CHAPTER 3

PROBLEM STATEMENT AND RESEARCH METHODOLOGY

“*Problem statement is the basic building block of any thesis. If the problem is defined properly, half of the work is assumed to be done. This chapter defines the research objective of this thesis. On the basis of the objectives the problem is formulated in various subparts. The formulation is discussed briefly in each section. An introduction to the computer simulation is provided, which will be used to create the simulator for solving the formulated problems. MATLAB is also introduced as a tool for implementing the simulators.*”

3.1 Introduction

If one wants to solve a problem, the first thing is to define it properly. A properly defined problem is half solved automatically. Therefore this chapter is one of the most important chapters in any thesis. Although it does appear that this is one of the simplest chapters to form. But in reality, it proves to be exactly opposite. One step in the wrong direction may lead to disastrous result. In this chapter the problem will be defined in terms of research objectives. Then the research methodology is discussed in brief.

3.2 Research Objectives

The aim of this research is to identify the impact of genetic behaviour in yield management. The overall objectives of this research can be defined as follows:

- To optimize (maximize) the revenue using yield management.
- To analyze the behaviour of genetic algorithm in yield management.
- To look into the impact of various arrival patterns in yield management.
- To prevent the premature convergence in yield management, if any.
- To estimate the cancellations and overbooking in yield management.
- To identify the best possible combination of various operators in GA for yield management.

3.3 Need of the Research

Yield Management as discussed in chapter 1 is not a very old technique. It was initiated in the late 70's. Since then a lot of research has undergone in this field. But one of the area which is found to be very less implemented in YM, is the use of evolutionary algorithms. As already specified in previous chapter, YM is a perfect candidate for application of evolutionary algorithm and especially GA, therefore it is quite interesting to see the impact of GA in YM. That is why this topic of research was chosen. This research may prove to be quite useful to the industries where YM can be implemented specifically airlines industry.

3.4 Problem Definition and Formulation

3.4.1 A Basic YM Problem

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. The period of reservation may be divided in a single or more than one time slices. Another assumption is to fix the fare of each class during each time slice and also assumed as known.

For finding the objective function, the purpose of which Following notations has been assumed for this problem:

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.

The purpose obviously is to maximize the revenue. For this one has to assume some constraints, which are shown below:

- First assumption is that there will be no cancellations, no-shows and overbooking. The total number of passenger travelling should be less than or equal to the capacity of the flight.
- 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:

Max. $\Sigma_\beta \Sigma_\alpha N_{\alpha, \beta} F_\beta$ ……………………………………………………………………………….. (3.1)

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,

$N_{\alpha, \beta} >= 0$, which indicates that number of customers in each time slice can be positive only.

3.4.2 Predicting Cancellations in YM

For estimating the cancellation, the following model is formulated.

Let ξ $(0 < \xi \leq 1)$ 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.

For estimating the expected number of show-ups for each booking level b, let the random variable s denote the number of consumers who show up at the service delivery time. Clearly, $s \leq b$, 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 ξ, hence it can also be written as $s(b; \xi)$.

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:

$\Pr\{s(b) = c\} = \frac{b!}{c!(b-c)!}\xi^{c}(1-\xi)^{(b-c)}$ ……………………………………………….(3.2)

Therefore the probability of finding no-shows or cancellations is simply

$\Pr\{Ns\} = 1 - \Pr\{s(b) = c\}$ …………………………………………………..…(3.3)

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

This model is added to the above defined model in the previous section.

3.4.3 Overbooking in YM

Let X be the number of no-shows with probability Pr{Ns}. Let Y be the number of seats that will be overbooked, i.e., if the airplane has *S* seats then the tickets will be sold upto S+Y tickets. Let the underage penalty be defined by C_{upen} and the overage penalty by C_{open}. In this case C_{open} represents the net penalties that are associated with refusing a seat to a passenger holding a confirmed reservation. Here, C_{upen} represents the opportunity cost of flying an empty seat. To explain further, if *X*>Y then the number of seats that could have been sold more are X-Y and those passengers would have seats on the plane. So C_{upen} equals the price of a ticket. If X<Y then the customer that needs to be bumped are Y-X and each has a net cost of C_{open}. Thus, the formula for optimal number overbooked seats takes following form:

$$F(Y^*) \geq \frac{\mathrm{Cupen}}{\mathrm{Cupen}+\mathrm{Copen}}$$

In simple words the number of seats that should be overbooked are the smallest possible value for F(Y*).

Therefore, the final objective function taken the following form:

$\text{Max. } (\Sigma_{\beta}\, \Sigma_{\alpha}\, N_{\alpha,\beta}\, F_{\beta} - C_{open})$ ……………………………………………………....(3.4)

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,

$N_{\alpha,\beta} >= 0,\ C_{open} >= 0.$

Where C_{open} = Penalty for bumping a customer.

This model is again used in combination with the model defined in section 3.4.1.

For solving the problem formulated in this section, simulators will be designed using MATLAB. Before taking an overview of MATLAB, simulator needs to be described briefly.

3.5 Simulator

A computer simulation or a computer model is a computer program that attempts to simulate an abstract model of a particular system. Computer simulations have become a useful part of mathematical modelling of many natural systems in physics, chemistry and biology, human systems in economics, psychology, and social science and in the process of engineering new technology, to gain insight into the operation of those systems.

Traditionally, the formal modelling of systems has been via a mathematical model, which attempts to find analytical solutions to problems which enable the prediction of the behaviour of the system from a set of parameters and initial conditions. Computer simulations build on, and are a useful adjunct to purely mathematical models in science, technology and entertainment. The reliability and the trust people put in computer simulations depends on the validity of the simulation model.

The simulator for present thesis has been designed using MATLAB.

3.6 MATLAB

In this research, Genetic Algorithms has been implemented using MATLAB. MATLAB stands for MATrix LABoratory and was developed by The MathWorks Inc. Dr. Cleve Moler, Chief Scientist at The MathWorks Inc. originally wrote MATLAB to provide easy access to matrix software developed in the LINPACK and EISPACK projects. These are carefully tested high-quality programming packages for solving linear equations and eigenvalue problems. Later, Jack Little harnessed the commercial potential of MATLAB and with permission rewrote MATLAB in C, added "M-files" and many new features and libraries in it and marketed it.

MATLAB is a software package used to perform scientific computations and visualization. It has an enhanced capability to analyze various scientific problems. It is highly flexible and has powerful graphics. It provides an Integrated Development Environment (IDE) for programming with numerous predefined functions for technical computations and visualizations. It can also incorporate user defined functions. MATLAB provides an excellent computational language, built-in state of art algorithms for mathematics and excellent visualization using readymade functions. Features and programming constructs available in C or C++ are also available in MATLAB.

Built in functions to handle matrices in complex mathematical computations are included in MATLAB, thus saving the programming time of the user. For visualization, large number of functions are available for two-dimensional and three-dimensional graphics. MATLAB provides Graphical User Interface (GUI) tools to handle graphics and perform specialized operations on graphics. MATLAB includes various toolboxes for specific applications such as Math analysis, Data acquisition, Control System Design, Digital Signal Processing, Neural Networks, Optimization, etc.

The programming environment of MATLAB is very interactive. Programs are interpreted rather than compiled. This makes debugging easier in it. Since it is interpreted language, it allows one to perform numerical calculations and visualize the results without the need of complex programming. It allows users to solve problems accurately, produce graphics easily and produce code efficiently. MATLAB is best for sophisticated Math, especially on large data sets and for things like matrix algebra.

Advantages of MATLAB are:

- ease of use,
- powerful graphics and plotting,
- platform independence,
- vast range of predefined functions.

Different modules have been developed in MATLAB for various operators of genetic algorithm. The source code is made general so that it will be used in more modules with little modifications. Source code is modular, structural and easy to use. Necessary documentations have been done to make it readable and easy to understand.

The aim of MATLAB is to enable us to solve complex numerical problems, without having to write programs in traditional languages like *C*. Thus, MATLAB interprets commands like Basic does, instead of compiling source code like *C* require. By using the relatively simple programming capability of MATLAB, it is very easy to create new commands and functions in MATLAB. In addition, these developed MATLAB programs (or scripts) can run without modification on different computers with MATLAB. Today, MATLAB has evolved into a very powerful programming environment by providing numerous toolboxes such as signal processing, image processing, and controls, optimization, and statistics computations.

3.7 Summary

This chapter defines and formulate the problem for the thesis. Initially research objectives are specified. Then the definition and formulation of the problem for the thesis is presented. Mainly three problems are defined, first being the basic YM problem, followed by the problem of cancellation of bookings and in the end the problem of overbooking. An introduction to the simulator is then provided. An overview of MATLAB is provided in the end. MATLAB has proved to be a great tool for optimization and especially for GA.