multi-criterion decision-making

Multi-criterion decision making is usually encountered in supply chain management. Minimizing operating cost is crucial to make profit. Analogously, customer service is important to maintain competition (Hoek & Chong, 2001). Customer service level can be measured by customer response time, ability to respond to market changes, consistent order cycle time, accuracy of order fulfillment rate, delivery time, flexibility in order quantity, flexibility in product specification, accuracy of information system, etc. (Ballou, 1999). The factor of short due dates is critical for winning customer orders (Song, Hicks, & Earl, 2002). However, it usually induces high operating costs, and unreliable due date because of the relatively shorter range of tolerance for tardiness.

Tardiness may cause penalty cost and negative impact on the company's reputation. Indeed, there is a trade-off between the length of lead time quoted to the customers and the fulfillment reliability (Van and Bertrand, 2001). Gordon, Proth, & Chu (2002) summarized the literatures in the area of due date assignment and scheduling, such as those applying the mean absolute deviation (MAD) of completion time about a common completion due date, or mean absolute lateness to minimize the penalty cost on single machine, or parallel machines problems. However, most of the research work only studied at the level of job-shop scheduling. There is a lack of studies in the analysis of demand due date factor at the higher decision level of distribution network problems.

Improving some factors may affect the others, such as customer service and operating cost. Thus, trade-offs have to be taken. Another hottest example is the trade-off between earliness and tardiness to the demand due date. Early completion enhances fulfillment reliability, but induces storage costs, while tardy completion may lower operating cost, but induces penalty costs in some situations. In other words, the optimization methodology should be capable of determining the trade-off between the benefits of due date fulfillment reliability and the tardiness according to the weightings predefined in a given business situation. This highly complex relationship further increases the difficulty of decision-making, and in this connection, the proposed optimization methodology should allow the modeling of importance weightings to make it more realistic, and practical.

Order fulfillment reliability

Order fulfillment reliability concerns the reliability of fulfilling an order on or before the due date required. Unreliability can seriously induce extra costs, including those of tangible or intangible penalty costs, product depreciation, high inventory level, interruption of production, poor customer satisfaction, and or even loss of market share. Reliability can be enhanced by a good scheduling and operation of supply chain activities. However, scheduling in a deterministic environment is not practical since uncertainties are inevitable. Conventionally, uncertainties are captured in normal probability distribution density function as stated in Equation (1). Probability representation, instead of fuzzy representation, is applied because it gives a numerical value that indicates the chance of happening of an event. Decision makers can obtain more information from probability representation.

$$f(x) = \mathrm{e}^{-(x-\mu)^2/2\sigma^2}/\sigma(2\pi)^{1/2} \tag{1}$$

Order fulfillment reliability is represented by the probability (PR) of supplying the demand on or before the defined due date. Figure 1 shows the normal probability distribution of the supply date $S(x_\mathrm{s})$ and the normal probability distribution of the required due date $D(x_\mathrm{d})$. By Equation (2), SL and dL are the lower boundaries, while SU and dU are the upper boundaries of the normal probability distribution of the supply date, and the due date, respectively, in which $Z = 3.4$ corresponding to probability of happening of 99.98%.

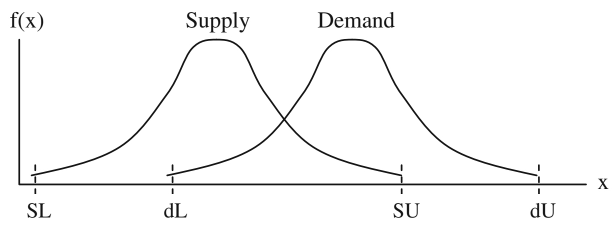

Fig. 1 A sample of supply date and demand due date in normal probability distribution

$$Z = (x - \mu)/\sigma \tag{2}$$

The probability of order fulfillment reliability is defined by Equation (3). The continuous density function $D(x_\mathrm{d})$ is considered as discrete function. An order is fulfilled if the supply arrives on or before the due date. The probability of happening equals to the probability of the order required on a particular date ($D(x_\mathrm{d})$) multiplied by the probability of the supply arriving on or before that particular date $\int S(x_\mathrm{s})dx$.

$$\mathrm{PR} = \sum_{x_c=dL}^{dU} \left[D(x_\mathrm{d}) \int_a^b S(x_\mathrm{s})dx \right] \tag{3}$$

where,

If $x_\mathrm{d} > \mathrm{SU}$, then $a = \mathrm{SL}$, and $b = \mathrm{SU}$, so that $\int S(x_\mathrm{s})\, dx = 1$.
If $x_\mathrm{d} > \mathrm{SL}$ and $x_\mathrm{d} < \mathrm{SU}$, then $a = \mathrm{SL}$, and $b = x_\mathrm{d}$.
If $x_\mathrm{d} < \mathrm{SL}$, then $\int S(x_\mathrm{s})dx = 0$.

The problem structure

A typical hypothetical three-tier distribution network model is testified. It consists of four manufacturing plants $M_i (i = 1, 2, 3, 4)$, four warehouses $W_j (j = 1, 2, 3, 4)$, and ten customer demands $D_k (k = 1, 2, \ldots, 10)$, with two different types of transportation modes ($m = 1, 2$) available between each link, as shown in Fig. 2. It is assumed that the transportation lead time in delivery mode 1 is shorter than that in delivery mode 2 for the same delivery arcs, while the delivery cost is higher. The problems here are to determine which demands should be produced in which manufacturing plants and supplied via which warehouses, the type of transportation mode for each delivery, and the production scheduling in the manufacturing plants. In this paper, five major criterions have been chosen—total system cost, total lead time, utilization, fulfillment reliability (in mean values), and fulfillment reliability (in probability), as shown in Fig. 3.

The decision variables are stated as follows:

- DC_{kj} – a share of DC_k allocated to warehouse j.
- DW_{ij} – a share of DW_j allocated to manufacturing plant i.
- DWM_{jkm} – transportation mode selected to delivery product from warehouse j to customer k.
- DMM_{ijm} – transportation mode selected to delivery product from manufacturing plant i to warehouse j.

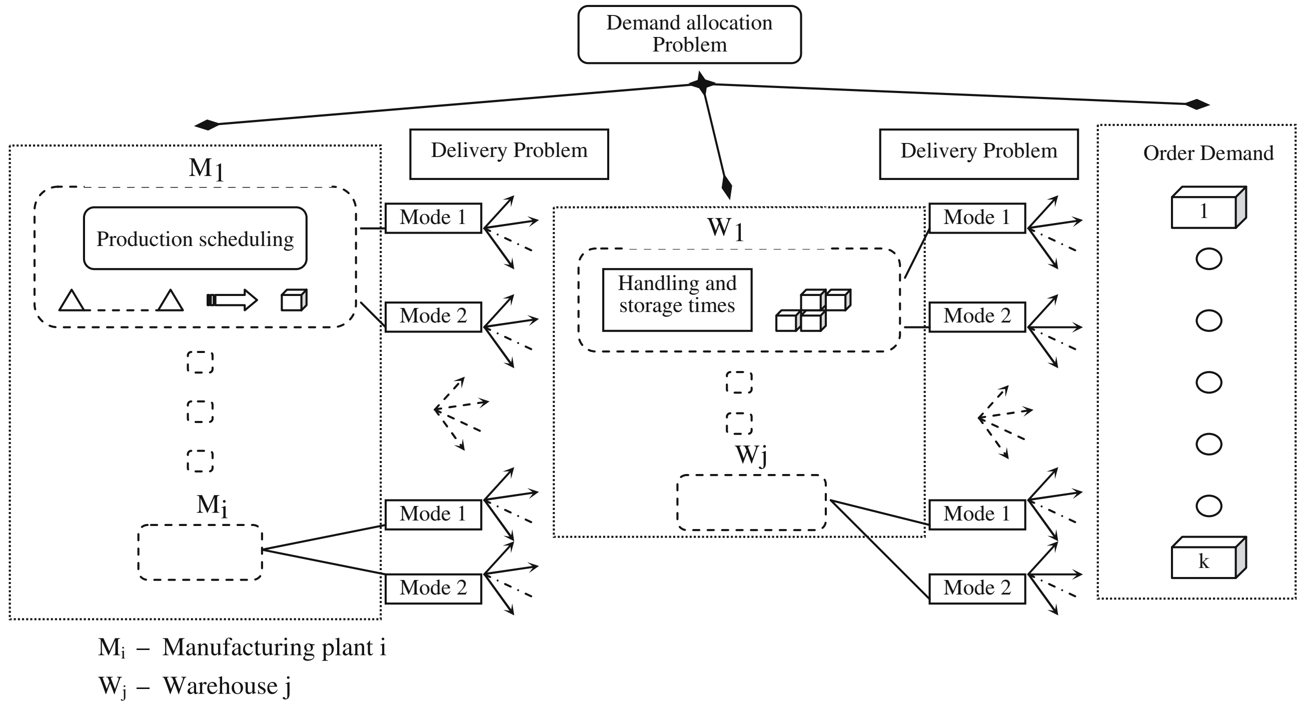

Fig. 2 A sample of distribution network model

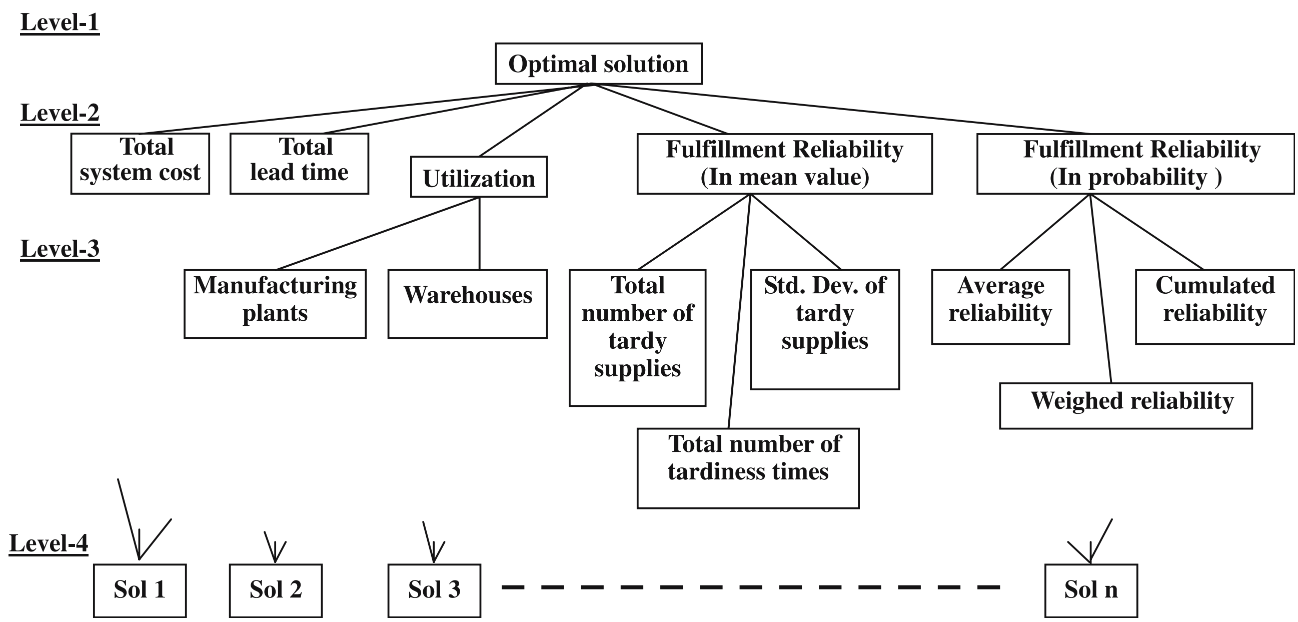

Fig. 3 The hierarchy structure of the proposed algorithm

WS_{jk} – storage unit time of demand k spent in warehouse j.

MS_{ik} – storage unit time of demand k spent in manufacturing plant i.

SCH_{ik} – ranking number of demand k in production scheduling in manufacturing plant i.

Objective functions

$$\text{Min}Z = \alpha\text{TC} + \beta\text{TL} + \chi U + \delta\text{FRM} + \varepsilon\text{FRP}, \quad (4)$$

where α, β, χ, δ, and ε are constants representing importance weightings.

Total system cost (TC)

$$\begin{aligned} TC = {} & \text{Total production cost in manufacturing plants} \\ & + \text{Total inventory handling cost in warehouses} \\ & + \text{delivery cost from manufacturing plants to warehouses} \\ & + \text{delivery cost from warehouses to customers} \\ & + \text{Total storage cost} \\ & + \text{Total penalty cost} \qquad (5) \\ = {} & \sum_{ij} \mathrm{DW}_{ij}\mathrm{MPC}_i + \sum_{ij} [(\mathrm{MPC}_i + \mathrm{CMD}_{ij})\mathrm{WHC}_j \mathrm{DW}_{ij}] \\ & + \sum_{ij} \mathrm{DW}_{ij}\mathrm{CMD}_{ij} + \sum_{jk} \mathrm{DC}_{jk}\mathrm{CWD}_{jk} \\ & + \sum_{ijk} (\mathrm{WSC}_j \mathrm{WS}_{jk} + \mathrm{MSC}_i \mathrm{MS}_{ik})\mathrm{DC}_k \\ & + \sum_k T_k \mathrm{DC}_k \mathrm{PC} \end{aligned}$$

where

$$\mathrm{CMD}_{ij} = \sum_m \mu(\mathrm{LM}_{ijm})\mathrm{MPC}_i \mathrm{DCR}_m \mathrm{DMM}_{ijm}$$
$$\mathrm{CWD}_{jk} = \sum_m \mu(\mathrm{LW}_{jkm})(\mathrm{MPC}_i + \mathrm{CMD}_{ij})\mathrm{WHC}_j \mathrm{DCR}_m$$

$$\begin{aligned} T_k &= \mu(\mathrm{AT}_k) - \mu(\mathrm{DU}_k), && \text{if } \mathrm{AT}_k > \mathrm{DU}_k \\ &= 0, && \text{if } \mathrm{AT}_k \le \mathrm{DU}_k \end{aligned}$$

Total lead time (TL)

This function aims to minimize the lead time of demands. A small value implies that the demands can to be delivered to customer earlier with a higher probability.

$$\mathrm{TL} = \sum_k \mu(\mathrm{AT}_k) \qquad (6)$$

$$\begin{aligned} \mathrm{TL}(\text{for} D_k) = {} & \text{Production lead time} + \text{Storage time in} \\ & \text{manufacturing plant} + \text{Delivery lead time} \\ & \text{from manufacturing plant to warehouse} \\ & + \text{Inventory handling time} + \text{Storage time in} \\ & \text{warehouse} + \text{Delivery lead time from} \\ & \text{warehouse to customer} \end{aligned}$$

$$\begin{aligned} \mu(\mathrm{AT}_k) = {} & \sum_{ijm} [\mu(\mathrm{MT}_{ik}) + \mathrm{MS}_{ik} + \mu(\mathrm{LM}_{ijm})\mathrm{DMM}_{ijm} \\ & + \mathrm{WLT}_j + \mathrm{WS}_{jk} + \mu(\mathrm{LW}_{jkm})\mathrm{DWM}_{jkm}] \\ \sigma(\mathrm{AT}_k) = {} & \sum_{ijm} [\sigma(\mathrm{MT}_{ik})^2 + \sigma(\mathrm{LM}_{ijm})^2\mathrm{DMM}_{ijm} \\ & + \sigma(\mathrm{LW}_{jkm})^2\mathrm{DWM}_{jkm}]^{1/2}, \end{aligned}$$

where the production lead time of demand $n (1 \le n \le k)$ in manufacturing plant i,

$$\begin{aligned} \mu(\mathrm{MT}_{in}) &= \text{Queuing time} + \text{Production lead time.} \\ &= \sum_{ij} \mu(\mathrm{MLT}_i)\mathrm{DW}_{ij} X_k \\ \sigma(\mathrm{MT}_{in}) &= \sum_{jk} \sigma(\mathrm{MLT}_i)\mathrm{DW}_{ij} X_k \\ X_k &= 1, \quad \text{if } \mathrm{SCH}_{ik} \le \mathrm{SCH}_{in}, \quad \text{else } X_k = 0. \end{aligned}$$

Utilization (U)

This function aims to improve the capacity utilizations equity between manufacturing plants and between warehouses, representing in standard deviation. The smaller the range, the more balanced is it.
Utilization deviation for manufacturing plants [σ(UM)],

$$\sigma(\mathrm{UM}) = \left\{\sum_i [\mathrm{UM}_i - \mu(\mathrm{UM})]^2/\text{no. of manufacturing plant}\right\}^{1/2}, \qquad (7)$$

where

$$\mathrm{UM}_i = \sum_j \mathrm{DW}_{ij}/\mathrm{MCA}_i$$
$$\mu(\mathrm{UM}) = \sum_j \mathrm{DW}_j / \sum_i \mathrm{MCA}_i$$

Utilization deviation for warehouses [σ(UW)],

$$\sigma(\mathrm{UW}) = \left\{\sum_j [\mathrm{UW}_j - \mu(\mathrm{UW})]^2/\text{no. of warehouse}\right\}^{1/2}, \qquad (8)$$

where

$$\mathrm{UW}_j = \sum_k \mathrm{DC}_{jk}/\mathrm{WCA}_j$$
$$\mu(\mathrm{UW}) = \sum_k \mathrm{DW}_k / \sum_j \mathrm{WCA}_j$$

Fulfillment reliability in mean value (FRM)

This function aims to minimize the tardiness of customer demands. The parameters considered are deterministic which take the mean value of the normal distribution. It is divided into three sub functions — total number of tardy supplies, total duration of tardiness, and standard deviation of tardy supplies. Indeed, these objective functions may conflict each other in some cases.

Total number of tardy supplies (TN) This sub-function aims to minimize the number of tardy supplies. It is independent of the demand quantity, and duration of tardiness, for example a total of three tardy supplies.

$$\mathrm{TN} = \text{number of tardy supplies.} \qquad (9)$$

Total duration of tardiness (TT) This sub-function aims to minimize the total duration of tardiness, in which the smallest value is the optimal solution. For example, three tardy supplies with a total of 9 days tardiness can be better than 1 tardy supply with 10 days tardiness in some situations.

$$\mathrm{TT} = \sum\nolimits_k T_k \tag{10}$$

Total deviation of tardy supplies (TD) This sub-function aims to minimize the long tardiness duration of supplies, which occur only in a few demands, representing in standard deviation. A large value indicates the difference of tardiness between demands is large.

$$\mathrm{TD} = \left\{\sum\nolimits_i [\mu(T_k) - \mu(T)]^2/k\right\}^{1/2}, \tag{11}$$

where average number of tardiness [$\mu(T)$]

$$\mu(T) = \sum\nolimits_k \mu(T_k)/k$$

Fulfillment reliability in probability (FRP)

This function aims to maximize order fulfillment reliability, representing in probabilities. It is subdivided into three sub functions—average reliability, weighed reliability, and standard deviation of reliability.

Average reliability (AR) This sub-function aims to maximize the average fulfillment reliability of the system. Every demand has the same importance of weighting.

$$\mathrm{AR} = \sum\nolimits_k \mathrm{PR}_k/k \tag{12}$$

Weighted reliability (WR) This sub-function aims to maximize the fulfillment reliability of the demands, but importance weightings are assigned.

$$\mathrm{WR} = \sum\nolimits_k \mathrm{DC}_k \mathrm{PR}_k/k \tag{13}$$

Standard deviation of reliability (STD R) This sub-function aims to maximize the balance of the fulfillment reliability between demands. It reduces the chance of sacrificing some particular demands for the high reliability of others. For example, to obtain a fulfillment reliability for two demands in probability 60% and 60% rather than in 99% and 20%.

$$\sigma(\mathrm{PR}) = \left\{\sum\nolimits_k [\mathrm{PR}_k - \mu(\mathrm{PR})]^2/k\right\}^{1/2} \tag{14}$$

Constraints

The problem is subject to the following constraints:
Production capacity constraints

$$\sum\nolimits_j \mathrm{DW}_{ij} \le \mathrm{MCA}_i \quad (\forall i) \tag{15}$$

This constraint ensures the total quantity of demand allocated from warehouses will not larger than the capacity of the manufacturing plant i.

Warehouse handling capacity constraints

$$\sum\nolimits_k \mathrm{DC}_{jk} \le \mathrm{WCA}_j \quad (\forall j) \tag{16}$$

This constraint ensures the total quantity of demand allocated will not larger than the capacity of the warehouse j.
Demand allocation constraints

$$\sum\nolimits_{jk} \mathrm{DC}_{jk} = \sum\nolimits_k \mathrm{DC}_k \tag{17}$$

This constraint ensures the total demand allocated to warehouses equal to the total demand.

$$\sum\nolimits_{ij} \mathrm{DW}_{ij} = \sum\nolimits_{jk} \mathrm{DC}_{jk} \tag{18}$$

This constraint ensures the total demand allocated to manufacturing plants equal to the total demand allocated to warehouses.
Transportation mode

$\mathrm{DWM}_{jkm} = 1$, if the delivery mode is selected, else $= 0$.
$\mathrm{DMM}_{ijm} = 1$, if the delivery mode is selected, else $= 0$.

These constraints ensure each demand will be transported by only 1 type of transportation between each link.

Structure parameters

The input values of the model are expressed as follows:

- Demand is generated between (250–500) units
- Production capacity (2000–2200) units
- Mean production lead time (0.01 − 0.025) unit times/unit
- Standard deviation of production lead time (0.001 − 0.0025) unit times/unit
- Production unit cost \$(1.0–1.5) /unit
- Manufacturing storage cost \$(0.2–0.4) /unit per unit times
- Warehouse capacity (2000–2500) units
- Warehouse handling lead time (1.0–2.0) unit times/order
- Warehouse handling cost \$(0.1–0.2) /unit of product value
- Warehouse storage cost \$(0.1–0.15) /unit
- Penalty cost \$(0.5) /unit per unit times
- Delivery cost of mode 1 \$(0.2) of product value
- Transportation cost of mode 2 \$(0.8) of product value
- Transportation lead time of mode 1 (0.3–0.5) of transportation lead time of mode 2 per unit times for the same delivery arc
- Standard deviation of transportation lead time (0.05–0.15) ratio of the corresponding mean value

Optimization methodology

Genetic algorithms

Genetic Algorithms (GAs) was developed by John Holland in 1960. It mimics the mechanism of genetic evolution in biological nature. To model the studied problem, a chromosome will be designed. It is composed of genes called alleles, and each gene represents 1 decision variable. A fixed number of chromosomes form an initial solutions pool. Through evolutions, new chromosomes (offspring) will be formed, and which is expected to be stronger than its parents, but this may not always be true. GA does not rely on analytical properties of the function to be optimized. In short, GA has two major characteristics. First, GA is re-iteratively and randomly generating new solutions. Second, these solutions are evaluated for the optimality according to predefined fitness functions.

This paper proposes an integrative optimization methodology to deal with the demand allocations, transportation, and production scheduling. Due to the complexity of the problem, the optimization process is divided into two parts, as shown in Fig. 4. Part I optimizes the demand allocations, and transportation, while Part II optimizes the production schedule. In detail, during each generation in Part I, each chromosome provides a partial solution for the distribution network, in which the demands have been allocated to warehouses and manufacturing plants, and the transportation mode has been decided. Then according to each chromosome, the production scheduling will be determined by Part II to form a completion solution. After that, these chromosomes will be feedback to Part I to calculate the fitness values, and to the remaining operators to complete one evolution. The iteration of Part I and Part II will carry on until the evolution completed.

Structure of the chromosomes

Two types of chromosomes are designed, Type A for Part I and Type B for Part II. Type A is a two-dimensional matrix, as shown in Fig. 5a. Region 1 shows the supply of the customer order from which warehouse, while Region 2 shows the assigned production plant, and the corresponding transportation mode adopted is shown in the last row. Type B is shown in Fig. 5b. The production-scheduling row indicates the ranking number of demand in the production scheduling in its assigned manufacturing plant.

Extended segment of chromosome

If a demand is larger than the maximum capacity of a warehouse or manufacturing plant, this order will be divided into two individual orders. One is in the quantity of the largest maximum absolute capacity of supply from the consequent layer, and the other is in the remaining quantity. If the remaining quantity is still too large to be absorbed, it will be further split until all absorbed. In practice, this distribution approach can benefit the supplier since capacity is fully utilized and also the operating and transportation cost could be reduced. For example, the administration cost required for each unit of product is relatively lower.

To model this feature, the chromosome further extends into two segments—(i) basic segment, and (ii) extended segment as shown in Fig. 6a, b. The basic segment is discussed in previous section. The splitting process is defined in the extended segment. For example, in Fig. 6a the customer order is larger than the largest absolute capacity of the available manufacturing plants. This customer order will then be split. The number of genes in the extended segment equals to the number of available manufacturing plants. Assuming D_1 with an order of 1000 units, which is greater than the largest maximum capacity of any individual manufacturing plant, say M_1 (800 units), D_1 will be split into two parts—800 units and 200 units. The first part will be allocated to M_1, and the other will be randomly allocated to other, for example M_3. Type B will also be extended corresponding to Type A for production scheduling, as shown in Fig. 6c.

Similar to the approach of dealing with manufacturing plant, if the customer order is larger than the largest absolute capacity of the warehouses, the extended segment will be set equal to the warehouse number. However, no change is required for Type B. In case of more than one order in excess of the largest maximum capacity of supply, each excess order will be split as mentioned previously. However, the priority of allocation among those excess orders is equal and arbitrary.

Genetic operators

Selector operator is applied roulette wheel selection, and the probability of crossover for optimization Part I is 20%, and Part II is 10%. Mutation operator deployed only in Part I with probability of 80%.

Elitist strategy aims to reserve the best chromosomes from the previous stage to the present stage without changing the gene structure (DeJong, 1975). This ensures the best chromosomes can survive. Similar idea is found in “Isbest strategy” (Onwubolu & Kumalo, 2001; Onwubolu & Muting, 2001). In this paper, after each evolution, the chromosome with the highest fitness value will be identified (more than one chromosome may exist). These chromosomes will be compared with those recorded in previous evolution. If these chromosomes are stronger, then they will be recorded. Otherwise, the recorded ones will be inserted in the mating pool during the next crossover. The insertion will be spread out evenly, i.e. each insertion will be separated equal to the number of insertion calculated. This ensures the survival of the best chromosomes and gives another chance to make them stronger.

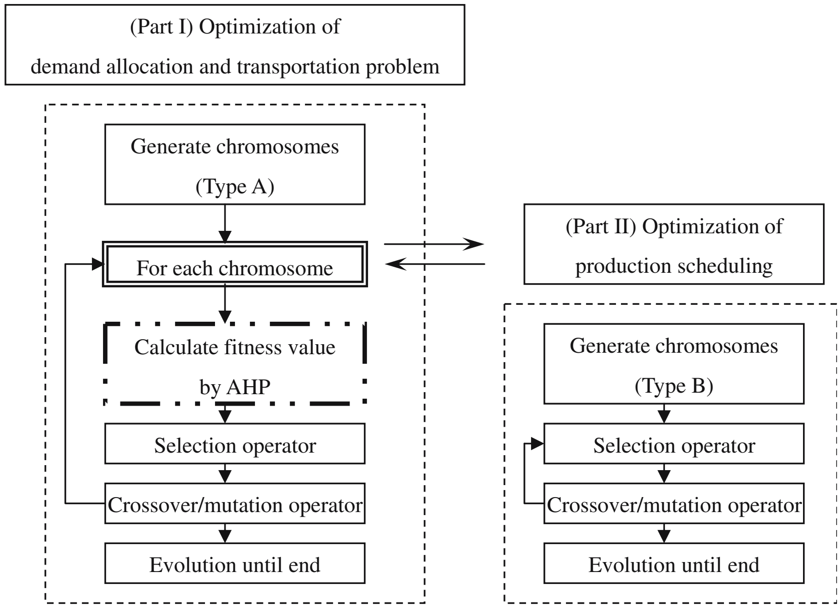

Fig. 4 An outline of the proposed optimization methodology

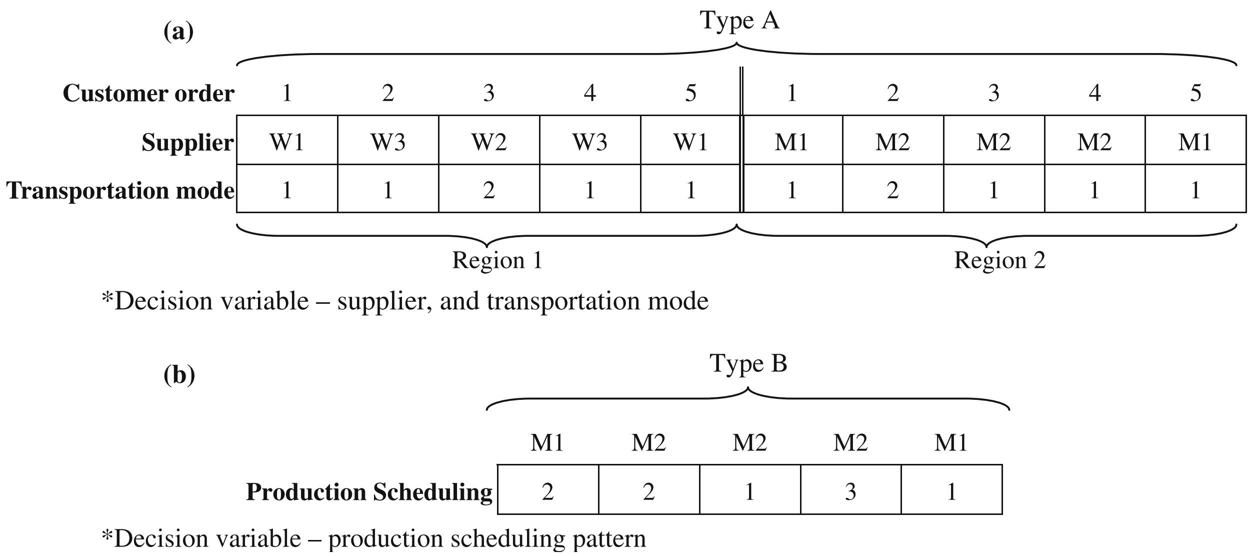

Fig. 5 A sample of chromosome structure. (**a**) A five demands, three warehouses, and three manufacturing plants allocation and transportation solution. (**b**) A five demands production scheduling solution

The number of insertion equals to approximately 10% of the total number of chromosome. Such a low percentage prevents the strongest chromosomes dominate the mating pool.

Analytic hierarchy process

Fitness value is a positive numerical value representing the strength/desirability of a chromosome according to the evaluation criteria. For example, in a problem of minimization of a cost function, if two solutions are found with A is $300, and B is $700, the expression of fitness values of A and B could be 0.7 (1–300/1000), and 0.3 (1–700/1000), respectively. In supply chain management, since multi-objective problems may usually be encountered, AHP will be applied to calculate fitness value.

AHP is developed by Thomas L. Saaty (1980). It is a well-proven MCDM methodology, especially powerful for those complex problems with a set of highly interrelated decision factors (Saaty, 1994). AHP organizes the complex interrelationships into a hierarchy or a network structure. It integrates all the criteria (objectives) into a hierarchy of weightings. These weightings and the hierarchy structure used influence the decision resulted. AHP needs numbers (quantitative) and a modicum of mathematics to formalize judgments and make tradeoffs. It also has the ability to determine which objective factor outweighs the others.

(a)

Basic Segment | Extended Segment

Supplier	W_1	W_3	W_2	W_3	W_2	$> M_1$	M_2	M_2	M_2	M_3	F	E	*R
Transportation mode	1	1	2	1	1	1	2	1	1	1	2	1	1
Customer order	1	2	3	4	5	1	2	3	4	5	1	2	3

Customer Number | Customer Number | Manufacturing plant Number

E – empty gene

****R*** – Remaining quantity

F – fully utilized

(b)

	Customer Demand						**Warehouse Demand**		
	1	**2**	**3**	**4**	**5**		**1**	**2**	**3**
Warehouse						**Manufacturer**			
1	1000	0	0	0	0	**1**	800	0	0
2	0	0	200	0	300	**2**	0	500	200
3	0	300	0	200	0	**3**	0	300	200

(c)

Basic Segment | Extended Segment

	M_1	M_2	M_2	M_2	M_3	M_1	M_2	M_3
Production Scheduling	0	3	1	4	2	1	2	1

Fig. 6 An example of splitting demand. (**a**) The sample of chromosome with the customer order 1 greater than the largest production capacity of manufacturing plant 1. (**b**) The allocation representation interpreted from the sample of chromosome as presented in (**a**). (**c**) The sample of an extended segment for chromosome type B with respect to chromosome type A shown in (**a**)

Another useful feature is the capability of importance weighting assignment, which means human knowledge and experience can directly be input into the system. Users can vary these weightings according to their desires. Instead of simply giving weightings to each criterion, AHP relates them in pair-wise comparison. In this connection, decision makers can specifically compare criterion A to criterion B and criterion C, and then specifically compare the criterion of B to C, or simply make changes in sub-criteria level without changing the weightings in the major criteria level. This weighting approach allows much more details, convenience, and flexibility on assigning weightings for the decision makers.

With decision maker's preference, AHP determines which solution is the most desirable in solution pool and ranks them in desirability. This value can be applied to represent the fitness value. The example problem in this paper can be modeled into four hierarchy levels, as shown in Fig. 3. Level 1 represents the objective, which is to determine the best solution among an options pool. Level 2 is the major criterions, and Level 3 is the sub-criteria under the major criterions. Lastly, level 4 contains the potential solutions in the solution pool. The calculation of fitness value is shown as following. For each potential solution, the values of the optimization functions will be represented as a relative value, shown in Equation 19. For example, the optimization function is the total system cost.

For minimization,

$$\text{Relative cost} = \frac{\text{Total system cost of the solution}}{\text{Highest total system cost of solution in the pool}} \tag{19}$$

For maximization,

$$\text{Relative cost} = 1 - \frac{\text{Total system cost of the solution}}{\text{Highest total system cost of solution in the pool}} \tag{20}$$

From the above expression, a smaller value represents that the solution has a better performance under that criterion. This relative value is then compared with all the others in the pool. Consider a pool with two solutions, CA and CB in the

Table 1 A sample of criterion matrix table for total system cost criterion

TC	CA	CB
CA	1	$RB-RA+1$
CB	$1/(RB-RA+1)$	1

relative cost values, if the relative value of CA (i.e. RA) is smaller than that of CB (i.e. RB), it means CA is better (with a lower cost) than CB. This implies CA should be weighted more. Therefore, the value of weighting of CA to CB is:

$$\text{Weighting of CA to CB} = \text{RB} - \text{RA} + 1 \tag{21}$$

In contrast, if RA is larger than RB,

$$\text{Weighting of CA to CB} = 1/(\text{RA} - \text{RB} + 1) \tag{22}$$

The numerical constant 1 in the Equations (21) and (22) serves two functions. First it makes the weighting equal to 1 when RA equals to RB, and also it forces the weighting to be larger than 1. These calculations form a criterion matrix table (see Table 1) for total system cost criterion in level-4 of Fig. 3. Detail calculation could be found in Chan and Chung (2004). Similar criterion matrix table will be evaluated for all the other criterions. The evaluation of the AHP value for each potential solution is completed by the multiplication of the normalized values found in level 4 with those found in level-2 and level-3.

Computation results and analysis

In this paper, three experiments have been established with the same model structure and input parameters as discussed in section 4. These experiments adopt different sets of weightings, emphasizing in different objective functions. The emphasis in Experiment 1 is on TC, Experiment 2 on TN, and Experiment 3 on AR. The proposed algorithm is implemented in a Java program on a PC, with 500 evolutions in the optimization Part I, and 10 evolutions in Part II to obtain a steady solution. The purpose of these experiments is to compare and analyze the different values of the total system cost, number of tardy supplies, and fulfillment reliability obtained.

In these experiments, if the mean completion time is later than the mean order due date defined, this customer order is classified as tardy. Penalty cost will be charged. From Experiments 1 and 2, the comparison shows the extra costs required to minimize the number of tardy supplies. In Table 2, Experiment 1 has the lowest TC ($16923.5) with two tardy supplies, which is the highest among the experiments, detail in Table 3. To minimize the number of tardy supplies in Experiment 2, the TC is required to increase by 2.45% to $17338.83. All demands can be satisfied by the assigned due date. This comparison also shows the ability of the proposed optimization algorithm to take the trade-off between the objective functions according to the importance weightings.

Table 2 Optimization results for experiments 1, 2 and 3

Experiment	1	2	3
Emphasis on	TC	TN	AR
Weighting Value	0.98	0.98	0.98
Results			
TC	16923.5	17338.83	27893.43
Lead time	200	194	150
UM	9.2	20.8	24.7
UW	24.9	20.8	29.5
TN	2	0	0
TT	4	0	0
TD	0.917	0	0
AR	0.645	0.74	0.99
WR	1932	2304	3150
STD R	0.237	0.176	0.017

Table 3 shows that each demand has a different probability of fulfillment reliability. For example, in Experiment 1, Orders 5 and 10 are estimated to have tardiness of 1 and 3 unit times, respectively. In numerical analysis, these two orders still have the probability of 44.23%, and 17.52% to be supplied on or before the due date required. However, AR obtained is only 64.5% (Table 4), which is relatively the lowest. This value is equivalent to have probability of 35.5% of tardiness. This implies high probability of increment in penalty cost from tardiness. Although Experiment 2 emphasizes on minimizing the number of tardiness, AR only increases to 74%, since it only considers the parameters in the mean values. With the emphasis on maximizing the average reliability function, Experiment 3 has the highest AR of 99%, and STD R has been improved as well.

In Experiments 1 and 2, tardiness could happen more seriously. In an extreme case, if all orders are required on the earliest possible due date, and all supplies are finished on the latest completion time according to their probability distribution densities, Table 4 shows the potential penalty cost and tardiness durations that could be incurred. Experiments 1 and 2 incur a relatively higher penalty cost. However, even with addition of these extra penalty costs in Experiment 1 (i.e. $16923.5 + $17591.6 = $34515.1), the total system costs are still lower than that of the Experiment 3 (i.e. $27893.43 + $8139.49 = $36032.9). However, the potential tardiness durations in Experiments 1 and 2 may be much longer than that in Experiment 3 because the supply dates in Experiment 3 are set relatively earlier to the mean due date required.

Table 3 Order fulfillment reliability of demands

Order	Due (μ, σ)		Experiment 1				Experiment 2				Experiment 3			
			Completion (μ, σ)		Early (+) / Tardy (−)	Prob.	Completion (μ, σ)		Early (+) / Tardy (−)	Prob.	Completion (μ, σ)		Early (+) / Tardy (−)	Prob.
1	21	2.10	15	0.99	6	0.9967	17	0.98	4	0.9725	9	0.41	6	1
2	22	2.20	22	1.87	0	0.5682	19	1.49	3	0.8183	18	1.33	4	0.9586
3	19	1.90	17	0.97	2	0.7779	17	1.14	2	0.8679	16	0.95	3	0.9485
4	24	2.40	24	1.32	0	0.5719	24	1.52	0	0.5693	17	1.63	7	0.9944
5	24	2.40	25	2.02	−1	0.4423	24	1.99	0	0.5633	18	1.17	6	0.9918
6	18	1.80	17	0.94	1	0.7681	16	0.86	2	0.8924	13	0.69	5	0.9972
7	20	2.00	20	1.15	0	0.5849	20	1.22	0	0.5836	15	0.91	5	0.9870
8	23	2.30	23	1.26	0	0.5676	23	1.38	0	0.5581	17	0.97	6	0.9947
9	19	1.90	13	0.78	6	0.9983	13	0.78	6	0.9983	12	0.67	7	0.9987
10	21	2.10	24	1.64	−3	0.1752	21	1.32	0	0.5793	15	0.98	6	0.9967

Table 4 Penalty cost incurred from possible earliest due date required and tardiest supplies

Order	Required due date	Experiment 1		Experiment 2		Experiment 3	
		Completion	Potential tardiness	Completion	Potential tardiness	Completion	Potential tardiness
1	14	18	−4	20	−6	10	0
2	15	28	−13	24	−9	23	−8
3	13	20	−7	21	−8	19	−6
4	16	29	−13	29	−13	23	−7
5	16	32	−16	31	−15	22	−6
6	12	20	−8	19	−7	15	−3
7	13	24	−11	24	−11	18	−5
8	15	27	−12	28	−13	20	−5
9	13	16	−3	16	−3	14	−1
10	14	30	−16	25	−11	18	−4
	Penalty cost induced		17591.6		16010.9		8139.47

Conclusions and future research

Reliability of order due date fulfillment is critical in customer service and customer retention. A reliable order fulfillment can be achieved by adequately scheduling different operations in different stages of the supply chain. However, scheduling in a deterministic environment is not practical because of supply chain uncertainties, sourcing from production, transportation, and forecasting of the coming order. Uncertainties may induce tardiness and consequently penalty cost. However, earliness may require higher operating cost and storage cost. Trade-off between earliness and tardiness should be determined.

Multi-criterion decision-making is usually required in supply chain management. This paper discussed the inter-relationship between total system cost, total lead time, utilization, fulfillment reliability (in mean value), and fulfillment reliability (in probability). To handle multi-criterion optimization, AHP is applied. The proposed integrative optimization methodology adopts GA for optimization, AHP to calculate the fitness values, and probabilistic representation to capture uncertainties. The proposed algorithm reiteratively optimizes the demand allocations and transportation, and production scheduling by separately applying GA in two different stages. A hypothetical three-echelon distribution network has been solved with three experimental runs. The comparison of optimal values obtained in each experiment demonstrated the influence of uncertainties. It also demonstrates the ability of the proposed algorithm to take the trade-off according to importance weightings.

Further research

The uncertainties in this distribution network problem only concerned with the production lead time, transportation lead time, and order due date. These uncertainties are related only with time. In practice, there are many other uncertainties exist in supply chain environment, such as the raw material supply quantity, production quantity, etc. These uncertainties increase the complexity of estimation of order fulfillment reliability of the system. Further research can be studied on the impact of various uncertainties on the reliability of order fulfillment.

References

Abdinnour-Helm, S. (1998). A hybrid heuristic for the uncapacitated hub location problem. *European Journal of Operational Research, 106*(2–3), 489–499.

Abdinnour-Helm, S. (1999). Network design in supply chain management. *International Journal of Agile Management Systems, 1*(2), 99–106.

Alshawi, S. (2001). Logistics in the Internet age: towards a holistic information and processes picture. *Logistics Information Management, 18*(8), 235–281.

Ballou, R.H. (1999). *Business logistics management*. 8th edition, Prentice Hall.

Beamon, B.M. (1999). Supply chain design and analysis: models and methods. *International Journal of Production Economics, 55*(1), 281–294.

Berry, L.M., Murtagh, B.A., McMahon, G.B., Sugden, S.J., & Welling, L.D. (1998). Genetic algorithms in the design of complex distribution networks. *International Journal of Physical Distribution and Logistics Management, 28*(5), 377–381.

Chakraborty, M., & Gupta, S. (2002). Fuzzy mathematical programming for multi objective linear fractional programming problem. *Fuzzy Sets and Systems, 125*(3), 335–382.

Chan, F.T.S., & Chung, S.H. (2004). Multi-criteria genetic optimization for distribution network problems. *International Journal of Advanced Manufacturing Technology*, 24, 517–532.

Chang, P.T., & Lo, Y.T. (2001). Modelling of job-shop scheduling with multiple quantitative and qualitative objectives and a GA/TS mixture approach. *International Journal Computer Integrated manufacturing, 14*(4), 367–384.

Chen, I.J., & Paulraj, A. (2004). Understanding supply chain management: critical research and a theoretical framework. *International Journal of Production Research, 42*(1), 131–163.

Cheung, R., Gen, M., & Tsujimura, Y. (1996). A tutorial survey of job-shop scheduling problems using genetic algorithms – I. *Computers and Industrial Engineering, 30*(4), 983–997.

Cheung, R., Gen, M., & Tsujimura, Y. (1999). A tutorial survey of job-shop scheduling problems using genetic algorithms—Part II. *Computers and Industrial Engineering, 37*(1), 51–55.

Cohen, M.A., & Lee, H.L. (1988). Strategic analysis of integrated production-distribution systems: models and methods. *Operations Research, 36*(2), 216–228.

DeJong, K.A. (1975). An analysis of the behavior of a class of genetic adaptive system. Doctoral dissertation, University of Michigan.

Glover, F. (1986). Future paths for integer programming links to artificial intelligence. *Computers and Operations Research, 13*(5), 533–589.

Glover, F. (1989). Tabu search part I. *ORSA Journal on Computing, 1*(3), 190–206.

González, E.L., & Fernández, M.A. (2000). Genetic optimization of a fuzzy distribution model. *International Journal of Physical Distribution and Logistic Management, 30*(7/8), 681–696.

Gordon, V., Proth, J.M., & Chu, C. (2002). A survey of the state-of-the-art of common due date assignment and scheduling research. *European Journal of Operational Research, 139*(1), 1–25.

Hoek, R. IV, & Chong, I. (2001). Epilogue: UPS Logistics – practical approaches to the e-supply chain. *International Journal of Physical Distribution and Logistics Management, 31*(6), 863–868.

Leão, M.T.P., & Matos, M.A. (1999). Multi-criteria distribution network planning using simulated annealing. *International Transactions in Operational Research, 6*(3), 371–391.

Lee, Y.H., Kim, S.H., & Moon, C. (2002). Production-distribution planning in supply chain using a hybrid approach. *Production Planning and Control, 13*(1), 35–46.

Lumsden, K., Dallari, F., & Ruggeri, R. (1999). Improving the efficiency of the Hub and Spoke system for the SKF European distribution network. *International Journal of Physical Distribution and Logistics Management, 29*(1), 60–64.

Marcel, M., & Evers, J.J.M. (1996). Distribution network design: an integrated planning support framework. *Logistics Information Management, 9*(1), 85–58.

Mármol, A.M., Puerto, J., & Fernández, F.R. (2002). Sequential incorporation of imprecise information in multiple criteria decision process. *European Journal of Operational Research, 137*(1), 123–133.

Milgate, M. (2001). Supply chain complexity and delivery performance: An international exploratory study. *Supply Chain Management: An International Journal, 6*(3), 106–118.

Miznuma, H., & Watada, J. (1995). Fuzzy mixed integer programming based on genetic algorithm and its application to resource distribution. *Japanese Journal of Fuzzy Theory and Systems, 7*(1), 97–116.

Miznuma, H., & Watada, J. (1995). Fuzzy mixed integer programming based on genetic algorithm and its application to resource distribution. *Japanese Journal of Fuzzy Theory and Systems, 7*(1), 97–116.

Onwubolu, G.C., & Kumalo, T. (2001). Optimization of multipass turning operations with genetic algorithms. *International Journal of Production Research, 39*(16), 3727–3745.

Onwubolu, G.C., & Muting, M. (2001). Optimizing the multiple constrained resources product mix problem using genetic algorithms. *International Journal of Production Research, 39*(9), 1897–1910.

Saaty, T.L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.

Saaty, T.L. (1994). *Fundamentals of decision making and priority theory with the analytic hierarchy process*. The Analytic Hierarchy Process Series, Vol. 6. RWS Publications.

Sakawa, M. (2002). *Genetic algorithms and fuzzy multi-objective optimization*. Kluwer Academic Publishers.

Schmidt, G., & Wilhelm, W.E. (2000). Strategic, tactical and operational decisions in multi-national logistics networks: A review and discussion of modelling issues. *International Journal of Production Research, 38*(7), 1501–1523.

Song, D.P., Hicks, C., & Earl, C. F. (2002). Product due date assignment for complex assemblies. *International Journal of Production Economics, 76*(1), 243–256.

Stank, T.P., & Goldsby, T.F. (2000). A framework for transportation decision making in an integrated supply chain. *Logistics Information Management, 5*(2), 71–77.

Thomas, D.J., & Griffin, P.M. (1996). Coordinated supply chain management. *European Journal of Operational Research, 94*(1), 1–15.

Van ooijen, H.P.G., & Bertrand, J.W.M. (2001). Economic due-date setting in job-shops based on routing and workload dependent flow time distribution functions. *International Journal of Production Economics, 74*(1), 261–268.

Vignaux, G.A., & Michalewica, Z. (1991). A genetic algorithm for the linear transportation problem. *IEEE Transactions on Systems, Man, and Cybernetics, 21*(2), 445–452.

Wilson, I. (1995). Distribution control systems within the supply chain. *Logistics Information Management, 8*(3), 80–82.

Wu, S.H., Fuh, J.Y.H., & Nee, A.Y.C. (2002). Concurrent process planning and scheduling in distributed virtual manufacturing. *IIE Transactions, 34*(1), 77–89.

Yrjölä, H. (2001). Physical distribution considerations for electronic grocery shopping. *International Journal of Physical Distribution and Logistic Management, 31*(10), 786–761.