

A SHORT SURVEY STUDY OF REDUCING WAITING TIME IN QUEUING SYSTEMS

S. Priya

Department of Mathematics, Sun Arts & Science College, Keeranoor, Tiruvannamalai – 606 755

Article History: Received 5th March,2017, Accepted 31st March,2017, Published 31st March,2017

ABSTRACT

Generally, queuing theory is used in modelling operations management and research problems of real life activities such as how customer check-out lines form (and how they can be minimized), how many calls a telephone switch can handle, how many computer users can share a mainframe. To minimize costs, if a management provides only one cashier, it forces every customer into a long, slow-moving line. Customers tiring of the wait would be likely to abandon their groceries and begin shopping at a new store. The question arises: what is an acceptable level of service at an acceptable cost to the provider. Feasible solutions and optimal solutions to the above can be traced by studying the performance indices and such studies have gained tremendous significance in most of real life situations. As waiting time of customers is one among the performance measures of a queuing system that influence its applicability, it is essential that waiting time of each potential customer joining a queuing system should be reduced. Through this survey, an attempt has been made to review the work done on the concept of reducing the waiting time in queuing models. The aim is to provide sufficient information to researchers, managers and industry people who are interested in the applications of queuing models in their managements. The most relevant and promising future applications of queuing theory are likely to occur in the areas of computer science and manufacturing systems. In computer science, queuing is a necessary consideration in contention for processing resources.

Keywords: M/M/S Model, M/G/1 Model, M/G/S Model, Queueing Models, Operation Research, Waiting Time

1.INTRODUCTION

Operations research is the science of developing and applying mathematical models to provide decision-makers with better strategies to plan and operate a system.

Queueing theory is an application of stochastic processes in O.R. The 'Queueing theory' is the probabilistic study of waiting lines. Even though it does not solve all types of waiting line problems, however it provides useful and vital information by forecasting or predicting the various characteristics and parameters of the particular waiting line under study.

In service systems, customer surveys demonstrate that the waiting time is one of the prime factors when the service level is too evaluated. There is an obvious trade off between waiting times and operational costs of the service system. Shortening waiting times usually requires the increasing of provision of staffs. Surveys and research advocate that customer satisfaction with waiting can be improved by

managing customer expectations of the waiting and customer perception of the waiting, even without reducing the waiting time itself. Since the prediction about the waiting times, the number of customers at any time, the time for which the server/servers remain busy or idle etc. rely heavily on the basic concept of stochastic processes, it can very well be taken as an application of stochastic processes. Also, queueing theory is generally considered a branch of operational research because the results are often used when making decisions about the resources needed to provide 'service'.

Erlang, Agner Krarup, (1918), an engineer at Copenhagen telephone exchange started to work on applying the theory of probabilities to problems of telephone traffic. Queueing theory has prevalent application in various fields, such as industrial problems (production, scheduling and maintenance), congestion in road traffic, scheduling of air traffic at the air port, waiting in a hospital, inventory control, epidemic process in biology, intelligent transportation systems, computer networks, telecommunications, advanced telecommunications systems etc. Formulas for each queueing model designate how the corresponding queueing system's performances; including the average amount of waiting that will occur, under a variety of situations. Little's theorem (1961) described the relationship between through put rate,

*Corresponding author: **S. Priya, M.Sc., M.Phil., SET.**
Assistant Professor of Mathematics, Sun Arts & Science College,
Keeranoor, Tiruvannamalai – 606 755

cycle time and work in process. He stated that the expected number of customers (N) for a system in steady state can be determined using the following equation: $L = \lambda T$ where, λ is the average customer arrival rate and T is the average service time for a customer. There fore, queuing theory is very helpful for determining how to operate a queuing system in the most efficient way. Provision of too much service capacity to operate the system involves excessive costs but not providing enough service capacity results in excessive waiting and all its unfortunate consequences. The models enable finding an appropriate balance between the cost of service and the amount of waiting.

Pruyn and Smidts (1993) studied the psychological factors involved in the reaction towards the waiting for service and customer delay. In the field experiment, they found amazingly small differences with respect to the evaluation of the single queue and multiple queue condition that customers prefer for a single queue system over multiple queues in general, because they perceive a single queue system to be fairer than a multiple queue system.

In a queuing model, minimization of time that customers have to wait and maximizing the utilization of the servers or resources are incompatible goals. In the research paper entitled mathematical analysis of single queue multi server and multi queue multi server queuing model Prasad and Badshah () proved that single queue multi server model is better than multi queue multi server model by discussing the relation between the performance measures of both queuing models in their case studies. They found that the expected waiting time of the customer in the queue (W_q) is less in the case of single queue as compare with case of S queues and the expected waiting time of the customer in the system (W_s) is less in the case of single queue as compared with case of S queues and derived the mathematical equations are as follows:

i) The expected number of customers waiting in the queue (L_q) is less in the case of single queue by comparing with case of S queues that is L_q (In the case of one queue) $<$ L_q (In the case of S Queues)

$$\frac{\rho^{S+1}}{\lambda(S-\rho)[(S-1)!(S-\rho)\sum_{n=0}^{S-1} \frac{\rho^n}{n!} + \rho^S]} < \frac{\rho^2}{S(S-\rho)}$$

ii) The expected waiting time of the customer in the queue (W_q) is less in the case of single queue as comparing with case of S queues that is

$$W_q \text{ (In the case of one queue)} < W_q \text{ (In the case of S Queues)}$$

$$\frac{\rho^{S+1}}{\lambda(S-\rho)[(S-1)!(S-\rho)\sum_{n=0}^{S-1} \frac{\rho^n}{n!} + \rho^S]} < \frac{\rho^2}{\lambda S(S-\rho)}$$

iii) The expected waiting time of the customer in the system (W_s) is less in the case of single queue as comparing with case of S queues that is

$$W_s \text{ (In the case of one queue)} < W_s \text{ (In the case of S Queues)}$$

$$\frac{\rho^{S+1}}{\lambda(S-\rho)[(S-1)!(S-\rho)\sum_{n=0}^{S-1} \frac{\rho^n}{n!} + \rho^S]} + \frac{1}{\mu} < \frac{\rho^2}{\lambda S(S-\rho)} + \frac{1}{\mu}$$

iv) The expected number of customers waiting in the system (L_q) is greater in the case of single queue as comparing with case of S queues that is

$$L_s \text{ (In the case of one queue)} > L_s \text{ (In the case of S Queues)}$$

$$\frac{\rho^{S+1}}{(S-\rho)[(S-1)!(S-\rho)\sum_{n=0}^{S-1} \frac{\rho^n}{n!} + \rho^S]} + \rho > \frac{\rho}{(S-\rho)}$$

From the above mathematical results they proved that the expected total cost is less for single queue multi server model as comparing with multi queue multi server model.

Maarten Schimmel (2013) considered the range of possible expected waiting times depends on the worst and best case situations on queuing time. The worst case is based on S times an M/G/1 queue, with S the amount of checkout counters. The best case assumes the perfect world, where customers always pick the line with the least amount of work, and go directly to an available counter, not leaving customers wait in another queue while a cashier is idle. The best case is based on a single M/G/S queue, with S again the amount of checkout counters. In practice, every counter has its own queue, suggesting that it functions as an M/G/1 queue. However, since customers join the shortest queue, or the queue with the least amount of work, it may also mimic an M/G/S queue. We have simulated a system where customers join the shortest queue. The results indicate that an M/G/S queue system is approached. With the models for best and worst case a limit on the amount of items a customer is allowed to bring to the express checkout can easily be calculated.

Srinivasan and Renganathan, (2002) analyzed that multiple vacation policy is better than single vacation policy regarding the waiting time of a customer in the system by considering some numerical results of the model. Uri Yechiali (2004) showed that the waiting time of a customer in $M^X/G/1$ queue with single vacation policy is less than the waiting time of a customer in $M^X/G/1$ queue with multiple vacation policy. Senthil Kumar and Arumuganathan (2009) presented some numerical results to study the effect of Bernoulli vacation probability 'p' and the effect of persistent probability α on mean waiting time W. Xiu-li Xu (2009) showed that the customer in the classical bulk input $M^{(X)}/M/1$ queue without vacation has the shortest mean waiting time and the customer in the bulk input $M^{(X)}/M/1$ queue with working vacations has the longest mean waiting time, which is identical with the practical instance. Maraghi et al. (2009), Khalaf et al. (2011) studied that the increasing of repair rate of the server or vacation rate causes the decreasing of mean waiting time of customer in the queue by considering some particular queuing models.

Reni Sagayaraj, et al., (2015) discussed about reduction of waiting time of a multi server queue model in fuel stations. The application of queue theory in cloud computing to reduce the waiting time. Bharkad and Durge (2014) proposed a queuing model in order to reduce the waiting time in cloud services.

2.CONCLUSION

The analysis of reducing the waiting time in queueing models helps a lot. This short survey reviews the work done in reduction of waiting time in various areas in queue systems. The ideas discussed in various papers have been synthesized. It can help statisticians, operations analyst, researchers, engineers, managers for using these models. A wide range of literature has been covered and proper references have been cited.

3.REFERENCES

- Erlang, A K. 1918. "Solution of some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges" The Post Office Electrical Engineers Journal, Vol. 10.
- Farzana Maraghi, A., Kailash C., Madan and Ken Darby-Dowman 2009. "Batch Arrival Queueing System with Random Breakdowns and Bernoulli Schedule Server Vacations having General Vacation Time Distribution", International Journal of Information and Management Sciences, Vol. 20, pp.55-70.
- Srinivasan, L and . Renganathan, N 2002. "Single Server, Bulk Arrival, Bernoulli Feedback Queue with Vacations: Some Performance Measures", Information and Management Sciences, Vol.13, No.4, pp.45-54.
- Little,J.D.C. 1961. "A Proof for the Queueing Formula: $L = \lambda W$," Operations Research, vol. 9(3), 383-387.
- Mathias Dharmawirya, Erwin Adi 2011. "Case Study for Restaurant Queueing Model Mathias", 2011 International Conference on Management and Artificial Intelligence IPEDR vol.6 (2011) © (2011) IACSIT Press, Bali, Indonesia.
- Maarten Schimmel 2013. "Deployment of express checkout lines at supermarkets Research paper Business Analytics".
- Senthil Kumar, M. and Arumuganathan, R. 2009. presented some numerical results to study the effect of Bernoulli vacation probability 'p' and the effect of persistent probability α on mean waiting time W.
- Vijay Prasad, S., Badshah, V.H. and Tariq Ahmad Koka, "Mathematical analysis of single queue multi server and multi queue multi server queueing models: comparison study"
- Pruyn, A.Th.H and Smidts, A. 1993. "Customers' Evaluations of Queues: Three Exploratory Studies", in E - European Advances in Consumer Research Volume 1, eds. W. Fred Van Raaij and Gary J. Bamossy, Provo, UT : Association for Consumer Research, Pages: 371-382.
- Uri Yechiali, 2004. "On the $M^X/G/1$ Queue with a waiting Server and Vacations", Sankhya: The Indian Journal of Statistics, Vol 66. Part1, pp.159-174.
- Senthil Kumar, M. and Arumuganathan, R. 2009. "Performance Analysis of an M/G/1 Retrial Queue with Non-Persistent Calls, Two Phases of Heterogeneous Service and Different Vacation Policies", Int. J. Open Problems Compt. Math., Vol. 2, No. 2, pp. 196-214.
- Xiu-li Xu , Zhong-jun Zhang and Nai-shuo Tian, 2009. "Analysis for the M/M/1 Working Vacation Queue", International Journal of Information and Management Sciences, Vol.20 , pp. 379-394
- Rehab, F. Khalaf, Kailash C. Madan and Lukas,C.A. 2011. "An $M^{[X]}/G/1$ Queue with Bernoulli Schedule, General Vacation Times, Random Breakdowns, General Delay Times and General Repair Times", Applied Mathematical Sciences, Vol. 5, No. 1, pp. 35 – 51
- Reni Sagayaraj M., Raguvaran R., and Daisy M. 2015. "A Study on Queueing Model to Reduce the Waiting Time in Fuel Stations", International Journal of Technical Research and Applications, Vo; 3, Issue 2, PP. 104-106
- Bharkad, N.N. and Durge M.H. 2014, "The application of queue theory in cloud computing to reduce the waiting time", Int. Journal of Engineering Research and Applications, Vol. 4, Issue 10 (Part -1), pp.64-69.
