



A SINGLE SERVER MARKOVIAN QUEUING SYSTEM WITH LIMITED BUFFER AND REVERSE BALKING

Girin Saikia

Gauhati University, India

E-mail: girin.stats@gmail.com

Amit Choudhury

Gauhati University, India

E-mail: achoudhury@rediffmail.com

Submission: 10/22/2020

Accept: 11/6/2020

ABSTRACT

The phenomena are balking can be said to have been observed when a customer who has arrived into queuing system decides not to join it. Reverse balking is a particular type of balking wherein the probability that a customer will balk goes down as the system size goes up and vice versa. Such behavior can be observed in investment firms (insurance company, Mutual Fund Company, banks etc.). As the number of customers in the firm goes up, it creates trust among potential investors. Fewer customers would like to balk as the number of customers goes up. In this paper, we develop an M/M/1/k queuing system with reverse balking. The steady-state probabilities of the model are obtained and closed forms of expression of a number of performance measures are derived.

Keywords: Finite capacity; Performance Measure; Queuing Model; Reverse Balking; Steady-State Solution



1. INTRODUCTION

In our day-to-day life, we commonly experience waiting lines at bus stops, bank counters, hospitals and the like. A queue or waiting line forms when we need some kind of service and arrive at a service channel that offers such facilities (Sharma, 2013). Queuing is essential to manage congestion in traffic of any type in the modern technological world. Queuing theory describes probabilistically and mathematically, the interaction between the arrival process of customers and the service provided to them in order to manage the system in an efficient manner. It started when Danish mathematician (Erlang, 1909) showed for the first time how probability theory can be used to mathematically model the task of managing telephone conversations.

Among the various aspects in the study of any queuing system, there is the phenomenon of customer attrition. Of late, this has gained importance especially in the cutthroat world of business and commerce. Serving a customer is the whole purpose of any business entity. Firms invest substantial portion of their time, effort, energy and money to attract customers. Consequently, should there be any attrition of customers from a queuing system, it is not only an opportunity lost to expand business (with consequent revenue loss) but the fair name of the firm also takes a beating in the marketplace. All this is detrimental to its interests.

One of the reasons behind attrition is long queues. In any real life queuing system, such long queues discourage customers in two ways. A customer may decide not to join the queue (which is defined as balking) or leave the queue after joining it (which is defined as reneging). In both the cases, the customer does not receive service. We can find such queuing systems almost all around us. In fact, attrition is very common in queues. Examples are firms operating in the domains of health, service systems, call centers just to name a few.

Balking is the subject matter of this paper. We can observe balking when a customer arrives into a queuing system but refuses to join it. Traditionally, balking occurred with longer queues - the higher the size of the queue, the higher was the probability of balking. However, a particular type of balking which has gained acceptability in recent years is reverse balking. It aims to model a phenomena where the probability of balking (of an arriving customer) is low when the queue size is large and vice-versa.

Clearly, this is the reverse of the traditional concept of balking – an antithesis perhaps but there are a number of real like service delivery systems where this phenomenon can be observed. For example, consider a financial entity say an insurance company, Mutual Fund



Company, commercial bank etc. If the number of investors is already large, it provides some sort of comfort to potential customers. In such cases, the (large) number of customers does not discourage potential (arriving) investors (customers) arriving into the queuing system.

On the contrary, the large size instills a sense of confidence among them with regard to the financial soundness of the firm (as a large number of customers are already patronizing it, the firm must be financially sound). As a result, the probability of an arriving customer joining the queue goes up (consequently, the probability of balking comes down).

In this paper, we therefore analyze a queuing system with reverse balking. Among the various models available in literature, we choose the single server Markovian queuing system with finite buffer. Symbolically, we denote it by M/M/1/k. We describe the model in a later section.

We arrange the subsequent sections of the paper as follows. After this introduction, section 2 reviews the literature. Section 3 and 4 contain model description and derivation of steady state probabilities. In section 4, we discuss a numerical example. We conclude in section 5. To unclutter the paper, we place the detailed mathematical derivations in Appendix.

2. LITERATURE SURVEY

There is a considerable body of literature on customer attrition. Among them Haight (1957) can be considered to be the earliest. Haghghi, Medhi & Mohanty (1986) analyzed a multi-server queuing model with balking and renege. They obtained the steady state distribution of the number of customers in the system. They also obtained an expression for average customer lost during a fixed duration of time.

Choudhury and Medhi (2011A) considered an M/M/k model with the restriction that customers may balk from a non-empty queue as well as may renege after they join the queue. Derivations of closed form expression of a number of performance measures are given. Choudhury and Medhi (2011B) analyzed a multi-server Markovian queuing system under the assumption that customers may balk as well as renege.

Explicit closed form expressions are given. Demonstration of results derived is through a numerical example with design connotations. Choudhury and Medhi (2011C) analyzed a single server finite buffer Markovian queuing model M/M/1/N with the additional restriction that customers may balk as well as renege.



Choudhury and Medhi (2012) considered a finite buffer multi server queuing system with balking along with position dependent renegeing. Derivations of explicit closed form expressions of a number of performance measures are given. Demonstration of usefulness of results derived is through a practical problem.

Som and Kumar (2017) considered a single server finite capacity queuing system with customer retention and balking in which the inter-arrival and service times follow negative-exponential distribution. The assumption is that the renegeing time is exponentially distributed. The steady state solution of the model is given. Derivation of some performance measures is given. A discussion on sensitivity analysis of the model is also given.

Jain, Kumar and Som (2014) perhaps developed and introduced the concept of reverse balking in a single server Markovian queuing system. Derivation of the steady-state solution of the model and different measures of effectiveness is given. Kumar, Som and Jain (2015) considered a single server finite capacity feedback queuing system with reverse balking. Feedback customer in queuing literature refers to a customer who is unsatisfied with incomplete, partial or unsatisfactory service.

They derived the steady-state solution of the model and obtained some important performance measures. Sasikala and Thaigaranjan (2016) considered the M/M/1/N interdependent queuing model with controllable arrival rates and reverse balking. They derived the steady state solutions along with the system characteristics of the model. Som and Kumar (2017) also considered a finite capacity Markovian queuing system with two heterogeneous servers, reverse balking and renegeing.

3. ASSUMPTIONS OF THE MODEL

In this paper, we shall deal with the M/M/1/k model with reverse balking. Specifically, the assumptions of the model are as follows:

- a) Arrivals follow Poisson probability distribution and the inter-arrival times follow exponentially distribution with parameter λ .
- b) There is only one server and the service times follow exponentially distribution with parameter μ .
- c) The system capacity is restricted to k i.e. the capacity of the system is finite. In other words, the maximum number of customers that can be present in the system at any point of time is k.



d) The queue discipline is FCFS (first come first served) i.e. the customer who arrives first into the queuing system will receive service before others.

e) The probability that an arriving customer will balk is $\{1 - \frac{1}{(a-nb)}\}$ and consequently, the probability that an arriving customer will join the queuing system (i.e. not balk) is $\frac{1}{(a-nb)}$. Here n is the state of the system (i.e. the number of customers present in the system) at the instant of time the test customer arrives into the system. Further, 'a' and 'b' are constants to be suitably chosen. Since $\frac{1}{(a-nb)}$ is a probability, the constants have to be chosen in such a manner that $(a - nb) > 0$ (for all $n = 1, 2, \dots k$). Essentially this means that 'k' should be less than $\frac{a}{b}$. Thus the buffer size should not exceed the ratio of the constants 'a' and 'b'. At this point, an obvious question is how to choose 'a' and 'b'. We shall demonstrate the same through an example in section 4. The probability of balking will increase as 'a' increases and probability of balking will decrease as 'b' increases. In other words, 'a' is directly proportion to probability of balking where as 'b' is inversely proportional to the balking probability.

4. THE STEADY STATE PROBABILITIES

In this section, we derive the steady state probabilities by using the Markov process method. Let p_n denote the probability that there are 'n' customers in the system. The steady state equations are

$$\mu p_1 = \frac{\lambda}{a} p_0 \tag{1}$$

$$\mu p_{n+1} + \left[\frac{\lambda}{a-(n-1)b} \right] p_{n-1} = \left[\frac{\lambda}{a-nb} + \mu \right] p_n \quad ; \quad n = 1, 2 \dots k-1 \tag{2}$$

$$\left[\frac{\lambda}{a-(k-1)b} \right] p_{k-1} = \mu p_k \tag{3}$$

Solving recursively, we get

$$p_n = \left[\frac{\lambda^n}{\mu^n \prod_{r=0}^{n-1} (a-rb)} \right] p_0 \quad ; \quad n = 0, 1, 2 \dots k \tag{4}$$

Where p_0 is obtained from the normalizing condition $\sum_{n=0}^k p_n = 1$ and is given as

$$p_0 = \left[\frac{1}{1 + \sum_{n=1}^k \left[\frac{\lambda^n}{\mu^n \prod_{r=0}^{n-1} (a-rb)} \right]} \right]$$



An important performance measure is the average number of customers in the system. It is traditionally denoted by L . We shall derive an expression for the same from the p.g.f. (probability generating function) of the steady state distribution.

Let $P(s)$ be the p.g.f. We recall that

$$L = P'(1) = \frac{d}{ds} P(s) |_{s=1}.$$

Under the assumptions of the model, (see appendix 1 for the derivation)

$$P'(s) = \frac{\lambda}{\mu} \sum_{n=0}^{k-1} \frac{(n+1)P_n s^n}{(a-nb)} \tag{5}$$

Substituting $s=1$ in equation (5), we obtain the average number of customers in the system.

$$L = P'(1) = \frac{\lambda}{\mu} \sum_{n=0}^{k-1} \frac{(n+1)P_n}{(a-nb)} \tag{6}$$

5. NUMERICAL EXAMPLE.

To illustrate the use of our results, we apply them to a queuing problem. We quote below an example from Choudhury and Medhi (2011A).

Traffic to a message switching Centre for Extraterrestrial Communications Corporation arrives in a random pattern (remember that ‘random pattern’ means exponential inter-arrival time) at an average rate of 240 messages per minute. The line has a transmission rate of 800 characters per second. The message length distribution (including control characters) is approximately exponential with an average length of 176 characters. Calculate the principal statistical measure of system performance assuming that a very large number message buffers is provide. Suppose, however that it is desired to provide only the minimum number of messages buffers required to guarantee that

$$p_k < 0.005$$

How many buffers should be provided?

This is a design problem. To be specific, we are required to determine the system capacity k (i.e. size of the buffer) in such a manner that the boundary condition $p_k < 0.005$ is satisfied. It is clear that we would like k to be as small as possible (so as to keep costs down – an additional buffer would mean additional expenditure). We shall therefore first check if $k=1$ enables us to attain the boundary condition. Then we shall examine for $k=2$ and so forth. An

additional issue is to fix the values of the constants ‘a’ and ‘b’ for which we shall examine various alternatives.

A reading of the problem statement tells us that the arrival rate is 4 per second and the service rate is 4.55 per second. Hence here, $\lambda = 4/s$ and $\mu = 4.55/s$. We analyze the problem for different choices of ‘a’ and ‘b’. In particular, we consider the following alternatives –

1. $a = 5 \ \& \ b = 1$
2. $a = 10 \ \& \ b = 1$
3. $a = 10 \ \& \ b = 2$
4. $a = 10 \ \& \ b = 3$
5. $a = 40 \ \& \ b = 1$
6. $a = 40 \ \& \ b = 19$

In what follows, we shall analyze different scenarios for different values of k and different alternatives of ‘a’ and ‘b’.

Table 1: Steady state probabilities for varying system capacity with $a = 5 \ \& \ b = 1$

Probabilities	System Capacity			
	k=1	k=2	k=3	k=4
p_0	0.85047	0.82341	0.81580	0.81250
p_1	0.14953	0.14477	0.14344	0.14286
p_2	undefined	0.03182	0.03152	0.03140
p_3	undefined	undefined	0.00924	0.00920
p_4	undefined	undefined	undefined	0.00404
Total	1.00000	1.00000	1.00000	1.00000

Table 1 displays a few steady state probabilities assuming $a=5 \ \& \ b=1$. For $k=1$, $p_1 = 0.14953 > 0.005$, for $k=2$, $p_2 = 0.03182 > 0.005$ and for $k=3$, $p_3 = 0.00924 > 0.005$. These do not satisfy the given design condition. On the other hand $k=4$ satisfies the given design restriction of $p_k < 0.005$. Therefore the number of buffers should be equal to 4. We need not check beyond $k=4$ because of our restriction that $k < \frac{a}{b}$ (in our case $\frac{a}{b} = 5$). Thus if we choose $a=5 \ \& \ b=1$, then the optimum choice of k is 4.

Table 2: Steady state probabilities for varying system capacity with $a = 10 \ \& \ b = 1$

Probabilities	System Capacity		
	k=1	k=2	k=3
p_0	0.91919	0.91199	0.91121
p_1	0.08081	0.08018	0.08011
p_2	undefined	0.00783	0.00782
p_3	undefined	Undefined	0.00086
Total	1.00000	1.00000	1.00000

Table 2 displays a few steady state probabilities assuming $a=10 \ \& \ b=1$. For $k=1$, $p_1 = 0.08081 > 0.005$, for $k=2$, $p_2=0.00783 > 0.005$. These do not satisfy the given design condition. On the other hand $k=3$ satisfies the given design restriction of $p_k < 0.005$. Therefore the number of buffers should be equal to 3. We need not check beyond $k=3$ because

in general, it is better to have as few buffers as possible. Thus if we choose $a=10$ & $b=1$, then optimum choice of k is 3.

Table 3: Steady state probabilities for varying system capacity with $a = 10$ & $b = 2$

Probabilities	System Capacity		
	k=1	k=2	k=3
p_0	0.91919	0.91110	0.90993
p_1	0.08081	0.08010	0.07999
p_2	undefined	0.00880	0.00879
p_3	undefined	Undefined	0.00129
Total	1.00000	1.00000	1.00000

Table 3 displays a few steady state probabilities assuming $a=10$ & $b=2$. For $k=1$, $p_1 = 0.08081 > 0.005$, for $k=2$, $p_2 = 0.00880 > 0.005$. These do not satisfy the given design condition. However $k=3$ satisfies the given design restriction of $p_k < 0.005$ ($p_3 = 0.00129$). Therefore the number of buffers should be equal to 3. Here we need not check beyond $k=3$ because in general, it is better to have as few buffers as possible. Thus if we choose $a=10$ & $b=2$, then the optimum choice of k is 3.

Table 4: Steady state probabilities for varying system capacity with $a = 10$ & $b = 3$

Probabilities	System Capacity		
	k=1	k=2	k=3
p_0	0.91919	0.90996	0.90795
p_1	0.08081	0.08000	0.07982
p_2	undefined	0.01005	0.01002
p_3	undefined	Undefined	0.00220
Total	1.00000	1.00000	1.00000

Table 4 displays a few steady state probabilities assuming $a=10$ & $b=3$. For $k=1$, $p_1 = 0.08081 > 0.005$, for $k=2$, $p_2 = 0.01005 > 0.005$. These do not satisfy the given design condition. However $k=3$ satisfies the given design restriction of $p_k < 0.005$ ($p_3 = 0.00220$). Thus if we choose $a=10$ & $b=3$, then the optimum choice of k is 3.

Table 5: Steady state probabilities for varying system capacity with $a = 40$ & $b = 1$

Probabilities	System Capacity	
	k=1	k=2
p_0	0.97849	0.97801
p_1	0.02151	0.02149
p_2	undefined	0.00049
Total	1.00000	1.00000

Table 5 displays a few steady state probabilities assuming $a=40$ & $b=1$. For $k=1$, $p_1 = 0.02151 > 0.005$. This does not satisfy the given design condition. However $k=2$ satisfies the given design restriction of $p_k < 0.005$. Therefore here the number of buffers should be equal to 2.

Table 6: Steady state probabilities for varying system capacity with $a = 40$ & $b = 19$

Probabilities	System Capacity	
	k=1	k=2
p_0	0.97849	0.97761
p_1	0.02151	0.02149
p_2	undefined	0.00090
Total	1.00000	1.00000

Table 6 displays a few steady state probabilities assuming $a=40$ & $b=19$. For $k=1$, $p_1 = 0.02151 > 0.005$. This does not satisfy the given design condition. However $k=2$ satisfies the given design restriction of $p_k < 0.005$ ($p_3=0.00090$). Therefore the optimal number of buffers is 2.

6. CONCLUSION:

In this paper, we have derived the steady state probabilities and a performance measure assuming a specific reverse balking rule. We have applied our results to a numerical problem. In the problem, we have examined the impact of different balking rules on the steady state probability. We recall that the assumption regarding balking probability is $\{1 - \frac{1}{(a-nb)}\}$. Therefore the probability of joining is $\frac{1}{(a-nb)}$.

We wanted to analyse the effect of change in ‘ a ’ and ‘ b ’. We have observed that if ‘ a ’ increases with ‘ b ’ held constant, then the probability of joining decreases. This is by comparing table 2 and table 3. In table 1, we have seen that the probability that there is no customer in the system is 0.81250. This is small when compared to the probability of 0.91121 in table 2. On other hand, we can examine the impact of increasing ‘ b ’ keeping ‘ a ’ constant with comparing table 2, table 3 and table 4.

The boundary condition is met with $k=4$ in table1, with $k=3$ in table 2 and $k=2$ in table 5. A clear pattern has emerged; ‘ k ’ decreases as ‘ a ’ increases (which is to be expected). While the optimum value of k will depend on the constants ‘ a ’ and ‘ b ’, it has to be emphasized that the values of ‘ a ’ and ‘ b ’ have to be chosen given the conditions of the queuing system and the assessment made by the management of the queuing system about nature of reverse balking probabilities. Therefore our reverse balking formulation, which is the main thrust of this paper, is apparently more flexible than Awasthi (2018).

REFERENCES:

Awasthi, B. (2018) Performance Analysis of $M / M / 1 / K$ Finite Capacity Queueing Model with Reverse Balking and Reverse Reneging. **Journal of Computer and Mathematical Sciences**, 9, 850–855.



- Choudhury, A., & Medhi, P. (2011A) Balking and reneging in multi-server Markovian queuing system. **International Journal of Mathematics in Operation Research**, 3(4), 377-394.
- Choudhury, A., & Medhi, P. (2011B) A simple analysis of impatience in multi-server queues. **International Journal of Applied Management Science**, 3, 294-315.
- Choudhury, A., & Medhi, P. (2011C). Some aspects of balking and reneging in finite buffer queue. **RAIRO-Operations Research**, 45(3), 223-240.
- Choudhury, A., & Medhi, P. (2012) A few words on M/E_k/1/1 queuing model with reneging customers. **International Journal of Research in Applied Natural and Statistical Sciences**, 1(1), 1-7
- Haghighi, A. M., Medhi, J., & Mohanty, S. G. (1986) On a multi-server Markovian queuing system with balking and reneging. **Computer and Operational Research**, 13(4), 421-425.
- Haight, F. A. (1957) Queuing with balking. **Biometrika**, 44(3/4), 360-369.
- Haight, F. A. (1959) Queuing with reneging. **Metrika**, 2(1), 186-197.
- Jain, N. K., Kumar, R., & Som, B. K. (2014) An M/M/1/N queuing system with reverse balking. **American Journal of Operational Research**, 2(2), 17-20.
- Kumar, R., Som, B. K., & Jain, S. (2015) An M/M/1/N feedback queuing system with reverse balking. **Journal of Reliability and Statistical Studies**, 8(1), 31-38.
- Sasikala, S., & Thaiparanjan, M. (2016) The M/M/1/N interdependent queuing model with controllable arrival rates and reverse balking. **International Journal of Current Research**, 8(1), 311-316.
- Som, B. K., & Kumar, R. (2012) An M/M/1/N queuing model with retention of reneged customers and balking. **American Journal of Operational Research**, 2(1), 1-5.
- Som, B. K., & Kumar, R. (2017) A heterogeneous queuing system with reverse balking and reneging. **Journal of Industrial and Production Engineering**, 2(2), 1-5.
- Sharma, J. K. (2013) **Operation research theory and applications**, 5 ed, Macmillan publishers India Ltd, Chennai.
- Swarup, K., Gupta, P. K., & Mohan, M. (2014) **An introduction to management science operation research**, 16 ed, Sultan Chand & Sons educational publishers, New Delhi.

APPENDIX.

Derivation of $P'(1)$

Let $P(s)$ denote the probability generating function, defined by

$$P'(s) = \sum_{n=0}^{\infty} P_n S^n$$

From equation (2) we have

$$\mu p_{n+1} + \left[\frac{\lambda}{a-(n-1)b} \right] p_{n-1} = \left[\frac{\lambda}{a-nb} + \mu \right] p_{n-1} \quad n = 1, 2, 3, \dots, k-1$$

Multiplying both sides of the above equation by S^n and summing over 'n'

$$\lambda s \left[\sum_{n=1}^{k-1} \frac{P_{n-1} S^{n-1}}{a-(n-1)b} \right] - \lambda \left[\sum_{n=1}^{k-1} \frac{P_n S^n}{a-nb} \right] = \mu \sum_{n=1}^{k-1} P_n S^n - \frac{1}{s} \mu \sum_{n=1}^{k-1} P_{n+1} S^{n+1} \quad (7)$$

From equation (3) we have



$$\left[\frac{\lambda}{a-(k-1)b} \right] p_{k-1} = \mu p_k$$

Multiplying both sides of the above equation by S^k

$$\frac{\lambda s}{a-(k-1)b} P_{k-1} S^{k-1} = \mu P_k S^k \tag{8}$$

Now adding (7) and (8)

$$\begin{aligned} \Rightarrow \lambda s \left[\sum_{n=1}^{k-1} \left\{ \frac{P_{n-1} S^{n-1}}{a-(n-1)b} \right\} + \frac{P_{k-1} S^{k-1}}{a-(N-1)b} \right] - \lambda \sum_{n=1}^{k-1} \frac{P_n S^n}{a-nb} \\ = \mu \left[\sum_{n=1}^{k-1} P_n S^n + P_k S^k - \frac{1}{s} \sum_{n=1}^{k-1} P_{n+1} S^{n+1} \right] \\ \Rightarrow \lambda s \sum_{n=0}^{k-1} \left\{ \frac{P_{n-1} S^{n-1}}{a-nb} \right\} - \lambda \sum_{n=0}^{k-1} \frac{P_n S^n}{a-nb} + \frac{\lambda P_0}{a} \\ = \mu \left[\sum_{n=1}^k P_n S^n - \frac{1}{s} \sum_{n=1}^{k-1} P_{n+1} S^{n+1} \right] \\ \Rightarrow \lambda (s-1) \sum_{n=0}^{k-1} \frac{P_n S^n}{a-nb} + \frac{\lambda P_0}{a} \\ = \mu \\ \Rightarrow \lambda (s-1) \sum_{n=0}^{k-1} \frac{P_n S^n}{a-nb} + \frac{\lambda P_0}{a} = \frac{\mu}{s} [sP(s) - sP_0 - P(s) + P_0 + sP_1] \\ \Rightarrow \frac{\lambda s}{\mu} (s-1) \sum_{n=0}^{k-1} \frac{P_n S^n}{a-nb} + \frac{\lambda s}{\mu a} P_0 = P(s)(s-1) - P_0(s-1) + \frac{\lambda s}{\mu a} P_0 \\ \Rightarrow P(s) = \frac{\lambda}{\mu} \sum_{n=0}^{k-1} \frac{P_n S^{n+1}}{a-nb} + P_0 \end{aligned}$$

Therefore, we have

$$\Rightarrow P'(s) = \frac{\lambda}{\mu} \sum_{n=0}^{k-1} \frac{(n+1)P_n S^n}{(a-nb)} \tag{9}$$

Putting $s=1$ in equation (9), we get the average number of customers in the system i.e. L

$$L = P'(1) = \frac{\lambda}{\mu} \sum_{n=0}^{k-1} \frac{(n+1)P_n}{(a-nb)}$$