J. Korean Math. Soc. **42** (2005), No. 3, pp. 517–527

DISCRETE-TIME QUEUE WITH VARIABLE SERVICE CAPACITY

Yutae Lee

Abstract. This paper considers a discrete-time queueing system with variable service capacity. Using the supplementary variable method and the generating function technique, we compute the joint probability distribution of queue length and remaining service time at an arbitrary slot boundary, and also compute the distribution of the queue length at a departure time.

1. Introduction

Recently, interests in the discrete-time queues have increased due to their numerous applications in the analysis of telecommunications systems and other related areas [12, 18, 19]. One of the reasons for this is that discrete-time queues fit the slotted nature of telecommunications systems better than the continuous-time counterparts, and hence they give more accurate performance measures of these systems[1, 2, 5, 16, 17, 22].

Bulk-service models are useful to investigate the performance of various telecommunications systems. Besides applications in telecommunications systems, bulk-service queues are also used in various other areas such as manufacturing, production, transportation, and other stochastic systems [3]. It may happen that the server has a fixed maximum capacity, or else the server may take customers according to variable service capacity. Such systems may serve as a model for a shuttle or automatic elevator.

Let us review some related papers. In the existing literature, there have been a number of contributions with respect to bulk-service queues.

Received December 5, 2003.

2000 Mathematics Subject Classification: 60K25, 90B22.

Key words and phrases: discrete-time queue, variable service capacity, moment, bound, queue length, performance analysis.

This research was supported by University IT Research Center Project.

Most of the investigations which employed bulk-service models have considered continuous-time queues. The readers are referred to Chaudhry and Templeton [4], Dshalalow [9] and the references therein for a detailed account of such models. This paper considers a discrete-time bulk-service queueing system. Recently, Dümmler and Schömig [10] considered $Geo/G^{a,b}/1$ queueing systems with $a = 1, 2, \cdots, b$, where a is the threshold value of activating the server and b is the service capacity. Gupta and Goswami [11] discussed analytic and computational aspects of $Geo/G^{a,b}/1/N$ queue. Chaudhry and Chang [3] considered a $Geo/G^{1,Y}/1/N$ queue with variable service capacity. They discussed the computational aspects of the distributions of the customers in the queue at various epochs.

This paper considers a discrete-time $Geo^X/G^Y/1$ queue with variable service capacity. Using an invariant relation, Kim et al. [13] also studied, but they just obtained the probability generating function of the stationary queue length at a random point on the continuous-time domain. In this paper, using the supplementary variable method and the generating function technique, we derive the joint distribution of queue length and remaining service time at an arbitrary slot boundary, and also compute the probability distribution of the queue length at a departure time. We calculate the mean queue length at an arbitrary slot boundary and a departure time, and provide their bounds.

2. Model

This paper considers a discrete-time single-server queue in which the time axis is divided into fixed-length contiguous intervals, referred to as slots. It is always assumed that service times can be started and completed only at slot boundaries and that their durations are integral multiples of slot durations [15]. Customers arrive to the system in accordance with a batch geometric process [6] and are accommodated in the buffer with infinite waiting-room. Let a_k be the number of customers that arrive during slot k. The numbers of customers entering the system during the consecutive slots are assumed to be i.i.d. non-negative discrete random variables with an arbitrary probability distribution, and are characterized by the probability generating function $A(z) \equiv E[z^{a_k}]$ with finite mean. The service times s of batches are assumed to constitute a set of i.i.d. positive random variables with a general discrete distribution, and are characterized by the probability generating function $S(z)$ with finite mean. It is also assumed that the service times and the arrival process are mutually independent. The customers are served in batches of variable capacity, the maximum service capacity being a finite integer value C. If c_n is the service capacity for the nth batch service, then the server can serve up to c_n customers at the beginning of the nth batch service. We define

$$u_j \equiv P\{c_n = C - j\}$$

and

$$U(y) \equiv \sum_{j=0}^{C} u_j y^j.$$

It is assumed that the service capacities are independent of the service times and the arrival process.

In this paper, even if there is no one in system to serve and/or the service capacity is zero, and hence the server can serve no customers, the server may start the service. It is also assumed that arriving customer cannot be accepted into the batch already undergoing service even if the capacity of server is available, but has to wait until the next service instant.

3. Queue length distribution

In this section, the queue lengths at the end of slots are analyzed. Before proceeding to the analysis of the queue length, we define some random variables. Let a random variable n_k indicate the number of customers in the queue at the end of slot k. A supplementary random variable r_k indicates the remaining service time at the end of slot k. Then $\{(n_k, r_k), k \geq 0\}$ constitutes a two dimensional Markov chain embedded at the end of each slot. If we denote by a_k the number of customers entering the system during slot k, then the system under consideration evolves as follows:

(a) If $r_k = 0$, then

$$\begin{aligned} n_{k+1} &= (n_k - C + j)^+ + a_{k+1} \text{ with probability } u_j, \quad j = 0, 1, \cdots, C, \\ r_{k+1} &= s - 1; \end{aligned}$$

(b) If $r_k > 0$, then

$$\begin{aligned} n_{k+1} &= n_k + a_{k+1}, \\ r_{k+1} &= r_k - 1. \end{aligned}$$

Let $P_k(x,y) \equiv E\left[x^{n_k} y^{r_k}\right]$. From the above state equations we have

$$\begin{aligned}
& P_{k+1}(x,y) \\
\equiv\ & E\left[x^{n_{k+1}} y^{r_{k+1}}\right] \\
=\ & E\left[I_{\{r_k=0\}} x^{n_{k+1}} y^{s-1} + I_{\{r_k>0\}} x^{n_k+a_{k+1}} y^{r_k-1}\right] \\
=\ & \frac{A(x)}{y}\left[P_k(x,y) - P_k(x,0)\right] + \frac{A(x)S(y)}{y} \sum_{j=0}^{C} u_j \left[x^{j-C} P_k(x,0)\right. \\
& \left. + E\left[I_{\{r_k=0, n_k<C-j\}}\right] - x^{j-C} E\left[I_{\{r_k=0,n_k<C-j\}} x^{n_k}\right]\right] \\
=\ & \frac{A(x)}{y}\left[P_k(x,y) - P_k(x,0)\right] \\
& + \frac{A(x)S(y)}{y}\left[\frac{U(x)P_k(x,0)}{x^C} + Q(1) - \frac{Q(x)}{x^C}\right] \\
=\ & \frac{A(x)}{y}\left[P_k(x,y) - \frac{x^C - U(x)S(y)}{x^C} P_k(x,0)\right. \\
& \left. + \frac{S(y)}{x^C}\left\{x^C Q(1) - Q(x)\right\}\right], \qquad (1)
\end{aligned}$$

where the function $Q_k(x)$ is defined as

$$\begin{aligned}
Q_k(x) &\equiv \sum_{j=0}^{C-1} u_j x^j E\left[I_{\{r_k=0,n_k<C-j\}} x^{n_k}\right] \\
&= \sum_{c=0}^{C-1} P\{r_k = 0, n_k = c\} \sum_{j=0}^{C-c-1} u_j x^{j+c}.
\end{aligned}$$

Now, we will find the ergodic condition of the Markov chain $\{(n_k, r_k)\}$. We assume that the Markov chain $\{(n_k, r_k)\}$ is irreducible, which is not a strong assumption. Note that this is true if $P(a_k = 1) > 0$. Obviously, $S'(1)A'(1) < C - U'(1)$ is the necessary condition for the Markov chain $\{(n_k, r_k)\}$ to be positive recurrent. Note that the average service capacity is $C - U'(1)$, while during a service time $S'(1)A'(1)$ customers will arrive on average. For the sufficient condition, Foster's criteria (see Lamperti [14]) is used. Appendix shows that $S'(1)A'(1) < C - U'(1)$ is also the sufficient condition for the Markov chain $\{(n_k, r_k)\}$ to be positive recurrent. Thus, assuming that the Markov chain $\{(n_k, r_k)\}$ is irreducible, the Markov chain is ergodic if and only if $S'(1)A'(1) < C - U'(1)$.

Assume that the Markov chain $\{(n_k, r_k)\}$ is ergodic. Then, there exists unique stationary distribution. Let

$$P(x,y) \equiv \lim_{k\to\infty} P_k(x,y)$$

denote the the stationary joint probability generating function of the Markov chain $\{(n_k, r_k)\}$. Letting $k \to \infty$ in (1) we obtain

$$\begin{aligned}(2) \qquad & [y - A(x)]\, P(x,y) \\ = \; & \frac{A(x)}{x^C}\left[S(y)\left\{x^C Q(1) - Q(x)\right\} - \left\{x^C - U(x)S(y)\right\} P(x,0)\right],\end{aligned}$$

where

$$Q(x) \equiv \lim_{k\to\infty} Q_k(x).$$

The left-hand side of (2) becomes zero at $y = A(x)$, at which the right-hand side must also be zero. Thus, if we choose $y = A(x)$ for $|x| \le 1$ in (2), then we determine $P(x,0)$ as

$$P(x,0) = \frac{S(A(x))\left[x^C Q(1) - Q(x)\right]}{x^C - U(x)S(A(x))}. \tag{3}$$

Equation (3) is of indeterminate form, that is, (3) has unknown term $Q(x)$, but the C unknowns

$$\lim_{k\to\infty} P\{r_k = 0, n_k = c\}, \quad c = 0, 1, \cdots, C-1,$$

can be determined by consideration of the zeros of the denominator in (3) that lie in the closed unit disk $\{x : |x| \le 1\}$. With Rouche's theorem, it can be shown that the equation $x^C - U(x)S(A(x)) = 0$ has exactly C zeros in the unit closed disc $\{x : |x| \le 1\}$. A detailed explanation can be found in Saaty[21]. Since $P(x,0)$ is a continuous function for $|x| \le 1$, the numerator $S(A(x))(x^C Q(1) - Q(x))$ of $P(x,0)$ should vanish at each of the zeros, yielding C equations. One of the zeros equals 1, and leads to a trivial equation[8]. However, the relation $P(1,0) = 1/S'(1)$ provides an additional equation. Using l'Hopital's rule, this relation is found to be

$$C - U'(1) - S'(1)A'(1) = S'(1)\left[CQ(1) - Q'(1)\right]. \tag{4}$$

The C roots of $x^C - U(x)S(A(x)) = 0$ in the closed unit disk $\{x : |x| \le 1\}$ are denoted by $x_0 = 1,\ x_1,\ \cdots,\ x_{C-1}$. If one of the roots is zero of $S(A(x))$, then it should be 0, which can not be true. Thus, the C

roots are zeros of $x^C Q(1) - Q(x)$, not of $S(A(x))$. Hence, by writing $x^C Q(1) - Q(x)$ in (3) as

$$(x-1)F\prod_{c=1}^{C-1}(x-x_c)$$

with F a constant, and using (4) to derive the value of F, it follows that

$$F = \frac{C - U'(1) - S'(1)A'(1)}{S'(1)\prod_{c=1}^{C-1}(1-x_c)}$$

and

$$x^C Q(1) - Q(x) = \frac{C - U'(1) - S'(1)A'(1)}{S'(1)}(x-1)\prod_{c=1}^{C-1}\frac{x-x_c}{1-x_c}, \tag{5}$$

so that (3) can be written as

$$P(x,0) = \frac{[C - U'(1) - S'(1)A'(1)]\,S(A(x))}{S'(1)\,[x^C - U(x)S(A(x))]}(x-1)\prod_{c=1}^{C-1}\frac{x-x_c}{1-x_c} \tag{6}$$

for $|x| \le 1$.

Therefore, using (5) and (6) in (2), we can determine

$$\begin{aligned} P(x,y) &= \frac{[C - U'(1) - S'(1)A'(1)]\,A(x)}{S'(1)\,[y - A(x)]}\cdot\frac{S(y) - S(A(x))}{x^C - U(x)S(A(x))} \\ &\quad \times (x-1)\prod_{c=1}^{C-1}\frac{x-x_c}{1-x_c}. \end{aligned} \tag{7}$$

Let $N(x)$ be the probability generating function of the queue length at an arbitrary slot boundary. Clearly, $N(x)$ is given by

$$\begin{aligned} N(x) &\equiv \lim_{k\to\infty} E\left[x^{n_k}\right] \\ &= P(x,1) \\ &= \frac{[C - U'(1) - S'(1)A'(1)]\,A(x)}{S'(1)\,[1 - A(x)]}\cdot\frac{1 - S(A(x))}{x^C - U(x)S(A(x))} \\ &\quad \times (x-1)\prod_{c=1}^{C-1}\frac{x-x_c}{1-x_c}. \end{aligned} \tag{8}$$

Let $D(x)$ be the probability generating function of the queue length at a departure time. The probability generating function $D(x)$ can be expressed by

$$D(x) = \frac{P(x,0)}{P(1,0)}$$

$$(9) \qquad = \frac{[C - U'(1) - S'(1)A'(1)]\, S(A(x))}{x^C - U(x)S(A(x))}(x-1)\prod_{c=1}^{C-1}\frac{x - x_c}{1 - x_c}.$$

The probability generating functions $N(x)$ and $D(x)$ are related by

$$N(x) = D(x) \cdot \frac{1}{S(A(x))} \cdot \frac{A(x)\,[1 - S(A(x))]}{S'(1)\,[1 - A(x)]},$$

where the denominator $S(A(x))$ of the second term represents the probability generating function of the number of customers that arrive during a service time, and the third term

$$\frac{A(x)\,[1 - S(A(x))]}{S'(1)\,[1 - A(x)]}$$

represents the probability generating function of the number of customers that arrive during an elapsed service time.

4. Moments

Now we can calculate the mean queue lengths at an arbitrary slot boundary and a departure time by using the differentiation of the respective probability generating functions $N(x)$ and $D(x)$ for $x = 1$. We obtain

$$\begin{aligned}
\mu_N &\equiv N'(1) \\
&= \frac{S''(1)\,[A'(1)]^2 + S'(1)A''(1) + S'(1)A'(1) - 2\,[S'(1)A'(1)]^2}{2\,[C - U'(1) - S'(1)A'(1)]} \\
&\quad + \frac{U''(1) + U'(1) + C\,[S'(1)A'(1) - U'(1)]}{2\,[C - U'(1) - S'(1)A'(1)]} - S'(1)A'(1) \\
(10) &\quad + \frac{A'(1)\,[S''(1) + S'(1)]}{2S'(1)} - \frac{1}{2}(C-1) + \sum_{c=1}^{C-1}\frac{1}{1 - x_c}
\end{aligned}$$

and

$$\mu_D \equiv D'(1)$$
$$= \frac{S''(1)\,[A'(1)]^2 + S'(1)A''(1) + S'(1)A'(1) - 2\,[S'(1)A'(1)]^2}{2\,[C - U'(1) - S'(1)A'(1)]}$$
$$+\frac{U''(1) + U'(1) + C\,[S'(1)A'(1) - U'(1)]}{2\,[C - U'(1) - S'(1)A'(1)]}$$
$$-\frac{1}{2}(C-1) + \sum_{c=1}^{C-1}\frac{1}{1-x_c}. \tag{11}$$

The expressions for the higher-order moments can be derived as well from the appropriate derivatives of the respective generating functions.

In this section, we are also interested in bounding μ_N and μ_D. In [8] the bound

$$\frac{1}{2}(C-1) \leq \sum_{c=1}^{C-1}\frac{1}{1-x_c} \leq \frac{1}{2}(C-1) + \frac{1}{2}(U'(1) + S'(1)A'(1), C-1)^-$$

has been shown to hold for

$$\sum_{c=1}^{C-1}\frac{1}{1-x_c},$$

where $(a,b)^-$ denotes the minimum of a and b. Hence,

$$B_N \leq \mu_N \leq B_N + \frac{1}{2}(U'(1) + S'(1)A'(1), C-1)^- \tag{12}$$

and

$$B_D \leq \mu_D \leq B_D + \frac{1}{2}(U'(1) + S'(1)A'(1), C-1)^-, \tag{13}$$

where

$$B_N \equiv \frac{S''(1)\,[A'(1)]^2 + S'(1)A''(1) + S'(1)A'(1) - 2\,[S'(1)A'(1)]^2}{2\,[C - U'(1) - S'(1)A'(1)]}$$
$$+\frac{U''(1) + U'(1) + C\,[S'(1)A'(1) - U'(1)]}{2\,[C - U'(1) - S'(1)A'(1)]}$$
$$-S'(1)A'(1) + \frac{A'(1)\,[S''(1) + S'(1)]}{2S'(1)}$$

and

$$B_D \equiv B_N + S'(1)A'(1) - \frac{A'(1)\,[S''(1) + S'(1)]}{2S'(1)}.$$

These bounds provide the approximation of the mean queue length at an arbitrary slot boundary and a departure time. The approximation does not require a complex numerical procedure because the bounds are represented by the first two moments of the distributions for the arrival-bulk size, the service capacity and the service time.

5. Concluding remarks

This paper considered a discrete-time queue with variable service capacity. By means of the probability generating functions and the supplementary variable method, we provided an analysis of the joint distribution for the queue length and the remaining service time at an arbitrary slot boundary, and also computed the distribution of the queue length at a departure time. We calculate the mean queue length at an arbitrary slot boundary and a departure time, and provide their bounds.

6. Appendix(Sufficient condition for $\{(n_k, r_k)\}$ to be positive recurrent)

We intend to show that $S'(1)A'(1) < C - U'(1)$ is the sufficient condition for the Markov chain $\{(n_k, r_k)\}$ to be positive recurrent. Now, we assume that $S'(1)A'(1) < C - U'(1)$. Foster's criteria is used in order to show that the Markov chain $\{(n_k, r_k)\}$ is positive recurrent. We choose a real number α such that $[S'(1) - 1]/[C - U'(1) - A'(1)] < \alpha < 1/A'(1)$ and choose the test function as follows: $f(i,j) \equiv \alpha i + j$. The the mean drift of the test function is

$$
\begin{aligned}
z_{ij} &\equiv E\left[f(n_{k+1}, r_{k+1}) - f(n_k, r_k) \mid (n_k, r_k) = (i,j)\right] \\
&= \begin{cases} \alpha A'(1) - 1 + S'(1) - \alpha i + \alpha \sum\limits_{k=C-i}^{C} (i - C + k) u_k, & \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } 0 \le i \le C-1, j = 0, & \\ \alpha A'(1) - 1 + S'(1) - \alpha[C - U'(1)], \ \text{if } i \ge C, j = 0, & \\ \alpha A'(1) - 1, \qquad\qquad\qquad\qquad \text{if } j >= 1. & \end{cases}
\end{aligned}
$$

Then, $z_{ij} \le \infty$ for all i and j, and

$$
z_{ij} \le \min\left(\alpha A'(1) - 1 + S'(1) - \alpha[C - U'(1)], \alpha A'(1) - 1\right) < 0
$$

for $i \ge C$ and/or $j \ge 1$. Let

$$
\epsilon = \min\left(\alpha A'(1) - 1 + S'(1) - \alpha[C - U'(1)], \alpha A'(1) - 1\right)/2.
$$

Then, $z_{ij} < -\epsilon$ for $i \geq C$ and/or $j \geq 1$. Hence, except finite subset $\{(i,j)|0 \leq i \leq C-1, j=0\}$ of the state space of $\{(n_k, r_k)\}$, we have $z_{ij} < -\epsilon$. Therefore, by Foster's criteria, we see that the Markov chain is positive recurrent.

References

[1] H. Bruneel, *Performance of discrete-time queueing systems*, Comput. Oper. Res. **20** (1993), 303–320.
[2] H. Bruneel and B. Kim, *Discrete-time Models for Communication Systems Including ATM*, Kluwer Academic Publishers, 1993.
[3] M. L. Chaudhry and S. H. Chang, *Analysis of the discrete-time bulk-service queue* $Geo/G^Y/1/n+B$, Oper. Res. Lett. 2003.
[4] M. L. Chaudhry and J. G. C. Templeton, *A First Course in Bulk Queues*, John Wiley & Sons, 1983.
[5] B. D. Choi, D. I. Choi, Y. Lee and D. K. Sung, *Priority queueing system with fixed-length packet-train arrivals*, IEE Proceedings-Communications **145** (1998), no. 5, 331–336.
[6] B. D. Choi, Y. Lee and D. I. Choi, $Geo^{X_1}, Geo^{X_2}/D/c$ *HOL priority queueing system with random order selection within each priority class*, Probab. Engrg. Inform. Sci. **12** (1998), no. 1, 125–139.
[7] E. Cinlar, *Introduction to Stochastic Processes*, Prentice-Hall, Inc., Englewood Cliff, New Jersey, 1975.
[8] T. J. J. Denteneer, J. S. H. van Leeuwaarden and J. A. C. Resing, *Bounds for a discrete-time multi server queue with an application to cable networks*, ITC 18, Berlin, 2003.
[9] J. H. Dshalalow, *Frontiers in queueing*, CRC Press, Boca Raton, Florida, 1997.
[10] M. A. Dümmler and A. K. Schömig, *Using discrete-time analysis in the performance evaluation of manufacturing systems*, 1999 International Conference on Semiconductor Manufacturing Operational Modeling and Simulation(SMOMS'99), San Francisco, California, January, 1999, 18–20.
[11] U. C. Gupta and V. Goswami, *Performance analysis of finite buffer discrete-time queu with bulk service*, Comput. Oper. Res. **29** (2002), 1331–1341.
[12] G. U. Hwang and K. Sohraby, *Power tail aymptotic results of a discrete time queue with long range dependent input*, J. Korean Math. Soc. **40** (2003), no. 1, 87–107.
[13] N. K. Kim and M. L. Chaudhry, *An invariance relation and a unified method to derive stationary queue-length distributions*, Oper. Res. **52** (2004), no. 5.
[14] J. Lamperti,, *Criteria for the recurrence or transience of stochastic process, I*, J. Math. Anal. Appl. **1** (1960) 314–330.
[15] Y. Lee, *Discrete Time Analysis of* $GI/G/1$ *Priority Queueing Systems*, M.S. dissertation, Korea Advanced Institute of Science and Technology(KAIST), Korea, 1993.
[16] ______, *Discrete-time* $Geo^X/G/1$ *queue with preemptive resume priority*, Math. Comput. Modelling **34** (2001), 243–250.

[17] Y. Lee and B. D. Choi, *Queueing system with multiple delay and loss priorities for ATM networks*, Inform. Sci. **138** (2001), no. 1–4, 7–29.
[18] Y. Lee and K. S. Lee, *Discrete-time $Geo^X/G/1$ queue with preemptive repeat different priority*, Queueing Syst. Theory Appl. **44** (2003) no. 4, 399–411.
[19] Y. Lee, Y. H. Kim and J. D. Huh, *Discrete-time $Geo^X/G/1$ queue with non-preemptive priority*, Comput. Math. Appl. **46** (2003), no. 10–11, 1625–1632.
[20] A. G. Pakes, *Some conditions for ergodicity and recurrence of Markov chains*, Oper. Res. **17** (1969), 1058–1061.
[21] T. L. Saaty, *Elements of queueing theory with applications*, McGraw-Hill Book company, 1961.
[22] H. Takagi, *Queueing Analysis Vol.3: Discrete-time systems*, North-Holland, Amsterdam, 1993.

Department of Information and Communication Engineering
Dongeui University
Busan 614-714, Korea
E-mail: ylee@deu.ac.kr