Iranian Journal of Operations Research
Vol. 3, No. 1, 2011, pp. 89-103

A Genetic Algorithm for Choice-Based Network Revenue Management

F. Etebari[1,*], A. Aaghaie[2], F. Khoshalhan[3]

In recent years, enriching traditional revenue management models by considering the customer choice behavior has been a main challenge for researchers. The terminology for the airline application is used as representative of the problem. A popular and an efficient model considering these behaviors is choice-based deterministic linear programming (CDLP). This model assumes that each customer belongs to a segment, which is characterized by a consideration set, which is a subset of the products provided by the firm that a customer views as options. Initial models consider a market segmentation, in which each customer belongs to one specific segment. In this case, the segments are defined by disjoint consideration sets of products. Recent models consider the extension of the CDLP to the general case of overlapping segments. The main difficulty, from a computational standpoint, in this approach is solving the CDLP efficiently by column generation. Indeed, it turns out that the column generation subproblem is difficult on its own. It has been shown that for the case of nonoverlapping segments, this can be done in polynomial time. For the more general case of overlapping segments, the column generation sub-problem is NP-hard for which greedy heuristics are proposed for computing approximate solutions.
Here, we present a new approach to solve this problem by using a genetic algorithm and compare it with the column generation method. We comparatively investigate the effect of using the new approach for firm's revenue.

Keywords: *Customer choice-based revenue management, Choice-based deterministic linear programming (CDLP), Segmentation, Genetic algorithm, Airline application.*

Manuscript received on 3/10/2010 and accepted for publication on 9/04/2012.

1. Introduction

Talluri and Van Ryzin [14] defined revenue management as demand management decisions and the methodologies and system required to make them. Quantity-based revenue management includes fixed and perishable capacity control during a specified period with objective of maximizing the obtained revenue.

Revenue management (RM) models were introduced by Little wood [12] who presented simple techniques to solve traditional revenue management models. He presented a solution approach to set a booking limit for the number of seats which should be assigned to low fares in airline networks. Traditionally, RM systems have been built upon the independent demand model assumption. This assumption views demand as a sequence of requests for products, which are insensitive to the capacity controls applied by the airline, and to market conditions like price offered by the competition, frequency of departures, brand preference of customers, etc. (for further details, see Talluri and Van Ryzin [14]). There is a wide agreement nowadays about the limitations of this assumption, based on the observation that the sale of a product is indeed the outcome of a customer's purchase decision subject to market conditions.

*Corresponding Author.
[1] Phd Candidate, N. 17, 4th floor, Pardis St., MollaSadra Ave., Vanak Sq., Tehran, Iran. Email: f_etebari@dena.kntu.ac.ir
[2] Associated professor, N. 17, 4th floor, Pardis St., MollaSadra Ave., Vanak Sq., Tehran, Iran. Email: AAghaie@kntu.ac.ir
[3] Associated professor, N. 17, 4th floor, Pardis St., MollaSadra Ave., Vanak Sq., Tehran, Iran. Email: khoshalhan@kntu.ac.ir

Furthermore, the development of low-cost airlines offering simplified, undifferentiated fare structure, and their usual strategy of saturating a market with several flights during the day, have raised the interest in formally capturing customer choice behavior in RM systems.

The earlier work on the choice behavior in network is the passenger origin and destination simulator (PODS) studies of Belobaba and Hopperstand [1]. This work focused on understanding the revenue management implications of passenger choice behavior on traditional RM methods. Talluri and Van Ryzin [4] provided an exact analysis of the optimal control policy for a single leg RM model under a general discrete choice model of demand. Zhang and Cooper [19] analyzed choice among parallel flights in the same O-D (origin-destination) market. That model assumed that customers choose among the same fare class on different flights but not among fare classes. They developed bounds and approximations for the resulting dynamic program.

Cooper and Gupta [5] showed which models that ignore customer choice may lead to the policies that, when used repeatedly, drive revenue down, a phenomenon they call the “spiral-down effect”. Chen and Homem-de-Mell [4] assumed that each customer has a preference order to describe her behavior regarding the order of the classes for which she tries to purchase tickets. If the customer’s first choice is not available, she either tries her second choice or decides not to purchase anything and so on. They modeled each customer’s decision at each step, trying the next choice or leaving the system, as a Bernoulli random variable with known probability.

Gallego et al. [6] provided a customer choice-based deterministic linear programming model (CDLP) for network revenue management. They supposed that with a flexible product offering, the firm had the ability to provide customers alternative products to serve the same market’s demand. One limitation of their market demand model is that it does not allow for any kind of segmentation. Van Ryzin and Liu [9] used the analysis of the model provided by Gallego et al. [6] to extend the concept of efficient sets. They proved that when capacity and demand are scaled up proportionally, the revenue obtained under the choice-based deterministic linear programming problem converges to the optimal revenue under the exact formulation. They presented a market segmentation model to describe choice behavior. The segments are defined by disjoint consideration sets of products, where a consideration set is a subset of the products provided by the firm which customers view as options. Bront et al. [3] considered the CDLP model of Gallego et al. [8] and further work done by Van Ryzin and Liu [13]. They extended the model to a more general case, where customers can belong to more than one segment according to a Multinomial logit model. A new deterministic linear program was offered by Kunnukal and Topaloglu [10] for the network revenue management problem with customer choice behavior. They also used randomized linear programming in choice based revenue management [14]. Vulcano et al. [18] developed a maximum likelihood estimation algorithm in discrete choice models for airline revenue management. Meissner and Strauss [14] offered a new heuristic method for specifying bid prices. Regarding the large number of variables in a real-size network, they developed a column generation algorithm to solve the CDLP model. The subproblem of the column generation algorithm is formulated as a 0-1 fractional programming problem, where the sum of several ratios were to be maximized. Because of the NP-hardness of the problem, they proposed implementing a greedy heuristic to solve the subproblem in polynomial time.

Ben-Akiva and Lerman[2] and Train [17] analyzed different discrete choice models. A comprehensive overview of discrete choice models and application of these models to the airline industry were provided by Garrow [17].

Inspired by of Bront et al.’s results [3] the extended model of CDLP with overlapping segments are considered here. As mentioned, the subproblem of this program is NP-hard, and there is no known algorithm for finding a global optimal solution in polynomial time. We present a new metaheuristic method with a good efficiency to overcome the complexity of the fractional linear programming subproblem. Of course, how to construct chromosomes and how to design highly efficient evolutionary operators, i.e., mutation and crossover, are crucial to a successful implementation of the GA.

The remainder of this paper is organized as follows. General structure of CDLP and its assumptions are described in Section 2. The proposed GA including the chromosomes structure and its operators is illustrated in Section 3. Extensive simulation study is reported in Section 4, and conclusions are given in Section 5.

2. Choice-Based Revenue Management

There are two main challenges we are faced with in implementing a choice-based revenue management:

- Modeling customer choice behavior and its estimation from available data.
- Using revenue optimization methods that can deal with complex, choice-based models of demand.

2.1. Customer Choice Model

To model customer choice behavior, we can assume that each customer wishes to maximize her utility while her utility for alternatives is a random variable. The firm is offering a set of alternatives $C=\{1,2,\ldots,m\}$ for the customer n who has a choice (or consideration) set C_n with the utility U_{in} for each alternative $i \in C_n$. This utility, without loss of generality, can be decomposed into a deterministic (also called expected utility) denoted by ϑ_{in} and a mean-zero random component ε_{in}. Hence, we have the utility function as follows:

$$U_{in} = \vartheta_{in} + \varepsilon_{in}. \tag{1}$$

In many cases, the representative component ϑ_{in} is modeled as a linear combination of several attributes,

$$\vartheta_{in} = \beta^T x_{in}, \tag{2}$$

where β is an unknown vector of weights to be computed from data and x_{in} is the vector of observable attributes for alternative i available to customer n at time of purchase, such as time and date of departure, price, departure airport, airline brand, and so on.

A most commonly used model to study how customers make their choice is the multinomial logit (MNL) model [16]. In this model, it is assumed that the ε_{in} in the utility functions are independent and identically-distributed random variables with a Gumbel distribution. The probability that customer n chooses alternative $i \in C_n$ in an MNL model is given by

$$P_n(i) = \frac{e^{\beta^T x_{in}}}{\sum_{j \epsilon C_n} e^{\beta^T x_{in}+1}}. \tag{3}$$

Bront el al. [3] considered the CDLP model of Gallego et al. [6] and the work of Van Ryzin and Liu [13]. They extended the model to a more general case, where customers can belong to more than one segment according to the MNL model.

2.2. Choice-Based Deterministic Linear Programming Model

In order to describe the problem and the corresponding model, we need some definitions:

- **Itinerary**: A specific sequence of legs on which passengers travel from their origin to their ultimate destinations.

- **Fare classes**: Different prices for the same travel service, usually distinguished from one another by the set of restrictions imposed by the firms.
- **Product**: Generally defined by an itinerary and fare-class combination.
- **Consideration set**: A subset of products provided by the firm that a customer views as an option.
- **Segment**: Customers, based on their preferences, are divided to different segments with each segment being defined by a consideration set of products.

The objective should be to find the set of alternative products for firms to decide to offer to customers at the time of the decision. The prices are fixed and the firm aims to maximize its revenue.

To define our model, consider a network with m resources (legs) providing n products. The set $N=\{1,2,\ldots,n\}$ denotes the set of products and r_j is the associated revenue (fare) for product $j \in N$. We study capacity usage by defining the vector $c = \{c_1, c_2, \ldots, c_m\}$ to denoteing the initial capacities of resources (legs). Resource usage according to the corresponding product is presented by defining an incidence matrix $A = [a_{ij}] \in B^{m\times n}$. The matrix entries are defined by

$$a_{ij} = \begin{cases} 1, & if\ resource\ i\ is\ used\ by\ product\ j \\ 0, & otherwise. \end{cases} \tag{4}$$

A_j, the jth column of A, denotes the incidence vector for product j and notation $i \in A_j$ indicates that product j uses resource i. Note that one product can use more than one resource. Time has discrete periods and runs forward until a finite number T, and it is assumed that we have at most one arrival for each period of time and each customer can buy only a single product.

We divide customers into L different segments. A consideration set C_l, $l = 1,2,\ldots,L$, is used to describe each segment. Here, we can make the difference of this model clearer with the existing works on customer choice-based modeling. Gallego et al. [6] considered a unique segment $C_1 = N$ and unlike Van Ryzin and Liu's approach [13], we can have overlapping segments, that is, $C_l \cap C_{l'} \neq \emptyset$, for certain $l \neq l'$.

If we have one arrival, p_l represents the probability that an arriving customer belongs to segment l with $\sum_{l=1}^{L} p_l = 1$. We consider a Poisson process of arriving streams of customers from segment l with rate $\lambda_l = \lambda p_l$ and total arriving rate of $\lambda = \sum_{l=1}^{L} \lambda_l$.

In each period of time t, the firm should decide about his set of offers (i.e, a subset $S \subset N$ of products that the firm makes available for customers). If set S is offered, the deterministic quantity $P_j(S)$ indicates the probability of choosing product $j \in S$, and $P_j(S) = 0$, otherwise. By total probability law, we have $\sum_{j\in S} P_j(S) + P_0(S) = 1$, where $P_0(S)$ indicates the no-purchase probability.

As already stated, we use a multinomial logit (MNL) model to find customer choice probabilities. According to an MNL choice model, the vector $\vartheta_l \geq 0$ is a customer's preference vector for available products in consideration set C_l and ϑ_{l0} represents the no-purchase preference. We let $P_{lj}(S)$ denote the probability of selling product $j \in C_l \cap S$ to a customer from segment l when set S is offered. So, customer choice probability can be expressed as follows:

$$P_{lj}(S) = \frac{\vartheta_{lj}}{\sum_{h\in C_l \cap S} \vartheta_{lh} + \vartheta_{l0}}. \tag{5}$$

In a general case, as a firm cannot recognize the corresponding segment of an arrival in advance, we consider $P_j(S)$, the probability that the firm sells product j to an arriving customer as

$$P_j(S) = \sum_{l=1}^{L} p_l P_{lj}(S). \tag{6}$$

The expected revenue, by offering set $S \subset N$ from an arriving customer is given by:

$$R(S) = \sum_{j \in S} r_j P_j(S). \tag{7}$$

Given that we offer set S, let $P(S) = (P_1(S), \dots, P_n(S))^T$ be the vector of purchase probabilities and A be the incidence matrix of resource used by products. Then, the vector of capacity consumption probabilities $Q(S)$ is given by

$$Q(S) = A.P(S), \tag{8}$$

where $Q(S) = (Q_1(S), \dots, Q_m(S))^T$, and $Q_i(S)$ indicates the probability of using a unit of capacity on leg $i, i = 1,2, \dots, m$.

The firm's decision consists of deciding that at any period of time t, which set of products should be offered, while it could not distinguish each customer's related segment in advance. However, as choice probabilities are time-homogeneous and demand is deterministic, it only matters how many times each set S is offered, knowing that exactly which period is not important, and the variable $t(S)$ represents the number of periods during which the set S is offered. Another assumption is that we let variable $t(S)$ be continuous as well (i.e., the firm could offer a set S for a whole or a fraction of a period of time). The model's objective is to maximize the firm's revenue by deciding the number of periods of time for each set of products. The formulation of the CDLP problem is:

$$\begin{aligned} V^{CDLP} &= max \sum \lambda R(S) t(S) \\ s.t. &\sum \lambda Q(S) t(S) \leq c \\ &\sum t(S) \leq T \\ &t(S) \geq 0. \end{aligned} \tag{9}$$

There are $m + 1$ constraints in this model, where the first m constraints are related to availability of capacity and the last one is for time availability. Because of the number of constraints $(m + 1)$, we could have a maximum of $m + 1$ variables with a positive value in the base. There are some remarks to be made here about the CDLP model and its optimal solution.

First, we should decide how to apply the solution of the CDLP model in our real problem and assign a starting and an ending time to offer each product. As mentioned before, the CDLP model's solution does not give us a sequence of products and the times. However, to order offer sets, various heuristic approaches can help us. Van Ryzin and Liu [13] developed an efficient decomposition heuristic to solve this problem. Second, in this problem there are an exponential number of primal variables. This means that a problem with n products has $2^n - 1$ possible non-empty subsets of products of set N. Regarding the large number of variables in practical networks, they developed a column generation algorithm to solve the CDLP model. However, the subproblem of the column generation algorithm is formulated as a 0-1 fractional programming one where the sum of several ratios should be maximized. Because of the NP-hardness of this problem, they proposed implementing a greedy heuristic algorithm to solve the subproblem in polynomial time.

3. Designing GA for CDLP Problem

Here, we present a genetic algorithm for solving the linear programming problem whose number of variables grow exponentially and the resulted subproblem of the column generation method is NP-hard.

3.1. Genetic Algorithm

Genetic algorithm was developed by Holland in the 1970s to understand the adaptive processes of natural systems [18]. Gas comprise a very popular class of population-based metaheuristics. The algorithms start from an initial population of solutions. Then, they iteratively incur the generation of a new population and the replacement of the current population. This replacement is based on selection methods.

Due to the large diversity of initial populations, P-metaheuristics are naturally more exploration search algorithms whereas S-metaheuristics are more exploitation search algorithms. This special characteristic of the P-metaheuristics leads us to improved solutions for the problem. Once the selection of individuals to form the parents is made, the role of the reproduction phase is the application of variation operators such as the mutation and crossover. Mutation operators are unary operators acting on a single individual. The probability P_m defines the probability to mutate each element (gene) of the representation. The crossover operator is binary and sometimes n-ary. The role of crossover operators is to inherit some characteristics of the two parents to generate the offsprings. The crossover rate represents the proportion of parents on which a crossover operator will act.

3.2. Structure of Chromosomes

To begin, the structure of chromosomes should be defined. Chromosomes in this problem are in matrix form. Because of the number of constraints $(m + 1)$, we will have a maximum of $m + 1$ variables with a positive value in the optimal solution. Therefore, chromosomes will form an $(m + 1) \times n$ matrix. Each row of the chromosome is representing a set of combination of n available products, and columns correspond to the products. A gene of the chromosome, denoted by $g(r, j)$, is associated with an entry of the matrix, where $g(r, j) = 1$ means that product j is in set r and $g(r, j) = 0$ means product j is not in set r.

Let us consider a very small airline network with three cities, e.g., Tehran, Tabriz and Mashhad, to make the problem clearer. A firm is offering two fare classes, low and high, for each flight (leg). Figure 1 illustrates the network. Eight products have been defined by an itinerary and the fare class combination. Table 1 represents a sample chromosome for this problem. This chromosome is decoded as offering $S_1 = \{1,3,7\}$, during $t(S_1)$ periods, $S_2 = \{1,2,3,7\}$, during $t(S_2)$ periods, $S_3 = \{1,3,4,6,7\}$, during $t(S_3)$ periods and $S_4 = \emptyset$, during $t(S_4)$ periods. Then, we have a finite set of products, corresponding to this chromosome. Other related inputs for this algorithm are:

- The capacity usage vector has $m + 1$ cells and the first m cells denote the initial capacities of legs and the last column represents the time horizon.
- The vector λ has l cells (number of segmentations) and each cell denotes the probability of having an arrival in a period of time.
- Resource usage matrix with the dimension $m \times n$, with each cell (*i*,*j*) denoting a usage of product j from resource i.
- Revenue vector which has n cells, each cell j denoting revenue (fare) for product j.
- Customer's preference matrix with the dimension $l \times (n + 1)$, with each cell (*i*,*j*) denoting preference of lth segment customers to jth product. The first column of this matrix indicates no-purchase preference.

Based on the information and the existing chromosome, we will formulate a related linear programming model. By solving this model, the optimal offering time periods for each set, being $t(S_1)$, $t(S_2)$, $t(S_3)$ and $t(S_4)$ for the above example, and the corresponding revenue is found.

3.3. Definition of the Operators

The basic crossover operator, the 1-point crossover, is used in this algorithm. To apply this operator, a crossover site k is randomly selected, and two offsprings are created by interchanging the segments of the parents. Figure 2 illustrate an example of the crossover operator with k being equal to 3.
Mutation introduces some diversification in the individuals by introducing some missing values in the current individuals of a population. Two different types of mutation operators are used. The first one, which operates intelligently, tries to diversify chromosomes while leading them toward an optimal solution. The second operator operates randomly and tries to cover the solution space.

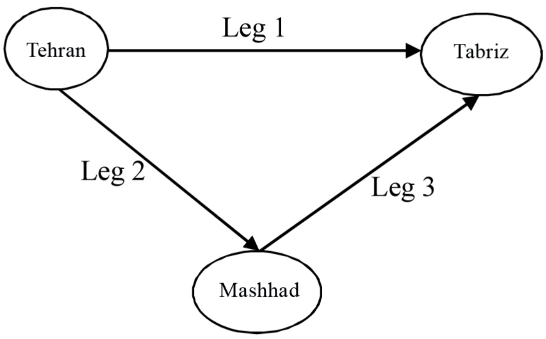

Figure 1. A small airline network

Table 1. Sample chromosome in a small airline

	P1	P2	P3	P4	P5	P6	P7	P8
S1	1	0	1	0	0	0	1	0
S2	1	1	1	0	0	0	1	0
S3	1	0	1	1	0	1	1	0
S4	0	0	0	0	0	0	0	0

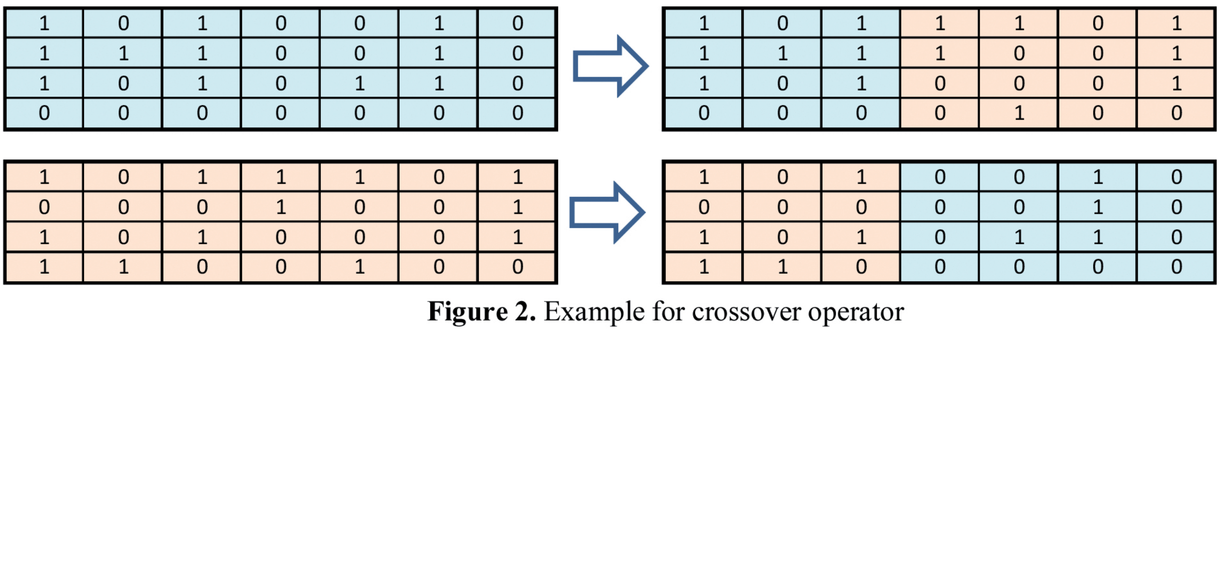

Figure 2. Example for crossover operator

3.3.1. Mutation Operator Type I

In order to define this operator, we use the concept of efficient sets [1]. An inefficient set T provides strictly less revenue $R(T)$ than do other sets and incurs at least as high a probability of consuming capacity $Q(T)$. After specifying parents on which a mutation operator will act (according to the mutation rate), revenue and capacity consumption for each set will be calculated accordingly and then the following ratio will be determined:

$$ER(S) = {R(S)}/{Q(S)}. \tag{11}$$

This ratio will be calculated for all sets in the chromosome. ER is related to efficiency ratio of each set. Then, the mutation operator will act on the set that has the minimum ratio. We try to eliminate sets that are inefficient. In the determined set, genes will be chosen randomly and their

1	0	1	0	0	1	0
1	1	1	0	0	1	0
1	0	1	0	1	1	0
0	0	0	0	0	0	0

1	0	1	0	0	1	0
1	1	1	0	0	1	0
0	1	0	1	0	0	1
0	0	0	0	0	0	0

Figure 3. Example for mutation operator

values will be changed to their complements. For instance, assume that theefficiency ratio for the 4 sets in Figure 3 are 1254, 1356, 897 and 987.

3.3.2. Mutation Operator Type II

Although operator type I has a specific characteristic, which is intelligence, characteristic of the type II operator is randomness. For specified parents on which a mutation operator will act, this operator chooses genes randomly and changes their values to their complements. An advantage of operator type I is that it acts intelligently and tries to move toward an optimal solution, but it has a drawback. The intelligence may cause a premature termination and do not give an opportunity for covering all the solution space by the chromosomes.

Experimental results show that the type I mutation operator produces revenue gains over the type II mutation operator in the tightly constrained capacity cases.

3.4. Selection Method

The selection method is a main search component in GAs. The selection strategy determines which individuals are chosen for mating (reproduction) and how many offsprings each selected individual produces. In this problem, the roulette wheel selection mechanism with elitism is used. We assign a selection probability to each individual that is proportional to its relative fitness and selection is done based on this probability. We consider an elitism rate in the beginning of the algorithm which denotes the percent of individuals that are directly transformed into the next generation. The population of the old generation is replaced with the new offsprings and the selected elites.

3.5. Stopping Criterion

We use an adaptive condition for the stopping criterion. Maximum number of iterations (generations) without improvement is used in this algorithm.

3.6. Parameter Tuning

A specific systematic approach is needed for tuning the parameters. The design of experiment methodology is used for the tunings. Design of experiments is a process of planning the specific experiments and running them for collecting appropriate data and analyzing the data by statistical methods. One most popular design for tuning parameters is the factorial design, which we adopt to use here. For this purpose, three critical values for each factor are selected. Fractional experiments are used to specify the level combinations to be used.

4. Computational Results

Here, we test our algorithm to solve the network revenue management problem when customers belong to overlapping segments in accordance with the MNL model and then compare the obtained results with the ones produced by the column generation algorithm. We assess the convenience of each model based on the quality of the solutions in term of the revenue obtained.

We consider different capacities by multiplying a scale factor α to the capacity of lags. Values of α = 0.6, 0.8, 1, 1.5 and 2 are used to solve the problem. We alter the preference in the choice behavior by varying the no-purchase performance vector $\vartheta_0 = (\vartheta_{10}, \vartheta_{20}, \dots, \vartheta_{l0})$.

Results for a small network are obtained with 50 instances having populations of size 200 for each instance as well as 40 instances with populations of size 100.

4.1. A Small Airline Network

First, we start evaluating the heuristic algorithm within column generation method and metaheuristic solution in a small network airline network. This example is also considered with different details by Liu and van Ryzin [13] and Bront et al. [3]. Consider a network with 4 airports and 7 flight legs. The capacities of the legs are $C = (100,150,150,150,150,80,80)$. The firm offers two high (H) and low (L) fares on each leg. Considering local and connecting itineraries, customers can choose among 22 available products defined by itineraries and fare class combinations. The problem consists of finding a policy which leads to prepare a set of products at any period of time during the booking horizon to offer to the customers while the revenue of the firm is to be maximized. The airline network is illustrated in Figure 4 and Tables 2 and 3 describe available products and customer segmentation in the network.

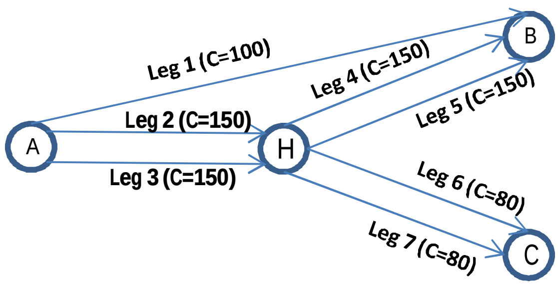

Figure 4. A small airline network

Table 2. Products definition in a small network

Products	Legs	Class	Fare	Products	Legs	Class	Fare
1	1	H	1000	12	1	L	500
2	2	H	400	13	2	L	200
3	3	H	400	14	3	L	200
4	4	H	300	15	4	L	150
5	5	H	300	16	5	L	150
6	6	H	500	17	6	L	250
7	7	H	500	18	7	L	250
8	{2,4}	H	600	19	{2,4}	L	300
9	{3,5}	H	600	20	{3,5}	L	300
10	{2,6}	H	700	21	{2,6}	L	350
11	{3,7}	H	700	22	{3,7}	L	350

Table 3. Customer segmentation in a small network

Segment	O-D	Con. Set	Observed utility	Landa
1	A-B	{1,8,9,12,19,20}	(10,8,8,6,4,4)	0.08
2	A-B	{1,8,9,12,19,20}	(1,2,2,8,10,10)	0.2
3	A-H	{2,3,13,14}	(10,10,5,5)	0.05
4	A-H	{2,3,13,14}	(2,2,10,10)	0.2
5	H-B	{4,5,15,16}	(10,10,5,5)	0.1
6	H-B	{4,5,15,16}	(2,2,10,8)	0.15
7	H-C	{6,7,17,18}	(10,8,5,5)	0.02
8	H-C	{6,7,17,18}	(2,2,10,8)	0.05
9	A-C	{10,11,21,22}	(10,8,5,5)	0.02
10	A-C	{10,11,21,22}	(2,2,10,10)	0.04

In Table 3, column 3 and 4 specify the corresponding consideration set and the preference values for the indicated products. The probability of a customer arrival for the corresponding segment is given in the last column. Booking horizon of this problem consists of 1500 periods of time. Table 4 summarizes the revenue obtained with two different algorithms.

Table 4. Computational results for parallel flights

α	ϑ_0	Column generation REV.	GA REV	%GA-CG	Average
0.6	(1,5)	215,793	213,980	-0.84	-0.53
	(5,10)	200,515	199,031	-0.74	
	(10,20)	170,137	170,137	0.00	
0.8	(1,5)	266,934	264,549	-0.89	-0.34
	(5,10)	223,173	223,172	0.00	
	(10,20)	188,574	188,334	-0.13	
1	(1,5)	281,967	282,325	0.13	0.59
	(5,10)	235,284	237,349	0.88	
	(10,20)	192,038	193,489	0.76	
1.2	(1,5)	284,772	285,834	0.37	0.61
	(5,10)	238,562	238,562	0.00	
	(10,20)	192,373	195,163	1.45	
1.4	(1,5)	287,076	290,354	1.14	1.03
	(5,10)	238,562	243,245	1.96	
	(10,20)	192,373	192,373	0.00	

First and second columns respectively correspond to the scale factor and the no-purchase preference vector. For instance, the first row in the table represents the case $C_{0.6} = 0.6 \times (100,150,150,150,150,80,80) = (60,90,90,90,90,48,48)$ and $\vartheta_0 = (5,10)$, namely the weight for the no-purchase option for first segments of each origin-destination are 5 and for second segments are 10. Third column is the obtained revenue from column generation algorithm and the next column represents the results of the genetic algorithm. Next and last columns are the gap of the results of the two algorithms and the average improvement obtained in applying the genetic algorithm. Generally, we can observe that that generic algorithm obtains nearly equivalent or better results in comparison with the column generation. In case that there is enough capacity, the genetic algorithm generates better results and when the capacity is tightened (lower values of C and no-purchase utility), both algorithms produce nearly the same results. Consider that the gain of 2% is significant in revenue management. Figure 6 shows the gap between obtained revenues from the two different algorithms and Figure 5 shows the average revenue gaps within different scale factors.

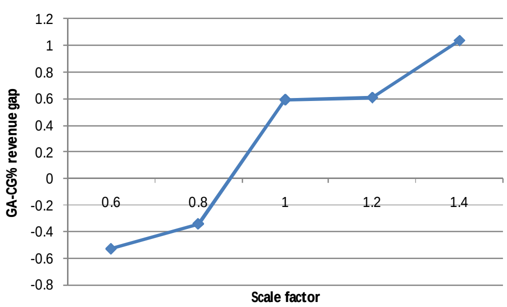

Figure 5. Average revenue gaps in different scale factors

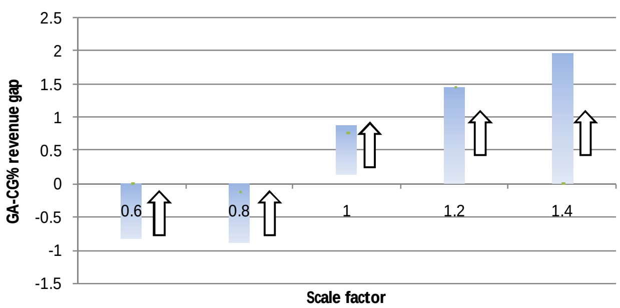

Figure 6. Revenue gaps in terms of scale factor and no-purchase preference

gives better results than the column generation method. This behavior of the genetic algorithm can be justified according to the solution space and its exploration characteristic. As the scale factor and customer no-purchase preferences are increasing, the solution space is expanding and because of the exploration strength (covering all solution space) of population-based algorithms, quality of the results is increasing.

4.2. A Railroad Network

Railroad network is based on specific railroads in Europe, of which we will consider a part of its network with five cities and four legs [19]. There are two high (H) and low (L) fare classes on each leg. Figure 7 illustrates this railroad network.

In this problem, there are 10 trains with a capacity of 100 passengers going from Paris to Amsterdam. Each train stops in Brussels, Rotterdam, Schiptol and Amsterdam. Thus, there are 10 markets shown in Table 4. Two fare classes and 10 markets produce a total of 200 products. Table 4 shows the price information associated with each market.

Customers are divided into 20 different segments based on their sensitivity to price and their origins and ultimate destinations. Table 5 shows each segment's definition according to our assumptions. We assume booking horizon including 1000 time periods. The experiments are done for three scale factors including 0.5, 1 and 1.5. The no-purchase preference is given in the last coordinate of the preference vector. Table 6 summarizes the results under different scenarios.

For this computation, different booking horizons are analyzed. Results indicate that similar to the previous network, our proposed algorithm has a better performance in comparison with the heuristic algorithm and specially the exploration capability of the method helps for obtainment of better results in special situations.

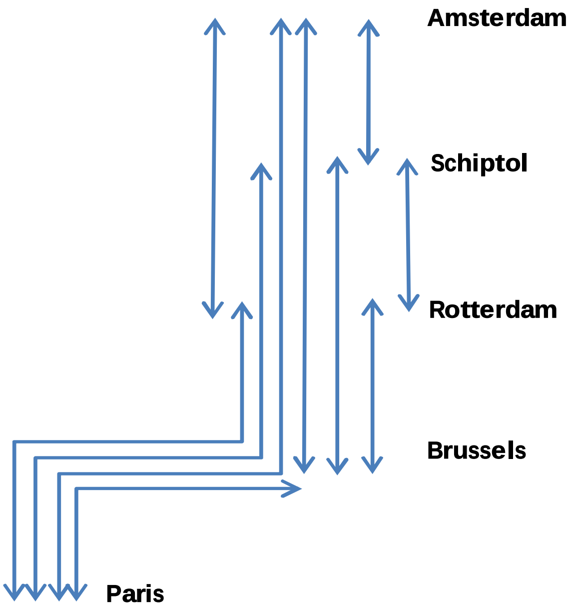

Figure 7. A railroad network

Table 4. Products definition in railroad network

O-D	Low fare	High fare
PAR-BRU	200	400
PAR-RTA	300	500
PAR-SCH	350	525
PAR-AMA	350	525
BRU-RTA	150	250
BRU-SCH	175	275
BRU-AMA	200	300
RTA-SCH	50	100
RTA-AMA	175	300
SCH-AMA	50	100

Table 5. Customer segmentation in railroad network

Segment	O-D	Cons. Set	Preference vector
1	PAR-BRU	{1,…,20}	{10,55,25,15,6,4,3,4,5,6,15,15,20,4,3,2,1,2,2,3,8}
2	PAR-BRU	{11,…,20}	{8,70,60,10,7,4,4,4,5,40,60}
3	PAR-RTA	{21,…,40}	(15,30,20,10,3,5,20,25,10,4,4,4,8,2,1,2,2,3,3,2,2}
4	PAR-RTA	{31,…,40}	{7,40,25,10,4,4,5,15,20,25,45}
5	PAR-SCH	{41,…,60}	{25,25,20,4,5,5,5,6,6,10,30,2,2,2,3,3,3,4,4,10}
6	PAR-SCH	{51,…,60}	{7,32,21,3,3,4,5,15,15,20,30}
7	PAR-AMA	{61,…,80}	{20,20,2,5,5,6,6,7,7,8,15,3,3,4,3,3,4,4,5,4,4}
8	PAR-AMA	{71,…,80}	{50,25,20,3,3,4,4,8,20,28,35}
9	BRU-RTA	{81,…,100}	{10,60,50,6,4,4,5,20,22,7,32,10,4,3,2,2,2,3,4,4,15}
10	BRU-RTA	{91,…,100}	{20,90,45,5,6,2,3,4,30,60,70}
11	BRU-SCH	{101,…,120}	{5,25,10,5,5,6,6,20,20,10,8,5,4,3,3,3,4,4,5,5,5}
12	BRU-SCH	{111,…,120}	{10,35,7,6,4,4,5,6,7,35,40}
13	BRU-AMA	{121,…,140}	{30,24,4,4,3,3,5,6,6,10,10,3,2,2,2,2,3,4,5,5,6}
14	BRU-AMA	{131,…,140}	{15,8,6,5,4,5,6,7,10,12,10}
15	RTA-SCH	{141,…,160}	{10,25,20,4,4,3,3,4,5,6,10,4,4,3,2,2,3,3,4,4,4}
16	RTA-SCH	{151,…,160}	{4,34,36,3,2,2,4,4,5,25,30}
17	PAR-AMA	{161,…,180}	{20,40,10,5,4,3,4,5,5,6,25,4,2,1,2,2,2,3,4,4,5}
18	PAR-AMA	{171,…,180}	{5,50,25,25,3,4,5,6,6,35,40}
19	SCA-AMA	{181,…,200}	{30,32,20,5,4,4,4,5,6,7,20,4,4,3,2,3,3,4,4,5,5}
20	SCA-AMA	{191,…,200}	{15,40,20,4,4,4,5,6,6,35,60}

Table 6. Computational results for railroad network

α	Booking horizon	Column generation REV.	GA REV	%GA-CG	Average
0.5	500	159,781	159,797	0.01	-0.06
	1000	285,425	285,400	-0.01	
	2000	407,336	407,301	-0.01	
	3000	452,403	451,389	-0.22	
1	500	166,163	168,548	1.44	0.68
	1000	319,562	321,876	0.72	
	2000	570,850	570,850	0.00	
	3000	727,754	731,923	0.57	
1.5	500	166,163	169,765	2.17	0.81
	1000	330,053	330,052	0.00	
	2000	616,783	618,234	0.24	
	3000	856,275	863,442	0.84	

5. Conclusion

We considered a more general form of the choice-based, deterministic, linear programming model for overlapping segments. A multinomial logit model for acquiring customer choice was presented and the more general case that customers belong to overlapping segments and choose according to the multinomial logit model was studied. A linear fractional programming problem was developed by introducing a column generation algorithm to solve the model for practical networks. The resulting model being NP-hard, a greedy heuristic was proposed to solve it. Experience showed that there were circumstances that the greedy heuristic could not find an entering column, and thus mixed integer fractional programming was proposed. We proposed a metaheuristic algorithm for solving this problem. Two types of mutation operators, according to the size of the solution space of the problem were used in the algorithm. According to the results, we observed that increasing the scale factor and customer no-purchase preferences simultaneously, the solution space was expanding. In accordance with the exploration characteristic of population-based algorithms, the obtained results were promising.

References

[1] Belobaba, P. and Hopperstad, C. (1999), Boeing/MIT simulation study: PODS results update, in AGIFORS Reservation and Yield Management Study Group Symposium Proceedings, London, 27-30

[2] Ben-Akiva, M.E. and Lerman, S.R. (1985), Discrete Choice Analysis: Theory and Application to Travel Demand, The MIT Press.

[3] Bront, J.J.M., I. Méndez-Díaz, and G. Vulcano (2009), A column generation algorithm for choice-based network revenue management, *Operations Research*, 57(3), 769-784.

[4] Chen, L. and Homem-de-Mello, T. (2010), Mathematical programming models for revenue management under customer choice, *European Journal of Operational Research*, 203(2),294-305.

[5] Cooper, W.L., Homem-de-Mello, T. and Kleywegt, A.J. (2006), Models of the spiral-down effect in revenue management, *Operations research-baltmmore then linthicum*, 54(5), 968.

[6] Gallego, G., et al., Iyengar, G., Phlilips, R. and Dubey, A. (2004), Managing flexible products on a network, Department of Industrial Engineering and Operations Research, Columbia University.

[7] Garrow, L.A. (2010), Discrete Choice Modelling and Air Travel Demand, Georgia Institute of Technology, USA, Ashgate Publishing Company.

[8] Holland, J. (1973), Genetic algorithm optimal allocation of trails SIAM, *Journal of Computing*, 2(2), 88-105.

[9] Hosseinalifam, M. (2009), A fractional programming approach for choice-based network revenue management, M.S. Thesis, Montreal University.

[10] Kunnumkal, S. and Topaloglu, H. (2008), A refined deterministic linear program for the network revenue management problem with customer choice behavior, *Naval Research Logistics (NRL)*, 55(6), 563-580.

[11] Kunnumkal, S. and H. Topalogl u,(2011), A randomized linear program for the network revenue management problem with customer choice behavior, *Journal of Revenue and Pricing management*, 10, 455-470.

[12] Littlewood, K. (1972), Forecasting and control of passenger bookings, in *AGIFORS Symposium, Israel*, 95-117

[13] Liu, Q. and Van Ryzin, G. (2008), On the choice-based linear programming model for network revenue management, *Journal of Manufacturing and Service Operations Management*, 10, 288-311

[14] Meissner, J. and Strauss, A. (2012), Improved bid prices for choice-based network revenue management, *European Journal of Operational Research*, 217, 417-427.

[15] Talluri, K.T. and Van Ryzin, G. (2004), Revenue management under a general discrete choice model of consumer behavior, *Management Science,* 50**,** 15-33.

[16] Talluri, K. and Van Ryzin, G. (2004), The Theory and Practice of Revenue Management, New York, Kluwer Academic Publishers.

[17] TRAIN, K. E. (2009). Discrete Choice Methods with Simulation, New York, Cambridge University Press.

[18] Vulcano, G., van Ryzin,G. and Chaar, W. (2010), Choice-based revenue management: An emprical study of estimation and optimization, *Manufacturing And Service Operations Management,* 12, 371-392.

[19] Zhang, D. and Cooper, W.L. (2005), Revenue management for parallel flights with customer-choice behavior, *Operations Research*, 53(3), 415-431.