HETEROGENEOUS SERVER MARKOVIAN QUEUE WITH RESTRICTED ADMISSIBILITY AND WITH RENEGING

R. Kalyanaraman
Department of Mathematics, Annamalai University, Annamalainagar-608 002
E-mail: r.kalyan24@rediffmail.com
R. Senthilkumar
Mathematics Wing, DDE, Annamalai University, Annamalainagar-608 002
E-mail: rsenthil1968@gmail.com

Abstract: We consider a two heterogeneous server $M/M/2$ queue with restricted admissibility. If both the servers are free, the head of the queue choose server 1 with probability α and server 2 with probability $1-\alpha(0<\alpha<1)$. If both the server are busy the customer waits in the queue for the first-free server to get service. In addition, the customer renege the system, if he become impatient while waiting in the queue. For this model, the steady state solution and some performance have been obtained. A numerical study has been calculated.

Keywords: Markovian Queue, Heterogeneous Server, Reneging and Steady State Solution.

1. Introduction: In most studies on multi server queueing systems, the servers are homogeneous. But, in real life situations, the homogeneous service mechanism is not true, except the service process in mechanically or electronically controlled.

In 1981, Neuts and Takahashi have pointed out that for the queueing system with two heterogeneous servers, analytical results are intractable. Even though, some researchers focused their studies on queue with two heterogeneous servers.

The equilibrium analysis for the general input and general service time and with n servers was given in Kendall (1953). A non-constructive existence theorem for the stationary distribution of a general input and general service time was presented in Kiefer and Wolfowitz (1955). Karlin and Mc Gregor (1958) obtained the busy period distribution for the $M/M/S$ queue. Krishnamoorthi (1963) considered a Poisson queue with two heterogeneous servers and with violation of the First-in-First out principle.

Heffer (1969) has analyzed waiting time distribution of $M/E_k/S$ queue. A Markovian queueing system with balking and two heterogeneous servers has been considered in Singh (1970). The author determines the capacity of the slower server and obtains the optimal service rates. Singh (1973) discussed a Markovian queue with the number of servers depending upon the queue length. Desmit (1983 a,b) presented an approach to identify the distribution of waiting times and queue lengths for the queue $GI/H_2/S$. He reduced the problem to the solution of the Wiener-Hopf-type equations and then used a factorization method to solve the system.

Lin and Kumar (1984) has analyzed the optimal control of a queueing system with two heterogeneous servers. Rubinovitch (1985 a,b) studied the problem of a heterogeneous two channels queueing systems. In his first paper he discussed three simple models and gave the condition when to discard to slower server depending on the expected number of customers in the system. In the second paper he studied a queueing model with a stalling concept. In 1999, Abou-El-Ata and Shawky introduced a simpler approach to find the condition when to discard the slower server in a heterogeneous two channels queue.

Barcelo (2003) has obtained an approximation for the mean waiting time $M/H_{2b}/S$ queue. Shin and Moon (2009) has carried out an approximate analysis of $M/G/C$ queue. Arkat and Farahani (2014), has used a partial-fraction decomposition approach to the $M/H_2/2$ queue. Models of queueing systems with different intensity of service is used for the study of telecommunication process. Zhernovyi (2011) analyzed queue with switching of service modes and threshold blocking of input flow. Kopytko and Zhernovyi (2011) investigated Markovian queue with switching of service mode.

In day to day life, there are numerous queueing situations in which all the arriving customers are not allowed to enter into the system by the administrator, if he feels that the customers arrive faster than they can be served, then he may adopt a policy of restricting the arriving customers. This will help him to prevent the system from becoming over loaded Rue and Rosenshine (1981) gavean optimal control policy on the arrival of an $M/M/1$ queue. Neuts (1984)considered an $M/G/1$ queue with a restriction on the time period at which the customers are admitted. Stidham (1985) considered an optimal arrival policy queue.

For the heterogeneous server model, if more than one server is free, the allocation of customer to the servers is a problem to be considered. Concerning the server allocation, one possibility is that the first job in the queue will be allocated as soon as any server becomes idle, but there are others. Indeed, there are four different allocation strategies among the most popular namely, (i) The Fastest Server First (FST) allocation, for which the fastest available server is always allocated first. (ii) the Randomly Chosen Server (RCS) allocation, for which the next job in the queue is randomly sent to anyone of the servers that is idle. (iii) the Slowest Server First (SSF) allocation, where the job is always first allocated to the slowest free server. (iv) if both the servers are idle, the head of the customer joins the first server with probability α or the second serverwithprobability$(1-\alpha)$, called Bernoulli Allocation (BA). Alvesetal (2011),analyzed heterogeneous $M/M/C$ queue with the rest three allocation policies Heterogeneous Server queue with fourth allocation policy have been studied by many researchers including Abou-El-Ata and Ibrahim (1992), Abou-El-Ata and Shawky (1999) and Al-Seedy-Etal (2009).

In real life time, many queueing situations arise in which the customer may be discouraged by a long queue. As a result, the customers either decide not to join the queue (called balk) or depend after joining the queue without getting service(called renege). Some real life examples are, inpatient telephone switchboard customer, inventory systems that store perishable goods. Queueing system with balking, reneging or both were studied by many researchers. An $M/M/1$ queue with customers reneging was proposed by Haight (1956).

2. Mathematical Model: This article, deal with an $M/M/2$ heterogeneous server queue with restricted admissibility of customers, reneging and with retention of reneged customers. The following assumptions have been considered:

a) Customers arrive at the system one by one according to a Poisson process with rate λ on arrival the admission of customer to the queue is based on a Bernoulli process. Let p be the probability that the arriving customer will be allowed to join the system and 1 - p be the probability that the customer is not-allowed to join the system.
b) After joining the queue, each customer will wait a random period of time for service to begin. If it has not be given by them, the customer will get impatient and leave the queue without getting service. This random period follows negative exponential distribution with rate ξ. The average reneging rate is given by
c) $\xi(n) = \begin{cases} 0 & 0 \le n \le 2 \\ (n-2)\xi, & n \ge 3 \end{cases}$

There are two servers, they serve the customers based on exponential distribution with rate μ_1 and μ_2 for server 1 and 2 respectively$(\mu_2 < \mu_1)$let $\mu = \mu_1 + \mu_2$. Each customer is served only by one server and the queue discipline is first come first served. If the system is empty, an arriving customer joins the first server with probability α and the second server with probability$1-\alpha$ $(0 < \alpha < 1)$. If an arriving customer finds both the server busy he waits for the first free server.

For the mathematical analysis, we define the following probabilities. Let $p_n(t)$ be the probability that there are n customers in the system at time t, $n \geq 0$
$p_{10}(t)$ be the probability that one customer in the system and the server one is busy at time t;
$p_{01}(t)$ be the probability that one customer in the system and the second server is busy at time t and $p_1(t) = p_{10}(t) + p_{01}(t)$.
The differential difference equations are derived by using the general birth death arguments.
The equations are

$$\frac{d}{dt}p_0(t) = -p\lambda p_{00}(t) + \mu_1 p_{10}(t) + \mu_2 p_{01}(t) \tag{1}$$
$$\frac{d}{dt}p_{10}(t) = -(p\lambda + \mu_1)p_{10}(t) + \alpha p\lambda p_{00}(t) + \mu_2 p_2(t) \tag{2}$$
$$\frac{d}{dt}p_{01}(t) = -(p\lambda + \mu_2)p_{01}(t) + (1-\alpha)p\lambda p_{00}(t) + \mu_1 p_2(t) \tag{3}$$
$$\frac{d}{dt}p_2(t) = -(p\lambda + \mu)p_2(t) + (\mu + \xi_q)p_3(t) + \lambda p p_1(t) \tag{4}$$
$$\frac{d}{dt}p_n(t) = -(p\lambda + \mu + (n-2)\xi_q)p_n(t) + \left[\left(\mu + \left((n+1)-2\right)\right)\xi_q\right]p_{n+1}(t) + \lambda p p_{n-1}(t), \quad 3 \leq n \leq N-1 \tag{5}$$
$$\frac{d}{dt}p_{N-1}(t) = -\lambda p p_{N-1}(t) + [\mu + (N-2)\xi_q]p_N(t), \qquad n \geq N \tag{6}$$

These equations are solved iteratively in steady-state in order to obtain the steady state solution. In steady state, $\lim_{n\to\infty} p_n(t) = p_n$. The equations (1) to (6) becomes

$$p\lambda p_{00} = \mu_1 p_{10} + \mu_2 p_{01} \tag{7}$$
$$(p\lambda + \mu_1)p_{10} = \alpha\, p\lambda p_{00} + \mu_2 p_2 \tag{8}$$
$$(p\lambda + \mu_2)p_{01} = (1-\alpha)\, p\lambda p_{00} + \mu_1 p_2 \tag{9}$$
$$(p\lambda + \mu)p_2 = (\mu + \xi_q)p_3 + \lambda p p_1 \tag{10}$$
$$[p\lambda + \mu + (n-2)\xi_q]p_n = [(\mu + ((n+1)-2))\xi_q]p_{n+1} + \lambda p p_{n-1}, \quad 3 \leq n \leq N-1 \tag{11}$$
$$\lambda p p_{N-1} = [\mu + (N-2)\xi_q]p_n, \qquad n \geq N \tag{12}$$

On solving equations (7) to (12), we have

$$p_{01} = \frac{p\lambda[p\lambda + (1-\alpha)\mu_1 + (1-\alpha)\mu_2]}{\mu_2(2p\lambda + \mu_1 + \mu_2)}p_{00} \tag{13}$$

$$p_{10} = \frac{p\lambda[p\lambda + \alpha(\mu_1 + \mu_2)]}{\mu_1(2p\lambda + \mu_1 + \mu_2)}p_{00} \tag{14}$$

$$p_1 = \frac{p\lambda(\mu_1 + \mu_2)[p\lambda + (1-\alpha)\mu_1 + \alpha\mu_2]}{\mu_1\mu_2(2p\lambda + \mu_1 + \mu_2)}p_{00} \tag{15}$$

$$p_2 = \frac{(p\lambda)^2[p\lambda + (1-\alpha)\mu_1 + \alpha\mu_2]}{\mu_1\mu_2(2p\lambda + \mu_1 + \mu_2)}p_{00} \tag{16}$$

$$p_3 = \frac{p\lambda}{\mu + \xi_q} L p_{00}$$

Where

$$L = \frac{(p\lambda)^2[p\lambda + (1-\alpha)\mu_1 + \alpha\mu_2]}{\mu_1\mu_2(2p\lambda + \mu_1 + \mu_2)}$$

$$p_n = \frac{(p\lambda)^{n-2}}{\prod_{i=1}^{n-2}(\mu + i\xi_q)} L p_{00}, \qquad n = 3,4,\ldots \tag{17}$$

Using the normalization condition $\sum_{n=0}^{\infty} p_n = 1$, we get

$$p_{00} = \frac{1}{\left[1 + \frac{L\mu}{p\lambda} + L + L\sum_{n=3}^{\infty}\frac{(p\lambda)^{n-2}}{\prod_{i=1}^{n=2}(\mu + i\xi_q)}\right]} \tag{18}$$

Equations (13) to (18) represents the steady state solution of the queueing model discussed in this article.

3. Some Performance Measures: In this section, we derive some performance measures of the queue definition section 2. Using the steady state probabilities presented in section 2, the following performance measures can be obtained:

a) Probability of down time = p_0
b) Probability that both servers be busy $p_B = 1 - p_0 - p_1$
c) Expected number of customers in the system

$$E(N) = \sum_{n=0}^{\infty} np_n$$

d) Expected number of customers in the queue

$$E(N_q) = \sum_{n=1}^{\infty} (n-1)p_n$$

e) Expected number of customers served

$$E(M) = \mu_1 p_{10} + \mu_2 p_{01} + (\mu_1 + \mu_2) \sum_{n=2}^{\infty} p_n$$

f) Average rate of loss of customers

$$E(R) = \sum_{n=1}^{\infty} (n-1)\xi p_n$$

4. Particular Cases: In this section, we derive some particular model by taking particular value to the parameters and particular distribution.

1. When $\mu_1 = \mu_2 = \mu$

$$p_1 = \frac{p\lambda}{\mu} p_{00}$$

$$p_2 = \frac{1}{2}\left(\frac{p\lambda}{\mu}\right)^2 p_{00}$$

Where $l = \frac{1}{2}\left(\frac{p\lambda}{\mu}\right)^2$

$$\Rightarrow p_1 = 2l \, \frac{\mu}{p\lambda} p_{00}$$

$$p_n = \frac{(p\lambda)^{n-2}}{\prod_{i=1}^{n-2}(\mu + i\xi_q)} lp_{00}, \qquad n = 3,4,5, \ldots$$

$$p_{00} = \frac{1}{\left[1 + \frac{l\mu}{p\lambda} + l + l\sum_{n=3}^{\infty} \frac{(p\lambda)^{n-2}}{\prod_{i=1}^{n=2}(\mu + i\xi_q)}\right]}$$

2. When $p = 1$

$$p_1 = \frac{p\lambda(\mu_1+\mu_2)[\lambda+(1-\alpha)\mu_1+\alpha\mu_2]}{\mu_1\mu_2(2\lambda+\mu_1+\mu_2)} p_{00}$$

$$p_2 = \frac{\lambda^2[\lambda+(1-\alpha)\mu_1+\alpha\mu_2]}{\mu_1\mu_2(2\lambda+\mu_1+\mu_2)} p_{00}$$

$$p_2 = Ap_{00}$$

Where

$$A = \frac{\lambda^2[\lambda + (1-\alpha)\mu_1 + \alpha\mu_2]}{\mu_1\mu_2(2\lambda + \mu_1 + \mu_2)}$$

$$p_3 = \frac{\lambda}{\mu + \xi_q} Ap_{00}$$

$$p_n = \frac{(\lambda)^{n-2}}{\prod_{i=1}^{n-2}(\mu + i\xi_q)} Ap_{00}, \qquad n = 3,4, \ldots$$

$$p_{00} = \frac{1}{\left[1 + \frac{A\mu}{\lambda} + A + A\sum_{n=3}^{\infty} \frac{(\lambda)^{n-2}}{\prod_{i=1}^{n-2}(\mu + i\xi_q)}\right]}$$

5. Numerical Study:

Table 1: $(\mu = 9, \mu_1 = 5, \mu_2 = 4, \xi_q = 2, p = 0.6, \alpha = 0.4)$

$\boldsymbol{\lambda}$	Pr. Down Time	Pr. Both Server Busy	$E(N)$	$E(N_q)$	$E(M)$	$E(R)$
0.1	0.986302	0.000091	0.013607	0.000091	0.059178	0.000181
0.2	0.972808	0.000358	0.026834	0.000358	0.116737	0.000716
0.3	0.959515	0.000794	0.039691	0.000794	0.172713	0.001588
0.4	0.946423	0.001392	0.052185	0.001392	0.227142	0.002783
0.5	0.933530	0.002144	0.064326	0.002144	0.280059	0.004288
0.6	0.920833	0.003045	0.076122	0.003045	0.331500	0.006090
0.7	0.908331	0.004087	0.087582	0.004087	0.381499	0.008174
0.8	0.896022	0.005265	0.098714	0.005265	0.430090	0.010529
0.9	0.883903	0.006572	0.109525	0.006572	0.477308	0.013143
1.0	0.871974	0.008002	0.120025	0.008002	0.523184	0.016003

Table 2: $(\mu = 5, \mu_1 = 3, \mu_2 = 2, \xi_q = 3, p = 0.7, \alpha = 0.5)$

$\boldsymbol{\lambda}$	Pr. Down Time	Pr. Both Server Busy	$E(N)$	$E(N_q)$	$E(M)$	$E(R)$
0.1	0.971275	0.000397	0.028329	0.000397	0.067989	0.001190
0.2	0.943426	0.001541	0.055033	0.001541	0.132080	0.004623
0.3	0.916443	0.003368	0.080189	0.003368	0.192453	0.010104
0.4	0.890313	0.005817	0.103870	0.005817	0.249288	0.017450
0.5	0.865021	0.008830	0.126149	0.008830	0.302757	0.026491
0.6	0.840548	0.012356	0.147096	0.012356	0.353030	0.037068
0.7	0.816877	0.016344	0.166779	0.016344	0.400270	0.049033
0.8	0.793987	0.020750	0.185264	0.020750	0.444633	0.062249
0.9	0.771858	0.025529	0.202613	0.025529	0.486271	0.076588
1.0	0.750469	0.030644	0.218887	0.030644	0.525328	0.091932

Table 3: $(\mu = 8, \mu_1 = 4, \mu_2 = 4, \xi_q = 2, p = 0.7, \alpha = 0.5)$

$\boldsymbol{\lambda}$	Pr. Down Time	Pr. Both Server Busy	$E(N)$	$E(N_q)$	$E(M)$	$E(R)$
0.1	0.982653	0.000150	0.017196	0.000150	0.068786	0.000301
0.2	0.965612	0.000591	0.033796	0.000591	0.135186	0.001183
0.3	0.948876	0.001308	0.049816	0.001308	0.199264	0.002615
0.4	0.932444	0.002284	0.065271	0.002284	0.261084	0.004569
0.5	0.916315	0.003508	0.080178	0.003508	0.320710	0.007016
0.6	0.900485	0.004964	0.094551	0.004964	0.378204	0.009928
0.7	0.884953	0.006640	0.108407	0.006640	0.433627	0.013280
0.8	0.869716	0.008523	0.121760	0.008523	0.487041	0.017046
0.9	0.854772	0.010602	0.134627	0.010602	0.538506	0.021204
1.0	0.840115	0.012864	0.147020	0.012864	0.588081	0.025729

Table 4: ($\mu = 9, \mu_1 = 5, \mu_2 = 4, \xi_q = 2, p = 1, \alpha = 0.5$)

λ	Pr. Down Time	Pr. Both Server Busy	$E(N)$	$E(N_q)$	$E(M)$	$E(R)$
0.1	0.977756	0.000244	0.022000	0.000244	0.097776	0.000489
0.2	0.956023	0.000956	0.043021	0.000956	0.191205	0.001912
0.3	0.934798	0.002103	0.063099	0.002103	0.280439	0.004207
0.4	0.914077	0.003656	0.802267	0.003656	0.365631	0.007313
0.5	0.893855	0.005587	0.100559	0.005587	0.446927	0.011173
0.6	0.871426	0.007867	0.118007	0.007867	0.524476	0.015734
0.7	0.854884	0.010472	0.134644	0.010472	0.598419	0.020945
0.8	0.836120	0.013378	0.150502	0.013378	0.668896	0.026756
0.9	0.817829	0.016561	0.165610	0.016561	0.736046	0.033122
1.0	0.800000	0.024000	0.180000	0.020000	0.800000	0.040000

References:

1. Abou-El-Ata, M.O., and Shawky, A.I., A simple approach for the slower server problem. *Commun. Fa. Sci. Univ*, v.48 pp. 1-6, 1999.
2. Abou-El-Ata, M.O., and Shawky, A.L., A simple approach for the slower server problem. *Commun Fac. Sci, uni. Anv series* A1, 88, 1-6, 1999.
3. Alves, F.S.Q., Yehia, H.C., Pedrosa, L.A.C., Cruz, F.R.B. and Laoucine Kerbache, Upper bounds on performance measures of Heterogeneous M/M/C Queues, *Mathematical problems in Engineering volume* 2011, 1-18.
4. Arkat, J., and Farahani, M.H., Partial Fraction recomposition approach to the M/H_2/2 queue. *Franian Jr. of Operation Research*, 5(1), 55-63, 2014.
5. Barcelo, F., An Engineering approximation for the mean waiting time in M/H_2/b=S queue. *In Proceeding of the International Network Optimization Conference, Paris France*, 27-19, October 2003.
6. Desmit, J.H.A., A numerical solution for the multi server queue with hyper exponential service times, *Oper. Res. Lett*, v.2, 217-224, 1983(a).
7. Desmit, J.H.A., The Queue GI=M=S with customers of di_erent types or the queue GI/HM/S, *Adv. Appl. Prob*., v.15, 392-419, 1983(b).
8. He_er, J.C., Steady state solution of the M/E_k/C (F1F0) queuing system, *CORSJ*., v.7, 16-30, 1969.
9. Kalyanaraman, R., and Suvitha, V., On a restricted admissibility single server compulsory vacation queue with essential and optional service, *Mathematica Applicanda*., v.44(1), 87-112, 2016.
10. Karlin, S., and Gregor, J. Mc., Many server queuing process with Poisson input and exponential service times, *Pacfic J. Math*., v.8, 87-118, 1958.
11. Kendall, D.G., Stochastic process in the theory of queues, *Ann. math. stat*., v.24, 333-354, 1953.
12. Kiefer, J., and Wolfowitz, J., On the theory of queues with many servers, *Trans Amer. math. soc.*, 781-18, 1955.
13. Kopytko, B. and Zhernovi, K., Stationary characteristics of M^Q/M/I queue with switching of service modes, *Scientific Research of the Ins. Mathematics and Computer Science*, v.2(10), 117-128, 2011.
14. Krishnamoorthi, B., On Poisson queue with two heterogeneous servers, *Oper. Res*., v.11(3) 321-330, 1963.
15. Lin, W., and Kumar, P., Optimal control of c queuing system with two heterogeneous servers, Automatic Control, *IEEE trans*., v.29(8), 696-703, 1984.
16. Neuts, M.F., The M/G/1 queue with limited number of admission or a limited admission period during each service time, *Technical report no 978, university of Delaware*, 1984.
17. Neuts, M.F., and Takhashi, Y., Asymptotic behavior of the stationary distribution in the GI/PH/C queue with heterogeneous server, *Z. Wahrscheinlickeirsth*, v.57, 441-452, 1981.
18. Rubinovitch, M., The slow server problem; A queue with stalling, *Jr. Appl. Prob*., v.22, 879-892, 1985(b).
19. Rue, R.C., and Rosenshine, M., Some properties of optimal control policies for entries to an M/M/1 queue. *Naval Research Logistics Quarterly*, v.28, 525-535, 1981.

20. Rubinovitch M., The slow server problem, *Jr. Appl. Prob*, v.22, 205-213, 1985(a).
21. Shin, Y.W., and Moon, D.H., Sensitivity and approximation of M=G=C queue: Numerical experiments *Proceedings of it 8th international symposium on operations Research and its applications*, 140-147, 2009.
22. Singh, V.P., Two servers Markovian queues with balking, Heterogeneous vs Homogeneous servers, *Oper. Res*., v.18(1), 145-159, 1970.
23. Singh, V.P., Queue dependence-servers, *Jr. of Eng. Maths*, v.7(2), 123-126, 1973.
24. Stidham, Jr.S., Optimal control of admission to the queueing systems, IEEE *Transactions on Automatic Control*, AC-30, 705-713, 1985.
25. Zyernovyi, K.Yu., Investigation of the $M^Q/G/I$ queue with switching of service modes and threshold blocking of input flow, *Information Process*, v.10(2), 159-180, 2011.

***