A Simulator To Estimate The Cancellations In YM

P.K.Suri[#1], Rakesh Kumar[#2], Pardeep Kumar Mittal[#3]

[#1]Dean(R&D), Chairman & Professor(CSE/IT/MCA), H.C.T.M., Kaithal(Haryana), India
[#2]Professor, Department of Computer Science & Applications, Kurukshetra University, Kurukshetra(Haryana), India
[#3]Assistant Professor, Department of Computer Science & Applications, Kurukshetra University, Kurukshetra(Haryana), India

Abstract- In case of Yield Management (YM), even if the yield is optimized using some technique, it is not guaranteed that the final revenue obtained will be maximum or optimized. Since in almost all the industries where YM can be implemented, some of the customers may either cancel their seats or some customers do not show-up even after their booking. These cases needs to be taken care of as lot of revenue loss may occur due to these cancellations or no-shows. This paper tries to estimate the cancellations via a simulator using GA in an airline YM. The simulator has been developed using MATLAB.

***Keywords- Advance Reservation, Cancellations,* Genetic Algorithm, No-shows, Yield Management**

I. Introduction

The drawback of the advance reservation (AR) system, to both consumers and service providers, is that consumers may either cancel their reservations or simply may not show up at the time when the contracted service is scheduled to be delivered. This will leave some capacity unused, thereby resulting in a loss to service providers. Clearly, this loss can be minimized if service providers do not provide any refund to consumers who either cancel or do not show up. But in this competitive environment, it may not be possible to use such type of measures to avoid loss of revenue.

The best way to understand about cancellation or no-shows is to observe airline passengers. Passengers can be divided into business travellers and leisure travellers. Business travellers are most likely to cancel or change their reservations because their travel arrangements depend on others' schedules and business opportunities. In contrast, students can be sure of their time of travel because they tend to travel during semester breaks and holidays that are not subjected to last-minute changes. All this means is that students are more likely to engage in an advance purchase of discounted non-refundable tickets, whereas business travellers are less likely to commit in advance, and therefore are more likely to either purchase fully refundable tickets or to postpone their ticket purchase to the last minute.

If the cancellation could be estimated in an efficient manner, the revenue of the firm can be increased significantly. The cancellations in airlines industry can be as high as 50%. If this happens the airlines is bound to lose a huge amount of revenue. Therefore it is very important to estimate the cancellations. In the present paper, an attempt has been made to model the estimation of cancellations in airlines booking.

II. Literature Review

It is vital for companies to sell a product before the expiration date, otherwise it is worthless. One cannot sell an empty seat on an airplane that has already taken off or you cannot sell last night's empty hotel room. On the other hand, it may be that all the products have sold out before the expiration and to other customers the company has to say no. In that case it can occur that the airlines may sell out the flight, but on the departure day the plane takes off with empty seats. These passengers, who book a seat, but do not show-up at the departure time, for whatever reason, are called no-show passengers. This is not the case for just the airlines, hotels and restaurants have the same problem with customers who make a reservation and do not show up. Another possibility is that they cancel the reservation within a few hours before expiration. The airlines and hotel industry have good cancellation policies to protect themselves. It is possible to develop some methods to avoid no-shows, like customer reminders, deposits, standby passengers, overselling or no money back guarantees. Standby passengers are arriving at the airport and they do not know before the take off time if they will be go with that flight [1]. The only way that they will travel is if there will be an empty seat, caused by a no-show passenger.

Subramanian et al. [2] studied a more general setting, where they considered the arrival of a cancellation, the arrival of a booking request and no arrival of any type as a combined stream and assume that at most one of these events can occur at any discrete time epoch. Under this setting they presented two models. In the first model, the cancellation and no-show probabilities do not depend on the fare classes. They showed that the resulting problem can be equivalently modelled as a queuing system discussed in the literature [3]. In their second model, they relaxed the class independence assumption and model a more general problem with class dependent cancellations and no-shows.

In an article by Pulugurtha & Nambisan [4], a decision-support tool is developed to estimate the number of seats to each fare class. Genetic algorithm is used as a technique to solve this problem. The decision support tool considers the effect of time-dependent demand, ticket cancellations and overbooking policies.

A closely related study is given by Feng et al. [5]. They considered a continuous-time model with cancellations and no-shows. They derived a threshold type optimal control policy, which simply states that a request should be admitted only if the corresponding fare is above the expected marginal seat revenue (EMSR).

Recently reported passenger-based predictive models employ explanatory variables extracted from databases containing specific information on each passenger and the passenger's itinerary. Hueglin *et. al.* [6] have applied classification trees and logistic regression to the problems of predicting both no-shows and cancellations at the passenger level throughout the booking phase. Kalka and Weber [7] at Lufthansa have used induction trees to compute passenger-level no-show probabilities, and compared their accuracy with conventional, historical-based methods. Continental Airlines [8] describes a decision-tree model based on a relatively small number of input passenger records and features. Selby [9] discusses an application of radial basis functions to passenger-based forecasting, but does not report specific results.

III. Problem Definition and Formulation

In this problem an assumption regarding a flight operating between a specified origin and destination has been made. The reservation for the flight starts form the first date of expected reservation up to the date of departure. Another assumption is to fix the fare of each class and also assumed as known.

The number of customers travelling in each class should be greater than or equal to lower bound and less than or equal to the upper bound.

On the basis of above assumptions, the objective function can be written as:

$\text{Max. } \Sigma_\beta \Sigma_\alpha N_{\alpha,\beta} F_\beta$ ……………… (1)

Subject to the constraints

$\Sigma_\beta \Sigma_\alpha N_{\alpha,\beta} <= C_t$ & $L_{\alpha,\beta} <= N_{\alpha,\beta} <= U_{\alpha,\beta}$ for all α and β,

$N_{\alpha,\beta} >= 0$,

Where C_t = Total capacity of a flight

$N_{\alpha,\beta}$ = Number of customers belonging to class β during time slice α.

F_β = Fare for class β.

$U_{\alpha,\beta}$ = Upper limit of demand for class β during time slice α.

$L_{\alpha,\beta}$ = Lower limit of demand for class β during time slice α.

For estimating the cancellation , the following model is formulated.

Let π ($0 < \pi \le 1$) denote the probability that a consumer with a confirmed reservation actually shows up at the service delivery time. In the technical language, this probability is often referred to as a consumer's *survival probability*. It is assumed that all consumers have the same show-up probability, and that a consumer's show-up probability is independent of all other consumers. That is, events such as last-minute *group* cancellations are ruled out.

For estimating the expected number of show-ups for each booking level, *b*. formally, let the random variable *s* denote the number of consumers who show up at the service delivery time. Clearly, $s \le b$, meaning that the number of show-ups cannot exceed the number of bookings. That is, our model does not allow for standby customers and only customers with confirmed reservations are provided with this service. In fact, because *s* depends on the number of bookings made, *s* is a function of *b* and will often be written as $s(b)$. Also, note that *s* is a random variable, which also depends on the individual's show-up probability π, hence it can also be written as $s(b; \pi)$. The general formula for computing the probability that exactly c consumers show up given that *b* consumers have confirmed reservations for the service is given by the following binomial distribution function:

$\Pr\{s(b) = c\} = \frac{b!}{c!(b-c)!}\pi^c(1-\pi)^{(b-c)}$ ………………..(2)

This formula provides the probability of show-ups for exactly c customers. Therefore the probability of finding no-shows or cancellations is simply

$\Pr\{Ns\} = 1 - \Pr\{s(b) = c\}$ ………………………(3)

where Ns indicates the probability of no-shows and/or cancellations.

IV. Implementation using Genetic Algorithm

For solving the above formulated problem, a genetic algorithm has been implemented using Matlab and is stated below:

Simulator using GA for Estimating the Cancellations

1. Init_pop = Randomly Generated population.
2. curr_pop = Init_pop.
3. While (!termination_criterion)
4. Evaluate Fitness of curr_pop using fitness function.
5. Select mating pool according to Roulette-wheel Selection OR Tournament Selection.
6. Apply Crossovers like One-point, Two-point & Uniform Crossovers on mating pool with probability 0.80.
7. Apply Mutation on mating pool with probability 0.03.
8. Replace generation with ($\lambda + \mu$)-update as curr_pop.
9. End While
10. π = Survival Probablity
11. For I = 1 to numofclasses do steps 12 to 16
12. For c = 1 to b[I] do steps 13 and 14
13. Prob_show[I,c] = $\frac{b[I]!}{c!(b[I]-c)!}\pi^c(1-\pi)^{(b-c)}$

    // Probabilty of customers showing up
14. End For.
15. Prob_can = 1 – Prob_show[I,c]

    // Cancellation Probability
16. End For
17. End

V. Results

In this case, a single flight is considered to operate between given origin and destination. The capacity of the flight is assumed to be 100.

Genetic algorithm is used as a solution technique using various combinations of different operators. The following GA parameters are taken into considerations:

Population size = 75

Maximum number of iterations = 50
Cross-over probability = 0.80
Mutation probability = 0.03
Tournament Selection parameter = 0.75
Number of simulations = 100

Using the above parameters and various combinations one can get the table 1 for the optimum results without cancellation, which is shown at the end of the paper as already explained in [10]. The graphs for the optimum results are also shown in fig.1 and fig.2.

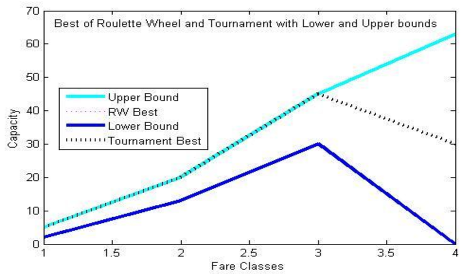

Fig.1 : Lower Bound, Upper Bound, and Estimated Fitness

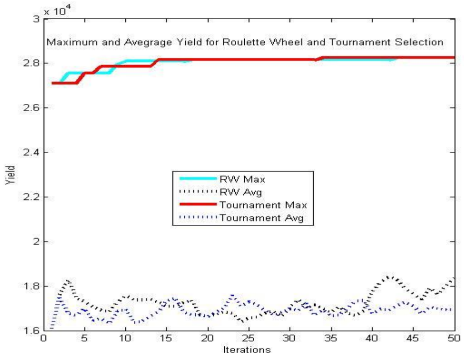

Fig.2: Average Fitness

For estimating cancellations, the survival probability is assumed to be 0.7 i.e. chances of showing up each customer are 70%. Upon simulating 100 times, the results obtained for each fare class are shown in table 2, 3 and 4. The graph for revenue comparison is shown in fig.3.

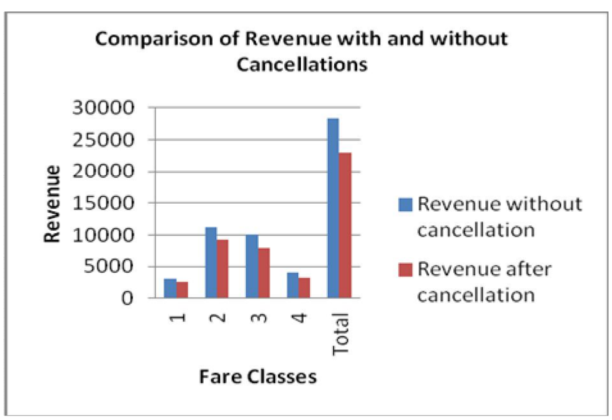

Fig.3 Comparison of Revenue with and without Cancellations

VI. Interpretation

The purpose of this paper is to estimate the cancellation in airline YM. The model has been formulated using binomial distribution. This distribution tells the probability of showing-up a customer in an efficient manner. It has been observed that is the survival probability for all the customer is assumed or calculated on the basis of historical data. Then this model helps to estimate the cancellations or no-shows.

The results show that a significant amount may be lost due to cancellations. Depending on the fare class the various probabilities are calculated and corresponding to that how many seats remains are shown in the tables 2 and 3. Table 4 shows the revenue obtained after cancellation.

The graph shown in 1, Cleary shows the impact of cancellation on various fare class and in totality as well.

These estimates have been obtained using specific example. These can also be used for other more complicated data sets as well and the companies can make some alternate strategy to avoid this loss of revenue. .

VII. Conclusion

In this paper, the estimate of cancellations in YM has been considered. The specific case study chosen was airlines. This paper presents a simulator to estimate the cancellations and no-shows in airlines. For the purpose of estimating the cancellation, the binomial distribution has been considered. First the GA is used to find the optimum seat allocation. Then the procedure for cancellation has been applied. The average probabilities in various classes have been calculated and correspondingly revenue is also evaluated.

This paper can be of very good help for the firms to apply various techniques to avoid the loss of revenue. For this purpose overbooking is the most popular technique which can be of great help to maximize profit. Other techniques could be last minute selling, no refunds, etc.

References

[1] Stephen Shaw, *Airline marketing and Management*, fifth edition, Ashgate publishing Limited, 2004.

[2] Subramanian, J., Stidham, S., and Lautenbacher, C. (1999), *Airline yield management with overbooking, cancellations, and no-shows,* Transportation Science, 33:147–167.

[3] Lippman, S. A. and Stidham, S. (1977), *Individual versus social optimization in exponential congestion systems,* Operations Research, 25:233–247.

[4] Srinivas S. Pulugurtha & Shashi S. Nambisan, "A Decision-Support Tool for Airline Yield Management Using Genetic Algorithm", *Computer Aided Civil and Infrastructure Engineering, 2003, 214-233.*

[5] Feng, Y., Lin, P., and Xiao, B. (2002), *An analysis of airline seat control with cancellations,* In Yao,D. D. and H. Zhang, X. Y. Z., editors, Stochastic Modeling and Optimization, with Applications in Queues, Finance, and Supply Chains. Springer-Verlag, Berlin.

[6] C. Hueglin and F. Vannotti. Data mining techniques to improve forecast accuracy in airline business. *Proc The*

Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 438–442, 2001.

[7] H-U. Kalka and K. Weber. PNR-based no-show forecast. *Presentation at 2000 AGIFORS meeting*, New York, NY, 2000.

[8] H. Feyen and C. Hueglin. Data mining techniques to improve forecast accuracy in airline business. P*resentation at 2001 AGIFORS meeting*, Bangkok, Thailand.

[9] D. Selby. Materialisation forecasting: a data mining perspective. *Operations Research in Space and Air*, 393–406, Kluwer Academic Publishers, Boston, 2003.

[10] Mittal, Pardeep Kumar, Kumar Rakesh, Suri, P.K., "Genetic Simulator for Airline Yield Management", *International Journal of Engineering Research & Technology(IJERT), ISSN 2278-0181, Vol3. Issue 9, September, 2013, pp. 2379-2386*

TABLE1: LOWER, UPPER AND BEST ESTIMATED DEMANDS IN EACH ASSUMED FARE CLASS

		Demand		
Fare Class	**Fare**	**Lower Limit**	**Upper Limit**	**Best Estimation**
1	100	0	63	30
2	250	30	45	45
3	500	13	20	20
4	800	2	5	5

TABLE 2
PROBABILITY OF CANCELLATION IN EACH FARE CLASS

Fare Class	Average % Probability of Cancellation	Standard Deviation	Range of Probability
4	10.9	6.69	4.21 - 17.59
3	18.8	8.39	10.41 - 27.19
2	17.9	6.81	11.09 - 24.71
1	16.7	8.06	8.64 - 24.75

TABLE 3
COMPARISON OF SEAT ALLOCATION WITHOUT CANCELLATION AND WITH CANCELLATION

Fare Class	Fare	Optimum Seat Allocation	Avg. Cancellations	Seats after cancellation
1	100	30	5	25
2	250	45	8	37
3	500	20	4	16
4	800	5	1	4

TABLE 4
COMPARISON OF REVENUE WITHOUT CANCELLATION AND WITH CANCELLATION

Fare Class	Fare	Optimum Revenue	Revenue after cancellation
1	100	3000	2500
2	250	11250	9250
3	500	10000	8000
4	800	4000	3200
Total		28250	22950