Available online at www.sciencedirect.com
SCIENCE DIRECT
Omega 34 (2006) 56 – 69
omega
The International Journal of Management Science
www.elsevier.com/locate/omega

A genetic algorithm approach to developing the multi-echelon reverse logistics network for product returns

Hokey Min[a,*], Hyun Jeung Ko[a], Chang Seong Ko[b]

[a]UPS Center for World Wide Supply Chain Management, University of Louisville, Burhans Hall, Shelby, Suite LL23, Louisville, KY 40292, USA

[b]Department of Industrial Engineering, Kyungsung University, Daeyeon-dong, Nam-Ku, Busan 608-736, Korea

Received 12 November 2003; accepted 23 July 2004
Available online 6 October 2004

Abstract

Traditionally, product returns have been viewed as an unavoidable cost of doing business, forfeiting any chance of cost savings. As cost pressures continue to mount in this era of economic downturns, a growing number of firms have begun to explore the possibility of managing product returns in a more cost-efficient manner. However, few studies have addressed the problem of determining the number and location of centralized return centers (i.e., reverse consolidation points) where returned products from retailers or end-customers were collected, sorted, and consolidated into a large shipment destined for manufacturers' or distributors' repair facilities. To fill the void in such a line of research, this paper proposes a nonlinear mixed-integer programming model and a genetic algorithm that can solve the reverse logistics problem involving product returns. The usefulness of the proposed model and algorithm was validated by its application to an illustrative example dealing with products returned from online sales.

Keywords: Reverse logistics; Location-allocation; Genetic algorithm

1. Introduction

As of 1999, the total value of returned merchandise was $62 billion, representing $10–$15 billion in losses to retailers in the United States, while the cost of handling these product returns was estimated to be $40 billion [1]. Faced with the mounting costs of managing product returns, some companies have begun to consider mapping the process of reverse logistics involving product returns and creating opportunities for cost savings and service improvements. These companies include e-tailers that have grown with increases in online sales, but were often overwhelmed by the scope and complexity of sending returned products back to their distributors or manufacturers for credit. According to ReturnBuy [1], return rates for online sales are substantially higher than traditional bricks-and-mortar retail sales, reaching 20–30% in certain categories of items. In an extreme case, Rogers and Tibben-Lembke [2] reported that an average return rate for the magazine publishing industry was 50%.

With ever-rising costs of product returns and dwindling profit margins, the optimal handling of product returns can be a competitive differentiator since a firm can save a substantial amount of transportation, inventory, and warehousing costs associated with product returns. Indeed, Shear

* Corresponding author. Tel.: +1-502-852-2818; fax: +1-502-852-2842.
E-mail addresses: h0min001@gwise.louisville.edu (H. Min), csko@ks.ac.kr (C.S. Ko).

0305-0483/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.omega.2004.07.025

Table 1
Comparison between reverse logistics and forward logistics

	Reverse logistics	Forward logistics
Quantity	Small quantities	Large quantities of standardized items
Information tracking	Combination of automated and manual information systems used to track items	Automated information systems used to track items
Order cycle time	Medium to long order cycle time	Short order cycle time
Product value	Moderate to low product value	High product value
Inventory control	Not focused	Focused
Priority	Low	High
Cost elements	More hidden	More transparent
Product flow	Two way ("push and pull")	One way ("pull")
Channel	More complex and diverse (multi-echelon)	Less complex (single or multi-echelon)

Source: Adapted and modified from Shear et al. [3], "The warehousing link of reverse logistics."

et al. [3] noted that handling costs associated with product returns could reach $50 per item and could be three times higher than outbound shipping costs. In addition, they observed that product returns often reduced current assets due to lower inventory values for returned products, increased short-term liabilities due to required repairs and refurbishment, lengthened order cycle time due to reshipment of ordered items, and decreased sales revenue due to lost sales. As such, those firms that are willing to implement an optimal strategy of handling product returns can bring in millions of dollars of potential cost savings. Poirier [4] recently observed that firms in the optimal (or efficient) supply chain network enjoyed 40% more cost savings, 33% more inventory reductions, and 44% higher customer services than those in the inefficient supply chain network.

Typically, a product return involves the collection of returned products at designated regional distribution centers or retail outlets, the transfer and consolidation of returned products at centralized return centers, the asset recovery of returned products through repairs, refurbishing, and remanufacturing, and the disposal of returned products with no commercial value. The product return process entails the determination of the number and location of initial collection points for returned products and the location/allocation of centralized return centers in such a way that total reverse logistics costs (e.g., inventory carrying and transportation costs) are minimized, capacity of initial collection points and centralized return centers are fully utilized, and the convenience of customers who return products is maximized. By nature, the product return process is more complicated than forward logistics operations due to the presence of multiple reverse distribution channels (direct return to manufacturers versus indirect return to regional collection points or centralized return centers), individualized returns with small quantities, extended order cycles associated with product exchanges, and a variety of disposition options (e.g., repair versus liquidation). Recognizing the inherent complexity of the product return process, this paper develops a mathematical model and its solution procedure that can optimally create the reverse logistics network linking initial collection points, centralized return centers, and manufacturing facilities.

2. Relevant literature

In a broader sense, reverse logistics refers to the distribution activities involved in product returns, source reduction/conservation, recycling, substitution, reuse, disposal, refurbishment, repair and remanufacturing (e.g., [5]). As shown in Table 1, reverse logistics differs significantly from forward logistics. Despite its differences, reverse logistics has drawn little attention from researchers and practitioners alike until recent years. For the last decade, increasing concerns over environmental degradation and increased opportunities for cost savings or revenues from returned products prompted some researchers to formulate more effective reverse logistics strategies. These researchers include Min [6] who developed a multiple objective mixed integer program that was designed to select the most desirable shipping options (direct versus consolidated) and transportation modes for product recall. Although he considered a tradeoff between transportation time and cost associated with reverse logistics, his model could not handle multi-modal situations. Caruso et al. [7] proposed a multiple objective mixed integer program and a heuristic solution procedure for solving the location-allocation of waste service users, processing plants, and sanitary landfills with capacity constraints. Considering a multiple planning horizon, Melachronoudis et al. [8] also developed a multiple objective integer program for the dynamic location of capacitated sanitary landfills. Del Castillo and Cochran [9] presented a pair of linear programs (one aggregated and another disaggregated) and a simulation model to optimally configure the reverse logistics network involving the return of reusable containers in such a way that the number of reusable containers was maximized. However, they did not take into account transportation issues related to reverse logistics. In an effort to recycle construction waste as sieved sand, Barros et al. [10] proposed a mixed integer program which determined the locations of regional depots for receiving the flow of sieved sand and treatment facilities for cleaning and storing polluted sand. Unlike other previous models, they considered two-echelon location problems with capacity constraints.

Initiating product recovery network design efforts, Thierry [11] introduced a linear program to design product-distribution and product recovery networks involving the collection of used copying machines. However, his model did not address the location issue of where the product recovery (resale of products after remanufacturing and refurbishment) process should be installed and at what capacity. Extending the work of Thierry [11], Krikke [12] proposed a network graph and a mixed integer program to optimize the degree of disassembly and evaluate product recovery options in collecting used copying machines and redistributing them after refurbishment, while determining the location and capacity of remanufacturing, central stocking, and disposal facilities. Similarly, Krikke et al. [13] developed a mixed integer program to determine the locations of shredding and melting facilities for the recovery and disposal of used automobiles, while determining the amount of product flows in the reverse logistics network. More recently, Jayaraman et al. [14] presented a mixed integer program to determine the optimal number and locations of remanufacturing facilities for electronic equipment. Jayaraman et al. [15] extended their prior work to solve the two-level hierarchical location problem involving the reverse logistics operations of hazardous products. They also developed heuristic concentration procedures combined with heuristic expansion components to handle relatively large problems with up to 40 collection sites and 30 refurbishment sites. Despite their success in solving large-sized problems, their model and solution procedures are still confined to a single period problem and are not designed to deal with the possibility of making trade-offs between freight rate discounts and inventory cost savings resulting from consolidation of returned products. For a detailed and updated review of product recovery network models, interested readers should refer to Fleschmann et al. [16].

As summarized above, a majority of existing reverse logistics models have, so far, focused on the environmental aspects (e.g., product recovery, recycling, reuse) of the reverse logistics network for used products, which ended their life cycles, and neglected various consolidation and channel selection decisions for product returns. The proposed model in this study will aim to design a reverse logistics network involving products returned due to either defects or changes in customers' needs/preferences.

3. Problem definition

As of 2000, product returns averaged approximately 6% of sales [17]. The rate of product returns is usually higher for books, magazines, apparel, greeting cards, CD-ROMs and electronics. In particular, mail catalogue or on-lines sales are more vulnerable to product returns. A work in Modern Material Handling estimated that 30% of online sales would be returned to e-retailers [18]. This high rate of returns would cost e-tailers $1.8 to $2.5 billion a year. Typical reasons for product returns may include: defects, in-transit damage, trade-ins, product upgrades, exchanges for other products, refunds, repair, recalls, and order errors. Regardless of the reasons for the returns, many e-tailers (84%) either absorb the cost of return shipment or offer a money-back guarantee for returned products, making product returns a major cost center. To control the cost of handling returns, a growing number of e-tailers and their third-party logistics providers (3PLs) have begun to examine ways to improve the efficiency of product returns. Examples of such ways are:

1. Reduction of return shipping costs by taking advantage of economies of scale. A number of separate consolidation points such as centralized return centers can be established to aggregate small shipments into a large shipment.
2. Enhancement of customer convenience for product returns. A number of initial collection points near to the customer population center can help customers reduce their travel time to the collection points for returns.
3. Reduction in in-transit inventory carrying costs associated with product returns. Since in-transit inventory carrying costs are proportionately related to transit time of transportation modes that are used for return shipment, one should consider the fastest mode of transportation while weighing its freight rate.

For an illustrative purpose, let us suppose that an e-tailer (called *Beta.com* hereafter) selling various computer equipment and peripherals has been inundated with returned products due to its liberal return policy. Although many customers prefer to return computers directly to original equipment manufacturers (OEMs), direct shipment is far more costly than indirect shipment due to frequent, small volume shipment that often requires a premium mode of transportation, such as UPS small package delivery services. In addition, many customers do not want to deal with the hassle of making shipping arrangements for returns through regional postal services. Instead, they would like to drop-off computers at one of the initial collection points located near their residence or office. Candidates for these initial collection points include: local pharmacies, video-rental stores, 24-h convenience shops, and gas stations. Since *Beta.com* does not own these collection points, the collection points will not incur fixed costs such as land purchase, lease, and property tax. However, the collection points will incur variables costs associated with renting limited space designated for "non-selling" returned products. Given the limited storage space of the initial collection points, returned products at the collection points should be quickly transshipped to centralized return centers where returned products are inspected for quality failure, sorted for potential repair or refurbishment, stored long enough to create volume for freight consolidation, and shipped to original manufacturers (or third-party logistics provider's repair depots).

From centralized return centers, some returned products, which are found to be defect- or damage-free, may be redistributed to customers after repackaging or re-labeling. Centralized return centers are dedicated to return handling and processing. Thus, they are owned and operated by *Beta.com* and separated from the typical warehousing functions. On the other hand, centralized return centers may play a critical role in linking the initial collection points to manufacturing or repair facilities within the reverse logistics network. One can bypass the centralized return center for returning products to manufacturers, if the initial collection points are closer to the location of given manufacturers than that of centralized return centers or consolidation at the centralized return center considerably delays the return process. With the above situations in mind, the main issues to be addressed by this study are:

1. Where to locate initial collection points in such a way that travel time (or distance) from existing and potential customers to the collection points is minimized?
2. Where to locate centralized return centers in a manner that costs of transshipment between the initial collection points and manufacturing (or repair facility) locations are minimized?
3. How to build the reverse logistics network in such a way that a timely pickup can be made between the initial collection point and the centralized return center? Considering hours-of-service-regulations stipulated by the *federal highway administration* (*FHA*), the locations of initial collection points should be within certain hours of driving time from the nearest centralized return center.
4. How frequently returned products at the initial collection point should be consolidated to minimize shipping costs, while delays in the return process are avoided?
5. How many initial collection points and centralized return centers are needed to minimize the customer hassles associated with product returns while minimizing the costs of handling returns?

To summarize, the reverse logistics problem facing *Beta.com* is primarily concerned with determining which retail stores should be chosen as initial collection points, which new centralized return centers to establish, and which reverse logistics channels to use. To deal with this problem, systematic decision-aid tools are needed which consider a multitude of conflicting factors affecting the reverse logistics network and an analysis of the tradeoffs among them. Such decision-aid tools include various mathematical programming techniques such as integer programming (see, e.g., [19]). Considering that *Beta.com*'s main objective is to maximize the potential cost savings accrued from the multi-echelon reverse logistics channel, we propose a single-objective, nonlinear, mixed-integer programming model as our decision-aid tool. The proposed model is designed to find the optimal location, number and size of both initial collection points and centralized return centers in the reverse logistics network under capacity limits and service requirements.

4. Model design

Prior to developing the nonlinear mixed-integer programming model for reverse logistics network design, we make the following underlying assumptions and simplifications: (1) The possibility of direct shipment from customers to a centralized return center is ruled out due to insufficient volume. (2) Given small volume of individual returns from customers, an initial collection point has sufficient capacity to hold returned products during the consolidation process. (3) The transportation costs between customers and their nearest collection points are negligible given short distances between customers and their nearest collection point. (4) The location/allocation plan covers a planning horizon within which no substantial changes are incurred in customer demands and in the transportation infrastructure.

4.1. Indices

i index for customers; $i \in I$
j index for initial collection points; $j \in J$
k index for centralized return centers; $k \in K$

4.2. Model parameters

a	annual cost of renting initial collection point j
b	daily inventory carrying cost per unit
w	annual working days
r_i	daily volume of products returned by customer i
h	handling cost of unit product per day
q_k	cost of establishing centralized return center k
m_k	maximum capacity of centralized return center k
d_{ij}	distance from customer i to initial collection point j
d_{jk}	distance from collection point j to centralized return center k
l	maximum allowable distance from a given customer to an initial collection point
$f(X_{j0}, d_{j0})$	$E\alpha\beta$ (= function for freight rate)

where α is a discount rate according to the volume of shipment between initial collection point j and centralized return center k; β is a penalty rate applied for the distance between collection point j and centralized return center k

$$\alpha = \begin{cases} 1 & \text{for } X_{j0} \leqslant p_1, \\ \alpha_1 & \text{for } p_1 < X_{j0} \leqslant p_2, \\ \alpha_2 & \text{for } x_{j0} > p_2, \end{cases}$$

$$\beta = \begin{cases} 1 & \text{for } d_{j0} \leqslant q_1, \\ \beta_1 & \text{for } q_1 < d_{j0} \leqslant q_2, \\ \beta_2 & \text{for } d_{j0} > q_2, \end{cases}$$

E	unit freight rate
p_1, p_2	volume of returned products for a discount
q_1, q_2	distance between collection point j and centralized return center k for penalties
z	minimum number of established initial collecton points
g	minimum number of established centralized return centers
M	arbitrarily set large number.

4.3. Decision variables

X_{jk} = volume of products returned from initial collection point *j* to centralized return center *k*,

T_j = length of a collection period (in days) at initial collection point j,

$$Y_{ij} = \begin{cases} 1, & \text{if customer } i \text{ is allocated to initial collection } j (i \in I,\ i \neq j), \\ 0, & \text{otherwise.} \end{cases}$$

$$Z_j = \begin{cases} 1, & \text{if an initial collection point is established at site } j (j \in J), \\ 0, & \text{otherwise.} \end{cases}$$

$$G_k = \begin{cases} 1, & \text{if a centralized return center is established at site } k (k \in K), \\ 0, & \text{otherwise.} \end{cases}$$

4.4. Mathematical formulation

(P) Minimize

$$\alpha \sum_j Z_j + bw \sum_j \left\{ \sum_i r_i Y_{ij} \frac{(T_j + 1)}{2} \right\} + hw \sum_i r_i + \sum_k q_k G_k + \sum_k \left\{ G_k \sum_j \left(X_{jk} \frac{w}{T_j} \right) \times f(X_{jk}, d_{jk}) \right\} \quad (1)$$

Subject to

$$\sum_j Y_{ij} = 1, \ \forall i \in I, \quad (2)$$

$$\sum_i Y_{ij} \leqslant M \cdot Z_j, \ \forall j \in J, \quad (3)$$

$$\sum_i r_i Y_{ij} T_j = \sum_k X_{jk}, \ \forall j \in J, \quad (4)$$

$$\sum_j X_{jk} \leqslant m_k G_k, \quad \forall k \in K, \quad (5)$$

$$d_{ij} Y_{ij} \leqslant l, \quad \forall i \in I, \ \forall j \in J, \quad (6)$$

$$z \leqslant \sum_j Z_j, \quad (7)$$

$$g \leqslant \sum_k G_k, \quad (8)$$

$$X_{jk} \geqslant 0, \quad \forall j \in J, \ \forall k \in K, \quad (9)$$

$$T_j \in (0, 1, 2, 3, 4, 5, 6, 7), \ \forall j \in J, \quad (10)$$

$$Y_{ij}, z_j, G_k \in (0, 1) \forall i \in I, \forall j \in J, \ \forall k \in K. \quad (11)$$

The objective function (1) minimizes total reverse logistics costs comprised of renting, inventory carrying, material handling, setup, and shipping costs. Notice that objective function (1) has a nonlinear form because both inventory carrying and shipping costs are affected by the length of a collection period. Constraint (2) assures that a customer is assigned to a single initial collection point. Constraint (3) prevents any return flows from the unopened initial collection point. Constraint (4) makes the incoming flow equal to the outgoing flow at an initial collection point. Constraint (5) ensures that the total volume of products returned from initial collection points does not exceed the maximum capacity of a centralized return center. Constraint (6) assures that each initial collection point should be located within a certain allowable proximity of customers. Constraints (7) and (8) maintain a minimum number of initial collection points and centralized return centers for product return. Constraint (9) preserves the non-negativity of decision variables X_{jk}. Constraint (10) limits a range of integrality of decision variables T_j. Constraint (11) assures the binary integrality of decision variables Y_{ij}, Z_j, and G_k.

5. Model application and results

The proposed model was applied to the hypothetical problem facing *Beta.com*. In order to handle an increasing volume of product returns and the subsequent repairs, *Beta.com* has been exploring the possibility of establishing initial collection points and centralized return centers, while maximizing easy and convenient returns from the customers. The potential locations of collection points and centralized return centers were summarized in Table 2. Once established, these facilities would serve a total of 30 clusters of customers, and daily demand of each cluster was shown in Table 3. The total daily demand is 848 units and an annual

Table 2
Potential sites of initial collection points and centralized return centers

Potenttial sites for initial collection points	Site coordinate		Potential sites for centralized return centers	Site coordinate	
	x	*y*		*x*	*y*
cp1	43.97	49.89	crc1	8.58	30.25
cp2	1.57	12.65	crc2	32.36	28.59
cp3	41.23	30.25	crc3	9.58	6.51
cp4	5.04	58.97	crc4	47.54	19.31
cp5	24.79	19.00	crc5	20.14	53.21
cp6	16.18	20.66			
cp7	30.18	45.30			
cp8	40.32	0.40			
cp9	6.94	33.58			
cp10	54.71	57.06			

Table 3
Locations and daily demands of customers

No.	Coordinates		Daily demand
	X	*Y*	
1	15.69	3.80	12
2	18.67	24.28	43
3	1.60	59.13	34
4	9.43	2.27	21
5	49.08	54.43	19
6	33.14	10.85	10
7	28.62	50.00	37
8	24.86	59.39	22
9	3.42	35.85	35
10	33.23	21.90	29
11	45.32	27.23	22
12	46.37	6.36	21
13	24.93	32.60	11
14	28.07	33.38	27
15	2.77	0.50	44
16	28.61	51.99	41
17	38.80	51.71	46
18	2.13	41.98	22
19	25.78	2.81	37
20	45.69	57.24	45
21	48.17	8.13	38
22	36.72	14.35	27
23	42.61	27.50	29
24	22.15	33.30	11
25	25.63	28.51	23
26	9.10	26.42	10
27	24.79	11.26	39
28	17.46	11.20	18
29	11.87	12.97	44
30	6.15	38.45	33

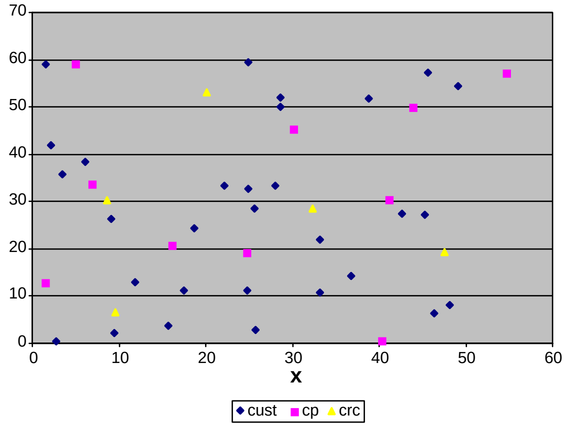

Fig. 1. Graphical representations of customers, initial collection points and centralized return centers.

working day is 250 (total annual demand =212,100 units). For simplicity, Euclidean distance is used for measuring travel distance, and the maximum allowable distance from customers to the nearest collection point is estimated to be 25 miles. Also, a graphical representation of those locations was shown in Fig. 1. Other input parameters associated with collection points and centralized return centers are summarized in Table 4. The *Beta.com* management team intended to determine the number of collection points and centralized return centers needed for handling products returned for repairs while considering easy and convenient access for customers within 25 miles of the nearest collection point. Also, the team considered an appropriate collection period for the tradeoff between distance and volume from a collection point to a centralized return center owing to small volume of returns. To aid the *Beta.com* management team

Table 4
Input parameters

Parameter	Index	Value
Annual cost of renting an initial collection point	a	\$200
Daily inventory carrying cost per unit	b	\$0.1
Working days per year	w	250
Unit handling cost at the collection point	h	\$0.1
Cost of establishing a centralized return center	q_k	\$3,000
Capacity of a centralized return center	m_k	1000 units
Service coverage	l	25 miles
Unit standard transportation cost	E	1
Discount rate with respect to shipping volume		
	α_1	0.8
	α_2	0.6
	p_1	200 units
	p_2	400 units
Penalty rate with respect to shipping distance		
	β_1	1.1
	β_2	1.2
	q_1	25 miles
	q_2	60 miles
Minimum number of the established collection points	z	1
Minimum number of the established centralized return centers	g	1

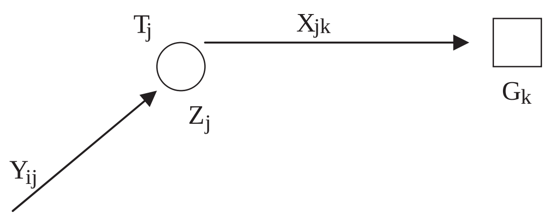

Fig. 2. Return network flows.

in finding the optimal alternative, we attempted to solve the model specified in Section 4 using a genetic algorithm (GA) since a reverse logistics network design belongs to a class of NP-complete problem (see, e.g., [20] for a definition of NP-completeness). In addition, GA can overcome computational complexity induced by the nonlinear objective function. Through model experimentation, the base-line model resulted in 315 integer variables, 60 continuous variables, and 355 constraints. Fig. 2 shows a proposed network flow for handling product returns.

5.1. Genetic algorithm development

In order to solve the reverse logistics network design problem for product returns, we propose a GA comprised of two sub-algorithms. GA is referred to as a stochastic solution search procedure that is designed to solve combinatorial problems using the concept of evolutionary computation imitating the natural selection and biological reproduction of animal species [21,22]. In the past, GA has been successfully applied to classical combinatorial problems such as capacitated plant location [23], fixed charge location [24], minimum spanning tree [25], network design [26], and warehouse allocation [27]. Given this proven effectiveness of GA for various combinatorial problems, GA is suitable for solving the reverse logistics network design problem. Another appeal of GA includes its flexible solution search process that can convert constrained problems into unconstrained problems and then cross the feasibility boundary to find near-optimal or optimal solutions in an “intelligent” (probabilistic) manner rather than relying on random enumerations or iterations. In particular, GA is chosen over other meta-heuristics procedures such as tabu search due to its ability to generate a collection of solutions rather than a single solution at each stage (see, e.g., [28] for an excellent discussion of tabu search algorithms).

Prior to the application of GA, we need to design the genetic representation (or chromosome) of the candidate solutions. Herein, a chromosome represents each solution in the initial solution set of solutions (population). The size of the population depends on the size and the nature of the problem at hand. The chromosome evolves through a crossover operator and a mutation operator to produce children, improving on the current set of solutions. The chromosomes in the population are then evaluated through a fitness function and the less fit chromosomes are replaced with better children. The processes of crossover, evaluation and selection are repeated for a predetermined number of iterations called generations, usually up to the point where the system ceases to improve or the population has converged to a few well performing chromosomes.

5.1.1. Encoding

The design of a suitable chromosome is the first step for a successful GA implementation because it applies probabilistic transition rule on each chromosome to create a population of chromosomes, representing a good candidate solution. Each chromosome developed in this study is based on single dimensional array which consists of binary values, representing decision variables related to initial collection points, centralized return centers, and collection periods (i.e., consolidation intervals or holding time for consolidation at the collection point). For example, the representation of a chromosome is illustrated in Fig. 3. The solution (chromosome) has 30 initial collection points, 7 days of collection periods at collection points, and five centralized return

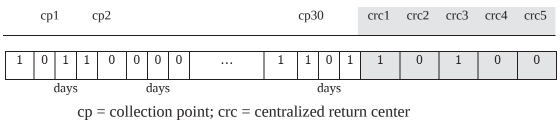

Fig. 3. A genetic representation scheme: cp = collection point; crc = centralized return center.

centers. Each collection point has four genes: the first gene represents opening (=1)/closing (=0) decisions; the remaining three genes represents collection periods for collection points so that eight possible values from 0 to 7 are obtained. As shown in Fig. 3, collection point 1 is open and has a 3 day-interval for consolidation of returned products for transshipment (e.g., 3 days $= 0 \times 4 + 1 \times 2 + 1 \times 1$); collection point 2 is closed; collection point 30 is open and has a 5 day-interval for consolidation (e.g., 5 days$=1\times4+0\times2+1\times1$). Each centralized return center has one gene representing an opening/closing decision to keep centralized return centers 1 and 3 open.

5.2. Genetic operators

The proposed GA solution procedure used four genetic operators described below.

5.2.1. Cloning operator

The cloning operator involves keeping the best solutions. In the proposed GA, the procedure works in such a way that it copies 20 percent of the current best chromosomes to a new population.

5.2.2. Parent selection operator

The parent selection operator is an important process that directs a GA search toward promising regions in a search space. Two parents are selected from the solutions of a particular generation by selection methods that assign reproduction opportunities to each individual parent in the population. There are a number of different selection methods, such as roulette wheel selection, tournament selection, rank selection, elitism selection, and random selection [22]. For this experimentation, we used a binary tournament selection method that began by forming two teams of chromosomes [29]. Each team consists of two chromosomes randomly drawn from the current population. The two best chromosomes that are taken from one of the two teams are chosen for crossover operations. As such, two offspring are generated and enter into a new population.

5.2.3. Crossover operator

The crossover operator generates new children by combining information contained in the chromosomes of the parents so that new chromosomes will have the best parts of the parents' chromosomes. The crossover probability indicates how often a crossover will be performed. There are several types of crossovers, including single-point crossover, multi-point crossover, and uniform crossover [22]. Herein, we applied the two-point crossover in which one is used for locating initial collection points and another for locating centralized return centers. The two locations of the crossover points are randomly selected in opening/closing decisions of facilities and then swap segments of the two parents' strings to produce two children.

5.2.4. Mutation operator

After recombination, some children undergo mutation. Mutation operates by inverting each bit in the solution with some small probability, usually from zero to 10 percent. The rationale is to provide a small amount of randomness, and to prevent solutions from being trapped at a local optimum. The type of mutation varies depending on the encoding as well as the crossover. In the proposed GA, the mutation operator first randomly selects a bit value of opening/closing decision variables on a chromosome, and then, flips a bit value from 0 to 1, or from 1 to 0. Next, in case of the decision variables for collection points, if the changed bit value is 0, make all zeros in the corresponding three bits for collection period; otherwise, randomly generate the three bit values. Hence, a good level of diversity in each generation is achieved.

5.2.5. Fitness function

Decoding the chromosome generates a candidate solution and its fitness value based on the fitness function. The fitness value is a measure of the goodness of a solution with respect to the original objective function and the "amount of infeasibility". The fitness function is formed by adding a penalty to the original objective function. To elaborate, the original objective function is comprised of various costs, such as the cost of renting initial collection points, the cost of carrying inventory at initial collection points, the cost of handling returns at initial collection points, the cost of establishing centralized return centers, and the cost of transshipment from initial collection points to centralized return centers.

In particular, we first generate decision variables from chromosomes using genetic operators such as opening/closing collection points, collection periods, and opening/closing centralized return centers. Then, based on the set of the variables, a fitness value of each chromosome can be obtained by applying two consecutive procedures. These procedures are coded in C + + and combined in the overall GA solution procedure.

The first step is used for obtaining total daily demand of the opened collection points. In other words, all customers should be allocated to the nearest collection points. To do this, we applied an assignment algorithm since we assumed that there is enough capacity of each collection point owing to small volume of returns. The mathematical representation is as follows:

$$\text{Minimize} \quad bw \sum_j \left\{ \sum_i r_i Y_{ij} \frac{(T_j+1)}{2} \right\} + a \sum_j Z_j$$

$$\text{Subject to} \quad \sum_j Y_{ij} = 1, \quad \forall i \in I,$$

$$\sum_i Y_{ij} \leqslant M Z_j, \quad \forall j \in J,$$

$$d_{ij} Y_{ij} \leqslant l, \quad \forall i \in I, \ \forall j \in J,$$

$$Y_{ij} \in (0,1), \quad \forall i \in I, \ \forall j \in J.$$

Next, the second step is used for assigning opened collection points to an appropriate centralized return center according to capacity limitation. In order to solve this, we applied a simplex method for a transportation problem. The mathematical representation is as follows:

$$\text{Minimize} \quad \sum_k \left\{ G_k \sum_j \left(X_{jk} \frac{w}{T_j} \right) f(X_{jk}, d_{jk}) \right\} + \sum_k q_k G_k$$

$$\text{Subject to} \quad \sum_i r_i Y_{ij} T_j = \sum_k X_{jk}, \quad \forall j \in J,$$

$$\sum_j X_{jk} \leqslant m_k G_k, \quad \forall k \in K,$$

$$0 \leqslant X_{jk}, \quad \forall j \in J, \ \forall k \in K.$$

The penalty function is mathematically expressed as:

$$\text{Penalty function} = \sum_j \sum_k pv \times g(X_{jk}, m_k, G_k) + \sum_i \sum_j pv \times h(d_{ij}, Y_{ij}, l),$$

where pv = penalty value.

$g(X_{jk}, m_k, G_k) = 1$ if $\sum_j X_{jk} > m_k G_k$; otherwise 0,

$h(d_{ij}, Y_{ij}, l) = 1$ if $d_{ij} Y_{ij} > l$; otherwise 0.

The penalty value is considerably larger than any possible objective function value corresponding to the current population of individuals.

5.3. An overall GA solution procedure

Once the representation scheme is selected, the overall algorithm of the proposed GA can be described as follows:

(1) Read the required data and generate an initial population based on population size, in which each chromosome is a one-dimensional array representing decision values. In each chromosome, first the opening/closing decision of any facility is randomly made using binary value (0 or 1). Second, if a collection point is open, a value of three genes for the collection period is randomly determined using binary values; If an initial collection point is closed, all three genes are zero.
(2) Set the generation zero and evaluate the fitness function of each chromosome in a population. The fitness function is the sum of the objective function of the original problem and the penalty function.
(3) Create a new population by repeating generation operations (cloning, parent selection, crossover, and mutation) until the new population is complete. The combined tournament and elitism method is used for selecting the parent. Two-point crossover and random mutation are used for positioning a chromosome.
(4) Replace new offspring in a new population.
(5) Stop the iteration if the end condition is satisfied; otherwise go to the next generation. Herein, the overall pseudo-code procedure for the proposed heuristic is outlined:

```
Read_Data( );
Initialize_Population( );
while (not terminate condition) do
        Evaluate_Fitnessfunction
        {
                Check_Feasibility( )
                Sol_Assignment( );
                Sol_Transportation( );
                Add_Penalty( );
        }
        Cloning( )
        Select_Parents( );
        Crossover( );
        Mutation( );
endwhile
Generate_Outputs;
```

6. Model experiments with sensitivity analysis

For illustrative purpose, a base-line model was solved by using the proposed GA that sets the parameter values through extensive experiments. These parameters are: population size = 400; maximum number of generations = 200; cloning = 20%; crossover rate = 80%; mutation rate varies from 5% to 10% as the number of generations increases. The GA solution procedure was executed on an IBM Pentium III computer equipped with a speed of 512 KB, and 256 MB of memory. The base-line solution required 12.48 min of CPU time. However, as Table 5 shows, larger values in population size and maximum number of generations tend to increase CPU time for model experiments. Fig. 4 shows the best fitness values at each generation as a function of the number of generations. Fig. 5 shows a graphical representation of the best solution with an objective function value

Table 5
The summary of model experiments with changes in GA parameters

Generation/population	100/300	200/300	300/300	400/300	500/300
Total cost	$216,520	$212,808	$213,500	$213,135	$213,850
Run time (in minutes)	4.24	8.30	12.62	16.91	21.30
Population/generation	200/200	300/200	400/200	500/200	600/200
Total cost	$213,135	$212,808	$211,570	$212,925	$213,195
Run time (in minutes)	5.64	8.30	12.48	15.58	18.58

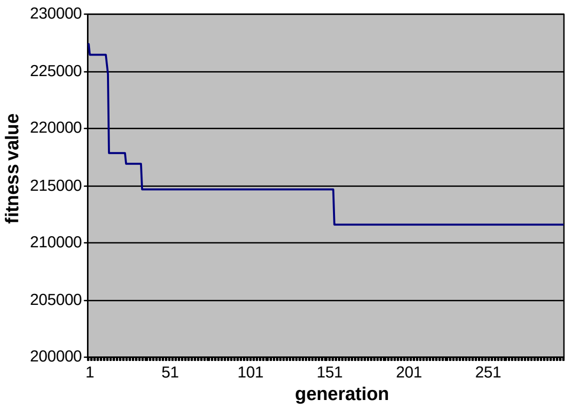

Fig. 4. The convergence of fitness values.

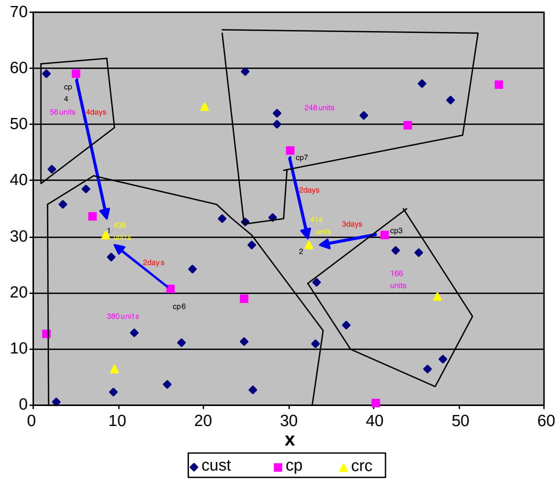

Fig. 5. A graphical display of the best base-line solution.

of $211,579, where four initial collection points and two centralized return centers are open. Table 6 summarized the base-line solution, where initial collection points 3, 4, 6, and 7 are selected and their collection periods are 3, 4, 2, and 2, respectively; centralized return centers 1 and 2 are open with their throughput capacity of 436 and 414, respectively.

6.1. *A sensitivity analysis of the maximum holding period to the model solution*

A longer maximum holding period for consolidation increases inventory carrying costs, but it can reduce total reverse logistics costs due to increased freight consolidation opportunities. To examine the sensitivity of the maximum holding period to total reverse logistics costs, we experimented on the model by changing the maximum holding period. Table 7 summarizes the results of sensitivity analysis with four different maximum holding periods (e.g., 1 day, 2 days, 3 days, and 4 days). The model experiments indicated that as the maximum holding period increased, the total reverse logistics cost decreased (e.g., 100%, 98%, 85%, and 84%), but the overall network structure remained stable, opening two centralized return centers and four initial collection points. In particular, we noticed a dramatic cost saving in total reverse logistics costs after setting the maximum holding period at three days (see Table 7).

6.2. *A sensitivity analysis of locations of initial collection points to the model solution*

The proximity of initial collection points to customer population centers can enhance the level of customer service due to easy access to customers' locations. However, to reduce distances between initial collection points and customers, *Beta-com* would require a larger number of initial collection points and thereby increase total reverse logistics costs. Thus, a sensitivity analysis was conducted to determine the desirable proximity of initial collection points to customers. Table 8 shows the summary of the sensitivity analysis results with four different possibilities of proximity (e.g., 17, 21, 25, and 29 miles). As expected, as the distance between initial collection points and customer locations increased, the total reverse logistics cost decreased (e.g., 100%, 95%, 88%, and 83%). Also, the longer distance between initial collection points and customer locations led to reduction in the total number of initial collection points and the subsequent savings in total reverse logistics cost. Compromising total reverse logistics cost and customer service, we recommend 25 miles between initial collection points and customers as

Table 6
The summary of the base-line model solution

Number of centralized return centers = 2	
Number of initial collection points = 4	
(Initial collection point; the number of collection frequency) = {(3; 3), (4; 4), (6; 2), (7; 2)}	
(Initial collection point; Customers allocated to the respective initial collection point)=	
{(3; 10, 11, 12, 21, 22, 23), (4; 3, 18), (6; 1, 2, 4, 6, 9, 15,19, 24, 25, 26, 27, 28, 29, 30), (7; 5, 7, 8, 13, 14, 16, 17, 20)}	
(Initial collection point; Volume of returned products per collection) = {(3; 498), (4; 224), (6;760), (7;496)}	
(Centralized return center; Initial collection point) = {(1 : 4, 6), (2; 3, 7)}	
(Centralized return center; throughput capacity) = {(1; 984), (2; 994)}	
Total annual cost of renting initial collection points	$800
Total cost of establishing centralized return centers	$6,000
Totalinventory carrying costs	$35,350
Total handling costs	$21,250
Total transportation costs	$148,170
Total annual reverse logistics costs	$211,570

Table 7
A sensitivity analysis with the varying maximum holding period

	Maximum holding periods			
	1 day	2 days	3 days	4 days
Total cost	$251,420	$245,803	$213,000	$211,570
Number of centralized return centers	2	2	2	2
Number of collection points	4	4	4	4
{(Centralized return center; Collection point)}	{(1;6,8),(5;4,7)}	{(1;6,8),(5;4,7)}	{(1; 6,7), (5;3,4)}	{(1;4,6),(2;3,7)}
{(Collection point; Holding period)}	{(4;1),(6;1), (7;1),(8;1)}	{(4;2),(6;1), (7;1), (8;2)}	{(3;3),(4;3), (6;2),(7;2)}	{(3;3),(4;4), (6;2), (7;2)}
Difference in total cost	100%	98%	85%	84%

Table 8
A sensitivity analysis with varying proximity of initial collection points

	Proximity			
	17 miles	21 miles	25 miles	29 miles
Total cost	$240,120	$228,130	$211,570	$198,140
Number. of centralized return centers	3	2	2	2
Number of collection points	8	6	4	3
{(Centralized return center; Collection point)}	{(2; 3,5,8), (3;2), (5;4,7,9,10)}	{(1; 1,4,5,10), (3;2,9)}	{(1; 4, 6), (2;3,7)}	{(1;9), (2;1,5)}
{(Collection point; Holding period)}	{(2;4), (3;6), (4;7), (5;3), (7,3),(8,5),(9,5), (10,6)}	{(1;3), (2;4), (3;4), (4;6),(8,3),(9,4)}	{(3;3), (4;4), (6;2), (7;2)}	{(1;2), (5;1), (9;4)}
Difference in total cost	100%	95%	88%	83%

Table 9
A sensitivity analysis with varying unit inventory carrying cost

	Unit inventory carrying cost			
	0.01	0.05	0.1	0.15
Total cost	\$175,248	\$196,925	\$211,570	\$214,345
Number of centralized return centers	2	2	2	2
Number of collection points	5	4	4	4
{(Centralized return center; collection point)}	{(1; 4,6,9), (2;3,10)}	{(1; 4,6), (4;1,8)}	{(1; 4, 6), (2;3,7)}	{(1; 4, 6), (2;3,7)}
{(Collection point; Holding period)}	{(3;5), (4;5), (6;3), (9;6), (10;5)}	{(1;3), (4;6), (6;2), (8;5)}	{(3;3), (4;4), (6;2), (7;2)}	{(3;4), (4;5), (6;2), (7;2)}
Difference in total cost	100%	112%	121%	122%

the ideal distance due to dramatic savings in total reverse logistics cost and substantial reduction in the number of initial collection points (from eight to four).

6.3. A sensitivity analysis of unit inventory carrying cost to the model solution

The longer the returned products were held at the initial collection point, the greater the saving in transportation cost, but the larger the inventory carrying cost that was incurred. Considering such a tradeoff, it would be worthwhile examining interplay between inventory carrying cost and transportation cost. As such, we experimented on the model with varying unit inventory carrying costs (e.g., \$0.01, \$0.05, \$0.1, and \$0.15), while keeping unit freight rate constant at one dollar. A result of the sensitivity analysis summarized in Table 9 shows that the total reverse logistics cost was very sensitive to changes in unit inventory carrying cost (e.g., 100%, 112%, 121%, and 122%). This result suggests that inventory control at the initial collection point is the key to successful reverse logistics operations for the returned product and a high-value product (e.g., an A-item in ABC classification) should require faster shipment and shorter holding time than a low-value product.

6.4. A sensitivity analysis of genetic algorithm parameters to the model solution

To reveal the effect that key parameters (i.e., population size and the maximum number of generations) of the proposed GA have on model solutions, we experimented the model with ten different sets of parameters. As summarized in Table 5 and Fig. 6, the model solution seems to be insensitive to changes in population size and the maximum number of generations.

None of differentials in total reverse logistics cost exceeded \$5,000. For example, Fig. 6 shows that six different combinations of GA parameters produced either identical or nearly-identical results. In other words, changes in population size and the maximum number of generations for GA do not affect the model solution significantly. This also implies that the proposed GA solution procedure is robust to changes in GA parameters.

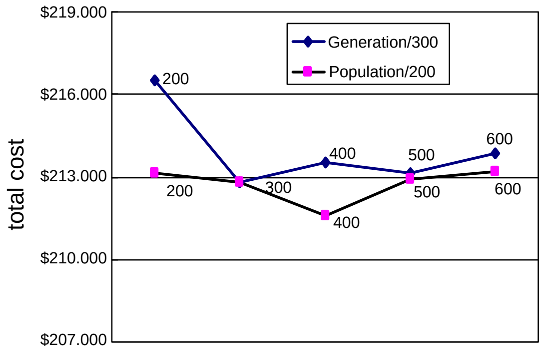

Fig. 6. A sensitivity to changes in GA parameters.

7. Concluding remarks and future research directions

Since both initial collection points and centralized return centers play a key role in successful operations of reverse logistics operations, the location/allocation decisions regarding the initial collection points and centralized return centers all but determine the success and failure of reverse logistics operations. This paper proposes a mathematical model and GA which aim to provide a minimum-cost solution for the reverse logistics network design problem involving product returns. The proposed model and solution procedure consider explicitly, trade-offs between freight rate discounts and inventory cost savings due to consolidation and transshipment. As such, the model and solution procedure enables reverse logisticians to determine the exact length of holding time for consolidation at the initial collection points and total reverse logistics costs associated with product returns. Computational experimentation reveals that GA presented a promise in solving practical-size problems with 30 customers, 10 potential sites of initial collection points, and five potential sites of centralized return centers. Also, the model and solution procedure produced the multi-echelon reverse logistics configuration that considers the interplays between initial collection points and centralized return centers. Despite numerous merits, the proposed model and solution procedure point to a number of directions for future work:

(1) The model can be expanded to include the element of risk and uncertainty involved in the reverse logistics network design problem.
(2) The theme of future research should include multi-objective treatments of the reverse logistics network design which explicitly analyze the tradeoffs among cost, response time, market potential, and speedy returns.
(3) The consideration of what-if scenarios involving changes in parameter values over time may be explored in the future.
(4) The comparisons of GA to other heuristics such as Lagrangian relaxation, heuristic concentration, and tabu search methods are worth investigating in future studies.
(5) The multi-echelon hierarchical network configuration, which considers the options of both direct shipment from customers to centralized return centers and indirect shipment through initial collection points, may be an intriguing subject for further studies.

Acknowledgements

The authors wish to thank the UPS Foundation for partly sponsoring this research. This research was also supported in part by Grant Number R01-2003-000-10077-0 available from the interdisciplinary program of the Korean Science and Engineering Foundation.

References

[1] ReturnBuy. The new dynamics of returns: the profit, customer and business intelligence opportunities in returns. Ashburn, Virginia; White Paper; 2000, ReturnBuy.com.
[2] Rogers DS, Tibben-Lembke RS. Going backward: reverse logistics trends and practices. Nevada, Reno: Reverse Logistics Executive Council; 1999.
[3] Shear H, Speh TW, Stock JR. The warehousing link of reverse logistics. Presented at the 26th annual warehousing education and research council conference, San Francisco, CA; 2003.
[4] Poirier CC. Using models to improve the supply chain. Boca Raton, FL: St. Lucie Press; 2004.
[5] Stock JR, Reverse logistics. White Paper, Oak Brook, IL: Council of Logistics Management; 1992.
[6] Min H. A bi-criterion reverse distribution model for product recall. Omega 1989;17(5):483–90.
[7] Caruso C, Colorni A, Paruccini M. The regional urban solid waste management system: a modeling approach. European Journal of Operational Research 1993;70:16–30.
[8] Melachronoudis E, Min H, Wu X. A multiobjective model for the dynamic location of landfills. Location Science 1995;3(3):143–66.
[9] Del Castillo E, Cochran JK. Optimal short horizon distribution operations in reusable container systems. Journal of the Operational Research Society 1996;47(1):48–60.
[10] Barros AI, Dekker R, Scholten V. A two-level network for recycling sand: a case study. European Journal of Operational Research 1998;110:199–214.
[11] Thierry M. An analysis of the impact of product recovery management on manufacturing companies. PhD thesis, Erasmus University, Rotterdam, The Netherlands; 1997.
[12] Krikke HR. Recovery strategies and reverse logistics network design. PhD thesis, University of Twente; Enschede, The Netherlands; 1998.
[13] Krikke HR, Kooi EJ, Schurr PC. Network design in reverse logistics: a quantitative model. In: Stähly P, editor. New trends in distribution logistics. Berlin, Germany: Springer; 1999. p. 45–62.
[14] Jayaraman V, Guide Jr. VDR, Srivastava R. A closed-loop logistics model for remanufacturing. Journal of the Operational Research Society 1999;50:497–508.
[15] Jayaraman V, Patterson RA, Rolland E. The design of reverse distribution networks: models and solution procedures. European Journal of Operational Research, 2003;150(2): 128–49.
[16] Fleschmann M, Krikke HR, Dekker R, Flapper DP. A characterization of logistics networks for product recovery. Omega 2000;28:653–66.
[17] Stock JR. The seven deadly sins of reverse logistics. Material Handling Management 2001;56(3):5–10.
[18] Anonymous, Return to sender, Modern Materials Handling 2000;55(6):64–5.
[19] Nemhauser GL, Wolsey LA. Integer and combinational optimization. Chichester: England: Wiley; 1988.
[20] Schrijver A. Combinatorial optimization: polyhedra and efficiency. Berlin: Germany: Springer; 2003.
[21] Goldberg DE. Genetic algorithms in search, optimization and machine learning. Reading, MA: Addison-Wesley; 1989.
[22] Gen M, Cheng R. Genetic algorithms and engineering optimizations. New York: Wiley; 2000.
[23] Gen M, Choi J, Tsujimura Y. Genetic algorithm for the capacitated plant location problem with single source constraints. Proceedings of seventh European congress on intelligent techniques and soft computing, Session CD-7; 1999.
[24] Jaramillo JH, Bhadury J, Batta R. On the use of genetic algorithms to solve location problems. Computers and Operations Research 2002;29:761–79.
[25] Zhou G, Gen M. Genetic algorithm approach on multi-criteria minimum spanning tree problem. European Journal of Operational Research 1999;114:141–52.

[26] Palmer CC, Kershenbaum A. An approach to a problem in network design using genetic algorithms. Network 1995; 26:151–63.
[27] Zhou G, Min H, Gen M. A genetic algorithm approach to the bi-criteria allocation of customers to warehouses. International Journal of Production Economics 2003;86(1):35–45.
[28] Glover F, Laguna M. Tabu search. Norwell, MA: Kluwer Academic Publishers; 1997.
[29] Chu PC, Beasley JE. A genetic algorithm for the generalized assignment problem. Computers and Operations Research 1997;24(1):17–23.