



Banking Queue Simulation and Optimization – A Review

Neeru¹ & Ms. Garima Garg²

¹ M. Tech scholar, SGI, Samalkha, Haryana

² Assistant Professor, Computer Science deptt, SGI Samlkha

ABSTRACT

Purpose – Many models have been proposed to counter Bulk customer congestion. Most of them are inefficient and does not reduce load on the main server. Proposed model calculates performance of a customer queue under varying servers and optimizes the performance of a queuing system. It also provides with a tool to automate the whole system.

Methodology used - *Simulating* system response to randomly generated customers with varying servers.

Findings – Proposed model gives methods that can be used to optimize the queue length and manage the average queue length of the costumers. It automates the behavior of queue according to varying servers for processing the customers. Model not only simulates a random and variable queue of customers but it also calculates average queue length of customers.

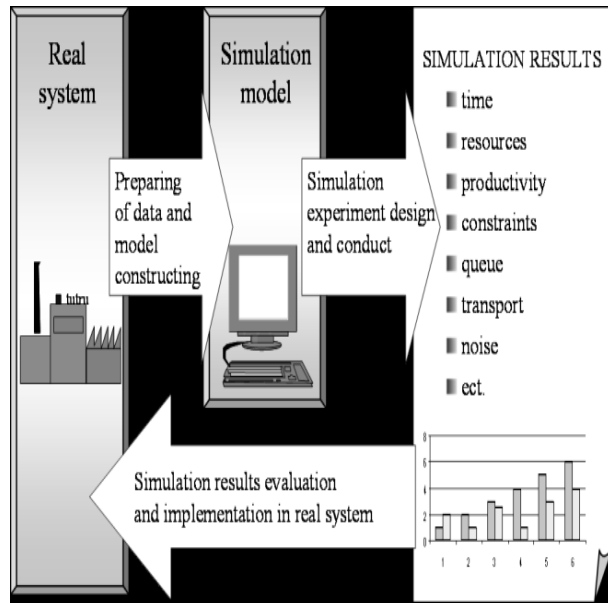
CHAPTER 1

1. INTRODUCTION -

Queuing theory plays an important role in modeling real life problems involving congestions in wide areas of applied sciences. Applications of queuing with impatience can be seen in traffic modeling, business and industries, computer-communication, health sectors and medical sciences etc. Queues with discouraged arrivals have applications in computers with batch job processing where job submissions are discouraged when the system is used frequently and arrivals are modeled as a Poisson process with state dependent arrival rate. The discouragement

affects the arrival rate of the queuing system. Morse considers discouragement in which the arrival rate falls according to a negative exponential law. We consider a single-server queuing system in which the customers arrive in a Poisson fashion with rate depending on the number of customers present in the system at that time i.e. $(n - 1)\lambda$. (Kumar et al., 2014)

Queuing with customer impatience has vast applications in computer communications, bio- medical modeling, service systems etc. It is important to note that the prevalence of the phenomenon of customer impatience has a very negative impact on the queuing system under investigation. If we talk from business point of view, the firms lose their potential customers due to customer impatience, which affects the business of firms as a whole. If firms employ certain customer retention strategies, then there are chances that a certain fraction of impatient customers can be retained in the queuing system. An impatient customer (due to reneging) may be convinced to stay in the service system for his service by utilizing certain convincing mechanisms. Such customers are termed as retained customers. When a customer gets impatient (due to reneging), he may leave the queue with some probability, say q and may remain in the queue for service with the probability $p(= 1 - q)$.



In modern computer communication networks, queuing theory is a useful tool to analyze node-to-node communication parameters. This is especially true in Packet Switched Computer Communication Systems. Nodes of many networks can be analyzed in terms of a standard M/G/1 queuing system. However, some situations require researchers to investigate complex M/G/1 queuing systems. Daigle. Illustrates how the M/G/1 paradigm can be used to obtain fundamental insight into the behavior of a slotted-time queuing system that represents a statistical multiplexing system.

Most queuing models assume that customers are served singly. But this assumption is far from the truth when we consider those numerous real-world situations in which customers are served in batches. In such queues, customers are served by a single server

(Multiple servers) in batches of maximum size “*b*” with a minimum threshold value “*a*.”

Such type of service rule is referred to as general bulk service (GBS) rule. The bulk service queues have potential applications in many areas, for example, in loading and unloading of cargoes at a seaport, in traffic signal systems, and in computer networks, jobs are processed in batches with a limit on the number of jobs taken at a time for processing. However, there are many instances where, after the completion of the service of a batch, if the server finds less than “*a*” customers in the queue, he leaves for a vacation. This time may be utilized by the server to carry out some additional work. On return from a vacation, if he finds “*a*” or more customers waiting, he takes them for service. Otherwise, he may remain idle (dormant) and continue to do so until the queue length reaches “*a*.” In queuing literature, such types of queues are known as bulk service queues with single vacation. Bulk service queues are, generally speaking, hard to analyze. Often the finite capacities in the bulk service queues increase the complexities of the solution and it becomes more complex if vacation(s) is taken into consideration.

The server works until all customers in the queue are served then takes a vacation; the server takes a second vacation if when he is back, there are no customers waiting, and so on, until he finds one or more waiting customers at which point he resumes service until all customers, including new arrivals, are served. A vacation is not initiated when there are customers in the system. Under the stochastic assumptions of stationary and *i.e.* inter-arrival and service times, it is known that the mean customer delay in the queue is the sum of two components: mean queue delay and mean vacation time. (Taha, 2011)



CHAPTER 2

LITERATURE REVIEW

Table 1 : Work done by different researchers on queue simulation over the years.

RESEARCHER	OBJECTIVE	METHODOLOGY	FINDINGS
Adham et al. (2012)	aims at developing a novel AQM algorithm to better QoS in terms of congestion prediction, queuing delay, packet loss and link utility, etc.	Employing input rate and current queue length to calculate the packet dropping/marketing probability.	The main limitation of this study is that all the simulations were merely under a single bottleneck network topology. Furthermore, the system stability was examined under just a few cases. Other cases like TCP connections mixed with HTTP connections, or UDP flows, etc. can also be tested.
Seal, (1995)	Demonstrates the application of spreadsheets in simulation queuing systems with arrivals from a finite population	Case study	Describes the approach for developing a generalized simulation model with any number of machines.
Proctor. 1994	Overwork and overcrowding in some periods was an important issue for the out- patient department of a local hospital in Chiai- Yi in Taiwan.	Case study	The waiting time was greatly reduced and the workload of the doctor was also reduced to a reasonable rate in the overwork and overcrowding periods.
Ghosal et al., 1995	Presents a method by which approximation is done through a quasi- isomorphic system which resembles the second queue in respect of one output, via delay time.	Extensive simulation experiments	This approach appears to have potentiality in studying intractable problems in communications and industrial management.
Lehaney, (1993)	Provides an introduction to simulation, and discusses the use of a modern simulation environment. Examples of the uses of simulation are queuing, scheduling and stock control. Simulation environments are now far more user- friendly, and software is more competitively priced, than ever before.	Case study	Simulation is a tool which can aid managers in policy making and decision making.
Jeong et al. (2008)	to provide a framework and prototype software	Queuing network analyzer (QNA) is	System performance measures generated by the open QNA are



	to use IDEF3 descriptions as a knowledge base from which a queuing network (QN) analysis is performed to compute system performance measures as part of quick response manufacturing	compared to the simulation through case studies.	reasonably close to the values obtained from simulation, particularly when the system utilization is low.
Proctor, (1994)	Considers computer simulation as a method for studying the behavior of business systems under a variety of assumed conditions for analyzing the simultaneous interaction of many variables to produce valuable insights into problems	Case study	Looks at how simulation can help to provide answers to complex decisions and problems which cannot be solved by conventional methods
Huang et al. (1996)	Describes a study which focused on the utilization of doctors and staff in the out-patient department, the time spent in the hospital by an out patient, and the length of the out patient queue	Case study	Results show that the waiting time was greatly reduced and the workload of the doctor was also reduced to a reasonable rate in the overwork and overcrowding periods.
Daniel et al. (1996)	Notes that patients attending public outpatient departments in Hong Kong spend a long time waiting for a short consultation, that clinics are congested and that both staff and patients are dissatisfied	Computerized simulation modeling	Demonstrates some ways in which managers in health care facilities can benefit from the use of computerized simulation modeling. Specifically, shows the effect of changing the duration of consultation and the effect of the application of an appointment system on patients' waiting time.
Warwick, (2009)	To reflect 40 years of queuing theory in application to library modeling and management. It suggests that these models have not had the impact that the early modeling's promised and suggests some reasons as to why.	snapshots of library queuing models	paper suggests that there is an urgent need to find ways to bridge the practitioner/researcher gap, broaden the application base of OR methodology within libraries, engage in constructive debate around library OR to build a consensus view as to the value of OR interventions, and identify directions for future collaborative work in libraries

2. Predicting Traffic Congestion

Mobility is an indispensable activity of our daily lives and road transport is one popular approach to mobility. However road

congestion occurrence can be irritating and costly. This work contributes to the modeling and therefore predicting road congestion of a Ghanaian urban road by way of queuing theory using stochastic process and initial



value problem framework. The approach is used to describe performance measure parameters, allowing the prediction of the level of queue built up at a signalized intersection as an insight into road vehicular congestion there and how such congestion occurrence can be efficiently managed.

It is aimed demonstrate a modeling of traffic evolution on an arterial road to Milam highway that serves surrounding suburbs and communities. Other possible benefit of this work is that it serves as a basis to other interesting investigations to characterize traffic congestion and the results obtained may serve as vital inputs to decisions that seek to improve traffic control and management. The objective therefore is to investigate the problem of congestion on the road segment and subsequently build upon this investigation to develop efficient tools capable of predicting and providing intelligent information on vehicular traffic flow.

CHAPTER 3

3. METHODOLOGY USED IN THE PAPER

1. FORMULATING A BLOCK TO GENERATE N RANDOM CUSTOMERS ARRIVING AT ANY INSTANT OF TIMES.
2. GENERATING A FIFO QUEUE
3. GENERATING A SINGLE SERVER TO PROCESS N CUSTOMERS
4. CALCULATING THE OUTPUT GIVING AVERAGE QUEUE LENGTH AT A GIVE INSTANT OF TIME USING SIMULINK
5. PLOTTING THE OUTPUT WITH SINGLE SERVER
6. INCREASING THE NUMBER OF SERVER TO REDUCE THE

AVERAGE QUEUE LENGTH AND LOAD ON THE SERVER

7. PLOTTING THE OPTIMIZED OUTPUT USING SIMULINK
8. AUTOMATING THE NUMBER OF SERVERS USING MATLAB CODE.

CHAPTER 4

4. PROBLEM FORMULATION

Customers in a queue are an important attribute of any real time system.

1. Simulation of a real time queuing system
2. calculation of average queue length of customer
3. plotting average queue length of the customer
4. varying the number of servers to manage server load
5. Optimizing queuing system
6. Automation of multiple servers

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

CONCLUSION

We have studied the single server queuing networks, when successive service times are different (or not) without breaking up the busy periods from stage-to-stage, leading to a great approximate simplification" the existence of an equivalent tandem queue effect.

It is very evident and concluded that a discrete system works well with a limited load for small amount of time. When the system is stress tested i.e. load is increased we observe that performance of system degrades considerably.

To overcome this number of servers serving the costumers in a discrete system are



increased which reduces load from the servers and hence optimizes the performance of discrete systems.

Due to the possible correlation between two successive local arrivals from the same incoming path, a curious model appears.. These theories assumed the local combination of distinguishable customers (with distinguishable queuing delay), instead of parts of upstream busy periods with indistinguishable

customers (and indistinguishable queuing delay depending on the maximum sojourn time initiating the new downstream busy period). These theories cannot be consistent with the concept of the equivalent tandem queue. Some significant consequences appear, since any link overload comes from a given incoming path which generates the tandem queue effect (i.e., correlation between local inter arrival time and upstream service time). The usual queuing standards (related to long queues) cannot protect against subsequent, significant overloads in the buffers due to some possible "agglutination phenomenon" (related to short queues). Usual network management methods should be revised, and should monitor the partial traffic streams loads (and not only the server load).

CHAPTER 6

REFERENCES

- [1.]Kumar (2009), A single server markovian queuing system with discouraged arrivals and retention of reneged customers., Yugoslav Journal of Operations Research, 24 (2014) Number 1, 119-126
- [2.]Muh (1993), A bulk queuing system under n policy with bi-level service delay discipline and start up time, Journal of Applied Mathematics and Stochastic Analysis, 6 Number 4, 359-384
- [3.]Gupta et al. (2004,)a finite capacity bulk capacity service queue with single vacation and markovian arrival process (2004).
- [4.]Taha (2011), Sample-Path Analysis of Single-Server Queue with Multiple Vacations, International Scholarly Research Network ISRN Applied Mathematics
- [5.]Lartey (2014), Predicting Traffic Congestion: A Queuing Perspective Open Journal of Modeling and Simulation, 2014, 2, 57-66
- [6.]Wang et. Al. (2009), On the Scalable Fairness and Efficient Active Queue Management of RED, I. J. Communications, Network and System Sciences, 1, 1-89
- [7.]Zadeh et al. (2012), A Batch Arrival Queue System with Coxian-2 Server Vacations and Admissibility Restricted, American Journal of Industrial and Business Management, 2, 47-54
- [8.]Baruah et. Al. (2013), A Two Stage Batch Arrival Queue with Reneging during Vacation and Breakdown Periods , American Journal of Operations Research, 3, 570-580
- [9.]Brill et. Al.(2013), Server Workload in an M/M/1 Queue with Bulk Arrivals and Special Delays , Applied Mathematics, 2012, 3, 2174-2177



- [10.] Peschansky et. Al. (2011), Stationary Characteristics of the Single-Server Queue System with Losses and Immediate Service Quality Control Applied Mathematics, 2011, 2, 403-409
- [11.] Gandole (2011), Computer Modeling and Simulation of Ultrasonic System for Material Characterization, Modeling and Numerical Simulation of Material Science, 1, 1-13
- [12.] Yu (2010), Steady-State Queue Length Analysis of a Batch Arrival Queue under N-Policy with Single Vacation and Setup Times, Intelligent Information Management, 2010, 2, 365-374
- [13.] Kukla (2008), Rationalization of foundry processes on the basis of simulation experiment, Archives of foundry engineering, Volume 8 Issue 3
- [14.] Kumar (2014), A single server markovian queuing sytem with discouraged arrivals and retention of renege costumers, Yugoslav Journal of Operations Research 24 , Number 1, 119-126
- [15.] Muh (1993), A bulk queuing system under n policy with bi-level service delay discipline and start up time, Journal of Applied Mathematics and Stochastic Analysis Number 4, 359-384
- [16.] Gupta et.Al. (2004), A finite capacity bulk service queue with single vacation and markovian arrival process
- [17.] Taha (2011), Sample-Path Analysis of Single-Server Queue with Multiple Vacations, International Scholarly Research Network ISRN Applied Mathematics, 12 pages
- [18.] vow Church, Andrew J. Newman, (2000) "Using simulations in the optimisation of fast food service delivery", British Food Journal, Vol. 102 Is: 5/6, pp.398 - 405