1998年度日本オペレーションズ・リサーチ学会
秋季研究発表会

Mean Sojourn Times of Cyclic Queues with Feedback

01505150 筑波大学 平山哲治 HIRAYAMA Tetsuji

1 Introduction

A single server serves customers at J stations with infinite buffer capacities. The server selects the stations in a cyclic order in order to serve customers. Switchover times are incurred when the server moves between the stations. Customers are served according to a predetermined scheduling algorithm. After receiving a service, each customer leaves the system, is routed to another station in the system, or rejoins the same station for another portion of service.

We would like to obtain mean sojourn times of customers at these stations.

2 Model description

Customers arrive at station i from outside the system according to a Poisson process with rate λ_i $(i = 1,\ldots,J)$. Service times S_i of i-customers (customers stay at station i) are independently and arbitrarily distributed with mean $E[S_i]$ and second moment $\overline{s^2}_i$. After receiving a service, an i-customer either is routed to station j with prob. p_{ij}, or departs from the system with prob. p_{i0}. Let $\mathbf{P} \equiv (p_{ij} : i,j = 1,\ldots,J)$. We assume that $\mathbf{P}^n \to 0$ as $n \to \infty$ and that ρ (resource utilization) is less than 1.

The server selects one of the stations in a cyclic order: Station 1 → Station 2 → ⋯ → Station J → Station 1 → ⋯. An arbitrarily distributed switchover time S_i^o with mean $\overline{S^o}_i$ and second moment $\overline{s^{o2}}_i$ is incurred when the server moves from station i to station $i+1$ $(i = 1,\ldots,J)$. (Let station $J+1$ denote station 1.) The arrival processes, the service times, the feedback processes and the switchover times are assumed to be independent of each other.

The system is separated into two parts: the 'service facility (SF)' and the 'waiting rooms (WRs)' of the stations. The server selects one of the stations at a time, and then opens its gate, which separates the SF from its WR, in order to admit some customers to the SF. Then he serves customers in the SF until he empties it. Every feedback customer enters one of the WRs.

Each time interval from when the server opens a gate of a station until the first time when he moves from the station is called a *service period*. Each time interval during which the server moves from one of the stations to another station is called *switchover period*. Let $\Pi = \{1,\ldots,J\}$ be a set of service periods, and let $\Pi^s = \{1^s,\ldots,J^s\}$ be a set of switchover periods where j^s is a switchover period from station j to station $j+1$.

Then let κ denote a *current period* where the server is selecting station κ if $\kappa \in \Pi$, or the server is moving from station j to station $j+1$ if $\kappa = j^s \in \Pi^s$. Let r be a remaining service time of a customer being served if $\kappa \in \Pi$, or a remaining length of a switchover period if $\kappa \in \Pi^s$. Number of i-customers in the SF (who are not being served) is denoted by g_i $(i = 1,\ldots,J)$, and the vector is denoted by $\mathbf{g} \equiv (g_1,\ldots,g_J)$. Number of i-customers in the WR is denoted by n_i $(i = 1,\ldots,J)$, and the vector is denoted by $\mathbf{n} \equiv (n_1,\ldots,n_J)$. Let $\kappa(t), r(t), g_i(t), n_i(t), \mathbf{n}(t)$ and $\mathbf{g}(t)$ be the values at time t of the corresponding variables.

We will prescribe the *scheduling algorithms* according to the following specifications: 1) selection orders of the stations, and 2) service disciplines of the stations. We consider a system where the server selects the stations in a cyclic order. A service discipline at each station is either 1) gated, or 2) exhaustive. The service order of customers in the SF is FCFS.

The e^{th} customer (denoted by c^e) arrives from outside the system at epoch σ_0^e $(e = 1,2,\ldots)$. Let σ_k^e be a time epoch just when he would arrive at one of the stations after completing his k^{th} service $(k = 1,2,\ldots)$. Let $X(t)$ denote the station of an arriving customer at the last transition epoch of the system before or on t $(t \geq 0)$; if it is not a customer arrival epoch, then $X(t) = 0$. Let $L(t)$ be informations of the system at time t. Then we define $\mathcal{Q} \equiv \{\mathbf{Y}(t) = (X(t), \kappa(t), r(t), \mathbf{g}(t), \mathbf{n}(t), L(t)) : t \geq 0\}$. The state space of $\mathcal{Q}$ is denoted by $\mathcal{E}$.

3 Performance Measures (PMs)

We would like to derive three types of performance measures (PMs). Let $e(= 1,2,\ldots)$ be a customer number. Then let

$$C_{Wi}^e(t) \equiv \begin{cases} 1, & \text{if } \mathrm{c}^e \text{ is in the WR at station } i \text{ at } t, \\ 0, & \text{otherwise,} \end{cases} \tag{3.1}$$

$$C_{Fi}^e(t) \equiv \begin{cases} 1, & \text{if } \mathrm{c}^e \text{ is in the SF as an } i\text{-customer at } t, \\ 0, & \text{otherwise,} \end{cases} \tag{3.2}$$

for $t \geq 0$ and $i = 1,\ldots,J$. For any event $\mathcal{K}$, let $1\{\mathcal{K}\}$ equals 1 if event $\mathcal{K}$ occurs, or it equals 0 otherwise. Then we define

$$W_i^e \equiv \int_0^\infty C_{Wi}^e(t)dt, \tag{3.3}$$

$$H_i^e(k) \equiv \int_0^\infty C_{Wi}^e(t)1\{\kappa(t) = k\}dt, \quad k \in \Pi \cup \Pi^s, \tag{3.4}$$

$$F_i^e \equiv \int_0^\infty C_{Fi}^e(t)dt, \tag{3.5}$$

for $i = 1, \ldots, J$. For $l = 0, 1, 2, \ldots$, we define

$$W_i(\mathbf{Y}, e, l) \equiv E\left[\int_{\sigma_l^e}^{\infty} C_{Wi}^e(t)dt | \mathbf{Y}(\sigma_l^e) = \mathbf{Y}\right], \quad (3.6)$$

$$H_i(\mathbf{Y}, e, l, k) \equiv E\left[\int_{\sigma_l^e}^{\infty} C_{Wi}^e(t)1\{\kappa(t) = k\}dt | \mathbf{Y}(\sigma_l^e) = \mathbf{Y}\right] (3.7)$$

$$F_i(\mathbf{Y}, e, l) \equiv E\left[\int_{\sigma_l^e}^{\infty} C_{Fi}^e(t)dt | \mathbf{Y}(\sigma_l^e) = \mathbf{Y}\right], \quad (3.8)$$

for $k \in \Pi \cup \Pi^s, i = 1, \ldots, J$ and $\mathbf{Y} \in \mathcal{E}$.

4 Expressions of the PMs.

We can show that the expected values of the performance measures (PMs) defined in the last section are given by

$$W_i(\mathbf{Y}, e, l) = r\varphi_i^0(j, \kappa, 0) + 1(r)\varphi_i^1(j, \kappa, 0) + (\mathbf{g}, \mathbf{n})w_i(j, \kappa, 0) + w_i(j, \kappa, 0), \quad (4.1)$$

$$H_i(\mathbf{Y}, e, l, k) = \begin{cases} r\varphi_i^0(j, \kappa, k) + 1(r)\varphi_i^1(j, \kappa, k) + \\ (\mathbf{g}, \mathbf{n})w_i(j, \kappa, k) + w_i(j, \kappa, k), & k \in \Pi, \\ r\varphi_i^0(j, \kappa, k) + w_i(j, \kappa, k), & k \in \Pi^s, \end{cases} \quad (4.2)$$

$$F_i(\mathbf{Y}, e, l) = r\eta_i^0(j, \kappa) + 1(r)\eta_i^1(j, \kappa) + (\mathbf{g}, \mathbf{n})f_i(j, \kappa) + f_i(j, \kappa), \quad (4.3)$$

for $\mathbf{Y} = (j, \kappa, r, \mathbf{g}, \mathbf{n}, L) \in \mathcal{E}, e = 1, 2, \ldots, l = 0, 1, 2, \ldots$ and $i \in \Pi$. ($1(r) = 1$ if $r > 0$, or $1(r) = 0$ otherwise.) The constants in these expressions can be obtained from the system parameters defined in the last section, and they may be different among the scheduling algorithms adopted. These expressions are linear in $(\mathbf{g}, \mathbf{n})$.

5 Steady state values.

In this section, we would like to evaluate the values:

$$\bar{w}_i(j) \equiv \lim_{N \to \infty} \frac{\sum_{e=1}^{N} E[(W_i^e + F_i^e)1\{X(\sigma_0^e) = j\}]}{\sum_{e=1}^{N} E[1\{X(\sigma_0^e) = j\}]}, \quad (5.1)$$

for $i, j = 1, \ldots, J$. $\bar{w}_i(j)$ denotes the average sojourn time that customers who arrive at station j from outside the system spend at station i. The time average values of the system states are defined by:

$$\tilde{\mathbf{Y}}^\kappa \equiv (\tilde{X}^\kappa, \kappa\tilde{q}^\kappa, \tilde{r}^\kappa, \tilde{\mathbf{g}}^\kappa, \tilde{\mathbf{n}}^\kappa, \tilde{L}^\kappa)$$

$$\equiv \lim_{t \to \infty} \frac{1}{t} \int_0^t E[\mathbf{Y}(s)1\{\kappa(s) = \kappa\}]ds, \quad (5.2)$$

$$\tilde{\mathbf{y}} \equiv (\tilde{\mathbf{g}}^1, \tilde{\mathbf{n}}^1, \tilde{\mathbf{g}}^2, \tilde{\mathbf{n}}^2, \ldots, \tilde{\mathbf{g}}^J, \tilde{\mathbf{n}}^J), \quad (5.3)$$

where $\tilde{\mathbf{g}}^\kappa \equiv (\tilde{g}_1^\kappa, \ldots, \tilde{g}_J^\kappa)$, $\tilde{\mathbf{n}}^\kappa \equiv (\tilde{n}_1^\kappa, \ldots, \tilde{n}_J^\kappa)$, and $\kappa \in \Pi \cup \Pi^s$.

Let $\{\Lambda_i : i = 1, \ldots, J\}$ be a set of values that satisfy the equations: $\Lambda_i = \lambda_i + \sum_{j=1}^J p_{ji}\Lambda_j, (i = 1, \ldots, J)$. Then we have

$$\tilde{q}^\kappa = \begin{cases} \Lambda_\kappa E[S_\kappa], & \kappa \in \Pi, \\ (1 - \rho)\overline{S^o}_{\kappa^-} / \sum_{j=1}^J \overline{S^o}_j, & \kappa \in \Pi^s. \end{cases} \quad (5.4)$$

$$\tilde{r}^\kappa = \begin{cases} \Lambda_\kappa \overline{s^2}_\kappa / 2, & \kappa \in \Pi, \\ (1 - \rho)\overline{s^{o2}}_{\kappa^-} / (2\sum_{j=1}^J \overline{S^o}_j), & \kappa \in \Pi^s. \end{cases} \quad (5.5)$$

Further we define following vectors:

$$\mathbf{s}_n(k) \equiv \sum_{\kappa \in \Pi \cup \Pi^s} \sum_{j=1}^J \lambda_j (s_{n1}(j, \kappa, k), \ldots, s_{nJ}(j, \kappa, k)), \quad (5.6)$$

$$\mathbf{s}_g(k) \equiv \sum_{\kappa \in \Pi \cup \Pi^s} \sum_{j=1}^J \lambda_j (s_{g1}(j, \kappa, k), \ldots, s_{gJ}(j, \kappa, k)) - (0, \ldots, 0, \underbrace{\tilde{q}^k}_{k^{th}\text{column}}, 0, \ldots, 0), \quad (5.7)$$

where

$$s_{ni}(j, \kappa, k) \equiv \begin{cases} \tilde{r}^\kappa \varphi_i^0(j, \kappa, k) + \tilde{q}^\kappa(\varphi_i^1(j, \kappa, k) + w_i(j, \kappa, k)), & k \in \Pi \cup \{0\} \\ \tilde{r}^\kappa \varphi_i^0(j, \kappa, k) + \tilde{q}^\kappa w_i(j, \kappa, k), & k \in \Pi^s, \end{cases}$$

$$s_{gi}(j, \kappa, k) \equiv \begin{cases} 0, & i \neq k, \\ \tilde{r}^\kappa \eta_k^0(j, \kappa) + \tilde{q}^\kappa(\eta_k^1(j, \kappa) + f_k(j, \kappa)), & i = k, \end{cases} \quad k \in \Pi.$$

Then we have

$$\tilde{\mathbf{n}}^k = \mathbf{s}_n(k), \qquad \tilde{\mathbf{g}}^k = 0, \qquad k \in \Pi^s. \quad (5.8)$$

Now we define the matrices:

$$\Phi(\kappa, k) \equiv (\textstyle\sum_{j=1}^J \lambda_j w_i(j, \kappa, k) : i = 1, \ldots, J) \in \mathcal{R}^{2J \times J},$$

$$\Theta(\kappa, k) \equiv (0, \ldots, 0, \underbrace{\sum_{j=1}^J \lambda_j f_k(j, \kappa)}_{k^{th}\text{column}}, 0, \ldots, 0) \in \mathcal{R}^{2J \times J},$$

for $\kappa \in \Pi \cup \Pi^s$ and $k \in \{0\} \cup \Pi$. Further we define the vectors in $\mathcal{R}^{1 \times 2J^2}$ and matrix in $\mathcal{R}^{2J^2 \times 2J^2}$.

$$\mathbf{s} \equiv (\mathbf{s}_g(1), \mathbf{s}_n(1), \mathbf{s}_g(2), \mathbf{s}_n(2), \ldots, \mathbf{s}_g(J), \mathbf{s}_n(J)),$$

$$\mathbf{s}_0 \equiv \sum_{\kappa \in \Pi^s} (0, \mathbf{s}_n(\kappa))(\Theta(\kappa, 1), \Phi(\kappa, 1), \ldots, \Theta(\kappa, J), \Phi(\kappa, J)),$$

$$\mathbf{S} \equiv \begin{pmatrix} \Theta(1,1) & \Phi(1,1) & \cdots & \Theta(1,J) & \Phi(1,J) \\ \Theta(2,1) & \Phi(2,1) & \cdots & \Theta(2,J) & \Phi(2,J) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Theta(J,1) & \Phi(J,1) & \cdots & \Theta(J,J) & \Phi(J,J) \end{pmatrix}.$$

Then we have

$$\tilde{\mathbf{y}} = (\mathbf{s} + \mathbf{s}_0)(\mathbf{I} - \mathbf{S})^{-1}. \quad (5.9)$$

Finally, sojourn times of customers are given by

$$\bar{w}_i(j) = \sum_{\kappa \in \Pi \cup \Pi^s} \{\tilde{r}^\kappa(\varphi_i^0(j, \kappa, 0) + \eta_i^0(j, \kappa)) + \tilde{q}^\kappa(\varphi_i^1(j, \kappa, 0) + w_i(j, \kappa, 0) + \eta_i^1(j, \kappa) + f_i(j, \kappa)) + (\tilde{\mathbf{g}}^\kappa, \tilde{\mathbf{n}}^\kappa)(w_i(j, \kappa, 0) + f_i(j, \kappa))\}. \quad (5.10)$$

References

[1] O.J. Boxma and W.P. Groenendijk, Pseudo-conservation laws in cyclic-service systems, J. Appl. Prob. 24 (1987) 949-964.

[2] S.J. Hong, T. Hirayama and K. Yamada, The mean sojourn time of multiclass M/G/1 queues with feedback, Asian-Pacific Journal of Operational Research 10 (1993) 233-249.

[3] H. Levy and M. Sidi, Polling systems: applications, modeling, and optimization, IEEE Trans. Commun. 38 (1990) 1750-1760.

[4] M. Sidi, H. Levy and S.W. Fuhrmann, A queueing network with a single cyclically roving server, Queueing Systems 11 (1992) 121-144. (See also: Queueing Systems 16 (1994) 193.)

[5] B. Simon, Priority queues with feedback, J.ACM 31 (1984) 134-149.