A MODEL OF OPERATIONS CAPACITY PLANNING AND MANAGEMENT FOR ADMINISTRATIVE SERVICE CENTERS

МОДЕЛ ЗА ПЛАНИРАНЕ И УПРАВЛЕНИЕ НА КАПАЦИТЕТА НА ОПЕРАЦИИТЕ В ЦЕНТЪР ЗА АДМИНИСТРАТИВНО ОБСЛУЖВАНЕ

Yulia Yorgova, PhD, Burgas Free University, Faculty of Business Studies,
62, San Stefano St., 8000 Burgas, Bulgaria, yulia@bfu.bg, tel. +359898446682

Abstract: The presented model of operations capacity planning allows obtaining quantitative dimensions of the service system parameters for Administrative Services Center. A methodology for the practical application of this model has also been presented here. The present paper aims at offering support for operations managers in the service sector for decision making regarding the operations capacity.
Key words: service operations capacity, Queuing Theory, arrival rate, service rate, public services, administrative services

Introduction

The main objective of all Administrative Services Centers (ASC), which are a formal part of some Bulgarian municipalities, is to provide public and administrative services of the highest quality. Public and administrative services are client services and are comprised of two key elements- the final service outcome achieved and the client's experience while obtaining the service. In order for administrative and public services to be perceived to be of high quality, the client's expectations and requirements should be met. These clients requirements include easy access, timely service, flexibility, high levels of competence and politeness, to name a few. In order to achieve a high quality and timely service, the ASC's operations capacity to meet demand is of particular importance. The management of service capacity is an integral part of the operations management decision making framework. These types of decisions in the services field are usually made under great uncertainty and often in the absence of adequate quantitative assessment of the relevant variables. Therefore, in order to increase effectiveness in the operations manager's work it is necessary to employ more precise quantitative models.

The purpose of this paper is to present a model of operations capacity planning and management for public and administrative mass services. The methodology for the application of this model is also suggested for services offered by ASC of the Burgas municipality.

The effective management of service operations requires that the relevant managers make decisions regarding the following:

- Forecasting of demand (number of customers)
- Identifying the necessary operations capacity (number of service channels and staff)
- Deciding on staff scheduling

Similar decisions in operations management are usually made after an examination of the service system, following the gathering and processing of the required data, an analysis of the final results and the formulation of conclusions and recommendations for practice. In order to improve the decision making process, it is useful to employ models, which provide quantitative dimensions of the service system.

The main interest of the present paper is the service system for public and administrative services, where queuing is a common occurrence and where it is necessary to always match operations capacity and demand.

The main components of any service system are: clients, service channels, and a queue. In this paper, operations is a term which encompasses the entering and accepting of customers into the service system, organizing customers into a queue, providing the customers required service, and the customers exiting from the service system. The main reason for the emergence of a queue is an insufficient operations capacity. Queuing is a common occurrence for mass services, where there is a large size of arrival units (service centers, transport centers, banks, hospitals, call centers, restaurants airports, etc).

A more detailed analysis of the service system shows that queuing occurs when the arrival rate (average number of arrivals for a given period of time) exceeds the service rate (average number of served customers per given period of time). Discriminating between the above mentioned components of the service system and examining the interactions between them, indicated that one may view the queuing phenomenon as a system and therefore an analytical queuing model may be formulated.

It is clear that in the services, the managing of a queue is the approach through which, the most optimal combination of demand and operations capacity may be achieved. The queue takes the form of client “inventory”, which acts as a buffer between demand and operations capacity. Of course, the aim is to minimize this virtual storage, because this is of no use to customers or staff providing the required service.

Based on the above discussion, it is clear that in order to develop a model of operations capacity planning and management for public and administrative services, it is necessary to study the interaction between the arrival rate, service rate and queue. The model and methodology of the present paper is focused specifically on identifying the operations capacity required to meet the demand. This paper does not deal with estimating the optimal operations capacity because the optimal operations capacity aims to keep the total cost to a minimum but does not necessarily meet demand.

Research methods

Queuing Theory applied for operations capacity planning and management of mass services.

Because the delivery and consumption of a service occur simultaneously, the demand needs to be met precisely as it arises. This requires service organizations to estimate and then provide the right capacity at the right time. Insufficient or greater than necessary capacity leads to losses in the short and long-term.

Queuing Theory offers the instruments most suitable for queuing cases. (Winston, 2004, Fitzsimmons, 2004)

The main objective of Queuing Theory is to establish the relationship between variables **λ, μ** и **M**, where:

λ – arrival rate (average number of arrivals for a given period of time);

μ – service rate (average number of served clients for a given period of time);

M – number of working channels;

These **λ, μ** and **M** are the main variables of a service system. When planning and managing services, it is necessary to identify the precise capacity of the service system, which is directly related to arrival rate **λ** and the service rate of the system **Mμ**.

In order to develop a model for the mass services delivery system, it is needed to precisely identify the main characteristics of the service system. In particular:

- Arrival process.
- Physical features of the queue.
- Queue discipline.
- Service process.
- Exit.

In order for the services system to be comprehensively defined, in addition to the study of its main characteristics, the relationship and interactions between these characteristics should also be studied. Once all of the above has been studied, it is possible to use Queuing Theory to develop a model for the system.

The application of Queuing Theory follows the following assumptions: (Andronov, Alexandrova, 2004)

1. Simple arrival process, which is a homogeneous Poisson process.
2. Clients represent an infinite population

3. Inter-arrival time is exponentially distributed, and this means that the distribution of arrival units for a given period of time is a Poisson distribution.
4. Service times is exponentially distributed
5. All clients are served on a first come-first served (FCFS) bases.

The above discussed assumptions define the commonly accepted system for mass service, which according to Kendall-Lee notation (Winston, 2004) should be presented as: M/M/n/FCFS/∞/∞.

Once all relevant information about the service system has been collected and the system has been described, mathematical modeling of all characteristics of the system is performed. It should be kept in mind that all mathematical modeling represents a certain simplification of any given real life situation. This means that all conclusions drawn from mathematical modeling should be viewed with caution.

Model of operation capacity planning and management of public and administrative services

This model is developed in four consecutive steps using the above discussed theory.

Step one*:* In order to define the service system model according to the theoretical framework it is necessary to identify its main components first and then establish whether or not Queuing Theory assumptions are satisfied

- *Arrival process*

Arrival process is created by client demand for a given service. Those using public services offered by ASC may be thought of as part of a group with an infinite population. A customer exiting a queue does not significantly affect the system probabilities.

Customers enter the system one by one, independent of one another and from a large number of sources. This makes requests entering the system random by nature, and they form a simple arrival process. The arrival process has unchangeable characteristics for any given interval. During normal working hours, the arrival process for public services has peak and non-peak arrival intensity, and can therefore appear as changeable. If time is divided into equal, one hour intervals, the arrival process may be treated as unchangeable. The discussed simple arrival process meets the assumptions of a homogeneous Poisson process. The statistical data is checked to ensure it corresponds to the theoretical distribution of clients entering the service system.

The degree of customer patience is difficult to measure objectively. It is safe to say that public service customers fall into two categories – patient and impatient.

The simple arrival process cannot be controlled the way some other types of services are controlled through specific decisions regarding the management of demand.

- *Physical features of clients and queue discipline*

Every one of the six working stations is designed to handle one or more similar types of services jobs. Any given working station consists of two working desks (channels). The queue of clients is one for each working station. Every time a client exits a queue and a service has been provided, another client could immediately take his/her place. This makes the queue infinite in population length. Queue discipline is characterized as “first come, first served”, or clients are serviced in the order in which they arrive

- *Service process*

Let us say that **M** is the number of working channels, which provide all types of public services. The service process is characterized by the following features:

- Average service time for one client **t** for any given interval, which has a negatively exponential distribution of probabilities;
- Maximum number of clients that can be served simultaneously corresponds to the number of working channels -**M**;
- Accessibility of servicing facilities – the system is completely accessible, because the channels may accept any subsequent client immediately after serving the previous client.

- *Exit*

Once a customer has been served there is a high probability of the customer returning to the source population and immediately becoming a competing candidate for service again, because the centre is the one offering the desired type of services.

Step two*:* When the system is clearly defined with all its components, it is then necessary to choose the most suitable mathematical model. Based on the characteristics identified in step one, the studied system is identified as multi-channel, single phase.

Step three*:* To clarify the functioning of the system and to draw conclusion regarding the system’s capacity, it is necessary to calculate the relevant operational parameters. For the purposes of this study, the relevant operational parameters should be calculated for each working station: probability of no clients, $\mathbf{P_0}$, average number of clients in the system, **Ls**, average total time in system, **Ws,** average number of clients waiting in queue, **Lq,** average time waiting in queue, **Wq** .

Step four*:* At this stage, based on the calculated parameters, conclusions should be drawn regarding the volume and utilization of the capacity at any given time interval. Once that is done, decisions need to be made for possible modifications of the system.

The volume of capacity is determined by the number of channels **M** working simultaneously**,** and the average service time per customer. In practice, the operation managers of mass services cannot influence the arrival rate, unless they completely change the intensity by including or excluding entire groups of clients (e.g. demand by some villages of the municipality) A change of the system's capacity may be achieved by modifying the service process in two directions:

- By increasing or decreasing the number of working channels, depending on forecasting of service demand
- By improving the operations for the purpose of decreasing the average service time, particularly during the periods when the arrival rate, **λ,** is grater than the service rate, **μ**.

In order to choose which operation parameters should be estimated and analyzed, it is important to identify the key indicator necessary for the functioning of the system. One such indicator may be the waiting time or queue length, and also average total time in system or number of clients in the system. The chosen indicator will be the one, which is most relevant for the specific nature of the type of service and the targets set by management. If management aims to decrease or eliminate the queue, **Lq** should be studied and calculated. In cases when waiting time is limited, data for **Wq**. should be estimated and analyzed and so on. For the specific example used in this study, it is suitable to use both of the above mentioned parameters. A starting point in the analysis could be to decide on a standard for a desired waiting time length, so that clients will not have a negative experience. Once a standard has been accepted, and based on the analyzed data operation managers may decide with greater precision the exact necessary system's capacity, which will allow demand to be met and for management targets to be achieved.

Results and Conclusions

Methodology for application of the model with actual data from the ASC.

Employing this model and bearing in mind the type of service requested, specific methodology may be offered, to allow the use of this model with actual data. In this particular case, observations and data regarding citizenship status and registration were tested for a period of five consecutive months, for one working station with two channels. The methodology includes:

1. *Identifying the main components of the service system.*

This is the stage for data collection regarding the number of arriving clients for every one hour interval, for a full working day and the average service time, **t** [min]**,** for every client, per minute. For every one of the five months and for every time interval, **λ** [number per hour] and service rate **μ** were calculated. (Due to the large volume, when testing the model with actual data, the calculated parameters are not shown here).

The statistical data employed in this study were put under Kolmogorov – Smirnov test of goodness of fit. The results show that the arrival process has a Poisson distribution.

Physical characteristics and queue discipline identify the queue as unlimited in capacity. Service time is exponentially distributed.

2. The choice for a mathematical model for the system.

All above mentioned assumptions for the use of Queuing Theory are satisfied. The system is identified as two channeled, single phase. Using Kendall-Lee's notation, it is presented as: M/M/2/FCFS/∞/∞.

3. Calculating the operational parameters of the system.

The operational parameters calculated for every one hour interval of the full working day, for every month, with a different number **M** =1 ÷ 4 working channels are as follows: probability of no clients, $\mathbf{P_0}$, average number of clients in the system, **Ls**, average total time in system,**Ws ,** average number of clients waiting in queue, **Lq,** average time waiting in queue, **Wq** .

4. Formulating conclusions regarding system capacity utilization and decisions for possible changes to the operations capacity (if necessary) through a modification of the service process.

In order to reach conclusions, the length of the queue, **Lq** and the average waiting time**, Wq** [min], should be analyzed for the various number of working channels. Similar analyses may also be performed for the remaining parameters of the system, depending on the objectives of the study.

In the case of this study, if it is accepted that the maximum waiting time for a client to begin receiving service is 10 minutes, the results show that it is sufficient to have a maximum of three working channels simultaneously in order to meet demand. Bearing in mind the various levels of demand during the course of a working day, the length of a working shift, and the need for staff work breaks, the maximum number of staff needed during the peak periods is four. In this study, two channels were used and it was shown that this capacity was insufficient. This could be fixed through the expansion of online public services of the ASC or by increasing the number of channels working during the peak periods.

Conclusions

1. This model allows the processing of quantitative data for the various parameters of the system. It also allows the operations manager to make the appropriate modifications of the system, depending on the key indicators decided upon.
2. This model may be applied for the purpose of operations capacity planning and management for all mass service systems, which allow the formation of queues, such as medical centers, supermarkets, airports, banks, call centers and others.
3. This model is suitable for emergency services. The difference in such services is that the operations capacity should be estimated to correspond to the maximum level of demand.
4. The testing of the model with actual data for administrative and public services from the ASC showed it to be applicable and useful in organizational practice.

References

1. Андронов, Ев., М.Александрова, Управление на операциите, УНСС, София, 2004
2. Fitzsimmons J.A., M.J. Fitzsimmons, Service Management: Operations, Strategy and Information Technology, Mc Graw Hill, Fourth Ed., 2004
3. Winston, W.L., Introduction to Probability Models. Operations Research, Volume Two, Fourth Edition, Thomson Brooks / Cole, 2004.