Discrete-Time Geo/G/2 Queue under a Serial and Parallel Queue Disciplines

Sivasamy Ramasamy, Onkabetse A. Daman and Sulaiman Sani *

Abstract—**This article discusses the steady state analysis of a heterogeneous server queuing system called the Geo/G/2 queue. We supposed that arrivals occur according to a geometric process to receive service on server-1 according to a geometric service time distribution with mean rate μ or on server-2 according to a general distribution $B(t)$ with mean rate μ_2. There are two queue disciplines employed for service namely; the serial queue discipline (queue discipline-I) and the parallel queue discipline (queue discipline-II). Using the embedded method when the discipline is serial and the supplementary variable technique when it is parallel, we present an exact analysis of the arrival distribution. Furthermore, the actual waiting time expectations are derived and approximated. Our analysis can be applied in managing service systems in many areas of communication, telecommunications, business and computer systems where services are discrete in nature for instance, in the performance evaluation and design of buffers for statistical multiplexers, traffic concentrators, switch modules, networks.**

Keywords: ***Geometric arrival, Geo/G/2 queue, Geo/(Geo+G)/2 queue, Geo/Geo,G/2 queue.***

1 Introduction

We study a geometric arriving and general service time queuing system[1] called the $Geo/G/2$ queue with two servers modeled as server-1 and server-2. This kind of queuing system is a model of slots with finite boundaries where at most only one arrival (customer) occurs in a given time (late arrival) for service. We suppose that arrivals (customers, packets, inputs) and service times (length of service) are integer multiples of a given slot. This implies that departures only occur at the other end of the slot. The service and inter-arrival durations between consecutive arrivals are measured as random numbers of slot durations. This give rise to a discrete-time queuing system of the form $Geo/Geo+G/2$ or $Geo/Geo, G/2$ depending on the schedule for service (queue discipline). Now, suppose that the proposed queuing systems have an infinite number of waiting positions with server-1 faster than server-2. Then, an interesting feature of this investigation is that, it derives the discrete results of this work to the corresponding continuous time version $M/G/2$ queues analyzed in Sivasamy, Daman & Sulaiman [15].

Normally, discrete-time scaling (small time scale) and their corresponding models often reflect an underlying application. For example, the clock time unit in a computer system, the fixed size data units (bits, bytes, fixed length packets) on a communication channel. Similarly, for sufficiently small slot lengths, discrete-time queuing models are approximations for the corresponding models where the time scale is continuous. Furthermore, asynchronous transfer mode (ATM) multiplexers and broadband integrated service digital network (B-ISDN) are used to transfer data sets, voice and video communications on discrete-time basis. In both the multiplexers and the B-ISDN for instance, the time axis is normally divided into slots (fixed-length of continuous intervals called slots of unit length (right-end boundary=left end boundary)). Takagi [17] indicated that this type of models is essentially the best choice for analyzing computer and communication systems. Bruneel & Kim [5] compared some features of discrete-time parallel queuing models with that of the continuous ones and indicated that in continuous-time parallel queuing models, the probability that an arrival and a departure occurring simultaneously in a very small time interval is zero. This is not so in discrete-time models where both events can occur simultaneously at a boundary epoch of a slot. Under this added advantage, it requires that the order of occurrence be taken care of; that is either arrival first (AF) or departure first (DF). Hence, the need to study such models for the benefit of computers or telecommunication systems such as the ones exemplified above.

The literature on discrete-time queuing models is enormous. Over the years, a lot of research has been carried out. For a survey of related ones, see Takagi [17], Bruneel [4], Bruneel & Kim [5], Ishizaki, Takine & Hasegawa [10], Bruneel, Walraevens, Claeys, & Wittevrongel [6], Briem, Theimer, & Kroner [3]. Precisely, Briem, Theimer & Kro-

*Sivasamy Ramasamy is a Professor of Statistics currently with the Department of Statistics, University of Botswana-Gaborone. His field of research is Stochastic modeling & Analysis. Onkabetse A. Daman is a senior Lecturer and the former HOD, Dept. of Mathematics, University of Botswana-Gaborone. His area of research is Stochastic Processes, Mathematical Analysis & Applications. Sulaiman Sani is a Research Scholar in the Department of Mathematics, University of Botswana-Gaborone. His area of research is on Stochastic modeling with bias in Queuing theory. Email: man15j@yahoo.com: **Manuscript submitted February 2015**.

[1]Under two distinct queue disciplines.

(Advance online publication: 14 November 2015)

ner [3] studied a discrete-time single server queuing system with a non-renewal customer process and provided an exact analysis of both the infinite and the finite capacity queue yields in terms of the state probabilities at departure instants as well as the loss probability. The results show that one can obtain interesting results for queues under discrete-time scaling assumption. For more on single server discrete queuing systems, see Bruneel [5], Ishizaki, Takine & Hasegawa [10]. Similarly, works on multi-server queuing systems with discrete-time scaling have grown tremendously in the last two decades. For a survey of related ones; see Goswami & Mund [8], Artalejo & Lopez Herrero [1], Goswami [7].

Now, what can be observed in these and many other literature sources on multi-server models under discrete time assumption is the adoption of the FCFS queue discipline. This assumption may not be realistic. For instance, if a telecommunication system or device provides service with varying speeds, then arriving packets might be preferred to be allocated the fastest device for service. On the other hand, if allocated the slowest device randomly, then there is a possibility that packets entering the system after the one allocated to the slow device to clear out earlier by obtaining service from the device with a faster working rate. Apparently in this case, the FCFS queue discipline is violated due to heterogeneity in service speeds of the two working devices. This and similar real life scenarios make the adoption of the FCFS queue discipline unrealistic in multi-server discrete-time queuing systems with embedded heterogeneity because of the high probability of violation therein. Hence, there is the need for designing alternative queue disciplines that can reduce the impact of the violation so that the resulting waiting times of customers in this kind of queuing systems are almost identical with that of the FCFS.

Our motivation to study these kind of models stemmed from the several applications of the models working under the proposed queue disciplines giving rise to either the $Geo/Geo+G/2$ or $Geo/Geo,G/2$ queuing models respectively. Specific examples of application scenarios include computer systems, communication systems, telecommunication networks, production management, Broadband Integrated Services Digital Network (BISDN), dynamic bandwidth allocation and flexible capacity allocation. In these systems, information is digitized and segmented into small packets arranged serially or in parallel. Therefore, analyzing models for such systems is an excellent tool for decision making relative to congestion management for better service delivery. Interestingly, unlike in many models designed with similar assumptions, here, the effects of the two queue disciplines can all be computed[2]. Most importantly, the analysis of the discrete time $Geo/Geo,G/2$ queue carried out is relatively close to that of Boxma, Deng & Zwart [2] and Krishnamoorthi [11] but more convenient for application in communications and modern computer systems. In this sense, each queue discipline here is an excellent alternative for use in our real life applications. Precisely, the model under queue discipline-I is called the discrete time $Geo/Geo+G/2$ with serialized servers and the model under queue discipline-II is called the discrete time $Geo/Geo,G/2$ queue with parallel servers. With this structuring, our work stands to benefit both the service provider and the customer through efficient resource management in the former and minimization of waiting times in the later.

[2]Both analytically and numerically.

1.1 Basic Assumptions

We suppose that a slot of unit length is given such that arrivals occur on slot boundaries according to an Arrival First (AF) policy. Marking the time axis by $0, 1, 2, .., t, ...$, and supposing that these arrivals occur at $0-, 1-, , t-$, time points such that a service starts only at slot boundaries with each service duration taking a number of slots. In addition, an arrival can leave the system upon service completion only[3]. Furthermore, we assumed that potential departures occur at slot boundaries at $0+, 1+, ..., t+, ...$ instants. Denote the time between successive arrivals (the inter-arrival time) by A. For a detail description of the discrete time concepts employed here, see Gupta & Goswami [9] who modeled a single-server bulk service queue with finite buffer space in a discrete-time environment and provided the analysis under both arrival first (AF) and departure first (DF) management policies and distributions of buffer content at various epochs. Such management policies play a significant role in the determination of steady-state probabilities relating to the number of arrivals in the system (queue) at special epochs (e.g., arrival, departure, and arbitrary epochs) and hence, they affect performance measures to a great extent. Now, assuming that the number of arrivals in successive slots are independent and identically distributed (i.i.d.) random variables (derived from a Geometric (or Bernoulli) Arrival Process:) subject to the condition that only one arrival occurs in a slot with probability λ; $(0 < \lambda < 1)$. This assumption ensures that the inter-arrival time A is geometrically distributed[4] with mean $\frac{1}{\lambda}$ and probability distribution $P(A = k\ slots) = \lambda(1-\lambda)^{k-1}$ for $k = 1, 2, ...$ Denote further by $A(z)$, the probability generating function (PGF) of inter-arrival times and by $L(z)$, the PGF of the number of arrivals in a slot such that

$$A(z) = \frac{\lambda z}{1-(1-\lambda)z}; |z| < 1,\ L(z) = 1-(1-z)\lambda \quad (1)$$

Now, if it is supposed that for arrivals serviced by server-1, the service time S_1 follows the geometric distribution

[3]No reneging, balking by arrivals in the system.

[4]We further suppose that S_1 and S_2 are mutually independent with each other and the inter-arrival time distribution.

with probability mass function $f_1(k) = P(S_1 = k) = \mu(1-\mu)^{k-1}$; $k = 1, 2, ...$ with mean $\frac{1}{\mu}$ per slot. Then the PGF of S_1 and the corresponding $L_s(z)$ representing the PGF of the number of departures in a slot are respectively given by[5]

$$F_1(z) = \sum_{k=1}^{\infty} \mu(1-\mu)^{k-1} z^k \frac{\mu z}{1-(1-\mu)z} \tag{2}$$

Let the service times S_2 of arrivals serviced by server-2 follows a general distribution $f_2(k) = P(S_2 = k)$ with PGF

$$F_2(z) = \sum_{k=1}^{\infty} f_2(z) z^k; \ |z| < 1 \tag{3}$$

If the mean service time is β with rate $\mu_2 = \frac{1}{\beta}$, then one can provide an analysis for this discrete queuing system similar to that carried out in Sivasamy, Daman & Sulaiman [15] in the continuous-time case.

1.2 The Serial Queue Discipline

This is a new proposal of connecting the servers in series[6] subject to the following:

1 If an arrival enters into an idle system, his service is immediately initiated by server-1. This arrival is served by server-1 at a constant rate μ if no other arrival occurs during his on-going service period.

2 Otherwise if at least one more arrival occurs before his service completed, then the first arrival is served jointly by both servers according to the service time distribution $f_{min}(k) = P(S_{min} \leq k)$, where $S_{min} = Min(S_1, S_2)$ and the PGF of S_{min} denoted by $F_{min}(z)$ given by[7]

$$F_{min}(z) = \sum_{k=1}^{\infty} f_{min}(k) z^k \tag{4}$$

The expected values of S_1, S_2 and S_{min} are respectively given by $E[S_1] = F_1'(z) = \frac{1}{\mu}$, $E[S_2] = F_2'(1) = \beta = \frac{1}{\mu_2}$ and $E[S_{min}] = F'_{min}(1)$. For example, if the service time S_2 of arrivals on server-2 is a geometric random variable with mean $\frac{1}{\mu_2}$, then the mean of S_{min} derived from (4) is equal to[8]

$$E[S_{min}] = F'_{min}(1) = \frac{1}{\mu + \mu_2 - \mu\mu_2}$$

Lemma 1.1 *If no arrival is served simultaneously and the service time for all arrivals is equivalent to some k slots, then the service (or departure) rate $r(k)$ offered by the combined service time distributions $f_1(k)$ and $f_2(k)$ of the serialized servers in the $Geo/Geo + G/2$ queuing system is given by*

$$r(k) = \frac{f_1(k)}{1 - P(S_1 < k)} + \frac{f_2(k)}{1 - P(S_2 < k)} \tag{5}$$

Proof Suppose $\rho = \frac{\lambda}{\mu} < 1$. Then a closed form expression for $r(k)$ can be obtained for particular cases of $f_2(k)$ like the Geometric, Negative-binomial, or Phase (PH) type distributions. Suppose that the service time distribution of arrivals served by the slow server (server-2) is any of the distributions mentioned above with a finite mean. Then our discussion follows viz;

1.3 Negative-binomial Distribution

Suppose that the service time S_2 is a Negative-Binomial $NB(\alpha, \beta)$ random variable with mass function $b(k; \alpha, \mu_2) = P(S_2 = k)$. Let S_2 denotes the number of slots required to complete a service by server-2 at the α^{th} success in a sequence of independent Bernoulli trials with probability of success $v \in (0 < v < 1)$, then

$$b(k; \alpha, v) = \binom{k-1}{\alpha-1} v^{\alpha} (1-v)^{k-\alpha}; \ k = \alpha, \alpha+1, ...$$

The mean of S_2 is denoted by $E[S_2] = \frac{\alpha}{v}$ and the variance by $Var[S_2] = \frac{\alpha(1-v)}{v^2}$. Similarly, the mean service rate μ_2 is given by $\frac{v}{\alpha}$. Finally, the service (or departure) rate $r(k)$ offered by the combined service time distributions $f_1(k)$ and $f_2(k)$ of serialized servers of the $Geo/Geo + NB(\alpha, v)/2$ queuing system is given by[9] $\mu + \frac{v}{\alpha}$.

Corollary 1.2 *Suppose that $\frac{1}{\mu_2} \to \infty$. Then the $Geo/Geo + G/2 \to Geo/Geo/1$.*

Queuing models operating under conditions (1) through (4) above are a discrete-time class of the $Geo/Geo + G/2$ queues.

1.4 The‘Parallel Queue‘Discipline

Here, we adopt the queue discipline in Krishnamoorthi [11] that minimizes the violation of the FCFS principle for

[5] In this case, $|z| < 1$, $L_s(z) = 1 - (1-z)\mu$.

[6] Queue Discipline-1

[7] Further simplification of (4) yields $F_{min}(z) = F_1(z) + F_2((1-\mu)z)(1 - F_1(z))$

[8] Similarly, the service rate of $S_{min} = \mu + \mu_2 - \mu\mu_2$, since there is a positive probability of serving an arrival simultaneously by both servers just before the end of a slot boundary.

[9] One can carryout similar analysis for the above mentioned distributions also.

a two heterogeneous server queue subject to the condition that the mean service rates of server-1 and server-2 are μ and μ_2 respectively. If an arrival occurs and find:

1. Both servers free; it occupies server-I (assuming that server-1 gives faster service on average).
2. Server-1 is engaged; it waits for service from server-I whether or not server-II is free. But if the number of arrivals waiting for service from server-I becomes m (a positive integer), it goes to server-2 if that server is free; otherwise it waits as the $(m+1)^{th}$ arrival in the queue. Note that the first m arrivals in the queue will be getting service from server-1 and the $(m+1)^{th}$ arrival in the queue will go to server-2 if that server becomes free prior to the finishing of service of the arrival in server-1. Otherwise he will move up as the m^{th} arrival in the queue. Hence may decide to take service from server-1.
3. Both servers are engaged and a queue of length n greater than or equal to m is formed; It joins the queue as the $(n+1)^{th}$ arrival. All arrivals after the m^{th} arrival in the queue take a decision only when they reach the $(m+1)^{th}$ position in the queue. The decision is taken according to the rule mentioned in 2 of server-I engaged above.

The positive integer m is to be chosen such that it is one less than the greatest integer in the ratio $\frac{\mu}{\mu_2}$. It is clear for this choice of m that: When there are m arrivals[10] waiting for service from server-I, an incoming arrival finds it profitable to go to server-2 if that server is free since $(m+2)\mu^{-1} < \mu_2^{-1}$. Similarly, when there are only $(m-1)$ arrivals waiting for service from server-1, an incoming arrival will find it profitable to join the queue for service from server-1, even if server-2 is free, since $(m+1)\mu^{-1} < \mu_2^{-1}$. In case $\frac{\mu}{\mu_2^{-1}}$ is an integer, then $\frac{\mu}{\mu_2^{-1}}-1$, so that joining the queue for service from server-1 is not any more or any less profitable than going to server-2 if the server is free. But there is no harm in assuming that even in this case, the arrival joins the queue for service from server-1 when there are only $(m-1)$ arrivals waiting for service.

Thus, this queue discipline achieves the objective that the least amount of waiting time is spent in the system according to the conditions present upon its arrivals and also reduces the violation of first-in first-out principle. In section 3, we consider the discrete-time $Geo/Geo, G/2$ queue under the above queue discipline with $m = 1$ and subject to the following time constraints that:

1. The service time is divided into slots numbered as $0, 1, 2, ..., ..., \eta$ with each slot of unit length.
2. A potential arrival occurs in an interval $(\eta-, \eta)$ while a potential departure served by either server-1 or by server-2 occurs in the interval $(\eta, \eta+)$ as considered in Gupta & Goswami [9].
3. The state of the queue length process is defined at departure epoch $\eta+$ by two random variables $N(\eta+)$ denoting the number of arrivals in the system at time $\eta+$ and η is past service time of the arrival served by server-2 if any.

[10] We analyze the case for $m = 1$ customer.

The rest of the work is organized as follows; in section two, we derive the PGFs of the number of arrivals present in the system, the waiting time distribution and their mean values. Section 3 highlights the various special features of the proposed methodology on the analysis of two paralleled heterogeneous servers $Geo/Geo, G/2$ system which does not violate the FCFS principle. Section 4 gives a concluding report and in section five, we give a conjecture that takes the results of this work to another model of interesting application similar to the ones discussed in this work.

2 The Geo/G/2 Queue under 1.2

Here, we consider 'the embedded time points' generated at departure instants of arrivals just after a service completion either by server-1 or by server-2. Hence the sequence of system states observed at these embedded points with state representation $N_k = N(t_k)$ denoting the number of arrivals left behind in the queue by the k^{th} departing arrival at departure epoch t_k forms a Markov Chain $\{N_k\}$ with state space $S = 0, 1, ...$

2.1 PGF for the Number of Arrivals

Let q_j be the steady state probability of finding j arrivals in the system as observed by a departing arrival with a z-transform $V(z) = \sum_{j=0}^{j=\infty} q_j z^j$. Similarly, let α_j denotes the probability that j arrivals occur in a service completion period with probability mass function $f_{min}(k)$ and δ_j denotes the probability that j arrivals occur in a geometrically distributed service time with probability mass function $f_1(k)$. Since these arrivals come from a geometric process at a steady rate λ, then for $j = 1, 2, 3, ...$, we have

$$\alpha_j = \sum_{k=1}^{\infty} f_{min}(k) \binom{k}{j} \lambda^j (1-\lambda)^{k-j} \tag{6}$$

and

$$\delta_j = \sum_{k=1}^{\infty} f_1(k) \binom{k}{j} \lambda^j (1-\lambda)^{k-j} \tag{7}$$

Now, denote the respective z-transforms of the probability distributions $\{\alpha_j\}$ and $\{\delta_j\}$ by $A_{min}(z)$ and $A_1(z)$ respectively such that

$$A_{min}(z) = \sum_{j=0}^{\infty} \alpha_j z^j = F_{min}(1 - \lambda + \lambda z) \tag{8}$$

and

$$A_1(z) = \sum_{j=0}^{\infty} \delta_j z^j = F_1(1 - \lambda + \lambda z) \tag{9}$$

Focusing on the embedded points under equilibrium conditions, let the unit step conditional transition probability of the system going from state i of the $(k-1)^{st}$ embedded point to state j in the k^{th} embedded point be $q_{ij} = P(N_k = j/N_{k-1} = i); i, j \in S$. These transition probabilities will form a unit step transition probability matrix $Q = q_{ij}$ as below:

$$Q = \begin{pmatrix} \delta_0 & \delta_1 & \delta_2 & \delta_3 & . & . & . & . \\ \delta_0 & \delta_1 & \delta_2 & \delta_3 & . & . & . & . \\ 0 & \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & . & . & . \\ 0 & 0 & \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & . & . \\ 0 & 0 & 0 & \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & . \end{pmatrix} \tag{10}$$

If we denote by $q_j = \lim_{n\to\infty} q_{ij}^n$, the equilibrium state probabilities at departure instants where q_{ij}^n represents the n-step probability of moving from state i to j such that $Q = q_{ij}$ with $q = (q_0, q_1, q_2, ...)$ and $e = (1, 1, 1, ...)^{'}$,. Assuming that $A'_{min}(1) < 1$, then the stationary distribution of the state transition matrix Q exits and is given by the unique solution of the following system of equations:

$$\mathbf{q}Q = \mathbf{q}, \mathbf{qe} = 1 \tag{11}$$

Now, multiplying the j^{th} equation of $\mathbf{q}Q = \mathbf{q}$ in (11) by z^j and summing all the left-hand sides and the right-hand sides from $j = 0$ to $j = \infty$, we get the PGF $V_{min}(z)$ of the queue length distribution q_j of the sequence $\{N_k\}$ below[11]

$$V_{min}(z) = \frac{q_1 z[A_{min}(z) - A_1(z)] + q_0[A_{min}(z) - D]}{A_{min}(z) - z} \tag{12}$$

Since $q_1 = \rho q_0$, $A'_{min}(1) < 1$, $A'_1(1) = \rho$ and $V(1) = 1$, we derive from (12) that

$$V_{min}(z) = \frac{(z-1)A_1(z) + (1+\rho z)[A_1(z) - A_{min}(z)]}{z - A_{min}(z)} \tag{13}$$

And[12] the mean number $E[N]$ of arrivals present in the system at a random point or at a departure epoch of time is[13]

$$E[N] = \frac{\left(\rho - A'_{min}(1)\right)\left[\rho + A'_{min}(1)(1+\rho)\right]}{1 + (\rho - A'_{min}(1))(1+\rho)} + K \tag{14}$$

The[14] discrete time equivalent of PASTA (Poison Arrivals See Time Averages) is referred to as BASTA (Bernoulli Arrivals See Time Averages) or GASTA (Geometric Arrivals See Time Averages). Using this property, we envisage that the distribution $V_{min}(z)$ in (13) also hold for the number of arrivals in the system as seen by an arbitrary arrival entering the system. Now, if we denote the PGF of the waiting time W representing the number of slots for which an arrival stays in the system with distribution $W(k) = P(W = k)$ by $W(z)$ for $0 < z < 1$. Then replacing $1 - (\lambda - \lambda z)$ by z in (13), one obtains that

$$V_{min}(z) = W(1 - \lambda + \lambda z) \tag{15}$$

And the mean waiting time of an arrival in the system obtained from (15) satisfies the well-known Little's formula. For a numerical illustration, suppose that λ varies from 5 to 15 as in Table-1 below while μ=8.0, μ_2=7.5, the following numerical values are obtained:

Table-1: Mean queue Length $E(N)$ and Mean waiting Time $\bar{W}$

$\mu = 8.0$ and $\mu_2 = 7.5$

λ	ρ	ρ_1	q_0	$E[N]$	$\bar{W}$
5	0.63	0.32	0.45	1.01	0.20
8	1.00	0.52	0.25	2.05	0.26
10	1.25	0.65	0.15	3.12	0.31
11	1.38	0.71	0.11	3.90	0.35
12	1.50	0.77	0.08	5.04	0.42
13	1.63	0.84	0.05	6.97	0.54
14	1.75	0.90	0.03	11.2	0.80
15	1.88	0.97	0.01	32.0	2.14

From table-1, it can be seen that both the mean queue length E[N] and the mean waiting time $\bar{W}$ steadily increase with increase in λ. Also, the stationary q_0 values decrease with increase in λ as expected.

Lemma 2.1 *Suppose server-2 breaks down during an operational period. Then the $Geo/Geo{+}G/2 \to Geo/Geo/1$ queue.*

Proof Suppose that the load $\rho = \frac{\lambda}{\mu} < 1$. Denote by $Q_0(z)$ and $W_0(z)$ the PGFs of the queue length and the waiting time distributions of the $Geo/Geo/1$ system such that

$$Q_0(z) = \frac{(1-z)(1-\rho)F_1(1 - \lambda = \lambda z)}{F_1(1 - \lambda = \lambda z) - z} \tag{16}$$

[11] $D = zA_1(z)$

[12] Here, $q_0 = \frac{1 - A'_{min}(1)}{1 + (\rho - A'_{min}(1))(1+\rho)}$

[13] This measure is obtained after differentiating the above equation at z=1.

[14] $K = \rho + \frac{[A'_{min}(1)]^2}{1 - A'_{min}(1)}$

with mean

$$E[N] = \rho + \frac{\lambda^2}{2(1-\rho}E(S_1(S_1-1)) \tag{17}$$

Similarly, let

$$W_0(z) = \frac{(1-z)(1-\rho)F_1(z)}{(1-z)-\lambda(1-F_1(z))} \tag{18}$$

with mean

$$\bar{W}_0 = \frac{E[N]}{\lambda} \tag{19}$$

Now, if server-2 breaks down, then the service situation is equivalent to a long service time of mean service time $\beta = \infty$ units. Consequently, $F_2(z)$ cannot exist. Putting μ_2=0 in (4) and an onward simplification, one obtains that $A_{min}(z) = F_1(1-\lambda+\lambda z) = A_1(z)$. Finally, the lemma follows if $A_{min}(z)$ is replaced by $A_1(z)$ in the PGF of the queue length distribution $V_{min}(z)$ in (13).

3 The $Geo/G/2$ Queue under 1.4

We will now discuss the steady state analysis of the $Geo/Geo, G/2$ queue under the parallel queue discipline outlined above for $m = 1$ arrival. The analysis is carried out using the past service time of the arrival being served by server-2 as a supplementary variable.

Denote by N the steady state number of arrivals in the system and by ζ the steady state past service time of the current arrival on server-2. Looking at the system at departure instants, then the bi-variate process $\{N, \zeta\}$ is a Markov process with state space $S = 0, 1, 2, ... \times [0, \infty)$. Suppose that P is a probability measure such that

$$R_0 = P[Both\ servers\ are\ idle]$$

$$R_{1,0} = [Only\ Server - 1\ is\ busy;\ N = 1]$$
$$R_{0,1}(\eta) = P[Only\ Server - 2\ is\ busy;\ N = 1]$$
$$R_{1,1}(\eta) = P[Both\ Servers\ are\ busy;\ N = 2]$$

$$R_{1,1,0} = P[Only\ Server - 1\ is\ busy;\ N = 2]$$

and that[15]

$$R_{1,1,1}(\eta) = P[Both\ Servers\ are\ busy,\ N = 3]$$

Remark We assign η to R_j only when server-2 is busy. Given that two or more customers are present in the system and that their past service time lies in $(\eta, \eta + d\eta)$ then in steady state, $R_j(\eta) \to R_j$.

We give the steady state difference equations (56) through (66) in terms of the above probability measures that describe the queue length process in Appendix A.

[15] For any R_j when both servers are busy, the supplementary variable ζ is such that $\eta \leq \zeta < \eta + d\eta$.

3.1 Difference Operator

Let Δ denotes the forward type finite difference operator that is; $\Delta f(x) = f(x+1) - f(x)$ for a real valued function $f(x)$. Analogous to rules for finding the derivative, we have: If c is a constant, then $\Delta c = 0$. Similarly, for two constants a and b define $\Delta(af + bg) = a\Delta f + b\Delta g$ and $\Delta(fg) = f\Delta g + g\Delta f + \Delta g \Delta f$.[16]

3.1.1 The Stationary PGF P(z)

Let, for $j = (0,1), (1,1), (1,1,1), 4, 5, ,$

$$Q_j(\eta) = \frac{R_j(\eta)}{1-B(\eta)} \tag{20}$$

and that

$$\tilde{R}_j = \sum_{\eta=0}^{\infty} Q_j(\eta)\Delta B(\eta); \quad \tilde{Q}_j = \beta \tilde{R}_j \tag{21}$$

Similarly, let

$$Q_j^*(z) = \sum_{\eta=0}^{\infty} Q_j(\eta) z^\eta \tag{22}$$

Substituting the quantities defined by (20), (21) and (22) into the equations provided by (56) through (66) of Appendix A coupled with the application of the product rule for two functions specified above, one obtains the following set of steady state equations:

$$\lambda R_0 = \mu R_{1,0} + \frac{1}{\beta}\tilde{Q}_{0,1} \tag{23}$$

$$(\lambda+\mu)R_{1,0} = \lambda R_0 + \mu R_{1,1,0} + \frac{1}{\beta}\tilde{Q}_{1,1} \tag{24}$$

$$(\lambda+\mu)R_{1,1,0} = \lambda R_{1,0} + \frac{1}{\beta}\tilde{Q}_{1,1,1} \tag{25}$$

$$\Delta Q_{0,1}(\eta) = -\lambda Q_{0,1}(\eta) + \mu Q_{1,1}(\eta) \tag{26}$$

$$Q_{0,1}(0+) = 0 \tag{27}$$

$$\Delta Q_{1,1}(\eta) = -(\lambda+\mu)Q_{1,1}(\eta) + \lambda Q_{0,1}(\eta) + L \tag{28}$$

$$Q_{1,1}(0+) = 0 \tag{29}$$

$$\Delta Q_{1,1,1}(\eta) = -(\lambda+\mu)Q_{1,1,1}(\eta) + \lambda Q_{1,1}(\eta) + M \tag{30}$$

[16] The $\Delta f \Delta g$ term is ignored in queuing applications.

$$Q_{1,1,1}(0+) = \lambda R_{1,1,0} + \frac{1}{\beta}\tilde{Q}_4 \tag{31}$$

And[17] for j=4,5,..., we have[18]

$$\Delta Q_j(\eta) = -(\lambda+\mu)Q_j(\eta) + \lambda Q_{j-1}(\eta) + H \tag{32}$$

$$Q_j(0+) = \frac{1}{\beta}\tilde{Q}_{j+1} \tag{33}$$

To solve (23) through (33), multiply each $Q_j(\eta)$ by an appropriate $z^\eta : |z| < 1$ and summing as in (22). One obtains following difference equations[19]

$$\left(\frac{1}{z}-1\right) Q^*_{0,1}(z) + \lambda Q^*_{0,1}(z) = \mu Q^*_{1,1}(z) \tag{34}$$

$$\left(\frac{1}{z}-1\right) Q^*_{1,1}(z) + (\lambda+\mu) Q^*_{1,1}(z) = O(z) \tag{35}$$

$$\left(\frac{1}{z}-1\right) Q^*_{1,1,1}(z) + (\lambda+\mu) Q^*_{1,1,1}(z) = T(z) \tag{36}$$

$$\left(\frac{1}{z}-1\right) Q^*_{4}(z) + (\lambda+\mu) Q^*_{4}(z) = \mu Q^*_5(z) + K \tag{37}$$

Finally[20], for j=5,6,7,..., we have

$$\left(\frac{1}{z}-1\right) Q^*_{j}(z) + (\lambda+\mu) Q^*_{j}(z) = \mu Q^*_{j+1}(z) + U \tag{38}$$

Lemma 3.1 *Given that the traffic condition $\lambda < \mu + (\mu_2 = \frac{1}{\beta})$ holds, then in a busy period*

$$Q^*_j(1) = \beta R_j \tag{39}$$

Moreover, for j=(0,1), (1,1), (1,1,1),4,5, $\tilde{Q}_j = R_j$.

Proof Denote by $X(t) = j$ the queue length at time t during a busy period when both servers in the $Geo/Geo, G/2$ system are busy. Let $t_1, t_2, ...$ be the departure epochs of arrivals served by server-2 at which the process restarts from the scratch.[21] Then the same property regenerates at $t_2, t_3, ...$ such that the sequence of time epochs t_n forms a renewal process. Consequently, $\{X(t)\}$ is a regenerative process on the state space $\Omega = \{(0,1),(1,1),(1,1,1),4,5,...\}$. Now, given that $\lambda < \mu + (\mu_2 = \frac{1}{\beta})$ holds, then upon service completion on server-2, the state probability is

$$R_j(t) = P[X(t) = j];\ j = (0,1),(1,1),... \tag{40}$$

In addition, if the past service time is η units for the arrival being served by server-2 at a time t, then the conditional probability that there are j arrivals in the system is

$$R_j(t) = \sum_{\eta=0}^{\infty} P[X(t) = j | t_1 = \eta]\Delta B(\eta) \tag{41}$$

$$= \sum_{\eta=0}^{\infty} P[X(t) = j, t_1 > t | t_1 = \eta]\frac{\Delta B(\eta)}{1-B(\eta)} \tag{42}$$

$$= Q_j(t) + \sum_{\eta=0}^{t} P[X(t) = j | t_1 > t-\eta | t_1 = \eta]\Delta B(\eta) \tag{43}$$

$$= Q_j(t) + \sum_{\eta=0}^{t} R_j(t-\eta)\Delta B(\eta) \tag{44}$$

Which is the unique solution of the discrete time renewal equation

$$R_j(t) = Q_j(t) + \sum_{\eta=0}^{t} Q_j(t-\eta)\Delta M(\eta) \tag{45}$$

Where $M(t) = E[X(t)]$ is the renewal function of the renewal process with distribution function $B(t)$. Since $Q_j(t)$ tends to zero as $t \to \infty$. Application of the key renewal theorem gives

$$R_j = \lim_{t\to\infty} R_j(t) \to \frac{1}{\beta}\sum_{t=0}^{\infty} Q_j(t) = \frac{1}{\beta}Q^*(1) \tag{46}$$

Remark Note that one cannot obtain $\{R_j\}$ completely for all general service time distributions. In case the service time distribution $B(\bullet)$ has a constant hazard rate, then a compact expression for each member of the sequence of $\{R_j\}$ can be computed.[22]

Now, if lemma 3.1 is applied in (34) through (38) and then simplified as $z \to 1$, one can obtain a compact expression for each member of the sequence R_j. Those simplified results are reported in Appendix-B. A summarized version

[17] Here, $L = \mu Q_{1,1}(\eta)$ and $M = \mu Q_4(\eta)$.

[18] $H = \mu Q_{j+1}(\eta)$

[19] Here also, $O = \lambda Q^*_{0,1}(z) + \mu Q^*_{1,1,1}(z)$ and $T = \lambda Q^*_{1,1} + \mu Q^*_4(z) + \lambda R_{1,1,0} + \frac{1}{\beta}\tilde{Q}_4$.

[20] $K = \frac{1}{\beta}\tilde{Q}_5$ and $U = \frac{1}{\beta}\tilde{Q}_{j+1}$

[21] That is, there exits the first time epoch t_1 beyond which with probability one the process is a probabilistic replica of the whole process starting at t=0.

[22] Since $\frac{\Delta B(\eta)}{1-B(\eta)} \to \frac{1}{\beta}$ and $\sum_{\eta=0}^{\infty} Q_j(t)\Delta B(t) = \frac{1}{\beta}R_j$

is given below[23] for two real values $a = (\mu^2 + \lambda\mu + \lambda^2)$ and $b = (a + \lambda\mu + 2\lambda^2)$.

$$R_1 = \left(\frac{\lambda}{b}\right)\left[\frac{\lambda^2}{\mu_2} + \frac{a + \lambda^2 + \lambda\mu}{\mu}\right] R_0 \quad (47)$$

$$R_2 = \left(\frac{\lambda}{b}\right)\left[\frac{\lambda}{\mu}\right]\left(\frac{1}{\mu_2}\right)\left[\frac{a\mu_2}{\mu} + \lambda^2\right] R_0 \quad (48)$$

$$R_{1,1,1} = \left(\frac{\lambda}{b}\right)\left[\frac{\lambda^2}{\mu_2}\right]\left(\frac{\lambda}{\mu}\right)^2 R_0 \quad (49)$$

$$R_4 = \left[\frac{\lambda^3 - \mu_2 a}{\mu + \mu_2}\right]\left[\frac{\lambda^3}{\mu_2 b \mu^2}\right] R_0 \quad (50)$$

In general, for $j \geq 4$

$$R_j = \left[\ (\rho_1)^{(j-4)}\right]\left(\frac{\lambda^3 - \mu_2 a}{\mu + \mu_2}\right)\left[\frac{\lambda^3}{\mu_2 b \mu^2}\right] R_0 \quad (51)$$

$$\sum_{j=4}^{\infty} R_j = \left[\frac{\lambda^3 - \mu_2 a}{\mu + \mu_2}\right]\left[\frac{\lambda^3}{\mu_2 b \mu^2}\right]\frac{R_0}{(1 - \rho_1)} \quad (52)$$

Similarly, the generating function $P(z)$, the mean queue length $E[N]$ and the mean waiting time $\bar{W}$ are respectively given by

$$P(z) = \sum_{j=0}^{\infty} R_j z^j = R_0 + .. + \frac{R_4 z^4}{(1 - \rho_1 z)} \quad (53)$$

$$E(N) = R_1 + 2R_2 + 3R_3 + R_4\left[\frac{4 - 3\rho_1}{(1 - \rho_1)^2}\right] \quad (54)$$

$$\bar{W} = \frac{(R_1 + 2R_2 + 3R_3) + R_4\left[\frac{4-3\rho_1}{(1-\rho_1)^2}\right]}{\lambda} \quad (55)$$

Lemma 3.2 *Suppose* $\lambda = \mu$. *Then the underlying Markov chain* $\{N = j\}$ *is ergodic if and only if* $\mu > 3\mu_2$.

Proof Under the stability condition $\lambda < \mu + \mu_2$ i.e. $\rho_1 < 1$, the underlying Markov chain $\{N = j\}$ is ergodic if and only if each $P(N = j) = R_j$ is positive inclusive of $R_4 = \left[\frac{\lambda^3 - \mu_2 a}{\mu + \mu_2}\right]\left[\frac{\lambda^3}{\mu_2 b}\right] R_0$. This implies that $\frac{\lambda^3 - \mu_2 a}{\mu + \mu_2} > 0$. Now, given that $\lambda = \mu$, then the lemma holds.

Lemma 3.3 *The stationary distribution* $\{R_j = P(N = j)\}$ *of the system size of the* $Geo/Geo, G/2$ *queue exists if and only if* $(\lambda^3 > \mu_2\ a)$ *holds where* $a = \lambda^2 + \lambda\mu + \mu^2$

[23] Detailed discussions and parallel results on the continuous time version M/G/2 queue is available in Sivasamy, Daman & Sulaiman [15].

Proof Since R_4 is proportional to $\lambda^3 - \mu_2 a$ and is positive definite (being a probability value), it is trivial that the stationary distribution $\{R_j = P(N = j)\}$ of the system size of the $Geo/Geo, G/2$ queue exist if $(\lambda^3 > \mu_2\ a)$ holds. Conversely, suppose $(\lambda^3 > \mu_2\ a) > 0$. Then R_4 is positive definite and so it is proportional to $(\lambda^3 - \mu_2 a)$.

4 Numerical Approximations

For a comparative study on the mean number of customers $E[N]$ and the mean waiting times $\bar{W}$ of the two models namely; the $Geo/(Geo + G)/2$ and the $Geo/Geo, G/2$ queues, we suppose that λ varies from 10.1 *to* 10.6 while $\mu = 5.5$ and $\mu_2 = 5.2$. Table-2 below summarizes the approximate values for $E(N)$ and $\bar{W}$ for the two models studied here.[24]

Table-2: Mean Performance Analysis

λ	$E[N]_{ser.}$	$\bar{W}_{ser.}$	$E[N]_{par.}$	$\bar{W}_{par.}$
10.1	18.87	1.87	18.10	1.79
10.2	22.46	2.20	22.24	2.18
10.3	27.83	2.70	27.98	2.92
10.4	36.77	3.54	37.18	3.58
10.5	54.62	5.20	55.23	5.23
10.6	108.14	10.20	108.89	10.27

5 Discussions & Remarks

Note that under equilibrium conditions, there is an insignificant difference between

1. $E[N]_{Geo/Geo+G/2}$ and $E[N]_{Geo/Geo,G/2}$.
2. $\bar{W}_{Geo/Geo+G/2}$ and $\bar{W}_{Geo/Geo,G/2}$.

Thus, one can conclude that though, some violations of the 'first-come first-served (FCFS)' principle occurred because of heterogeneity of servers in Geo/G/2 queues generally as pointed out by Krishnamoorthi [11] , the two alternative queue disciplines here minimize such violations in the long run. This is because the steady state characteristics for the $Geo/Geo, G/2$ queue under the parallel queue discipline and that of the $Geo/Geo + G/2$ queue under serial queue discipline differ insignificantly as observed in table-2 above. Similarly, we infer from these results that, if the arrival rate is far away from the combined service rate, then it is operationally better to allocate an arrival to a server instead of joint service when

[24] The stability condition $\lambda < (\mu + \mu_2)$ holds since $(\mu + \mu_2) = 10.7 > 10.6 = \lambda_{max}$. Also, as in table-2, $ser.$ means serial and $par.$ means parallel.

another arrival is present. As can be seen in the table-2 above where in this case both the mean queue length and the waiting time of the model under the parallel queue discipline is stationary smaller than that of the model under serial service. However, the serial queue discipline is a better alternative especially in high-speed service systems with arrival rates approaching the combined service rates.

The results obtained in this work can be applied in service systems where customer distribution is required for better practice. Comparing the model with similar ones discussed in the literature, the ones here are simpler and results compact, hence a good alternative for use. Most importantly is the new result of our work that under the serial queue discipline applied on the two-serially connected servers as in the $Geo/Geo+G/2$ and the parallel queue discipline applied when the servers are in parallel as in the $Geo/Geo,G/2$ queue, the $Geo/Geo+G/2$ and $Geo/Geo,G/2$ models are identical if and only if $\lambda^3 > \mu_2(\mu^2+\lambda\mu+\lambda^2)$. This ensures that the associated Markov chain for the arrival distribution is ergodic.

There is a scope in studying the models discussed here via Markov-renewal theory as in Senthamaraikannan and Sivasamy [14].

The importance of control in the design and analysis of models of service systems cannot be over emphasized. For a detail discussion, see Mohammad & Ali [13]. Recently, Sulaiman & Daman [16] have presented an exact analysis of both the arrival distribution and waiting time expectation for a continuous time M/G/2 queuing system working under a control queue discipline.[25] The analysis shows that the $M/G/2$ queue with the embedded control performs better than the continuous time $M/G/2$ queue under both the serial and the parallel queue disciplines. Under similar assumptions employed for the model with control, we conjectured that

Conjecture 1 *Under heavy traffic conditions, the $Geo/G(control)/2$ model will perform better than the $Geo/G/2$ queues studied in this work.*

Acknowledgment: This work has been supported by the authorities of University of Botswana.

Appendices

Appendix A

If $R_j = P[N=j]$ is the probability that there are j arrivals in the system, then

$$R_0 = P[N=0] = R_{0,0}$$

$$R_1 = P[N=1] = R_{1,0} + R_{0,1}$$

$$R_2 = P[N=2] = R_{1,1,0} + R_{1,1}$$

$$R_3 = P[N=3] = R_{1,1,1}$$

$$R_j = P(N=j);\ j=4,\ j=5,....,$$

Thus when $\zeta > 0$, it can easily be verified that the steady state probability functions of the queue length distribution $\{R_j, \eta < \zeta < \eta + d\eta\}$ satisfy the below difference equations[26]

$$\lambda R_0 = \mu R_{1,0} + \sum_{\eta=0}^{\infty} R_{0,1}(\eta)\frac{\Delta B(\eta)}{1-B(\eta)} \tag{56}$$

$$(\lambda+\mu)R_{1,0} = I + \sum_{\eta=0}^{\infty} R_{1,1}(\eta)\frac{\Delta B(\eta)}{1-B(\eta)} \tag{57}$$

$$\Delta R_{0,1}(\eta) = -(\lambda + \frac{\Delta B(\eta)}{1-B(\eta)})R_{0,1}(\eta) + \mu R_{1,1}(\eta) \tag{58}$$

$$R_{0,1}(0+) = 0, \qquad j=1 \tag{59}$$

$$(\lambda+\mu)R_{1,1,0} = \lambda R_{1,0} + \sum_{\eta=0}^{\infty} R_{1,1,1}(\eta)\frac{\Delta B(\eta)}{1-B(\eta)} \tag{60}$$

$$\Delta R_{1,1}(\eta) = -\left(\lambda+\mu+\frac{dB(\eta)}{1-B(\eta)}\right)R_{1,1}(\eta) + Y \tag{61}$$

$$R_{1,1}(0+) = 0, \qquad j=2 \tag{62}$$

$$\Delta R_{1,1,1}(\eta) = -\left(\lambda+\mu+\frac{\Delta B(\eta)}{1-B(\eta)}\right)R_{1,1,1}(\eta) + Z \tag{63}$$

$$R_{1,1,1}(0+) = \lambda R_{1,1,0} + \sum_{\eta=0}^{\infty} R_4(\eta)\frac{\Delta B(\eta)}{1-B(\eta)} \tag{64}$$

For $j \geq 4$, we have[27]

$$\Delta R_j(\eta) = -\left(\lambda+\mu+\frac{\Delta B(\eta)}{1-B(\eta)}\right)R_j(\eta) + \theta \tag{65}$$

$$R_j(0+) = \sum_{\eta=0}^{\infty} R_{j+1}(\eta)\frac{\Delta B(\eta)}{1-B(\eta)} \tag{66}$$

Appendix B

For $a = (\mu^2+\lambda\mu+\lambda^2)$, $b = (a+\lambda\mu+2\lambda^2)$, $\rho = \frac{\lambda}{\mu}$ and $\rho_1 = \frac{\lambda}{\mu+\mu_2}$, we have

$$R_0 = R_0$$

[25] Details of this queue discipline is found in Sulaiman & Daman [16].

[26] Y$=\lambda R_{0,1}(\eta) + \mu R_{1,1,1}(\eta)$. Also, $Z = \lambda R_{1,1}(\eta) + \mu R_4(\eta)$. Finally, $\theta = \lambda R_{j-1}(\eta) + \mu R_{j+1}(\eta)$.

[27] In (57), $I = \lambda R_0 + \mu R_{1,1,0}$

$$R_{1,1} = \rho R_{0,1}$$

$$R_{1,0} = \left[\frac{\mu_2}{\lambda}\left(\frac{\lambda^2+\lambda\mu+a}{\lambda\mu}\right)\right] R_{0,1} \tag{67}$$

$$R_{1,1,0} = \left[\left[\frac{\mu_2}{\lambda}\right]\left(\frac{a}{\mu^2}\right)\right] R_{0,1} \tag{68}$$

$$R_{0,1} = \left[\frac{\lambda^3}{\mu_2[a+2\lambda^2+\lambda\mu]}\right] R_0 \tag{69}$$

$$R_2 = \left[\frac{\lambda^2}{\mu\mu_2[a+2\lambda^2+\lambda\mu]}\left(\frac{a}{\mu}+\lambda^2\right)\right] R_0 \tag{70}$$

$$R_{1,1,1} = \left(\frac{\lambda}{\mu}\right)^2\left[\frac{\lambda^3}{\mu_2[a+2\lambda^2+\lambda\mu]}\right] R_0 \tag{71}$$

$$R_4 = \left[\frac{\lambda^3-\mu_2 a}{\mu+\mu_2}\right]\left[\frac{\lambda^3}{\mu_2[a+2\lambda^2+\lambda\mu]}\right] R_0 \tag{72}$$

References

[1] Artalejo, J.R., & Lopez-Hererro, M.J., A Simulation study of a Discrete-Time Multiserver Retrial Queue with Finite Population, *Journal of Statistical Planning and Inference,* **137**, 2536-2542, 2007.

[2] Boxma, O.J., Deng, Q., & Zwart, A.P., Waiting time asymtotics for the M/G/2 queue with heterogeneous servers, *Queuing Systems,* **40**, 5-31, 2002.

[3] Briem, U., Theimer, T.H., & Kroner, H., A General Discrete-Time Queuing Model: Analysis and Applications, *International Teletraffic Congress ITC-13,* **14**, 13-19, 1991.

[4] Bruneel, H., Performance of Discrete-Time Queuing Systems, *Computers in Operations Research,* **20**(3), 303-320, 1993.

[5] Bruneel, H., & Kim, B.G., *Discrete time models for communication systems including ATM*, Kluwer Academic Publishers, Norwell, MA, 1999.

[6] Bruneel, H., Walraevens, J., Claeys, D., & Wittevrongel, S., Analysis of a Discrete-Time Queue with Geometrically Distributed Service Capacities, *LNCS 7314, Springer-Verlag-Berlin- Heidelberg,* 121135, 2012.

[7] Goswami, V., Analysis of discrete-time multiserver queue with balking, *International Journal of Management Science and Engineering Management,* **9**(1), 21-32, 2013.

[8] Goswami, V., & Mund, G.B., Multiserver bulk service discrete-time queue with finite buffer and renewal input, *Computers and Mathematics with Applications,* **57**, 1377-1388, 2009.

[9] Gupta, U.C., & Goswami, V., Performance analysis of finite buffer discrete-time queue with bulk service, *Computers & Operations Research,* **29**, 1331-1341, 2002.

[10] Ishizaki, F., Takine, T., & Hasegawa, T., Analysis of a Discrete-Time Queue with geometrically Distributed Gate Opening Intervals, *Journal of the Operations Research Society of Japan,* **39**(1), 1-24, 1996.

[11] Krishnamoorthi, B., On Poisson Queue with two Heterogeneous Servers, *Operations Research,* **2**(3), 321-330, 1963.

[12] Mirtchev, S., & Goleva, R., Discrete-Time Single server model with a Multimodal packet size distribution, *Proceedings of a Conjoint Seminar in Modeling and Control of Information Processes,* Sofia, Bulgaria, CTP, Sofia, 83-101, 2009.

[13] Mohammad, K. & Ali, M., A Hybrid Method for Solving Optimal Control Problems, *IAENG International Journal of Applied Mathematics,* **42**(2), 80-86, 2012.

[14] Senthamaraikannan, K., & Sivasamy, R., Embedded Processes of a Markov Renewal Bulk Service Queue, *Asia Pacific Journal of Operational Research,* **11**(1), 51-65, 1997.

[15] Sivasamy, R., Daman, O.A., & Sulaiman, S., An M/G/2 Queue where Customers are Served subject to a Minimum Violation of the FCFS Queue Discipline, *European Journal of Operational Research,* **240**, 140146, 2014.

[16] Sulaiman, S., & Daman, O.A., An M/G/2 Queue with Heterogeneous Servers working Under a Control Service Schedule: Stationary Performance Analysis. *IAENG International Journal of Applied Mathematics,* **45**(1), 21-30, 2015.

[17] Takagi, H., *Queuing Analysis: A Foundation of Performance Evaluation in Discrete-Time systems,* Volume 3, North-Holland Publishing, Amsterdam, 1993.