ON THE MAP/G/1 QUEUE WITH LEBESGUE–DOMINATED SERVICE TIME DISTRIBUTION AND LCFS PREEMPTIVE REPEAT SERVICE DISCIPLINE

Lothar Breuer

Department of Computer Science, University of Trier,
54286 Trier, Germany. email: breuer@info04.uni-trier.de

ABSTRACT

The present paper contains an analysis of the MAP/G/1 queue with last come first served (LCFS) preemptive repeat service discipline and Lebesgue–dominated service time distribution. The transient distribution is given in terms of a recursive formula. The stationary distribution as well as the stability condition are obtained by means of Markov renewal theory via a QBD representation of the embedded Markov chain at departures and arrivals.

Key Words: MAP, last come first served, preemptive repeat, QBD

1 INTRODUCTION

This short note gives simple solutions for the transient and stationary distributions of the MAP/G/1 queue with last come first served (LCFS) preemptive repeat service discipline and Lebesgue–dominated service time distribution. The characteristic feature which makes the queue under consideration especially tractable is that after an arrival expels the user in service from the service facility, the latter's remaining service time does not have any effect on the future of the queueing process and hence may be forgotten. This allows a model with a countable state space (namely the number of users in the system and the phase of the arrival process). The transient distribution is given in terms of a recursive formula. The stationary distribution as well as the stability condition are obtained by means of Markov renewal theory.

For a short account of applications which are modelled by the type of queue considered here, see He and Alfa [1]. Former research of queues which are similar to that under consideration here can further be found in Gaur [2, 3], Kumar et al. [4], Miyazawa [5], Pechinkin [6], Machihara [7, 8] as well as He and Alfa [9, 10].

The rest of this section contains basic definitions and notations. In section 2, the transient distributions of the queue are given as recursive formulae. Section 3 contains an analysis of the stationary regime by means of Markov renewal theory and a QBD representation of the queueing process at arrival and departure times.

Let G denote the distribution function of the service time and $G^c := 1 - G$ its complement. For Lebesgue–dominated service time distributions there is an integrable function g such that $G^c(t) = \int_t^\infty g(u)du$ for all $t \in I\!R_0^+$. It is well known (e.g. Davis [11], p.37) that in this case we further have a representation $G^c(t) = e^{-\int_0^t h(u)du}$ for all $t \in I\!R_0^+$ with a so–called hazard function $h(u)$ defined by $h(u) = g(u)/G^c(u)$ for all $u \in I\!R_0^+$.

Markovian Arrival Processes (MAPs) and Batch Markovian Arrival Processes (BMAPs) have been introduced by Neuts [12] and Lucantoni [13] in order to provide input streams for queueing systems which are Markovian (and hence analytically more tractable) on the one hand but very versatile (even dense in the class of point processes, see Asmussen and Koole [14]) on the other hand. For generalizations of BMAPs see Breuer [15]. A procedure for statistical model fitting of BMAPs is introduced in Breuer [16].

Let (D_0, D_1) denote the representation of the MAP that models the arrival stream of the queue. D_0 and D_1 are two matrices of dimension $d \times d$. An entry $D_{n;i,j}$ (with $n = 1$ or $i \neq j$) denotes the infinitesimal rate of a so–called phase change from i to j while observing n arrivals ($n = 0, 1$). The entries $D_{0;i,i} = -\sum_{j \neq i} D_{0;i,j} - \sum_{j=1}^d D_{1;i,j}$ are simply the negative sum of the other row entries in order to make $D = D_0 + D_1$ a generator matrix. The number of phases shall be denoted by $d \geq 1$.

The last come first served (LCFS) preemptive repeat service discipline means that whenever a new user enters the queue, it is admitted to the service facility immediately. If another user is in service at the arrival time (i.e. if its service has not been completed yet), then this user must leave the service facility and go back to the waiting facility. Once it reenters the service facility, the service is repeated completely, i.e. another whole service time is required for the user. Denote the queueing process (i.e. the two–dimensional process indicating the number of users in the system as well as the phase of the arrrival process) by $Q = (Q_t : t \in I\!R_0^+)$. Thus the state space of Q is $I\!N_0 \times \{1, \dots, d\}$, and $Q_t = (k, i)$ means that at time t there are k users in the system (including the service facility) and the arrival process is in phase i.

2 TRANSIENT DISTRIBUTION

Let $P_t((k,i),(m,j))$ denote the conditional probability of observing the arrival phase process being in state $j \in \{1, \dots, d\}$ and $m \in I\!N_0$ users in the system at time $t \in I\!R_0^+$ given that at time 0 the arrival phase process was in state $i \in \{1, \dots, d\}$, there were $k \in I\!N_0$ users in the system and, if $k > 0$, the service process had just begun. Define the matrices $P_t(k,m) := (P_t((k,i),(m,j)))_{1 \le i,j \le d}$ for all $t \in I\!R_0^+$ and $k, m \in I\!N_0$.

Let $(T_n : n \in I\!N_0)$ denote the sequence of jump (i.e. arrival or departure) times of the queue length process. These are stopping times and can be defined by $T_0 := 0$ and $T_n := \inf\{t > T_{n-1} : pr_1(Q_t) \neq pr_1(Q_{t-})\}$ for all $n \in I\!N$, with pr_1 denoting the projection mapping onto the first dimension (i.e. on the number of users in the system). Summing up over the number of jumps of the queue length process until time t, we have

$$P_t(k,m) = \sum_{n=0}^{\infty} P_t^{(n)}(k,m)$$

with

$$P_t^{(0)}(k,m) = \begin{cases} e^{D_0 t} & k = m = 0 \\ e^{\int_0^t (D_0 - h(s)I)ds} & k = m \ge 1 \\ 0 & \text{else} \end{cases} \tag{1}$$

and recursively

$$P_t^{(n+1)}(0,m) = \int_0^t e^{D_0 u} D_1 P_{t-u}^{(n)}(1,m)\, du \tag{2}$$

and

$$P_t^{(n+1)}(k,m) = \int_0^t e^{\int_0^u (D_0 - h(s)I)ds} \Big(D_1 P_{t-u}^{(n)}(k+1,m) + h(u)I \cdot P_{t-u}^{(n)}(k-1,m) \Big) \, du \quad (3)$$

for $k \geq 1$. Since the queueing process is skip–free with respect to the number of users in the system, the probability $P_t^{(n)}(k,m)$ will vanish for $n < |k-m|$.

Assume that the queueing process starts idle, i.e. that at time 0 there are no users in the system. Then the transient distribution $(P_t(k) : k \in I\!N_0)$ of the number of users in the system at time $t \in I\!R^+$ under the initial phase distribution p_0 is given by

$$P_t(k) = p_0 P_t(0,k) 1_d$$

with p_0 being a row vector and 1_d denoting a d–dimensional column vector with all entries being 1.

The computation of the transition probability matrices $P_t(k,m)$ is of the same complexity as the computation of a Markov renewal function (cf. Çinlar [17], pp.317ff.). The following iteration is similar to the one given in Bellman [18], p.168: Write $P(t;k,m) := P_t(k,m)$, including t as a variable. Define the operator f by $f(P^{(n)}) := P^{(n+1)}$. Then the transition probability kernel P can be computed as the limit

$$P = \lim_{n \to \infty} P_n$$

with

$$P_0 = P^{(0)} \qquad \text{and} \qquad P_{n+1} = f(P_n) + P_0$$

for all $n \in I\!N_0$. Here, P_n can be interpreted as the transition probability matrix with paths restricted to at most n jumps of the queue length process until time t. In comparison, $P^{(n)}$ would be the transition probability matrix with paths restricted to exactly n jumps of the queue length process until time t.

More explicitly, in order to derive the block matrix $P_t = (P_t(k,m))_{k,m \in I\!N_0}$ we first identify $P_0(u;k,m) := P_u^{(0)}(k,m)$ for all $u \in [0,t]$ and $k,m \in I\!N_0$ according to equation (1). Then we determine $P_1(u;k,m) = P_u^{(1)}(k,m) + P_0(u;k,m)$ for all $u \in [0,t]$ and $k,m \in I\!N_0$ according to equations (2) and (3). Having obtained the values of the matrix function P_n in such a way, we can iteratively determine the values of P_{n+1} by

$$P_{n+1}(u;0,m) = \int_0^t e^{D_0 u} D_1 P_n(t-u;1,m) \, du + P_0(u;k,m)$$

for all $u \in [0,t]$ and $m \in I\!N_0$ according to equation (2) and similarly the values $P_{n+1}(u;k,m)$ for all $u \in [0,t]$, $k \in I\!N$ and $m \in I\!N_0$ according to equation (3). The iteration can be stopped as soon as

$$\inf_{k \in I\!N_0, 1 \le i \le d} \sum_{m \in I\!N_0, 1 \le j \le d} P_n(t; (k,i), (m,j)) > 1 - \varepsilon$$

i.e. when the lower bound of the probability mass reaches some threshold value $1 - \varepsilon$ with small $\varepsilon > 0$.

3 STATIONARY DISTRIBUTION

The queue can be modelled as a Markov renewal process (see Çinlar [17], pp.313ff) in the following way. Let $(T_n : n \in I\!N_0)$ denote the sequence of jump times of $pr_1(Q)$ as defined in section 2. Further define $X = (X_n : n \in I\!N_0)$ by $X_n := Q_{T_n}$ for all $n \in I\!N_0$. Then (X,T) is a Markov renewal process. The transition probabilities for the Markov chain X shall be denoted by $\tilde{P}((k,i),(m,j))$, indicating the conditional probability that $X_1 = (m,j)$ given that $X_0 = (k,i)$. Define further the matrices $\tilde{P}(k,m) = (\tilde{P}((k,i),(m,j)))_{1 \le i,j \le d}$ for all $k,m \in I\!N_0$. Then we can write

$$\tilde{P}(k,m) = \begin{cases} B = -D_0^{-1} D_1 & (k,m) = (0,1) \\ A_0 = \int_0^\infty e^{D_0 t} D_1 e^{-\int_0^t h(s)ds} dt & k = m-1 \ge 1 \\ A_2 = \int_0^\infty e^{-\int_0^t h(s)ds} h(t) e^{D_0 t} dt & k = m+1 \ge 1 \end{cases}$$

for all $k,m \in I\!N_0$. The distributions of the Markov renewal intervals $T_n - T_{n-1}$ are given as

$$\begin{aligned} F_{km}(t) = F_k(t) &= P(T_n - T_{n-1} > t | X_{n-1} = k) \\ &= \begin{cases} e^{D_0 t} & k = 0 \\ e^{\int_0^t (D_0 - h(s)I)ds} & k \ge 1 \end{cases} \end{aligned}$$

independently of $m \in I\!N_0$. Since there are only finitely many phases and the service time is strictly positive with probability 1, this implies $\lim_{n \to \infty} T_n = \infty$ almost certainly.

The transition probability matrix

$$\tilde{P} = \begin{pmatrix} 0 & B & 0 & 0 & \dots \\ A_2 & 0 & A_0 & 0 & \ddots \\ 0 & A_2 & 0 & A_0 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

represents a discrete time QBD process, which can be solved by standard methods (see Latouche, Ramaswami [19, 20]). Define $A := A_0 + A_2$. By assumption, the phase process is irreducible and independent of the service. Hence, the transition matrix A is irreducible and has a unique invariant measure α. The chain X is ergodic iff $2\alpha A_0 1_d < 1$. Assume that this is satisfied. Denote the stationary distribution of X by π: π is a row vector, indexed as $\pi = (\pi(k, j) : k \in I\!N_0, j \in \{1, \dots, d\})$, and satisfying $\pi = \pi\tilde{P}$.

The queueing process Q is the minimal semi–Markov process associated with (X, T), i.e. $Q_t = X_n$ for $T_n \leq t < T_{n+1}$ (see Çinlar [17], pp.337ff). Let $P_t^Q(k, j)$ denote the probability of Q being in state (k, j) at time t, i.e. of observing k users in the system and the arrival phase being j. Then the stationary distribution ν of Q is independent of the initial distribution and given by

$$\nu(k, j) := \lim_{t\to\infty} P_t^Q(k, j) = \frac{\pi(k, j) \cdot \sigma(k, j)}{\pi\sigma}$$

with $\sigma = (\sigma(k, j) : k \in I\!N_0, j \in \{1, \dots, d\})$ denoting the column vector of mean sojourn times, i.e. $\sigma(k, j) = E_{(k,j)}(T_1)$ with $E_{(k,j)}$ denoting the conditional expectation under initial state $(k, j) \in I\!N_0 \times \{1, \dots, d\}$.

The vector σ can be determined in a straightforward manner as

$$\sigma(0, j) = \int_0^\infty e_j^T e^{D_0 t} 1_d \, dt = -e_j^T D_0^{-1} 1_d$$

and

$$\sigma(k, j) = \int_0^\infty e_j^T e^{\int_0^t (D_0 - h(s)I)ds} 1_d \, dt$$

for $k \geq 1$ and $j \in \{1, \dots, d\}$, with e_j^T denoting the jth canonical row base vector.

ACKNOWLEDGEMENT

The author wishes to thank the referees for many helpful comments which contributed to an improvement of the first version of the paper.

References

[1] Qi-Ming He and Attahiru Sule Alfa. The $MMAP[K]/PH[K]/1$ queues with a last-come-first-served preemptive service discipline. *Queueing Syst.*, 29(2-4):269–291, 1998.

[2] K.N. Gaur. Preemptive repeat-different priority queue. *Pure Appl. Math. Sci.*, 3:73–80, 1976.

[3] K.N. Gaur. Time-independent solution of the preemptive-repeat priority queue with bulk arrivals. *Pure Appl. Math. Sci.*, 9:43–51, 1979.

[4] Vijay Kumar, T.P. Singh, and Mohd Iqbal. k-phase service mechanism under pre-emptive priority service rule. *Pure Appl. Math. Sci.*, 17:41–48, 1983.

[5] Masakiyo Miyazawa. On the system queue length distributions of LCFS-P queues with arbitrary acceptance and restarting policies. *J. Appl. Probab.*, 29(2):430–440, 1992.

[6] A.V. Pechinkin. Service characteristic in the system $GI/GI/1/\infty$ with a last come, first serve preemptive repeat priority discipline. *Autom. Remote Control*, 54:978–988, 1993.

[7] Fumiaki Machihara. A preemptive priority queue as a model with server vacations. *J. Oper. Res. Soc. Japan*, 39(1):118–131, 1996.

[8] Fumiaki Machihara. A preemptive priority queue with nonrenewal inputs. In Srinivas R. Chakravarthy and Attahiru S. Alfa, editors, *Matrix-analytic methods in stochastic models*, pages 21–42. Marcel Dekker, New York, NY, 1997.

[9] Qi-Ming He and Attahiru Sule Alfa. Computational analysis of $MMAP[K]/PH[K]/1$ queues with a mixed FCFS and LCFS service discipline. *Nav. Res. Logist.*, 47(5):399–421, 2000.

[10] Qi-Ming He and Attahiru Sule Alfa. The Discrete Time $MMAP[K]/PH[K]/1/LCFS-GPR$ Queue and Its Variants. In Guy Latouche and Peter Taylor, editors, *Advances in Algorithmic Methods for Stochastic Models*, pages 167–190, New Jersey, USA, 2000. Notable Publications.

[11] M.H.A. Davis. *Markov models and optimization.* London: Chapman & Hall, 1993.

[12] Marcel F. Neuts. A versatile Markovian point process. *J. Appl. Probab.*, 16:764–774, 1979.

[13] David M. Lucantoni. New results on the single server queue with a batch Markovian arrival process. *Commun. Stat., Stochastic Models*, 7(1):1–46, 1991.

[14] Soeren Asmussen and Ger Koole. Marked point processes as limits of Markovian arrival streams. *J. Appl. Probab.*, 30(2):365–372, 1993.

[15] Lothar Breuer. On Markov–Additive Jump Processes. *Queueing Systems*, 40(1):75–91, 2002.

[16] Lothar Breuer. An EM Algorithm for Batch Markovian Arrival Processes and its Comparison to a Simpler Estimation Procedure. *Annals of Operations Research*, 112, 2002. to appear.

[17] Erhan Çinlar. *Introduction to stochastic processes.* Englewood Cliffs, N. J.: Prentice-Hall, 1975.

[18] Richard Bellman. *Introduction to matrix analysis*. Philadelphia, PA: SIAM, 1997.

[19] Guy Latouche and V. Ramaswami. A logarithmic reduction algorithm for quasi-birth-death processes. *J. Appl. Probab.*, 30(3):650–674, 1993.

[20] G. Latouche and V. Ramaswami. *Introduction to matrix analytic methods in stochastic modeling.* Philadelphia, PA: SIAM, 1999.