Journal of the Operations Research
Society of Japan
2005, Vol. 48, No. 3, 226-255

MULTICLASS POLLING SYSTEMS WITH MARKOVIAN FEEDBACK: MEAN SOJOURN TIMES IN GATED AND EXHAUSTIVE SYSTEMS WITH LOCAL PRIORITY AND FCFS SERVICE ORDERS

Tetsuji Hirayama
University of Tsukuba

(Received November 27, 2003; Revised June 13, 2005)

Abstract We consider polling systems in which a single server visits stations in cyclic order and serves customers at each station according to either the gated rule or the exhaustive rule of the station. There are multiple classes of customers at each station, and they are served in either the priority order or the first-come-first-served (FCFS) order. After completing a service at a station, each customer may be routed to one of the stations or leave the system according to the Markovian feedback mechanism. In this paper, a new approach to mean sojourn times in multiclass queues, developed in [11, 14, 15], is extended to the feedback polling systems as follows. We define the conditional expected sojourn times and find their linear functional expressions by solving some equations. The steady state average sojourn times are derived from these expressions by simple limiting procedures, and their values are obtained by solving a set of linear equations. We also consider composite scheduling algorithms and calculate mean path times.

Keywords: Queues with feedback, multiclass polling systems, mean sojourn times, linear functional expressions

1. Introduction

A polling system with multiple classes of customers and their feedback is analyzed in this paper. A single server selects (or visits) stations in cyclic order and admits customers at the selected station into the service facility according to either the gated rule or the exhaustive rule of the station. Since there are multiple customer classes at each station, customers in the service facility are served in either the first-come-first-served (FCFS) or the fixed priority order. In previous research, Sidi, Levy and Fuhrmann [25] analyzed a polling system with feedback. We extend their system to include multiple customer classes and local priorities. We can easily calculate the mean path time, which is the mean amount of time spent by an arbitrary customer traversing a specific path. The initial version of our feedback polling model is given in [13]. A new approach to analyzing the polling systems, which was developed in [15], is extended to the feedback model in this paper.

Numerous studies and techniques have been developed for computing the mean waiting times in polling systems. The *buffer occupancy method* has been considered by many researchers (e.g., Cooper [4], Cooper and Murray [5], Eisenberg [9], and Takagi [28]). Levy and Sidi [22] analyzed polling systems with simultaneous arrivals by using the method. An iterative algorithm for this method was discussed by Levy [20]. The descendant set (DS) technique proposed by Konheim, Levy and Srinivasan [19] is a variation of the method. It has been shown that the buffer occupancy method and the DS technique can be applied to a polling system with feedback ([19, 25]). Another method called the *station time method* was considered by Ferguson and Aminetzah [10]. Takine and Hasegawa [31] analyzed a symmetric system under Bernoulli scheduling by using the *stochastic decomposition* property.

Polling systems are used to analyze computer communication systems. Their applications to data link control and to token ring networks were discussed by Takagi [29]. Levy and Sidi [21] provided feedback polling models of the SR-ARQ scheme in a communication network. Takagi [27] discussed feedback polling models of file transfer networks and error-prone transmission systems. An application to call processing in switching systems was discussed by Katayama [17]. In recent years, various types of multimedia messages composed of multiple packets have been transmitted through a network. Therefore, scheduling policies and priorities for transmitting messages have become essential issues. We may model messages at a network node as customers with multiple services and priorities, and then obtain their whole message delays (path times) rather than their packet delays.

Several variations of feedback polling systems have been studied. Katayama [16, 18] investigated multi-stage tandem queues served by a single cyclic server. The mean sojourn times in polling systems with Bernoulli feedback were obtained by Takine, Takagi and Hasegawa [32]. Priority queueing systems with feedback have also been considered. A Bernoulli feedback queue was considered by Disney [7], and the queue with multiple customer classes was considered by Doshi and Kaufman [8]. Deterministic feedback systems were investigated by Van den Berg, Boxma and Groenendijk [1], and Daigle and Houstis [6]. General feedback queues with multiple classes of customers and priorities were considered by Simon [26], and Paterok and Fischer [23]. A general multiclass feedback queue with gated disciplines was considered by Hirayama [14].

The first step in our method is to define stochastic process $\mathcal{Q}$ that represents an evolution of the system states and is to define conditional expected sojourn times such as (2.7), (2.11) and (2.15). Then we make 'feedback equations' (equations (2.9), (2.13) and (2.17)) in Section 2. The second step is to consider the conditional expected sojourn times per service stage $W^I_{i,\alpha}, H^I_{i,\alpha}$ and $F^I_{i,\alpha}$, whose expressions are derived from the analysis of customers at polling instants in Section 3. We can obtain their linear functional expressions (equations (4.22)–(4.25)) in Section 4. In the third step, the conditional expected sojourn times $W_{i,\alpha}, H_{i,\alpha}$ and $F_{i,\alpha}$ defined in Section 2 are obtained by solving the 'feedback equations.' They also have the linear functional expressions (equations (5.11)–(5.13) in Section 5). Finally their steady state average values are obtained by a simple limiting procedure, and can be calculated by solving a set of linear equations (equations (6.15)–(6.16) in Section 6).

2. Model Description

In this section, we describe our model of the multiclass polling system with Markovian feedback. J *groups* of customers arrive at the system, and customers belonging to group i, called i-*customers*, stay at station i ($i = 1, \ldots, J$). Group i consists of L_i *classes* of customers, and customers belonging to class α in group i, called (i, α)-*customers*, arrive from outside the system according to a Poisson process with rate $\lambda_{i,\alpha}$ ($\alpha = 1, \ldots, L_i$). Let $\lambda \equiv \sum_{i=1}^{J} \sum_{\alpha=1}^{L_i} \lambda_{i,\alpha}$ be the total arrival rate. The e^{th} customer arrives from outside the system at epoch σ_0^e and is denoted by $\boldsymbol{c}^e$ ($e = 1, 2, \ldots$)[1]. Let $\mathcal{S} \equiv \{(i, \alpha) : i = 1, \ldots, J \text{ and } \alpha = 1, \ldots, L_i\}$, and the total number of classes is denoted by $J_c \equiv \sum_{i=1}^{J} L_i$.

A single server serves customers at these stations. Service times $S_{i,\alpha}$ of (i, α)-customers are independently, identically and arbitrarily distributed with mean $E[S_{i,\alpha}] > 0$ and second moment $\overline{s^2_{i,\alpha}}$. Customers are served according to the predetermined *scheduling algorithm*

[1]The customer $\boldsymbol{c}^e$ arrives from outside the system according to a Poisson process with rate λ, and when it arrives, it becomes an (i, α)-customer with probability $\lambda_{i,\alpha}/\lambda$.

defined below. After completing a service at each service stage, an (i,α)-customer either returns to the system as a (j,β)-customer with probability $p_{i,\alpha,j,\beta}$, or departs from the system with probability $1-\sum_{j=1}^{J}\sum_{\beta=1}^{L_j} p_{i,\alpha,j,\beta}$[2]. The feedback probability matrix is given by $\boldsymbol{P} \equiv (p_{i,\alpha,j,\beta} : (i,\alpha),(j,\beta) \in \mathcal{S})$. Since we assume that $\boldsymbol{P}^n \to \boldsymbol{O}$ as $n \to \infty$, all arriving customers eventually leave the system. For any $(i,\alpha) \in \mathcal{S}$, let $T_{i,\alpha}$ be the total amount of service times received by a customer arriving as an (i,α)-customer until it departs from the system, and let $\overline{T}_{i,\alpha}$ be its expected value. The set of the values $\overline{T}_{i,\alpha}$, $(i,\alpha) \in \mathcal{S}$, satisfies:

$$\overline{T}_{i,\alpha} \quad = \quad E[S_{i,\alpha}] + \sum_{k=1}^{J}\sum_{\gamma=1}^{L_k} p_{i,\alpha,k,\gamma}\overline{T}_{k,\gamma}, \quad (i,\alpha) \in \mathcal{S}. \tag{2.1}$$

Then, let $\rho \equiv \sum_{i=1}^{J}\sum_{\alpha=1}^{L_i} \lambda_{i,\alpha}\overline{T}_{i,\alpha} < 1$ be the resource utilization of the system. The server selects (or visits) the stations in the cyclic order: $1,2,\ldots,J,1,2,\ldots$. An arbitrarily distributed switchover time with mean $\overline{s_i^o}$ and second moment $\overline{s_i^{o2}}$ is incurred when the server switches services from station i to station $i+1$.[3] 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, which are called the 'service facility' and the 'waiting room.' There is a gate at each station between the set of its queues in the waiting room and the set of its queues in the service facility, which intercepts the migration of customers between them. Each customer arriving at each station from outside the system or by the feedback mechanism enters its queue in the service facility when its gate is opened; otherwise, it enters its queue in the waiting room[4]. The server selects one of the stations at a time, and then opens its gate in order to admit some customers at the station to its queues in the service facility. Then, the server serves the customers in the service facility until the server empties it, and then selects another station and opens its gate. Since the gates of the stations that are not selected by the server are closed, all customers at such stations must wait for service in the queues in the waiting room. Once a customer begins a service, its service is not interrupted by other customers (that is, the service is non-preemptive).

Each time interval from when the server selects a station until the first time the server empties the service facility is called a *service period*. Each time interval when the server switches selection of the stations is called a *switchover period*. Let $\Pi = \{1,\ldots,J\}$ be the set of (indices of) the service periods, where $i \in \Pi$ denotes the service period of station i. And let $\Pi^s = \{1^s,\ldots,J^s\}$ be the set of (indices of) the switchover periods, where $i^s \in \Pi^s$ denotes the switchover period from station i to station $i+1$. The subsets of Π and Π^s are defined as:

$$\Pi_{k,j} \quad \equiv \quad \begin{cases} \{k,k+1,\ldots,j-1\}, & k<j, \\ \{k,k+1,\ldots,J\}, & k>j=1, \\ \{k,k+1,\ldots,J,1,\ldots,j-1\}, & k>j\neq 1, \\ \emptyset, & k=j, \end{cases} \tag{2.2}$$

$$\Pi^s_{k,j} \quad \equiv \quad \begin{cases} \{k^s,(k+1)^s,\ldots,(j-1)^s\}, & k<j, \\ \{k^s,(k+1)^s,\ldots,J^s\} & k>j=1, \\ \{k^s,(k+1)^s,\ldots,J^s,1^s,\ldots,(j-1)^s\}, & k>j\neq 1, \\ \emptyset, & k=j. \end{cases} \tag{2.3}$$

[2]The customer's return or departure occurs immediately after the service completion without any time lag.
[3]For notational convenience, station 0 and station $J+1$ denote station J and station 1, respectively.
[4]Which queue the customer enters depends on its group and class (see the scheduling algorithms below).

For any period $k \in \Pi \cup \Pi^s$, let k^- denote the period just before the period k:

$$k^- = (i-1)^s \text{ for } k = i \in \Pi, \text{ or } \ k^- = i \text{ for } k = i^s \in \Pi^s. \tag{2.4}$$

Customers in the system are served according to a predetermined *scheduling algorithm.* We will prescribe it according to the following specifications:

- Selection orders of the stations by the server.
- Customer selection rules used when the server admits customers into the service facility.
- Service orders of customers in the service facility.

The selection order of the stations is the *cyclic*, as described before. The customer selection rule functions at every station only when it is selected by the server, and is either *gated* or *exhaustive.* When the server selects one of the stations with the gated rule, all customers staying at the station just when it is selected by the server enter its queues in the service facility, and then the gate is immediately closed. The service period of the station continues until the first time when all of the customers admitted into the service facility complete their services. $\mathcal{H}_g$ denotes the set of stations with the gated rule (the gated groups). When the server selects one of the stations with the exhaustive rule, the gate of the station remains open (that is, customers arriving at the station later may still enter the service facility) and the server continues to serve all customers until the station is cleared of customers for the first time. The service period of the station finishes at this time, and its gate is closed. $\mathcal{H}_e$ denotes the set of stations with the exhaustive rule (the exhaustive groups).

The service order of customers in the service facility is either *FCFS order*, or *fixed priority order.* If the server selects a group with the FCFS order, it serves all customers in the service facility in a first-come-first-served order. This group has a queue in the waiting room and a queue in the service facility, both of which are common to all classes in the group. $\mathcal{H}_{gF}$ denotes the set of gated groups with the FCFS order (the gated FCFS groups), and $\mathcal{H}_{eF}$ denotes the set of exhaustive groups with the FCFS order (the exhaustive FCFS groups). If the server selects a group with the fixed priority order, it serves all customers in the service facility according to a fixed priority order, where class α customers in the group have priority over class β customers in the group if $\alpha < \beta$. All customers in each class in the service facility are served in a first-come-first-served order. Each class in the group has an individual queue in the waiting room and an individual queue in the service facility. $\mathcal{H}_{gP}$ denotes the set of gated groups with the fixed priority order (the gated priority groups), and $\mathcal{H}_{eP}$ denotes the set of exhaustive groups with the fixed priority order (the exhaustive priority groups).

Note In the feedback queueing system, each customer may have many arrival epochs. One is the customer's arrival epoch from outside the system and the others are the customer's feedback arrival epochs[5]. Hence, at each time the server selects a customer according to any of the 'first-come-first-served' orders, it must be specified for each customer under consideration which arrival epoch of the customer is its 'coming' epoch. In this paper, we assume that it is the last epoch when the customer arrived in its current group and class at this time. Then the server selects a customer whose last arrival epoch (or coming epoch) is the earliest among the customers under consideration. In other words, each (j, β)-customer arriving from outside or by feedback joins one of the following: (1) the tail of the queue of group j if $j \in \mathcal{H}_{gF} \cup \mathcal{H}_{eF}$, or (2) the tail of the queue of class β in group j if $j \in \mathcal{H}_{gP} \cup \mathcal{H}_{eP}$. ■

[5]We use the term 'arrival' to denote both an arrival from outside and an arrival by feedback.

2.1. Relation to the existing models

We now summarize the relationship between our model and other existing polling models described in the references. We first note that we can analyze exhaustive and gated disciplines only in contrast to some existing models that analyze limited service disciplines. We also note that our model is used to analyze the systems with switchover times, whereas some existing models may be used to analyze the systems without switchover times or the systems with another type of overhead called a setup time or a vacation. For simplicity, we sometimes omit these differences in the following discussion.

A polling system with Markovian feedback (with switchover times) was studied by Sidi, Levy and Fuhrmann [25]. They considered a system with gated and exhaustive service disciplines, in which each station has a single class of customers. They derived the mean sojourn times by the buffer occupancy method and derived a pseudo-conservation law. This system is a special case of our system with $L_i = 1$ $(i = 1, \ldots, J)$. The relationship between the buffer occupancy method and our method is given in Section 6.4.

In Bernoulli feedback queues, customers may return only to the same station at which they arrived with the same probability at every feedback epoch. The mean sojourn time in the symmetric Bernoulli feedback polling system with a 1-limited service discipline was derived by Takagi [27], and the mean sojourn times in the symmetric system with exhaustive, gated and 1-limited disciplines were derived by Takine, Takagi and Hasegawa [32]. (They analyzed the systems with switchover times.) We can analyze the mean sojourn times in the asymmetric Bernoulli feedback system by setting

$$p_{i,\alpha,j,\beta} = \begin{cases} p_{i,\alpha}, & (j, \beta) = (i, \alpha), \\ 0, & \text{otherwise.} \end{cases}$$

Multi-stage tandem queueing systems served by a single server were considered by Katayama [16–18]. In [16], an N-stage system (without switchover times) with exhaustive, K-limited and gated service disciplines was considered, and the mean sojourn times were obtained. Each customer arriving from outside at stage 1 receives exactly N services. Each customer goes to stage $i + 1$ after completing a service at stage i $(i = 1, \ldots, N - 1)$. The server visits the stages in cyclic order (stage 1 $\to$ stage 2 $\to \cdots \to$ stage N $\to$ stage 1 $\to \cdots$). We can analyze the mean sojourn times in the system by setting $J = N$ and

$$\lambda_{1,\alpha} > 0; \qquad \lambda_{i,\alpha} = 0, \quad (i = 2, \ldots, N);$$
$$p_{i,\alpha,j,\beta} = \begin{cases} 1, & j = i + 1 \ \& \ \beta = \alpha, \\ 0, & \text{otherwise,} \end{cases} \qquad (i = 1, \ldots, N - 1).$$

In [17], the 2-stage multi-class system (without switchover times) with exhaustive disciplines was considered, and the mean sojourn times of all types of customers were obtained. After completing a service at station 0 (the common station), each type-n customer arriving at the station returns to station 0 with probability p or proceeds to station n with probability $q = 1 - p$ $(n = 1, \ldots, N)$. A single server visits the stations in cyclic order. We can analyze the mean sojourn times in the tandem system by setting $J = N + 1$, $L_i = 1$, $(i = 1, \ldots, N)$, $L_{N+1} = N$ (the common station is modeled as station $N + 1$ in our system), and

$$\lambda_{N+1,\alpha} > 0, \quad (\alpha = 1, \ldots, N); \qquad \lambda_{i,1} = 0, \quad (i = 1, \ldots, N);$$
$$p_{i,\alpha,j,\beta} = \begin{cases} p, & j = i = N + 1 \ \& \ \beta = \alpha, \\ 1 - p, & i = N + 1, \ j = \alpha \ \& \ \beta = 1, \quad (\alpha = 1, \ldots, N), \\ 0, & \text{otherwise.} \end{cases}$$

2.2. Definitions of the system states and the performance measures

Let us consider the system operating under some fixed scheduling algorithm. Let M^e be the total number of service stages of customer $\boldsymbol{c}^e$ from its arrival from outside the system at epoch σ_0^e until its departure from the system $(e = 1, 2, \ldots)$. Then let σ_k^e be the time just when, after completing its k^{th} service stage, it arrives (by a feedback) at one of the stations or departs from the system $(k = 1, 2, \ldots, M^e)$. For convenience, let $\mathcal{R}, \mathcal{R}_+, \mathcal{I}_+$ be respectively a set of real numbers, a set of nonnegative real numbers, and a set of nonnegative integers.

Let $(\kappa(t), a(t))$ denote that a $(\kappa(t), a(t))$-customer is being served at time t if $\kappa(t) \in \Pi$, or denote that the server is switching from station i to station $i+1$ if $\kappa(t) = i^s \in \Pi^s$ $(a(t) = 0$ in this case). Then let $r(t)$ denote the remaining service time of a customer being served at time t if $\kappa(t) \in \Pi$, or the remaining length of a switchover period if $\kappa(t) \in \Pi^s$. Let $\mathcal{S}_A \equiv \mathcal{S} \cup \{(1^s, 0), (2^s, 0), \ldots, (J^s, 0)\}$ be the set of the status of the server. The number of (i, α)-customers in the service facility at time t (who are not being served) is denoted by $g_{i,\alpha}(t)$, and the number of (i, α)-customers in the waiting room at time t is denoted by $n_{i,\alpha}(t)$. Let $\boldsymbol{g}(t) \equiv (g_{i,\alpha}(t) : (i, \alpha) \in \mathcal{S}) \in \mathcal{I}_+^{J_c}$, and let $\boldsymbol{n}(t) \equiv (n_{i,\alpha}(t) : (i, \alpha) \in \mathcal{S}) \in \mathcal{I}_+^{J_c}$. The sample paths of these processes are assumed to be left-continuous with right-hand limits. We specify information of the system at time t: $L(t) \equiv \{(j_m(t), \beta_m(t), s_m(t)) : m = 1, 2, \ldots\}$ where $(j_m(t), \beta_m(t))$ denotes (station, class) and $s_m(t)$ denotes the status of a customer who has arrived the m^{th} earliest of all customers in the system at time t. Let us consider transition epochs of these processes consisting of customer arrival epochs, service completion epochs and switchover period completion epochs. Then let $X(t)$ and $\Gamma(t)$ respectively denote the station and the class of a customer arriving at the last transition epoch before or on t; if it is not a customer arrival epoch, then $(X(t), \Gamma(t)) = (0, 0)$. $(X(t), \Gamma(t))$ and $L(t)$ are right continuous with left-hand limits. Then we define the stochastic process

$$\mathcal{Q} \equiv \{\boldsymbol{Y}(t) \equiv (X(t), \Gamma(t), \kappa(t), a(t), r(t), \boldsymbol{g}(t), \boldsymbol{n}(t), L(t)) : t \geq 0\} \tag{2.5}$$

that represents an evolution of the system. For any scheduling algorithm defined above, $\mathcal{Q}$ may embed a Markov process. Possible values of $\boldsymbol{Y}(t)$ are called *states*, and the state space of $\mathcal{Q}$ is denoted by $\mathcal{E}$.

We define three types of the system performance measures of customer $\boldsymbol{c}^e$ $(e = 1, 2, \ldots)$. First type of the performance measures is related to the waiting times of customer $\boldsymbol{c}^e$ in the waiting room. We define for any $t \geq 0$ and $(i, \alpha) \in \mathcal{S}$,

$$C_{Wi,\alpha}^e(t) \equiv \begin{cases} 1, & \text{if } \boldsymbol{c}^e \text{ stays in the waiting room as an } (i,\alpha)\text{-customer at time } t, \\ 0, & \text{otherwise.} \end{cases}$$

The $\boldsymbol{c}^e$'s waiting time spent in the waiting room as an (i, α)-customer is defined by

$$W_{i,\alpha}^e \equiv \int_0^\infty C_{Wi,\alpha}^e(t)dt, \quad (i, \alpha) \in \mathcal{S}. \tag{2.6}$$

For $l = 0, 1, 2, \ldots$, we define the two expected waiting times conditioned on the state of the system at time σ_l^e which are spent by customer $\boldsymbol{c}^e$ in the waiting room as an (i, α)-customer after time σ_l^e.

$$W_{i,\alpha}(\boldsymbol{Y}, e, l) \equiv E\left[\int_{\sigma_l^e}^\infty C_{Wi,\alpha}^e(t)dt \,\middle|\, \boldsymbol{Y}(\sigma_l^e) = \boldsymbol{Y}\right], \quad (i, \alpha) \in \mathcal{S}, \tag{2.7}$$

$$W_{i,\alpha}^I(\boldsymbol{Y}, e, l) \equiv E\left[\int_{\sigma_l^e}^{\sigma_{l+1}^e} C_{Wi,\alpha}^e(t)dt \,\middle|\, \boldsymbol{Y}(\sigma_l^e) = \boldsymbol{Y}\right], \quad (i, \alpha) \in \mathcal{S}, \tag{2.8}$$

for $\boldsymbol{Y} \in \mathcal{E}$. $W_{i,\alpha}(\boldsymbol{Y}, e, l)$ is the overall expected waiting time after time σ_l^e whereas $W_{i,\alpha}^I(\boldsymbol{Y}, e, l)$ is the expected waiting time during a service stage in $[\sigma_l^e, \sigma_{l+1}^e)$. Then the following 'feedback equation' holds.

$$W_{i,\alpha}(\boldsymbol{Y}, e, l) \;=\; W_{i,\alpha}^I(\boldsymbol{Y}, e, l) + E[W_{i,\alpha}(\boldsymbol{Y}(\sigma_{l+1}^e), e, l+1) | \boldsymbol{Y}(\sigma_l^e) = \boldsymbol{Y}], \tag{2.9}$$

for $\boldsymbol{Y} \in \mathcal{E}$ and $l = 0, 1, 2, \ldots$ $((i, \alpha) \in \mathcal{S})$.

Second type of the performance measures is related to pieces of the $\boldsymbol{c}^e$'s waiting times in the waiting room[6]. Let

$$H_{i,\alpha}^e(k) \;\equiv\; \int_0^\infty C_{Wi,\alpha}^e(t) \mathbf{1}\{\kappa(t) = k\} dt, \quad (i, \alpha) \in \mathcal{S}, \quad k \in \Pi \cup \Pi^s. \tag{2.10}$$

$H_{i,\alpha}^e(k)$ denotes the $\boldsymbol{c}^e$'s waiting time spent in the waiting room as an (i, α)-customer while the system is in period k. For $l = 0, 1, 2, \ldots$, we define the two expected waiting times conditioned on the state of the system at time σ_l^e which are spent by customer $\boldsymbol{c}^e$ in the waiting room as an (i, α)-customer after time σ_l^e while the system is in period k.

$$H_{i,\alpha}(\boldsymbol{Y}, e, l, k) \;\equiv\; E\left[\int_{\sigma_l^e}^\infty C_{Wi,\alpha}^e(t) \mathbf{1}\{\kappa(t) = k\} dt \,\middle|\, \boldsymbol{Y}(\sigma_l^e) = \boldsymbol{Y} \right], \quad (i, \alpha) \in \mathcal{S}, \tag{2.11}$$

$$H_{i,\alpha}^I(\boldsymbol{Y}, e, l, k) \;\equiv\; E\left[\int_{\sigma_l^e}^{\sigma_{l+1}^e} C_{Wi,\alpha}^e(t) \mathbf{1}\{\kappa(t) = k\} dt \,\middle|\, \boldsymbol{Y}(\sigma_l^e) = \boldsymbol{Y} \right], \quad (i, \alpha) \in \mathcal{S}, \tag{2.12}$$

for $\boldsymbol{Y} \in \mathcal{E}$, $k \in \Pi \cup \Pi^s$ and $l = 0, 1, 2, \ldots$. $H_{i,\alpha}(\boldsymbol{Y}, e, l, k)$ is the overall expected waiting time after time σ_l^e whereas $H_{i,\alpha}^I(\boldsymbol{Y}, e, l, k)$ is the expected waiting time during a service stage in $[\sigma_l^e, \sigma_{l+1}^e)$. Then the following 'feedback equation' holds.

$$H_{i,\alpha}(\boldsymbol{Y}, e, l, k) \;=\; H_{i,\alpha}^I(\boldsymbol{Y}, e, l, k) + E[H_{i,\alpha}(\boldsymbol{Y}(\sigma_{l+1}^e), e, l+1, k) | \boldsymbol{Y}(\sigma_l^e) = \boldsymbol{Y}], \tag{2.13}$$

for $\boldsymbol{Y} \in \mathcal{E}$ and $l = 0, 1, 2, \ldots$ ($k \in \Pi \cup \Pi^s$ and $(i, \alpha) \in \mathcal{S}$).

Third type of the performance measures is related to the sojourn times of customer $\mathbf{c}^e$ in the service facility including its service times. We define for any $t \geq 0$ and $(i, \alpha) \in \mathcal{S}$,

$$C_{Fi,\alpha}^e(t) \equiv \begin{cases} 1, & \text{if } \boldsymbol{c}^e \text{ stays in the service facility or receives} \\ & \text{a service as an } (i,\alpha)\text{-customer at time } t, \\ 0, & \text{otherwise.} \end{cases}$$

The $\boldsymbol{c}^e$'s sojourn time (i.e., sum of its waiting times and its service times) spent in the service facility as an (i, α)-customer is defined by

$$F_{i,\alpha}^e \;\equiv\; \int_0^\infty C_{Fi,\alpha}^e(t) dt, \quad (i, \alpha) \in \mathcal{S}. \tag{2.14}$$

For $l = 0, 1, 2, \ldots$, we define the two expected sojourn times conditioned on the state of the system at time σ_l^e which are spent by customer $\boldsymbol{c}^e$ in the service facility as an (i, α)-customer after time σ_l^e.

$$F_{i,\alpha}(\boldsymbol{Y}, e, l) \;\equiv\; E\left[\int_{\sigma_l^e}^\infty C_{Fi,\alpha}^e(t) dt \,\middle|\, \boldsymbol{Y}(\sigma_l^e) = \boldsymbol{Y} \right], \quad (i, \alpha) \in \mathcal{S}, \tag{2.15}$$

$$F_{i,\alpha}^I(\boldsymbol{Y}, e, l) \;\equiv\; E\left[\int_{\sigma_l^e}^{\sigma_{l+1}^e} C_{Fi,\alpha}^e(t) dt \,\middle|\, \boldsymbol{Y}(\sigma_l^e) = \boldsymbol{Y} \right], \quad (i, \alpha) \in \mathcal{S}, \tag{2.16}$$

[6]For any event $\mathcal{K}$, let $\mathbf{1}\{\mathcal{K}\} = 1$ if event $\mathcal{K}$ is true, or $\mathbf{1}\{\mathcal{K}\} = 0$ otherwise. The following pieces of the waiting times are necessary mainly to the analytical purposes.

for $\boldsymbol{Y} \in \mathcal{E}$. $F_{i,\alpha}(\boldsymbol{Y}, e, l)$ is the overall expected sojourn time after time σ_l^e whereas $F_{i,\alpha}^I(\boldsymbol{Y}, e, l)$ is the expected sojourn time during a service stage in $[\sigma_l^e, \sigma_{l+1}^e)$. Then the following 'feedback equation' holds.

$$F_{i,\alpha}(\boldsymbol{Y}, e, l) = F_{i,\alpha}^I(\boldsymbol{Y}, e, l) + E[F_{i,\alpha}(\boldsymbol{Y}(\sigma_{l+1}^e), e, l+1)|\boldsymbol{Y}(\sigma_l^e) = \boldsymbol{Y}], \qquad (2.17)$$

for $\boldsymbol{Y} \in \mathcal{E}$ and $l = 0, 1, 2, \dots$ ($(i, \alpha) \in \mathcal{S}$).

3. System States at Polling Instants

In this section, we consider the system at any polling instant, which is a time epoch just when the server selects a station. Let us consider the system with any scheduling algorithm defined in the last section. For any arrival epoch τ, the number of (k, γ)-customers at the first polling instant of station i after τ is denoted by $\nu_{k,\gamma}^i(\tau)$. For any $\boldsymbol{Y} \in \mathcal{E}$, we define

$$\overline{\nu}_{k,\gamma}^i(\boldsymbol{Y}) \equiv E[\nu_{k,\gamma}^i(\tau)|\boldsymbol{Y}(\tau) = \boldsymbol{Y}], \qquad (k, \gamma) \in \mathcal{S} \text{ and } i = 1, \dots, J, \qquad (3.1)$$

$$\overline{\boldsymbol{\nu}}^i(\boldsymbol{Y}) \equiv (\overline{\nu}_{k,\gamma}^i(\boldsymbol{Y}) : (k, \gamma) \in \mathcal{S}) \in \mathcal{R}^{1 \times J_c}, \qquad i = 1, \dots, J, \qquad (3.2)$$

$$\overline{\boldsymbol{\nu}}(\boldsymbol{Y}) \equiv (\overline{\nu}_{k,\gamma}^k(\boldsymbol{Y}) : (k, \gamma) \in \mathcal{S}) \in \mathcal{R}^{1 \times J_c}. \qquad (3.3)$$

3.1. Derivation of the equations for the conditional expected values

Let $N_{i,\alpha,k,\gamma}$ be the number of (k, γ)-customers who arrive during a service period of station i starting with an (i, α)-customer, and who still stay at station k at the service period completion epoch. Then its expected value $\overline{N}_{i,\alpha,k,\gamma}$ and its expected value $\overline{N}_{i,\alpha,k,\gamma}(r)$ conditioned on the remaining service time r of the initial (i, α)-customer, respectively, satisfy

$$\overline{N}_{i,\alpha,k,\gamma} = \begin{cases} \lambda_{k,\gamma} E[S_{i,\alpha}] + p_{i,\alpha,k,\gamma}, & i \in \mathcal{H}_g, \\ \lambda_{k,\gamma} E[S_{i,\alpha}] + p_{i,\alpha,k,\gamma} + \sum_{\beta=1}^{L_i} (\lambda_{i,\beta} E[S_{i,\alpha}] + p_{i,\alpha,i,\beta}) \overline{N}_{i,\beta,k,\gamma}, & k \neq i \in \mathcal{H}_e, \\ 0, & k = i \in \mathcal{H}_e, \end{cases} \qquad (3.4)$$

$$\overline{N}_{i,\alpha,k,\gamma}(r) = \begin{cases} \lambda_{k,\gamma} r + p_{i,\alpha,k,\gamma}, & i \in \mathcal{H}_g, \\ \lambda_{k,\gamma} r + p_{i,\alpha,k,\gamma} + \sum_{\beta=1}^{L_i} (\lambda_{i,\beta} r + p_{i,\alpha,i,\beta}) \overline{N}_{i,\beta,k,\gamma}, & k \neq i \in \mathcal{H}_e, \\ 0, & k = i \in \mathcal{H}_e. \end{cases} \qquad (3.5)$$

For any arrival epoch τ and $\kappa(\tau) \in \Pi$, $\nu_{k,\gamma}^i(\tau)$ is composed of the following customers.

- $\nu_{k,\gamma}^{\kappa(\tau)+1}(\tau)$ is the number of (k, γ)-customers at the first polling instant of the next polled station $\kappa(\tau) + 1$, which is composed of [7]
  - (k, γ)-customers staying in the waiting room at τ,
  - (k, γ)-customers arriving during the current service period $\kappa(\tau)$, and
  - (k, γ)-customer arriving during the following switchover period $\kappa(\tau)^s$.
- $\nu_{k,\gamma}^i(\tau)$ ($i \neq \kappa(\tau) + 1$) is composed of
  - (k, γ)-customers at the polling instant of station $i - 1$ (they vanish if $k = i - 1$),
  - (k, γ)-customers arriving during the service period of station $i - 1$, and
  - (k, γ)-customers arriving during the following switchover period $(i - 1)^s$.

Let $\boldsymbol{Y} = (j, \beta, \kappa_0, a_0, r, \boldsymbol{g}, \boldsymbol{n}, L) \in \mathcal{E}$ be the system state at any arrival epoch where $\boldsymbol{g} = (g_{i,\alpha} : (i, \alpha) \in \mathcal{S})$ and $\boldsymbol{n} = (n_{i,\alpha} : (i, \alpha) \in \mathcal{S})$. We define

$$\mathbf{1}_{i,\alpha}(j, \beta) \equiv \begin{cases} 1, & (j, \beta) = (i, \alpha), \\ 0, & \text{otherwise;} \end{cases} \qquad \mathbf{1}(r) \equiv \begin{cases} 1, & r > 0, \\ 0, & r = 0. \end{cases} \qquad (3.6)$$

[7] If $k = \kappa(\tau) \in \mathcal{H}_e$, the first two terms of the following explanations vanish.

Then for $\kappa_0 \in \Pi$ and $(k,\gamma) \in \mathcal{S}$, we have

$$\overline{\nu}^{\kappa_0+1}_{k,\gamma}(\boldsymbol{Y}) = \begin{cases} n_{k,\gamma} + \mathbf{1}_{k,\gamma}(j,\beta) + \mathbf{1}(r)\overline{N}_{\kappa_0,a_0,k,\gamma}(r) \\ \quad + \sum_{\alpha=1}^{L_{\kappa_0}} g_{\kappa_0,\alpha}\overline{N}_{\kappa_0,\alpha,k,\gamma} + \lambda_{k,\gamma}\overline{s^o_{\kappa_0}}, & \kappa_0 \in \mathcal{H}_g, \\ n_{k,\gamma} + \mathbf{1}_{k,\gamma}(j,\beta) + \mathbf{1}(r)\overline{N}_{\kappa_0,a_0,k,\gamma}(r) \\ \quad + \sum_{\alpha=1}^{L_{\kappa_0}} (g_{\kappa_0,\alpha} + \mathbf{1}_{\kappa_0,\alpha}(j,\beta))\overline{N}_{\kappa_0,\alpha,k,\gamma} + \lambda_{k,\gamma}\overline{s^o_{\kappa_0}}, & k \neq \kappa_0 \in \mathcal{H}_e, \\ \lambda_{\kappa_0,\gamma}\overline{s^o_{\kappa_0}}, & k = \kappa_0 \in \mathcal{H}_e, \end{cases} \tag{3.7}$$

$$\overline{\nu}^{i}_{k,\gamma}(\boldsymbol{Y}) = \begin{cases} \overline{\nu}^{i-1}_{k,\gamma}(\boldsymbol{Y}) + \sum_{\alpha=1}^{L_{i-1}} \overline{\nu}^{i-1}_{i-1,\alpha}(\boldsymbol{Y})\overline{N}_{i-1,\alpha,k,\gamma} + \lambda_{k,\gamma}\overline{s^o_{i-1}}, & k \neq i-1, \\ \sum_{\alpha=1}^{L_{i-1}} \overline{\nu}^{i-1}_{i-1,\alpha}(\boldsymbol{Y})\overline{N}_{i-1,\alpha,i-1,\gamma} + \lambda_{i-1,\gamma}\overline{s^o_{i-1}}, & k = i-1, \end{cases} \tag{3.8}$$

$(i \neq \kappa_0 + 1)$. In a similar manner, for $\kappa_0 \in \Pi^s$ and $(k,\gamma) \in \mathcal{S}$, we have

$$\overline{\nu}^{\kappa_0^-+1}_{k,\gamma}(\boldsymbol{Y}) = (n_{k,\gamma} + \mathbf{1}_{k,\gamma}(j,\beta)) + \lambda_{k,\gamma} r, \tag{3.9}$$

$$\overline{\nu}^{i}_{k,\gamma}(\boldsymbol{Y}) = \begin{cases} \overline{\nu}^{i-1}_{k,\gamma}(\boldsymbol{Y}) + \sum_{\alpha=1}^{L_{i-1}} \overline{\nu}^{i-1}_{i-1,\alpha}(\boldsymbol{Y})\overline{N}_{i-1,\alpha,k,\gamma} + \lambda_{k,\gamma}\overline{s^o_{i-1}}, & k \neq i-1, \\ \sum_{\alpha=1}^{L_{i-1}} \overline{\nu}^{i-1}_{i-1,\alpha}(\boldsymbol{Y})\overline{N}_{i-1,\alpha,i-1,\gamma} + \lambda_{i-1,\gamma}\overline{s^o_{i-1}}, & k = i-1, \end{cases} \tag{3.10}$$

$(i \neq \kappa_0^- + 1)$.

3.2. Solutions of the equations

First we obtain equation (3.11) for $\overline{\nu}^k_{k,\gamma}(\boldsymbol{Y})$ in the following manner.

- Case: $\kappa_0 \in \Pi$. For $k \neq \kappa_0 + 1$, by adding (3.8) for $i \in \Pi_{\kappa_0+2,k+1}$, eq. (3.11) can be obtained. For $k = \kappa_0 + 1$, eq. (3.11) is an identity.
- Case: $\kappa_0 \in \Pi^s$. For $k \neq \kappa_0^- + 1$, by adding (3.10) for $i \in \Pi_{\kappa_0^-+2,k+1}$, eq. (3.11) can be obtained. For $k = \kappa_0^- + 1$, eq. (3.11) is an identity.

Then we have the equations of $\overline{\nu}^k_{k,\gamma}(\boldsymbol{Y})$ $((k,\gamma) \in \mathcal{S})$ for any $\boldsymbol{Y} = (j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n},L) \in \mathcal{E}$.

$$\overline{\nu}^k_{k,\gamma}(\boldsymbol{Y}) = \begin{cases} \overline{\nu}^{\kappa_0+1}_{k,\gamma}(\boldsymbol{Y}) + \sum_{i\in\Pi_{\kappa_0+1,k}} \sum_{\alpha=1}^{L_i} \overline{\nu}^i_{i,\alpha}(\boldsymbol{Y})\overline{N}_{i,\alpha,k,\gamma} + \lambda_{k,\gamma} \sum_{i\in\Pi_{\kappa_0+1,k}} \overline{s^o_i}, & \kappa_0 \in \Pi, \\ \overline{\nu}^{\kappa_0^-+1}_{k,\gamma}(\boldsymbol{Y}) + \sum_{i\in\Pi_{\kappa_0^-+1,k}} \sum_{\alpha=1}^{L_i} \overline{\nu}^i_{i,\alpha}(\boldsymbol{Y})\overline{N}_{i,\alpha,k,\gamma} + \lambda_{k,\gamma} \sum_{i\in\Pi_{\kappa_0^-+1,k}} \overline{s^o_i}, & \kappa_0 \in \Pi^s. \end{cases} \tag{3.11}$$

Let us express these expressions in matrix forms. Let

$$\boldsymbol{N}(\kappa_0) \equiv (N_{i,\alpha,k,\gamma}(\kappa_0) : (i,\alpha),(k,\gamma) \in \mathcal{S}) \in \mathcal{R}^{J_c\times J_c}, \quad \kappa_0 = 1,\dots,J;$$

$$N_{i,\alpha,k,\gamma}(\kappa_0) \equiv \begin{cases} \overline{N}_{i,\alpha,k,\gamma}, & i \in \Pi_{\kappa_0+1,k},\ \alpha = 1,\dots,L_i, \\ 0, & i \notin \Pi_{\kappa_0+1,k},\ \alpha = 1,\dots,L_i, \end{cases} \quad (k,\gamma) \in \mathcal{S};$$

$$\boldsymbol{s}^o_\lambda(\kappa_0) \equiv \left(\lambda_{k,\gamma} \sum_{i\in\Pi_{\kappa_0+1,k}} \overline{s^o_i} : (k,\gamma) \in \mathcal{S}\right) \in \mathcal{R}^{1\times J_c}, \quad \kappa_0 = 1,\dots,J.$$

Then equation (3.11) can be written as

$$\overline{\boldsymbol{\nu}}(\boldsymbol{Y}) = \begin{cases} \overline{\boldsymbol{\nu}}^{\kappa_0+1}(\boldsymbol{Y}) + \overline{\boldsymbol{\nu}}(\boldsymbol{Y})\boldsymbol{N}(\kappa_0) + \boldsymbol{s}^o_\lambda(\kappa_0), & \kappa_0 \in \Pi, \\ \overline{\boldsymbol{\nu}}^{\kappa_0^-+1}(\boldsymbol{Y}) + \overline{\boldsymbol{\nu}}(\boldsymbol{Y})\boldsymbol{N}(\kappa_0^-) + \boldsymbol{s}^o_\lambda(\kappa_0^-), & \kappa_0 \in \Pi^s, \end{cases} \tag{3.12}$$

or[8],

$$\overline{\boldsymbol{\nu}}(\boldsymbol{Y}) = \begin{cases} (\overline{\boldsymbol{\nu}}^{\kappa_0+1}(\boldsymbol{Y}) + \boldsymbol{s}^o_\lambda(\kappa_0))(\boldsymbol{I} - \boldsymbol{N}(\kappa_0))^{-1}, & \kappa_0 \in \Pi, \\ (\overline{\boldsymbol{\nu}}^{\kappa_0^-+1}(\boldsymbol{Y}) + \boldsymbol{s}^o_\lambda(\kappa_0^-))(\boldsymbol{I} - \boldsymbol{N}(\kappa_0^-))^{-1}, & \kappa_0 \in \Pi^s. \end{cases} \tag{3.13}$$

[8]The following inverse matrices can be shown to exist under the assumptions stated in Section 2.

Further from (3.8) and (3.10), we have

$$\overline{\nu}^i_{k,\gamma}(\boldsymbol{Y}) \tag{3.14}$$
$$= \begin{cases} \sum_{\ell\in\Pi_{k,i}} \sum_{\alpha=1}^{L_\ell} \overline{\nu}^\ell_{\ell,\alpha}(\boldsymbol{Y})\overline{N}_{\ell,\alpha,k,\gamma} + \lambda_{k,\gamma}\sum_{\ell\in\Pi_{k,i}} \overline{s^o_\ell}, & \kappa_0 \in \Pi_{i,k}\cup\Pi^s_{i,k}, \\ \overline{\nu}^{\kappa_0+1}_{k,\gamma}(\boldsymbol{Y}) + \sum_{\ell\in\Pi_{\kappa_0+1,i}} \sum_{\alpha=1}^{L_\ell} \overline{\nu}^\ell_{\ell,\alpha}(\boldsymbol{Y})\overline{N}_{\ell,\alpha,k,\gamma} + \lambda_{k,\gamma}\sum_{\ell\in\Pi_{\kappa_0+1,i}} \overline{s^o_\ell}, & \kappa_0 \in \Pi_{k,i}, \\ \overline{\nu}^{\kappa_0^-+1}_{k,\gamma}(\boldsymbol{Y}) + \sum_{\ell\in\Pi_{\kappa_0^-+1,i}} \sum_{\alpha=1}^{L_\ell} \overline{\nu}^\ell_{\ell,\alpha}(\boldsymbol{Y})\overline{N}_{\ell,\alpha,k,\gamma} + \lambda_{k,\gamma}\sum_{\ell\in\Pi_{\kappa_0^-+1,i}} \overline{s^o_\ell}, & \kappa_0 \in \Pi^s_{k,i}, \end{cases}$$

for $k \neq i$, $\boldsymbol{Y} = (j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n},L) \in \mathcal{E}$, $(k,\gamma)\in\mathcal{S}$ and $i = 1,\dots,J$.

Then we have the following linear functional expressions of $\overline{\boldsymbol{\nu}}(\boldsymbol{Y})$ and $\overline{\boldsymbol{\nu}}^i(\boldsymbol{Y})$.

Proposition 1 For any $\boldsymbol{Y} = (j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n},L) \in \mathcal{E}$ and $i = 1,\dots,J$,

$$\overline{\boldsymbol{\nu}}(\boldsymbol{Y}) = (r,\mathbf{1}(r))\boldsymbol{b}^0(\kappa_0,a_0) + (\boldsymbol{g},\boldsymbol{n})\boldsymbol{B}^0(\kappa_0) + \boldsymbol{b}^0_0(\kappa_0,j,\beta), \tag{3.15}$$
$$\overline{\boldsymbol{\nu}}^i(\boldsymbol{Y}) = (r,\mathbf{1}(r))\boldsymbol{b}^i(\kappa_0,a_0) + (\boldsymbol{g},\boldsymbol{n})\boldsymbol{B}^i(\kappa_0) + \boldsymbol{b}^i_0(\kappa_0,j,\beta), \tag{3.16}$$

where the constants $\boldsymbol{B}^m(\kappa_0) \in \mathcal{R}^{2J_c\times J_c}, \boldsymbol{b}^m(\kappa_0,a_0) \in \mathcal{R}^{2\times J_c}, \boldsymbol{b}^m_0(\kappa_0,j,\beta) \in \mathcal{R}^{1\times J_c}$ $(m = 0,1,\dots,J)$ can be determined by the expressions (3.7), (3.9), (3.13) and (3.14). ■

4. Quantities at Each Service Stage

In this section we first derive expressions of $W^I_{i,\alpha}(\cdot,e,l)$, $H^I_{i,\alpha}(\cdot,e,l,k)$ and $F^I_{i,\alpha}(\cdot,e,l)$ of customer $\boldsymbol{c}^e$ spent during its l^{th} service stage. Then we consider the expected numbers of customers at its l^{th} service completion epoch. For any state $\boldsymbol{Y} = (j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n},L) \in \mathcal{E}$ at the arrival epoch σ^e_l, $\boldsymbol{c}^e$ is a (j,β)-customer at this epoch. Hence $W^I_{i,\alpha}(\boldsymbol{Y},e,l) = 0, H^I_{i,\alpha}(\boldsymbol{Y},e,l,k) = 0$ and $F^I_{i,\alpha}(\boldsymbol{Y},e,l) = 0$ for $(i,\alpha) \neq (j,\beta)$, and we consider only the case $(i,\alpha) = (j,\beta)$.

Now let us consider an (i,α)-customer staying at station i ($(i,\alpha)\in\mathcal{S}$). Let $T^\delta_{i,\alpha}$ be the total amount of service times the customer receives until the first time it departs from the set of classes $(i,1),\dots,(i,\delta)$ at station i after at least receiving its initial service as an (i,α)-customer ($(i,\delta)\in\mathcal{S}$). Let $\overline{T}^\delta_{i,\alpha}$ be its expected value and $\overline{T}^\delta_{i,\alpha}(r)$ be its expected value conditioned on its initial remaining service time r as an (i,α)-customer. Then we have[9]:

$$\begin{aligned} \overline{T}^\delta_{i,\alpha} &= E[S_{i,\alpha}] + \textstyle\sum_{\beta=1}^\delta p_{i,\alpha,i,\beta}\overline{T}^\delta_{i,\beta}; \\ \overline{T}^\delta_{i,\alpha}(r) &= r + \textstyle\sum_{\beta=1}^\delta p_{i,\alpha,i,\beta}\overline{T}^\delta_{i,\beta}, \end{aligned} \tag{4.1}$$

for $(i,\alpha)\in\mathcal{S}$ and $\delta = 0,1,\dots,L_i$. We define the following quantities:

$$\varrho^+_{i,\delta} \equiv \textstyle\sum_{\alpha=1}^\delta \lambda_{i,\alpha}\overline{T}^\delta_{i,\alpha}, \qquad (i = 1,\dots,J \text{ and } \delta = 0,1,\dots,L_i), \tag{4.2}$$

$$D(\boldsymbol{Y}) \equiv \begin{cases} r + \sum_{\alpha=1}^{L_{\kappa_0}} g_{\kappa_0,\alpha}E[S_{\kappa_0,\alpha}], & \kappa_0\in\mathcal{H}_g, \\ \dfrac{\mathbf{1}(r)\overline{T}^{L_{\kappa_0}}_{\kappa_0,a_0}(r) + \sum_{\alpha=1}^{L_{\kappa_0}}(g_{\kappa_0,\alpha} + \mathbf{1}_{\kappa_0,\alpha}(j,\beta))\overline{T}^{L_{\kappa_0}}_{\kappa_0,\alpha}}{1-\varrho^+_{\kappa_0,L_{\kappa_0}}}, & \kappa_0\in\mathcal{H}_e, \end{cases} \tag{4.3}$$

$$\delta(i,\alpha) \equiv \begin{cases} E[S_{i,\alpha}], & i\in\mathcal{H}_g, \\ \dfrac{\overline{T}^{L_i}_{i,\alpha}}{1-\varrho^+_{i,L_i}}, & i\in\mathcal{H}_e, \end{cases} \qquad (i,\alpha)\in\mathcal{S}. \tag{4.4}$$

for $\boldsymbol{Y} = (j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n},L) \in \mathcal{E}$ ($\boldsymbol{g} = (g_{i,\alpha} : (i,\alpha)\in\mathcal{S})$ and $\boldsymbol{n} = (n_{i,\alpha} : (i,\alpha)\in\mathcal{S})$). $D(\boldsymbol{Y})$ is the expected remaining length of current service period κ_0 conditioned on the system state $\boldsymbol{Y}$. $\delta(i,\alpha)$ is the expected length of a service period i starting with an (i,α)-customer. Let us consider the system with any scheduling algorithm defined before.

[9]The empty sum which arises when $\delta = 0$ is equal to 0.

4.1. Gated groups

We first derive expressions of the performance measures for $\boldsymbol{c}^e$ arriving at station j with the gated rule ($j \in \mathcal{H}_g$). Let $\boldsymbol{Y} = (j, \beta, \kappa_0, a_0, r, \boldsymbol{g}, \boldsymbol{n}, L) \in \mathcal{E}$ be (a realized value of) the state of the system at the $\boldsymbol{c}^e$'s arrival epoch σ_l^e where $\boldsymbol{g} = (g_{i,\alpha} : (i,\alpha) \in \mathcal{S})$ and $\boldsymbol{n} = (n_{i,\alpha} : (i,\alpha) \in \mathcal{S})$). Hence $\boldsymbol{c}^e$ becomes a (j,β)-customer at this epoch.

Its expected waiting time $W^I_{j,\beta}(\boldsymbol{Y}, e, l)$ is the expected length of time between the arrival epoch and the first polling instant of station j, which is composed of the following periods.

- Case: $\kappa_0 \in \Pi$.
  - The current service period κ_0 and the following switchover period, whose expected remaining length is $D(\boldsymbol{Y}) + \overline{s^o_{\kappa_0}}$.
  - The first service period of station k and the following switchover period, whose expected length is $\sum_{\gamma=1}^{L_k} \overline{\nu}^k_{k,\gamma}(\boldsymbol{Y})\delta(k,\gamma) + \overline{s^o_k}$, for $k = \kappa_0 + 1, \ldots, j-1$.
- Case: $\kappa_0 \in \Pi^s$.
  - The current switchover period, whose expected remaining length is r.
  - The first service period of station k and the following switchover period, whose expected length is $\sum_{\gamma=1}^{L_k} \overline{\nu}^k_{k,\gamma}(\boldsymbol{Y})\delta(k,\gamma) + \overline{s^o_k}$, for $k = \kappa_0^- + 1, \ldots, j-1$.

Then we have

$$W^I_{j,\beta}(\boldsymbol{Y}, e, l) = \begin{cases} D(\boldsymbol{Y}) + \overline{s^o_{\kappa_0}} + \sum\limits_{k \in \Pi_{\kappa_0+1,j}} \left(\sum\limits_{\gamma=1}^{L_k} \overline{\nu}^k_{k,\gamma}(\boldsymbol{Y})\delta(k,\gamma) + \overline{s^o_k} \right), & \kappa_0 \in \Pi, \\ r + \sum\limits_{k \in \Pi_{\kappa_0^-+1,j}} \left(\sum\limits_{\gamma=1}^{L_k} \overline{\nu}^k_{k,\gamma}(\boldsymbol{Y})\delta(k,\gamma) + \overline{s^o_k} \right), & \kappa_0 \in \Pi^s. \end{cases} \tag{4.5}$$

For $k \in \Pi$, $H^I_{j,\beta}(\boldsymbol{Y}, e, l, k)$, which is a piece of $W^I_{j,\beta}(\boldsymbol{Y}, e, l)$, is the expected waiting time while the system is in service period k. Then we have

$$H^I_{j,\beta}(\boldsymbol{Y}, e, l, k) = \begin{cases} \sum_{\gamma=1}^{L_k} \overline{\nu}^k_{k,\gamma}(\boldsymbol{Y})\delta(k,\gamma), & (k \in \Pi_{\kappa_0+1,j} \ \& \ \kappa_0 \in \Pi) \text{ or} \\ & \quad (k \in \Pi_{\kappa_0^-+1,j} \ \& \ \kappa_0 \in \Pi^s), \\ D(\boldsymbol{Y}), & k = \kappa_0 \in \Pi, \\ 0, & \text{otherwise.} \end{cases} \tag{4.6}$$

For $k \in \Pi^s$, $H^I_{j,\beta}(\boldsymbol{Y}, e, l, k)$, which is a piece of $W^I_{j,\beta}(\boldsymbol{Y}, e, l)$, is the expected waiting time while the system is in switchover period k. Then we have

$$H^I_{j,\beta}(\boldsymbol{Y}, e, l, k) = \begin{cases} \overline{s^o_{k^-}}, & (k \in \Pi^s_{\kappa_0+1,j} \ \& \ \kappa_0 \in \Pi) \text{ or } (k \in \Pi^s_{\kappa_0^-+1,j} \ \& \ \kappa_0 \in \Pi^s), \\ \overline{s^o_{k^-}}, & (k = \kappa_0^s \ \& \ \kappa_0 \in \Pi), \\ r, & k = \kappa_0 \in \Pi^s, \\ 0, & \text{otherwise.} \end{cases} \tag{4.7}$$

The expected sojourn time $F^I_{j,\beta}(\boldsymbol{Y}, e, l)$ depends on the service order at station j.

- Case: $j \in \mathcal{H}_{gF}$. $\boldsymbol{c}^e$ starts service when the following customers complete services: j-customers already in the waiting room at the $\boldsymbol{c}^e$'s last arrival epoch σ^e_l.
- Case: $j \in \mathcal{H}_{gP}$. $\boldsymbol{c}^e$ starts service when the following customers complete services: 1) (j,α)-customers staying at station j at its polling instant ($\alpha \leq \beta - 1$); 2) (j,β)-customers already in the waiting room at the $\boldsymbol{c}^e$'s last arrival epoch σ^e_l.

Then we have

$$F^I_{j,\beta}(\boldsymbol{Y},e,l) \;=\; \begin{cases} \displaystyle\sum_{\alpha=1}^{L_j} n_{j,\alpha}E[S_{j,\alpha}] + E[S_{j,\beta}], & j \in \mathcal{H}_{gF},\ \kappa_0 \in \Pi \cup \Pi^s, \\ \displaystyle\sum_{\alpha=1}^{\beta-1} \overline{\nu}^j_{j,\alpha}(\boldsymbol{Y})E[S_{j,\alpha}] + (n_{j,\beta}+1)E[S_{j,\beta}], & j \in \mathcal{H}_{gP},\ \kappa_0 \in \Pi \cup \Pi^s. \end{cases} \tag{4.8}$$

Finally, we would like to obtain the conditional expected values of $(\boldsymbol{g}(\sigma^e_{l+1}), \boldsymbol{n}(\sigma^e_{l+1}))$. It also depends on the service order at station j. The number $n_{k,\gamma}(\sigma^e_{l+1})$ is essentially a sum of the following (k,γ)-customers:

- (k,γ)-customers staying in the system at the first polling instant of station j $(k \neq j)$, and
- (k,γ)-customers arriving from outside or by feedback during the sojourn time (related to) $F^I_{j,\beta}(\boldsymbol{Y},e,l)$.

Then we have

$$\begin{aligned} &E[n_{k,\gamma}(\sigma^e_{l+1})|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}] \qquad\qquad (4.9)\\ &= \begin{cases} \overline{\nu}^j_{k,\gamma}(\boldsymbol{Y}) + \lambda_{k,\gamma}F^I_{j,\beta}(\boldsymbol{Y},e,l) + \sum_{\alpha=1}^{L_j} n_{j,\alpha}p_{j,\alpha,k,\gamma}, & j \in \mathcal{H}_{gF}\ \&\ k \neq j, \\ \overline{\nu}^j_{k,\gamma}(\boldsymbol{Y}) + \lambda_{k,\gamma}F^I_{j,\beta}(\boldsymbol{Y},e,l) + \sum_{\alpha=1}^{\beta-1} \overline{\nu}^j_{j,\alpha}(\boldsymbol{Y})p_{j,\alpha,k,\gamma} + n_{j,\beta}p_{j,\beta,k,\gamma}, & j \in \mathcal{H}_{gP}\ \&\ k \neq j, \\ \lambda_{j,\gamma}F^I_{j,\beta}(\boldsymbol{Y},e,l) + \sum_{\alpha=1}^{L_j} n_{j,\alpha}p_{j,\alpha,j,\gamma}, & j \in \mathcal{H}_{gF}\ \&\ k = j, \\ \lambda_{j,\gamma}F^I_{j,\beta}(\boldsymbol{Y},e,l) + \sum_{\alpha=1}^{\beta-1} \overline{\nu}^j_{j,\alpha}(\boldsymbol{Y})p_{j,\alpha,j,\gamma} + n_{j,\beta}p_{j,\beta,j,\gamma}, & j \in \mathcal{H}_{gP}\ \&\ k = j. \end{cases} \end{aligned}$$

The number $g_{k,\gamma}(\sigma^e_{l+1})$ consists of the following (k,γ)-customers.

- Case: $k \neq j$. None of customers is in the service facility at this epoch.
- Case: $k = j$. Customers staying at the first polling instant of station j and not served before customer $\boldsymbol{c}^e$.

Then we have

$$\begin{aligned} &E[g_{k,\gamma}(\sigma^e_{l+1})|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}] \qquad\qquad (4.10)\\ &= \begin{cases} 0, & k \neq j, \\ \overline{\nu}^j_{j,\gamma}(\boldsymbol{Y}) - (\mathbf{1}_{j,\gamma}(j,\beta) + n_{j,\gamma}), & j \in \mathcal{H}_{gF},\ (k=j), \\ 0, & j \in \mathcal{H}_{gP}\ \&\ \gamma < \beta,\ (k=j), \\ \overline{\nu}^j_{j,\beta}(\boldsymbol{Y}) - (1 + n_{j,\beta}), & j \in \mathcal{H}_{gP}\ \&\ \gamma = \beta,\ (k=j), \\ \overline{\nu}^j_{j,\gamma}(\boldsymbol{Y}), & j \in \mathcal{H}_{gP}\ \&\ \gamma > \beta,\ (k=j). \end{cases} \end{aligned}$$

4.2. Exhaustive groups

We derive expressions of the performance measures for $\boldsymbol{c}^e$ arriving at station j with the exhaustive rule $(j \in \mathcal{H}_e)$. Let $\boldsymbol{Y} = (j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n},L) \in \mathcal{E}$ be (a realized value of) the state of the system at the $\boldsymbol{c}^e$'s arrival epoch σ^e_l where $\boldsymbol{g} = (g_{i,\alpha} : (i,\alpha) \in \mathcal{S})$ and $\boldsymbol{n} = (n_{i,\alpha} : (i,\alpha) \in \mathcal{S})$). Hence $\boldsymbol{c}^e$ becomes a (j,β)-customer at this epoch.

For $j \in \mathcal{H}_{eP}$, let a '(j,δ)-busy period' denote a period until the first time when all customers belonging among classes $(j,1),\ldots,(j,\delta)$ clear $(\delta = 0,1,\ldots,L_j)$. (A $(j,0)$-busy period is defined to be a period until the first time when a group j customer being served currently completes its service.) Then the expected length of a (j,δ)-busy period starting with a (j,α)-customer is given by $\overline{T}^{\delta}_{j,\alpha}/(1-\varrho^+_{j,\delta})$, $(\alpha = 1,\ldots,L_j$ and $\delta = 0,1,\ldots,L_j)$.

The main difference between the exhaustive rules and the gated rules is in the case $\kappa_0 = j$. Then in a manner similar to the gated rules, we have

$$W^I_{j,\beta}(\boldsymbol{Y}, e, l) = \begin{cases} 0, & j = \kappa_0 \in \Pi, \\ D(\boldsymbol{Y}) + \overline{s^o_{\kappa_0}} + \displaystyle\sum_{k \in \Pi_{\kappa_0+1,j}} \left(\sum_{\gamma=1}^{L_k} \overline{\nu}^k_{k,\gamma}(\boldsymbol{Y}) \delta(k,\gamma) + \overline{s^o_k} \right), & j \neq \kappa_0 \in \Pi, \\ r + \displaystyle\sum_{k \in \Pi_{\kappa_0^-+1,j}} \left(\sum_{\gamma=1}^{L_k} \overline{\nu}^k_{k,\gamma}(\boldsymbol{Y}) \delta(k,\gamma) + \overline{s^o_k} \right), & \kappa_0 \in \Pi^s. \end{cases} \tag{4.11}$$

$H^I_{j,\beta}(\boldsymbol{Y}, e, l, k)$ is also obtained in a similar manner. For $\kappa_0 = j$,

$$H^I_{j,\beta}(\boldsymbol{Y}, e, l, k) = 0, \qquad k \in \Pi \cup \Pi^s. \tag{4.12}$$

For $\kappa_0 \neq j$ and $k \in \Pi$, $H^I_{j,\beta}(\boldsymbol{Y}, e, l, k)$, which is a piece of $W^I_{j,\beta}(\boldsymbol{Y}, e, l)$, is the expected waiting time while the system is in service period k. Then we have

$$H^I_{j,\beta}(\boldsymbol{Y}, e, l, k) = \begin{cases} \sum_{\gamma=1}^{L_k} \overline{\nu}^k_{k,\gamma}(\boldsymbol{Y}) \delta(k,\gamma), & (k \in \Pi_{\kappa_0+1,j} \ \&\ \kappa_0 \in \Pi) \text{ or} \\ & \quad (k \in \Pi_{\kappa_0^-+1,j} \ \&\ \kappa_0 \in \Pi^s), \\ D(\boldsymbol{Y}), & k = \kappa_0 \in \Pi, \\ 0, & \text{otherwise.} \end{cases} \tag{4.13}$$

For $\kappa_0 \neq j$ and $k \in \Pi^s$, $H^I_{j,\beta}(\boldsymbol{Y}, e, l, k)$, which is a piece of $W^I_{j,\beta}(\boldsymbol{Y}, e, l)$, is the expected waiting time while the system is in switchover period k. Then we have

$$H^I_{j,\beta}(\boldsymbol{Y}, e, l, k) = \begin{cases} \overline{s^o_{k^-}}, & (k \in \Pi^s_{\kappa_0+1,j} \ \&\ \kappa_0 \in \Pi) \text{ or } (k \in \Pi^s_{\kappa_0^-+1,j} \ \&\ \kappa_0 \in \Pi^s), \\ \overline{s^o_{k^-}}, & (k = \kappa^s_0 \ \&\ \kappa_0 \in \Pi), \\ r, & k = \kappa_0 \in \Pi^s, \\ 0, & \text{otherwise.} \end{cases} \tag{4.14}$$

The expected sojourn time $F^I_{j,\beta}(\boldsymbol{Y}, e, l)$ depends on the service order at station j.

- Case: $j \in \mathcal{H}_{eF}$. $\boldsymbol{c}^e$ starts service when the following customers complete services:
  - j-customers already in the service facility at epoch σ^e_l, (if $\kappa_0 = j$), or
  - j-customers already in the waiting room at epoch σ^e_l, (if $\kappa_0 \neq j$).
- Case: $j \in \mathcal{H}_{eP}$. $\boldsymbol{c}^e$ starts service when a $(j, \beta - 1)$-busy period with the initial works of the following customers completes:
  - a customer being served at σ^e_l, and j-customers belonging among classes $(j, 1)$, ..., (j, β) staying in the service facility at σ^e_l, (if $\kappa_0 = j$), or
  - j-customers belonging among classes $(j, 1), \ldots, (j, \beta - 1)$ staying at station j at its polling instant, and (j, β)-customers staying at the waiting room at σ^e_l, (if $\kappa_0 \neq j$).

Then we have

$$F^I_{j,\beta}(\boldsymbol{Y}, e, l) = \begin{cases} r + \displaystyle\sum_{\alpha=1}^{L_j} g_{j,\alpha} E[S_{j,\alpha}] + E[S_{j,\beta}], & \kappa_0 = j \in \mathcal{H}_{eF}, \\ \displaystyle\sum_{\alpha=1}^{L_j} n_{j,\alpha} E[S_{j,\alpha}] + E[S_{j,\beta}], & \kappa_0 \neq j \in \mathcal{H}_{eF}, \\ \dfrac{\mathbf{1}(r) \overline{T}^{\beta-1}_{j,a_0}(r) + \sum_{\alpha=1}^{\beta} g_{j,\alpha} \overline{T}^{\beta-1}_{j,\alpha}}{1 - \varrho^+_{j,\beta-1}} + E[S_{j,\beta}], & \kappa_0 = j \in \mathcal{H}_{eP}, \\ \dfrac{\sum_{\alpha=1}^{\beta-1} \overline{\nu}^j_{j,\alpha}(\boldsymbol{Y}) \overline{T}^{\beta-1}_{j,\alpha} + n_{j,\beta} \overline{T}^{\beta-1}_{j,\beta}}{1 - \varrho^+_{j,\beta-1}} + E[S_{j,\beta}], & \kappa_0 \neq j \in \mathcal{H}_{eP}. \end{cases} \tag{4.15}$$

Finally, we would like to obtain the conditional expected values of $(\boldsymbol{g}(\sigma^e_{l+1}), \boldsymbol{n}(\sigma^e_{l+1}))$. It also depends on the service orders of customers at station j. Then for $j \in \mathcal{H}_{eF}$, (k,γ)-customers in the waiting room at the completion epoch of the sojourn time (related to) $F^I_{j,\beta}(\boldsymbol{Y}, e, l)$ are composed of

- (k,γ)-customers staying at the beginning of the sojourn time, and
- (k,γ)-customers arriving from outside or by feedback during the sojourn time.

None of customers is in the waiting room of station j at the completion epoch of the sojourn time, since the gate of station j still opened at this time. Then we have

$$E[n_{k,\gamma}(\sigma^e_{l+1})|\boldsymbol{Y}(\sigma^e_l) = \boldsymbol{Y}] \tag{4.16}$$
$$= \begin{cases} n_{k,\gamma} + \lambda_{k,\gamma} F^I_{j,\beta}(\boldsymbol{Y}, e, l) + \mathbf{1}(r) p_{j,a_0,k,\gamma} + \sum_{\alpha=1}^{L_j} g_{j,\alpha} p_{j,\alpha,k,\gamma}, & \kappa_0 = j \ \& \ k \neq j, \\ \overline{\nu}^j_{k,\gamma}(\boldsymbol{Y}) + \lambda_{k,\gamma} F^I_{j,\beta}(\boldsymbol{Y}, e, l) + \sum_{\alpha=1}^{L_j} n_{j,\alpha} p_{j,\alpha,k,\gamma}, & \kappa_0 \neq j \ \& \ k \neq j, \\ 0, & k = j, \end{cases}$$

for $j \in \mathcal{H}_{eF}$. Then (j,γ)-customers in the service facility at the completion epoch of the sojourn time are composed of

- (j,γ)-customers staying at the beginning of the sojourn time and not served before $\boldsymbol{c}^e$, and
- (j,γ)-customers arriving from outside or by feedback during the sojourn time.

Obviously none of k-customers $(k \neq j)$ is in the service facility at the completion epoch of the sojourn time. Then we have

$$E[g_{k,\gamma}(\sigma^e_{l+1})|\boldsymbol{Y}(\sigma^e_l) = \boldsymbol{Y}] \tag{4.17}$$
$$= \begin{cases} 0, & k \neq j, \\ \lambda_{j,\gamma} F^I_{j,\beta}(\boldsymbol{Y}, e, l) + \mathbf{1}(r) p_{j,a_0,j,\gamma} + \sum_{\alpha=1}^{L_j} g_{j,\alpha} p_{j,\alpha,j,\gamma}, & \kappa_0 = j \ \& \ k = j, \\ \overline{\nu}^j_{j,\gamma}(\boldsymbol{Y}) - n_{j,\gamma} - \mathbf{1}_{j,\gamma}(j,\beta) + \lambda_{j,\gamma} F^I_{j,\beta}(\boldsymbol{Y}, e, l) + \sum_{\alpha=1}^{L_j} n_{j,\alpha} p_{j,\alpha,j,\gamma}, & \kappa_0 \neq j \ \& \ k = j, \end{cases}$$

for $j \in \mathcal{H}_{eF}$.

For $j \in \mathcal{H}_{eP}$, let $N^\delta_{j,\alpha,k,\gamma}$ be the number of (k,γ)-customers who arrive (from outside or by feedback) during a (j,δ)-busy period starting with a (j,α)-customer, and who still stay at station k at the end of the period (($k \neq j$ and $\gamma = 1, \ldots, L_k$) or ($k = j$ and $\gamma = \delta+1, \ldots, L_j$)). Then let $\overline{N}^\delta_{j,\alpha,k,\gamma}$ be its expected value and let $\overline{N}^\delta_{j,\alpha,k,\gamma}(r)$ be its expected value conditioned on the remaining service time r of the initial (j,α)-customer. Then

$$\overline{N}^\delta_{j,\alpha,k,\gamma} = \lambda_{k,\gamma} E[S_{j,\alpha}] + p_{j,\alpha,k,\gamma} + \sum_{\delta'=1}^{\delta} (\lambda_{j,\delta'} E[S_{j,\alpha}] + p_{j,\alpha,j,\delta'}) \overline{N}^\delta_{j,\delta',k,\gamma}, \tag{4.18}$$

$$\overline{N}^\delta_{j,\alpha,k,\gamma}(r) = \lambda_{k,\gamma} r + p_{j,\alpha,k,\gamma} + \sum_{\delta'=1}^{\delta} (\lambda_{j,\delta'} r + p_{j,\alpha,j,\delta'}) \overline{N}^\delta_{j,\delta',k,\gamma}, \tag{4.19}$$

for $(j,\alpha) \in \mathcal{S}$; $\delta = 0, 1, \ldots, L_j$, and $(k,\gamma) \in \mathcal{S}$ ($\gamma > \delta$ if $k = j$). Note that the $\boldsymbol{c}^e$'s waiting times related to $F^I_{j,\beta}(\cdot, e, l)$

- from its arrival epoch to the start of its service (for $\kappa_0 = j$), and
- from the polling instant of station j to the start of its service (for $\kappa_0 \neq j$)

are $(j, \beta-1)$-busy periods with initial works of customers with higher priorities than the $\boldsymbol{c}^e$'s staying in the service facility at the beginning of the waiting times. (Recall the explanations of eq. (4.15) for $j \in \mathcal{H}_{eP}$.) Then in a manner similar to eq. (4.16), we have

$$E[n_{k,\gamma}(\sigma^e_{l+1})|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}] \tag{4.20}$$
$$= \begin{cases} n_{k,\gamma} + \mathbf{1}(r)\overline{N}^{\beta-1}_{j,a_0,k,\gamma}(r) + \sum_{\alpha=1}^{\beta} g_{j,\alpha}\overline{N}^{\beta-1}_{j,\alpha,k,\gamma} + \lambda_{k,\gamma}E[S_{j,\beta}], & k \neq j \ \&\ \kappa_0 = j, \\ \overline{\nu}^j_{k,\gamma}(\boldsymbol{Y}) + \sum_{\alpha=1}^{\beta-1} \overline{\nu}^j_{j,\alpha}(\boldsymbol{Y})\overline{N}^{\beta-1}_{j,\alpha,k,\gamma} + n_{j,\beta}\overline{N}^{\beta-1}_{j,\beta,k,\gamma} + \lambda_{k,\gamma}E[S_{j,\beta}], & k \neq j \ \&\ \kappa_0 \neq j, \\ 0, & k = j, \end{cases}$$

for $j \in \mathcal{H}_{eP}$. (Note that none of customers arrive by feedback during $\boldsymbol{c}^e$'s service.) Further the explanation for (j,γ)-customers in the service facility at the completion epoch of the sojourn time (related to) $F^I_{j,\beta}(\boldsymbol{Y},e,l)$ is similar to that for $j \in \mathcal{H}_{eF}$ except that in this case (j,γ)-customers $(\gamma < \beta)$ are cleared from the system when $\boldsymbol{c}^e$ starts service. Then we have

$$E[g_{k,\gamma}(\sigma^e_{l+1})|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}] \tag{4.21}$$
$$= \begin{cases} 0, & k \neq j, \\ \lambda_{j,\gamma}E[S_{j,\beta}], & k = j \ \&\ \gamma < \beta, \\ \mathbf{1}(r)\overline{N}^{\beta-1}_{j,a_0,j,\beta}(r) + \sum_{\alpha=1}^{\beta} g_{j,\alpha}\overline{N}^{\beta-1}_{j,\alpha,j,\beta} + \lambda_{j,\beta}E[S_{j,\beta}], & k = j \ \&\ \gamma = \beta \ \&\ \kappa_0 = j, \\ g_{j,\gamma} + \mathbf{1}(r)\overline{N}^{\beta-1}_{j,a_0,j,\gamma}(r) + \sum_{\alpha=1}^{\beta} g_{j,\alpha}\overline{N}^{\beta-1}_{j,\alpha,j,\gamma} + \lambda_{j,\gamma}E[S_{j,\beta}], & k = j \ \&\ \gamma > \beta \ \&\ \kappa_0 = j, \\ \overline{\nu}^j_{j,\beta}(\boldsymbol{Y}) - n_{j,\beta} - 1 & \\ \quad + \sum_{\alpha=1}^{\beta-1} \overline{\nu}^j_{j,\alpha}(\boldsymbol{Y})\overline{N}^{\beta-1}_{j,\alpha,j,\beta} + n_{j,\beta}\overline{N}^{\beta-1}_{j,\beta,j,\beta} + \lambda_{j,\beta}E[S_{j,\beta}], & k = j \ \&\ \gamma = \beta \ \&\ \kappa_0 \neq j, \\ \overline{\nu}^j_{j,\gamma}(\boldsymbol{Y}) & \\ \quad + \sum_{\alpha=1}^{\beta-1} \overline{\nu}^j_{j,\alpha}(\boldsymbol{Y})\overline{N}^{\beta-1}_{j,\alpha,j,\gamma} + n_{j,\beta}\overline{N}^{\beta-1}_{j,\beta,j,\gamma} + \lambda_{j,\gamma}E[S_{j,\beta}], & k = j \ \&\ \gamma > \beta \ \&\ \kappa_0 \neq j, \end{cases}$$

for $j \in \mathcal{H}_{eP}$.

4.3. Linear functional expressions of the quantities

From the analysis in this section, we can easily see the following important properties:

- The component $(j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n})$ of state $\boldsymbol{Y}=(j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n},L) \in \mathcal{E}$ at $\boldsymbol{c}^e$'s arrival epoch σ^e_l is sufficient to derive $W^I_{j,\beta}(\boldsymbol{Y},e,l)$, $H^I_{j,\beta}(\boldsymbol{Y},e,l,k)$, $F^I_{j,\beta}(\boldsymbol{Y},e,l)$ and the expected vector of the numbers of customers $E[(\boldsymbol{g}(\sigma^e_{l+1}),\boldsymbol{n}(\sigma^e_{l+1}))|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}]$.
- These performance measures and the conditional expected numbers of customers at σ^e_{l+1} are linear with respect to the component $(\boldsymbol{g},\boldsymbol{n})$.

Proposition 2 Let $\boldsymbol{Y}=(j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n},L) \in \mathcal{E}$, $e=1,2,\ldots$, and $l=0,1,2,\ldots$. Then

$$W^I_{j,\beta}(\boldsymbol{Y},e,l) = (r,\mathbf{1}(r))\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,0) + (\boldsymbol{g},\boldsymbol{n})\boldsymbol{w}^{j,\beta}(\kappa_0,0) + w^{j,\beta}(\kappa_0,0), \tag{4.22}$$

$$H^I_{j,\beta}(\boldsymbol{Y},e,l,k) = \begin{cases} (r,\mathbf{1}(r))\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,k) + (\boldsymbol{g},\boldsymbol{n})\boldsymbol{w}^{j,\beta}(\kappa_0,k) + w^{j,\beta}(\kappa_0,k), & k \in \Pi, \\ (r,\mathbf{1}(r))\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,k) + w^{j,\beta}(\kappa_0,k), & k \in \Pi^s, \end{cases} \tag{4.23}$$

$$F^I_{j,\beta}(\boldsymbol{Y},e,l) = (r,\mathbf{1}(r))\boldsymbol{\eta}^{j,\beta}(\kappa_0,a_0) + (\boldsymbol{g},\boldsymbol{n})\boldsymbol{f}^{j,\beta}(\kappa_0) + f^{j,\beta}(\kappa_0), \tag{4.24}$$

$$W^I_{i,\alpha}(\boldsymbol{Y},e,l)=0,\ H^I_{i,\alpha}(\boldsymbol{Y},e,l,k)=0,\ F^I_{i,\alpha}(\boldsymbol{Y},e,l)=0, \quad ((i,\alpha)\neq(j,\beta)). \tag{4.25}$$

Further we have

$$E[(\boldsymbol{g}(\sigma^e_{l+1}),\boldsymbol{n}(\sigma^e_{l+1}))|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}] = (r,\mathbf{1}(r))\boldsymbol{v}^{j,\beta}(\kappa_0,a_0) + (\boldsymbol{g},\boldsymbol{n})\boldsymbol{U}^{j,\beta}(\kappa_0) + \boldsymbol{u}^{j,\beta}(\kappa_0). \tag{4.26}$$

The above coefficients:

$$\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,k), \boldsymbol{w}^{j,\beta}(\kappa_0,k), w^{j,\beta}(\kappa_0,k), \boldsymbol{\eta}^{j,\beta}(\kappa_0,a_0), \boldsymbol{f}^{j,\beta}(\kappa_0), f^{j,\beta}(\kappa_0),$$
$$\boldsymbol{v}^{j,\beta}(\kappa_0,a_0), \boldsymbol{U}^{j,\beta}(\kappa_0), \boldsymbol{u}^{j,\beta}(\kappa_0), \qquad ((j,\beta)\in\mathcal{S},\ (\kappa_0,a_0)\in\mathcal{S}_A \text{ and } k \in \{0\}\cup\Pi\cup\Pi^s)$$

for every scheduling algorithm can be determined from the given system parameters by using the expressions given in this section. ■

5. Expressions of the Performance Measures

In this section, we obtain the expressions of the performance measures defined in Section 2.

Let $\hat{\boldsymbol{w}}_{i,\alpha}(j,\beta,k), \hat{\boldsymbol{f}}_{i,\alpha}(j,\beta) \in \mathcal{R}^{2J_c\times 1}$ $((i,\alpha),(j,\beta)\in\mathcal{S};\ k\in\{0\}\cup\Pi)$ be the solutions of the following set of equations:

$$\begin{aligned}\hat{\boldsymbol{w}}_{i,\alpha}(j,\beta,k) &= p_{j,\beta,i,\alpha}\boldsymbol{w}^{i,\alpha}(j,k)+\sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\boldsymbol{U}^{m,\delta}(j)\hat{\boldsymbol{w}}_{i,\alpha}(m,\delta,k), &(5.1)\\ \hat{\boldsymbol{f}}_{i,\alpha}(j,\beta) &= p_{j,\beta,i,\alpha}\boldsymbol{f}^{i,\alpha}(j)+\sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\boldsymbol{U}^{m,\delta}(j)\hat{\boldsymbol{f}}_{i,\alpha}(m,\delta), &(5.2)\end{aligned}$$

where $\boldsymbol{w}^{i,\alpha}(j,k), \boldsymbol{f}^{i,\alpha}(j)$ and $\boldsymbol{U}^{m,\delta}(j)$ are given in (4.22), (4.23), (4.24) and (4.26). Further let $\hat{w}_{i,\alpha}(j,\beta,k)$, $\hat{f}_{i,\alpha}(j,\beta)\in\mathcal{R}$ $((i,\alpha),(j,\beta)\in\mathcal{S};\ k\in\{0\}\cup\Pi\cup\Pi^s)$ be the solutions of

$$\hat{w}_{i,\alpha}(j,\beta,k) = \begin{cases}\sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\hat{w}_{i,\alpha}(m,\delta,k)+p_{j,\beta,i,\alpha}w^{i,\alpha}(j,k)\\ \qquad +\sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\boldsymbol{u}^{m,\delta}(j)\hat{\boldsymbol{w}}_{i,\alpha}(m,\delta,k), \quad k\in\{0\}\cup\Pi,\\ \sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\hat{w}_{i,\alpha}(m,\delta,k)+p_{j,\beta,i,\alpha}w^{i,\alpha}(j,k), \quad k\in\Pi^s,\end{cases} \tag{5.3}$$

$$\hat{f}_{i,\alpha}(j,\beta) = p_{j,\beta,i,\alpha}f^{i,\alpha}(j)+\sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\left\{\boldsymbol{u}^{m,\delta}(j)\hat{\boldsymbol{f}}_{i,\alpha}(m,\delta)+\hat{f}_{i,\alpha}(m,\delta)\right\}, \tag{5.4}$$

where $w^{i,\alpha}(j,k), f^{i,\alpha}(j)$ and $\boldsymbol{u}^{m,\delta}(j)$ are given in (4.22), (4.23), (4.24) and (4.26). Let define constants:

$$\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,k) \equiv \begin{cases}\boldsymbol{1}_{i,\alpha}(j,\beta)\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,k)+\boldsymbol{v}^{j,\beta}(\kappa_0,a_0)\hat{\boldsymbol{w}}_{i,\alpha}(j,\beta,k),\\ \qquad\qquad k\in\{0\}\cup\Pi,\\ \boldsymbol{1}_{i,\alpha}(j,\beta)\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,k), \qquad k\in\Pi^s,\end{cases} \tag{5.5}$$

$$\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,k) \equiv \boldsymbol{1}_{i,\alpha}(j,\beta)\boldsymbol{w}^{j,\beta}(\kappa_0,k)+\boldsymbol{U}^{j,\beta}(\kappa_0)\hat{\boldsymbol{w}}_{i,\alpha}(j,\beta,k), \quad k\in\{0\}\cup\Pi, \tag{5.6}$$

$$w_{i,\alpha}(j,\beta,\kappa_0,k) \equiv \begin{cases}\boldsymbol{1}_{i,\alpha}(j,\beta)w^{j,\beta}(\kappa_0,k)+\boldsymbol{u}^{j,\beta}(\kappa_0)\hat{\boldsymbol{w}}_{i,\alpha}(j,\beta,k)+\hat{w}_{i,\alpha}(j,\beta,k),\\ \qquad\qquad k\in\{0\}\cup\Pi,\\ \boldsymbol{1}_{i,\alpha}(j,\beta)w^{j,\beta}(\kappa_0,k)+\hat{w}_{i,\alpha}(j,\beta,k), \qquad k\in\Pi^s,\end{cases} \tag{5.7}$$

$$\begin{aligned}\boldsymbol{\eta}_{i,\alpha}(j,\beta,\kappa_0,a_0) &\equiv \boldsymbol{1}_{i,\alpha}(j,\beta)\boldsymbol{\eta}^{j,\beta}(\kappa_0,a_0)+\boldsymbol{v}^{j,\beta}(\kappa_0,a_0)\hat{\boldsymbol{f}}_{i,\alpha}(j,\beta), &(5.8)\\ \boldsymbol{f}_{i,\alpha}(j,\beta,\kappa_0) &\equiv \boldsymbol{1}_{i,\alpha}(j,\beta)\boldsymbol{f}^{j,\beta}(\kappa_0)+\boldsymbol{U}^{j,\beta}(\kappa_0)\hat{\boldsymbol{f}}_{i,\alpha}(j,\beta), &(5.9)\\ f_{i,\alpha}(j,\beta,\kappa_0) &\equiv \boldsymbol{1}_{i,\alpha}(j,\beta)f^{j,\beta}(\kappa_0)+\boldsymbol{u}^{j,\beta}(\kappa_0)\hat{\boldsymbol{f}}_{i,\alpha}(j,\beta)+\hat{f}_{i,\alpha}(j,\beta), &(5.10)\end{aligned}$$

for $(i,\alpha),(j,\beta)\in\mathcal{S}$, and $(\kappa_0,a_0)\in\mathcal{S}_A$. Then it can be shown that

$$\begin{aligned}&\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,0)=\textstyle\sum_{k\in\Pi\cup\Pi^s}\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,k);\\ &\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,0)=\textstyle\sum_{k\in\Pi}\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,k); \qquad (\kappa_0,a_0)\in\mathcal{S}_A,\ (i,\alpha),(j,\beta)\in\mathcal{S}.\\ &w_{i,\alpha}(j,\beta,\kappa_0,0)=\textstyle\sum_{k\in\Pi\cup\Pi^s}w_{i,\alpha}(j,\beta,\kappa_0,k);\end{aligned}$$

Then we define

Expressions of the performance measures:

$$\hat{W}_{i,\alpha}(\boldsymbol{Y},e,l) = (r,\boldsymbol{1}(r))\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,0)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,0)+w_{i,\alpha}(j,\beta,\kappa_0,0), \tag{5.11}$$

$$\hat{H}_{i,\alpha}(\boldsymbol{Y},e,l,k) = \begin{cases}(r,\boldsymbol{1}(r))\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,k)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,k)+w_{i,\alpha}(j,\beta,\kappa_0,k), & k\in\Pi,\\ (r,\boldsymbol{1}(r))\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,k)+w_{i,\alpha}(j,\beta,\kappa_0,k), & k\in\Pi^s,\end{cases} \tag{5.12}$$

$$\hat{F}_{i,\alpha}(\boldsymbol{Y},e,l)=(r,\boldsymbol{1}(r))\boldsymbol{\eta}_{i,\alpha}(j,\beta,\kappa_0,a_0)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{f}_{i,\alpha}(j,\beta,\kappa_0)+f_{i,\alpha}(j,\beta,\kappa_0), \tag{5.13}$$

for $\boldsymbol{Y} = (j, \beta, \kappa_0, a_0, r, \boldsymbol{g}, \boldsymbol{n}, L) \in \mathcal{E}$, $e = 1, 2, \ldots$, $l = 0, 1, 2, \ldots$ and $(i, \alpha) \in \mathcal{S}$. ∎

Then we have

Proposition 3 Let us consider the system defined in Section 2. Then $\hat{W}_{i,\alpha}(\cdot)$ defined in (5.11) satisfies equation (2.9), $\hat{H}_{i,\alpha}(\cdot, k)$ defined in (5.12) satisfies equation (2.13), and $\hat{F}_{i,\alpha}(\cdot)$ defined in (5.13) satisfies equation (2.17) ($(i, \alpha) \in \mathcal{S}$; $k \in \Pi \cup \Pi^s$).

Proof: See Appendix B. ∎

Note It can be shown that equations (2.9), (2.13) and (2.17) have unique solutions. Then $W_{i,\alpha}(\cdot) = \hat{W}_{i,\alpha}(\cdot)$, $H_{i,\alpha}(\cdot, k) = \hat{H}_{i,\alpha}(\cdot, k)$ and $F_{i,\alpha}(\cdot) = \hat{F}_{i,\alpha}(\cdot)$, and equations (5.11)–(5.13) give the linear functional expressions of the performance measures defined in Section 2. (Since the proof of uniqueness is similar to that in [14], we omit it.) ∎

6. Steady State Values

Let us consider the system operating under any scheduling algorithm defined in Section 2. In this section, we would like to evaluate the values:

$$\bar{w}_{i,\alpha}(j, \beta) \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{e=1}^{N} E[W^e_{i,\alpha} + F^e_{i,\alpha} | (X(\sigma^e_0), \Gamma(\sigma^e_0)) = (j, \beta)], \tag{6.1}$$

which denotes the average sojourn time of (i, α)-customers arriving from outside the system as (j, β)-customers ($(i, \alpha), (j, \beta) \in \mathcal{S}$)[10].

For analytical convenience, we define

$$\bar{W}_{i,\alpha}(j, \beta, \kappa_0, a_0) \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{e=1}^{N} E[W^e_{i,\alpha} \cdot \mathbf{1}^e(\kappa_0, a_0) | (X(\sigma^e_0), \Gamma(\sigma^e_0)) = (j, \beta)], \tag{6.2}$$

$$\bar{H}_{i,\alpha}(j, \beta, \kappa_0, a_0, k) \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{e=1}^{N} E[H^e_{i,\alpha}(k) \cdot \mathbf{1}^e(\kappa_0, a_0) | (X(\sigma^e_0), \Gamma(\sigma^e_0)) = (j, \beta)], \tag{6.3}$$

$$\bar{F}_{i,\alpha}(j, \beta, \kappa_0, a_0) \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{e=1}^{N} E[F^e_{i,\alpha} \cdot \mathbf{1}^e(\kappa_0, a_0) | (X(\sigma^e_0), \Gamma(\sigma^e_0)) = (j, \beta)], \tag{6.4}$$

for $(i, \alpha), (j, \beta) \in \mathcal{S}$, $(\kappa_0, a_0) \in \mathcal{S}_A$, and $k \in \Pi \cup \Pi^s$ where $\mathbf{1}^e(\kappa_0, a_0) = \mathbf{1}\{(\kappa(\sigma^e_0), a(\sigma^e_0)) = (\kappa_0, a_0)\}$. The time average values of the state of the system are defined by:

$$\begin{aligned} \tilde{\boldsymbol{Y}}^{k,\gamma} &\equiv (\tilde{X}^{k,\gamma}, \tilde{\Gamma}^{k,\gamma}, k\tilde{q}^{k,\gamma}, \gamma\tilde{q}^{k,\gamma}, \tilde{r}^{k,\gamma}, \tilde{\boldsymbol{g}}^{k,\gamma}, \tilde{\boldsymbol{n}}^{k,\gamma}, \tilde{L}^{k,\gamma}) \\ &\equiv \lim_{t\to\infty} \frac{1}{t} \int_0^t E[\boldsymbol{Y}(s)\mathbf{1}\{(\kappa(s), a(s)) = (k, \gamma)\}]ds, \end{aligned} \tag{6.5}$$

$$\tilde{\boldsymbol{Y}}^{k} \equiv (\tilde{X}^{k}, \tilde{\Gamma}^{k}, k\tilde{q}^{k}, \tilde{a}^{k}, \tilde{r}^{k}, \tilde{\boldsymbol{g}}^{k}, \tilde{\boldsymbol{n}}^{k}, \tilde{L}^{k}) \equiv \lim_{t\to\infty} \frac{1}{t} \int_0^t E[\boldsymbol{Y}(s)\mathbf{1}\{\kappa(s) = k\}]ds, \tag{6.6}$$

for $(k, \gamma) \in \mathcal{S}_A$, where $\tilde{\boldsymbol{g}}^{k,\gamma} \equiv (\tilde{g}^{k,\gamma}_{i,\alpha} : (i, \alpha) \in \mathcal{S})$, $\tilde{\boldsymbol{n}}^{k,\gamma} \equiv (\tilde{n}^{k,\gamma}_{i,\alpha} : (i, \alpha) \in \mathcal{S})$, $\tilde{\boldsymbol{g}}^{k} \equiv (\tilde{g}^{k}_{i,\alpha} : (i, \alpha) \in \mathcal{S})$ and $\tilde{\boldsymbol{n}}^{k} \equiv (\tilde{n}^{k}_{i,\alpha} : (i, \alpha) \in \mathcal{S})$.

6.1. Derivation of the mean sojourn times

Let $\Lambda_{i,\alpha}$ be the composite arrival rate of (i, α)-customers. The set of the values satisfies

$$\Lambda_{i,\alpha} = \lambda_{i,\alpha} + \sum_{j=1}^{J} \sum_{\beta=1}^{L_j} \Lambda_{j,\beta} p_{j,\beta,i,\alpha}, \qquad (i, \alpha) \in \mathcal{S}. \tag{6.7}$$

[10]All related limits (time averages and customer averages) defined in this section are assumed to exist.

Then steady state values $\tilde{q}^{\kappa_0,a_0}$, which is the long-run fraction of time that the server's status is (κ_0, a_0), are calculated as follows:

$$\tilde{q}^{\kappa_0,a_0} = \begin{cases} \Lambda_{\kappa_0,a_0} E[S_{\kappa_0,a_0}], & (\kappa_0, a_0) \in \mathcal{S}, \\ (1-\rho)\overline{s^o_{\kappa_0^-}} \Big/ \left(\sum_{\ell=1}^J \overline{s^o_\ell}\right), & \kappa_0 \in \Pi^s,\ a_0 = 0. \end{cases} \tag{6.8}$$

The steady state values $\tilde{r}^{\kappa_0,a_0}$ can be calculated as follows:

$$\tilde{r}^{\kappa_0,a_0} = \begin{cases} \Lambda_{\kappa_0,a_0} \overline{s^2_{\kappa_0,a_0}}/2, & (\kappa_0, a_0) \in \mathcal{S}, \\ (1-\rho)\overline{s^{o2}_{\kappa_0^-}} \Big/ \left(2\sum_{\ell=1}^J \overline{s^o_\ell}\right), & \kappa_0 \in \Pi^s,\ a_0 = 0. \end{cases} \tag{6.9}$$

From the results in the last section, equations (6.2)–(6.5), and the PASTA property, we obtain

$$\begin{aligned} \bar{W}_{i,\alpha}(j,\beta,\kappa_0,a_0) &= (\tilde{r}^{\kappa_0,a_0}, \tilde{q}^{\kappa_0,a_0})\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,0) \\ &\quad + (\tilde{\boldsymbol{g}}^{\kappa_0,a_0}, \tilde{\boldsymbol{n}}^{\kappa_0,a_0})\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,0) + \tilde{q}^{\kappa_0,a_0} w_{i,\alpha}(j,\beta,\kappa_0,0), \end{aligned} \tag{6.10}$$

$$\bar{H}_{i,\alpha}(j,\beta,\kappa_0,a_0,k) = \begin{cases} (\tilde{r}^{\kappa_0,a_0}, \tilde{q}^{\kappa_0,a_0})\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,k) \\ \quad + (\tilde{\boldsymbol{g}}^{\kappa_0,a_0}, \tilde{\boldsymbol{n}}^{\kappa_0,a_0})\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,k) + \tilde{q}^{\kappa_0,a_0} w_{i,\alpha}(j,\beta,\kappa_0,k), & (k \in \Pi), \\ (\tilde{r}^{\kappa_0,a_0}, \tilde{q}^{\kappa_0,a_0})\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,k) + \tilde{q}^{\kappa_0,a_0} w_{i,\alpha}(j,\beta,\kappa_0,k), & (k \in \Pi^s), \end{cases} \tag{6.11}$$

$$\begin{aligned} \bar{F}_{i,\alpha}(j,\beta,\kappa_0,a_0) &= (\tilde{r}^{\kappa_0,a_0}, \tilde{q}^{\kappa_0,a_0})\boldsymbol{\eta}_{i,\alpha}(j,\beta,\kappa_0,a_0) \\ &\quad + (\tilde{\boldsymbol{g}}^{\kappa_0,a_0}, \tilde{\boldsymbol{n}}^{\kappa_0,a_0})\boldsymbol{f}_{i,\alpha}(j,\beta,\kappa_0) + \tilde{q}^{\kappa_0,a_0} f_{i,\alpha}(j,\beta,\kappa_0). \end{aligned} \tag{6.12}$$

Now we use the generalized Little's formula ($H = \lambda G$) [33, 34] to obtain

$$\tilde{n}^k_{i,\alpha} = \sum_{(j,\beta)\in\mathcal{S}} \lambda_{j,\beta} \sum_{(\kappa_0,a_0)\in\mathcal{S}_A} \bar{H}_{i,\alpha}(j,\beta,\kappa_0,a_0,k), \tag{6.13}$$

$$\tilde{g}^k_{i,\alpha} = \begin{cases} \sum_{(j,\beta)\in\mathcal{S}} \lambda_{j,\beta} \sum_{(\kappa_0,a_0)\in\mathcal{S}_A} \bar{F}_{i,\alpha}(j,\beta,\kappa_0,a_0) - \tilde{q}^{i,\alpha}, & (k = i), \\ 0, & (k \neq i), \end{cases} \tag{6.14}$$

$((i,\alpha) \in \mathcal{S},\ k \in \Pi \cup \Pi^s)$. Hence we can get the following set of equations of $\tilde{\boldsymbol{g}}^k$ and $\tilde{\boldsymbol{n}}^k$.

$$\tilde{n}^k_{i,\alpha} = \begin{cases} \tilde{\varphi}_{i,\alpha}(k) + \sum_{(j,\beta)\in\mathcal{S}} \lambda_{j,\beta} \sum_{\kappa_0\in\Pi\cup\Pi^s} (\tilde{\boldsymbol{g}}^{\kappa_0}, \tilde{\boldsymbol{n}}^{\kappa_0})\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,k), & (k \in \Pi), \\ \tilde{\varphi}_{i,\alpha}(k), & (k \in \Pi^s), \end{cases} \tag{6.15}$$

$$\tilde{g}^k_{i,\alpha} = \begin{cases} \tilde{\eta}_{i,\alpha} + \sum_{(j,\beta)\in\mathcal{S}} \lambda_{j,\beta} \sum_{\kappa_0\in\Pi\cup\Pi^s} (\tilde{\boldsymbol{g}}^{\kappa_0}, \tilde{\boldsymbol{n}}^{\kappa_0})\boldsymbol{f}_{i,\alpha}(j,\beta,\kappa_0), & (k = i), \\ 0, & (k \neq i \ \text{ or } k \in \Pi^s), \end{cases} \tag{6.16}$$

where

$$\tilde{\varphi}_{i,\alpha}(k) \equiv \sum_{(j,\beta)\in\mathcal{S}} \lambda_{j,\beta} \sum_{(\kappa_0,a_0)\in\mathcal{S}_A} \left\{(\tilde{r}^{\kappa_0,a_0}, \tilde{q}^{\kappa_0,a_0})\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,k) + \tilde{q}^{\kappa_0,a_0} w_{i,\alpha}(j,\beta,\kappa_0,k)\right\};$$

$$\tilde{\eta}_{i,\alpha} \equiv \sum_{(j,\beta)\in\mathcal{S}} \lambda_{j,\beta} \sum_{(\kappa_0,a_0)\in\mathcal{S}_A} \left\{(\tilde{r}^{\kappa_0,a_0}, \tilde{q}^{\kappa_0,a_0})\boldsymbol{\eta}_{i,\alpha}(j,\beta,\kappa_0,a_0) + \tilde{q}^{\kappa_0,a_0} f_{i,\alpha}(j,\beta,\kappa_0)\right\} - \tilde{q}^{i,\alpha},$$

for $(i,\alpha) \in \mathcal{S}$ and $k \in \Pi \cup \Pi^s$.

Proposition 4 The average sojourn times are given by

$$\begin{aligned} \bar{w}_{i,\alpha}(j,\beta) &= \sum_{(\kappa_0,a_0)\in\mathcal{S}_A} (\tilde{r}^{\kappa_0,a_0}, \tilde{q}^{\kappa_0,a_0}) \left(\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,0) + \boldsymbol{\eta}_{i,\alpha}(j,\beta,\kappa_0,a_0)\right) \\ &\quad + \sum_{\kappa_0\in\Pi\cup\Pi^s} (\tilde{\boldsymbol{g}}^{\kappa_0}, \tilde{\boldsymbol{n}}^{\kappa_0}) \left(\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,0) + \boldsymbol{f}_{i,\alpha}(j,\beta,\kappa_0)\right) \\ &\quad + \sum_{(\kappa_0,a_0)\in\mathcal{S}_A} \tilde{q}^{\kappa_0,a_0} \left(w_{i,\alpha}(j,\beta,\kappa_0,0) + f_{i,\alpha}(j,\beta,\kappa_0)\right), \quad (i,\alpha),(j,\beta) \in \mathcal{S}, \end{aligned} \tag{6.17}$$

where $\tilde{r}^{\kappa_0,a_0}$ and $\tilde{q}^{\kappa_0,a_0}$ are given by (6.9) and (6.8), respectively, and where $\tilde{\boldsymbol{g}}^{\kappa_0}$ and $\tilde{\boldsymbol{n}}^{\kappa_0}$ are given by solving equations (6.15) and (6.16). ■

6.2. A pseudo-conservation law

The work decomposition result and the pseudo-conservation law for polling systems were given by Boxma and Groenendijk [3], and they were extended by Boxma [2]. The pseudo-conservation law for a polling system with local priority was considered by Shimogawa and Takahashi [24] and Takahashi and Kumar [30]. Sidi et al. [25] considered it for a polling system with feedback. We can obtain a pseudo-conservation law for our polling system.

Let $\overline{T^2_{j,\beta}}$ be the second moments of $T_{j,\beta}$. Then we have

$$\overline{T^2_{j,\beta}} = \overline{s^2_{j,\beta}} + 2E[S_{j,\beta}](\overline{T}_{j,\beta} - E[S_{j,\beta}]) + \sum_{k=1}^{J}\sum_{\gamma=1}^{L_k} p_{j,\beta,k,\gamma}\overline{T^2_{k,\gamma}}, \qquad (j,\beta) \in \mathcal{S}.$$

Further let $\bar{n}^{Cj}_{i,\alpha}$ be the average number of (i,α)-customers at a completion epoch of service period j. Since we have defined $\overline{N}_{j,\beta,i,\alpha}$ to be the expected value of (i,α)-customers who arrive during a service period of station j starting with a (j,β)-customer and who still stay at station i at its completion epoch, we have the following equation for $\bar{n}^{Cj}_{i,\alpha}$.

$$\bar{n}^{Cj}_{i,\alpha} = \begin{cases} \bar{n}^{Cj-1}_{i,\alpha} + \lambda_{i,\alpha}\overline{s^o_{j-1}} + \sum_{\beta=1}^{L_j}(\bar{n}^{Cj-1}_{j,\beta} + \lambda_{j,\beta}\overline{s^o_{j-1}})\overline{N}_{j,\beta,i,\alpha}, & i \neq j, \\ 0, & i = j \in \mathcal{H}_e, \\ \sum_{\beta=1}^{L_j}(\bar{n}^{Cj-1}_{j,\beta} + \lambda_{j,\beta}\overline{s^o_{j-1}})\overline{N}_{j,\beta,j,\alpha}, & i = j \in \mathcal{H}_g. \end{cases} \tag{6.18}$$

Pseudo-conservation law We have the following pseudo-conservation law for a weighted sum of the average sojourn times.

$$\sum_{(i,\alpha)\in\mathcal{S}}\sum_{(j,\beta)\in\mathcal{S}} \lambda_{j,\beta}\bar{w}_{i,\alpha}(j,\beta)\overline{T}_{i,\alpha} = \frac{\lambda\overline{T^2}}{2(1-\rho)} + \sum_{i=1}^{J}\sum_{\alpha=1}^{L_i}\left(\tilde{q}^{i,\alpha}E[S_{i,\alpha}] - \tilde{r}^{i,\alpha}\right) + \tilde{Y} \tag{6.19}$$

where $\overline{T^2} = \sum_{j=1}^{J}\sum_{\beta=1}^{L_j}(\lambda_{j,\beta}/\lambda)\overline{T^2_{j,\beta}}$, and

$$\tilde{Y} = \sum_{j=1}^{J}\left(\frac{\overline{s^o_j}}{\sum_{\ell=1}^{J}\overline{s^o_\ell}}\right)\left\{\sum_{i=1}^{J}\sum_{\alpha=1}^{L_i}\bar{n}^{Cj}_{i,\alpha}\cdot\overline{T}_{i,\alpha} + \rho\frac{\overline{s^{o2}_j}}{2\overline{s^o_j}}\right\}. \quad ■ \tag{6.20}$$

Note The first term $\lambda\overline{T^2}/2(1-\rho)$ in the right-hand side in equation (6.19) is the mean amount of work in the corresponding ordinary feedback system (without switchover times). The third term $\tilde{Y}$ in the equation is the mean amount of work at an arbitrary epoch in a switchover period. The sum of these two mean values is the mean work in our multiclass polling system with feedback, which comes from the work decomposition result given in [2].

The above pseudo-conservation law can be arranged as

$$\sum_{i=1}^{J}\sum_{\alpha=1}^{L_i}\overline{T}_{i,\alpha}\Lambda_{i,\alpha}E[W_{i,\alpha}] = \frac{\lambda\overline{T^2}}{2(1-\rho)} - \sum_{i=1}^{J}\sum_{\alpha=1}^{L_i}\left(\tilde{r}^{i,\alpha} + \{\overline{T}_{i,\alpha} - E[S_{i,\alpha}]\}\tilde{q}^{i,\alpha}\right) + \tilde{Y}$$

where $E[W_{i,\alpha}] \equiv \sum_{k\in\Pi\cup\Pi^s}\left(\tilde{n}^k_{i,\alpha} + \tilde{g}^k_{i,\alpha}\right)/\Lambda_{i,\alpha}$ is the mean waiting time per service stage of (i,α)-customers. This equation corresponds to equation (6.3) in [25] if we consider the single class model. ■

6.3. Numerical examples

We give two numerical examples. Since we can consider multiple classes of customers at each station, we investigate various types of scheduling algorithms and compare them.

Model 1. There are 4 stations in the system, and 4 types of customers arrive at each station. Each customer returns only to the same station at which it arrived from outside. For example, our model may represent a message switching node in a network where each arriving message contains multiple packets. The service parameters are given in Table 1.

Each number at the 'Stage' row denotes a service stage number. Each customer receives services in the order: stage 1 $\rightarrow$ stage 2 $\rightarrow \cdots$. Type 1 customers receive at most 5 stages of service, type 2 and type 3 customers receive at most 3 stages of service, and type 4 customers receive 1 stage of service. The variances of all service times are set at 0.5. The mean and the variance of all switchover times are 0.5 and 0.2, respectively. We plot the graphs by varying the arrival rates (or the resource utilization ρ) in Figures 1–3.

We consider the three types of the scheduling algorithms of customers in each station.

1. *Priority 1*: The following (local) priority order is given to the customer types: Type 1 $>$ Type 2 $>$ Type 3 $>$ Type 4.[11]
2. *Priority 2*: The following (local) priority order is given to the customer types: Type 1 $<$ Type 2 $<$ Type 3 $<$ Type 4.
3. *Round-robin (R-R)*: All customers are served in FCFS order regardless of their service stages and their types. All arriving customers join at the tail of the queue of the station.

Figure 1 depicts the average sojourn times in the systems with the exhaustive rule. Since the 'Priority 1' algorithm gives the higher priority to the lower numbered type, the average sojourn time of the lower numbered type is almost smaller than that of the higher numbered type (Figure 1.1). On the other hand, the 'Priority 2' algorithm gives the opposite results (Figure 1.2). Since the 'round-robin' algorithm gives indifferent services to all customers, the average sojourn times of customers increase as their overall average service times increase (Figure 1.3). Figure 2 depicts the average sojourn times in the systems with the gated rule. Since the gate is certainly closed at every time when each customer returns to the station after completing its service stage, it must wait until its next service period. Hence the average sojourn times of customers extremely depend on the number of their service stages (Figures 2.1–2.3). Figure 3 depicts the overall average sojourn times of all customers. It is shown that the exhaustive rule is superior to the gated rule, and that the 'Priority 2' algorithm is better than the other scheduling algorithms.

Table 1: Service parameters for all types of customers

Type 1					
Stage (i)	1	2	3	4	5
M.S.T. at i	1.5	1.5	1.5	1.5	1.5
F.P. ($i \rightarrow i+1$)	0.8	0.7	0.6	0.5	0.0

Type 2					
Stage (i)	1	2	3	4	5
M.S.T. at i	1.5	1.5	1.5	—	—
F.P. ($i \rightarrow i+1$)	0.7	0.6	0.0	—	—

Type 3					
Stage (i)	1	2	3	4	5
M.S.T. at i	1.0	1.0	1.0	—	—
F.P. ($i \rightarrow i+1$)	0.7	0.6	0.0	—	—

Type 4					
Stage (i)	1	2	3	4	5
M.S.T. at i	1.0	—	—	—	—
F.P. ($i \rightarrow i+1$)	0.0	—	—	—	—

M.S.T. at i: mean service time at stage i
F.P. ($i \rightarrow i+1$): feedback probability from stage i to stage $i+1$

[11] These inequalities mean that type i customers have priority over type $i+1$ customers ($i = 1, 2, 3$).

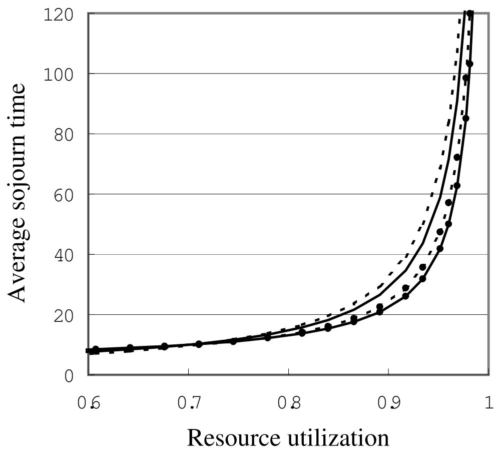

Figure 1.1: Priority 1 (exhaustive)

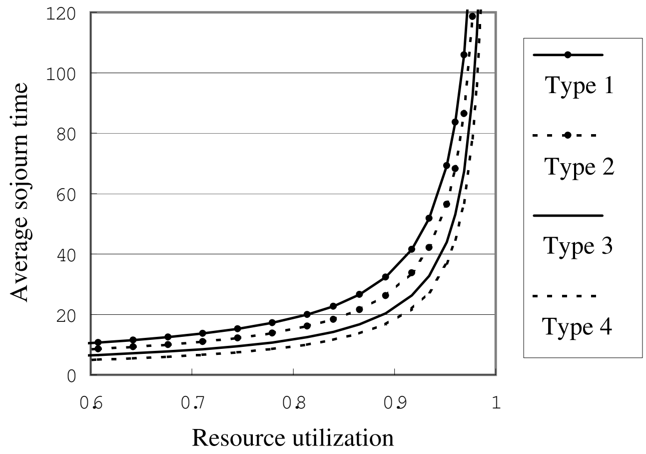

Figure 1.2: Priority 2 (exhaustive)

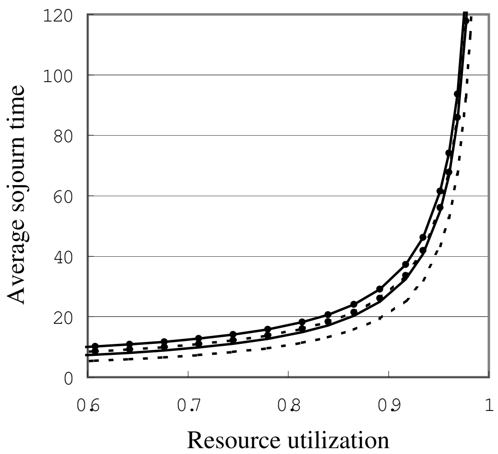

Figure 1.3: Round-robin (exhaustive)

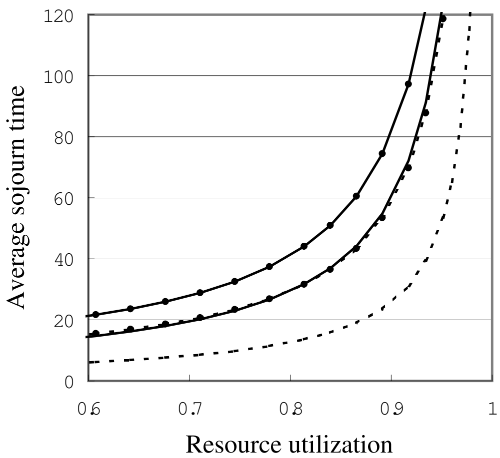

Figure 2.1: Priority 1 (gated)

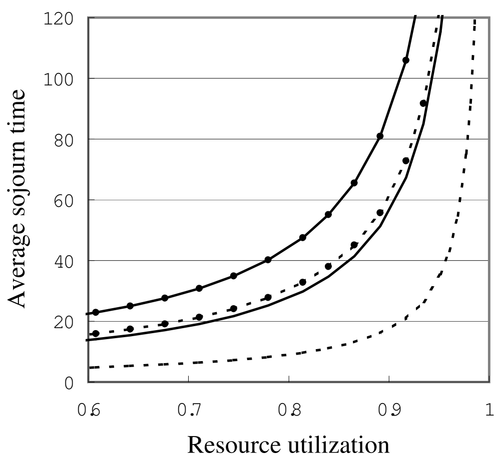

Figure 2.2: Priority 2 (gated)

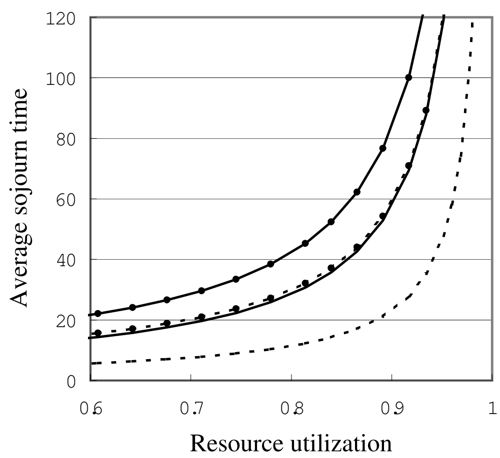

Figure 2.3: Round-robin (gated)

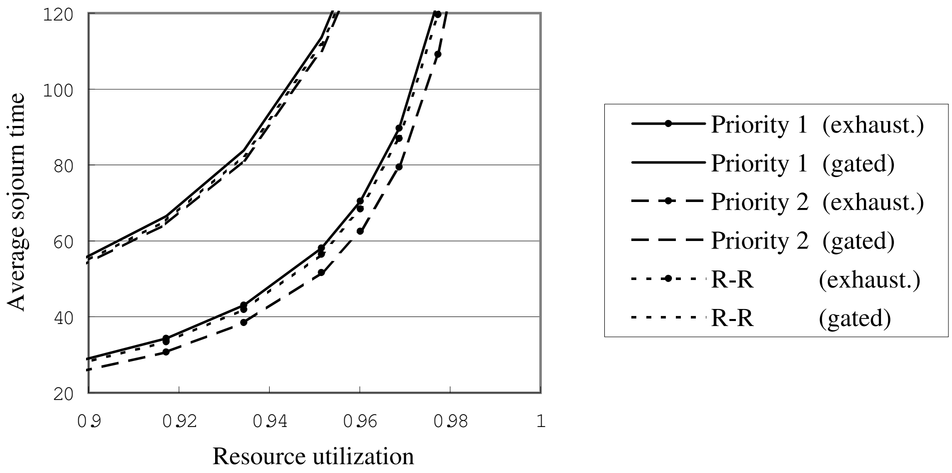

Figure 3: All scheduling algorithms (overall average sojourn times)

Model 2. There are 6 stations in the system, and 4 types of customers arrive from outside at the 'client nodes' (stations $3, 4, 5, 6$). Stations 1 and 2 are the 'server nodes' at which some common services of customers coming from the client nodes are executed. The paths of services and their parameters for each type of customers are given in Table 2.

The arrival rates listed in the table are relative values whose actual values vary with the resource utilization ρ. The variances of all service times are 1.0. The mean switchover time from station 1 to station 2 is 2.0, the other mean switchover times are 4.0, and their variances are 0.5. All types of customers arrive at each station. Type 1 customers receive four stages of services, Type 2 and 3 customers receive three stages of services. Type 4 customers receive 2 stages of services where their second stage branches probabilistically.

We consider the two types of the scheduling algorithms of customers at each client node.

1. *Priority 1*: The following (local) priority order is given to the customer types: Type 1 $>$ Type 2 $>$ Type 3 $>$ Type 4.
2. *Priority 2*: The following (local) priority order is given to the customer types: Type 1 $<$ Type 2 $<$ Type 3 $<$ Type 4.

Table 2: Paths of services and their parameters for all customer types

Type 1:	Station	X	→	1	→	X	→	2	→
A.R.=1	M.S.T.	4.0	→	2.0	→	4.0	→	3.0	→
	F.P.		1.0		1.0		1.0		1.0
Type 2:	Station	X	→	1	→	X	→		
A.R.=2	M.S.T.	5.0	→	2.5	→	5.0	→		
	F.P.		1.0		1.0		1.0		
Type 3:	Station	X	→	2	→	X	→		
A.R.=2	M.S.T.	5.0	→	2.5	→	5.0	→		
	F.P.		1.0		1.0		1.0		
Type 4:	Station	X	→	1	→				
			↘	2	→				
A.R.=3	M.S.T.	6.0	→	2.0	→				
			↘	3.0	→				
	F.P.		0.5		1.0				
			0.5		1.0				

X (= $3, 4, 5, 6$): any client node, A.R. : (relative) arrival rate,
M.S.T. : mean service time, F.P. : feedback probability

We assume that the service rules at the server nodes are the exhaustive FCFS and the service rules at the client nodes are the exhaustive priority.

The graphs of the mean sojourn times are given in Figures 4–6. We show in Figure 4 that the average sojourn time at station 6 is the least of all, and that the average sojourn time at station 3 is the largest of all. As the length of the path of services of customers becomes longer, their sojourn times become larger. The relative difference between the average sojourn time for type 2 customers and that for type 3 customers depends on the scheduling algorithms. The overall average sojourn time for 'priority 1' scheduling is slightly better than that for 'priority 2' scheduling (within 1 % difference).

6.4. Relation to the buffer occupancy method

In this section we consider a single class system where each station has a single class of customers ($L_k = 1$ for $k = 1, \ldots, J$), and obtain the relationship between our method and the buffer occupancy method. Hence we omit the class index from our quantities (e.g., $p_{k,i}$ is the feedback probability from station k to station i, and $\tilde{n}_i^k$ and $\tilde{g}_i^k$ are the average numbers of waiting i-customers in the waiting room and the service facility, respectively, while the system is in period k).

Let define the moments of the buffer occupancy variables ([25]):

$$f_i^*(j) \equiv E[X_i^j]; \qquad f_i^*(j,k) \equiv \begin{cases} E[X_i^j X_i^k], & j \neq k, \\ E[X_i^j(X_i^j - 1)], & j = k, \end{cases}$$

where X_i^j is the number of customers residing at station j when station i is polled. Let

$$\hat{b}_k \equiv E[T_k^{L_k}] = \frac{E[S_k]}{1 - p_{k,k}}; \qquad \hat{b}_k^{(2)} \equiv E[(T_k^{L_k})^2] = \frac{\overline{s_k^2}}{1 - p_{k,k}} + \frac{2p_{k,k}E[S_k]^2}{(1 - p_{k,k})^2};$$

$$\hat{p}_{k,i} \equiv \frac{p_{k,i}}{1 - p_{k,k}}, \quad (i \neq k); \qquad E[C] \equiv \frac{\sum_{l=1}^J \overline{s_l^o}}{1 - \rho};$$

for $k \in \mathcal{H}_e$. Then we have the following relationships.

For $i \neq k \in \mathcal{H}_g$:

$$\begin{aligned}
\tilde{n}_i^k &= \frac{1}{E[C]} \left\{ f_k^*(i,k)E[S_k] + f_k^*(k)\frac{\lambda_i \overline{s_k^2}}{2} + f_k^*(k,k)\frac{E[S_k]}{2}(\lambda_i E[S_k] + p_{k,i}) \right\}; \\
\tilde{n}_i^{k^s} &= \frac{1}{E[C]} \left\{ f_k^*(i)\overline{s_k^o} + f_k^*(k)(\lambda_i E[S_k] + p_{k,i})\overline{s_k^o} + \frac{\lambda_i \overline{s_k^{o2}}}{2} \right\}; \\
\tilde{g}_i^k &= 0.
\end{aligned}$$

For $i \neq k \in \mathcal{H}_e$:

$$\begin{aligned}
\tilde{n}_i^k &= \frac{1}{E[C]} \left\{ f_k^*(i,k)\frac{\hat{b}_k}{1 - \lambda_k \hat{b}_k} + f_k^*(k,k)\frac{(\lambda_i \hat{b}_k + \hat{p}_{k,i})\hat{b}_k}{2(1 - \lambda_k \hat{b}_k)^2} \right. \\
&\qquad \left. + f_k^*(k)\left(\frac{\lambda_k (\hat{b}_k)^2 \hat{p}_{k,i}}{(1 - \lambda_k \hat{b}_k)^2} + \frac{(\lambda_i + \lambda_k^2 \hat{b}_k \hat{p}_{k,i})\hat{b}_k^{(2)}}{2(1 - \lambda_k \hat{b}_k)^3} \right) \right\}; \\
\tilde{n}_i^{k^s} &= \frac{1}{E[C]} \left\{ f_k^*(i)\overline{s_k^o} + f_k^*(k)\frac{\lambda_i \hat{b}_k + \hat{p}_{k,i}}{1 - \lambda_k \hat{b}_k}\overline{s_k^o} + \frac{\lambda_i \overline{s_k^{o2}}}{2} \right\}; \\
\tilde{g}_i^k &= 0.
\end{aligned}$$

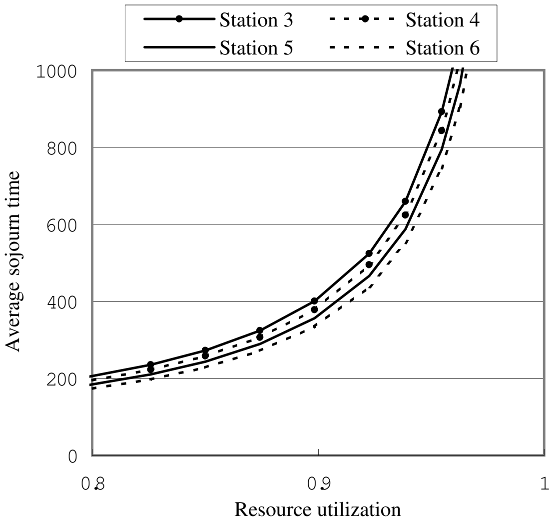

Figure 4.1: Priority 1 (all stations)

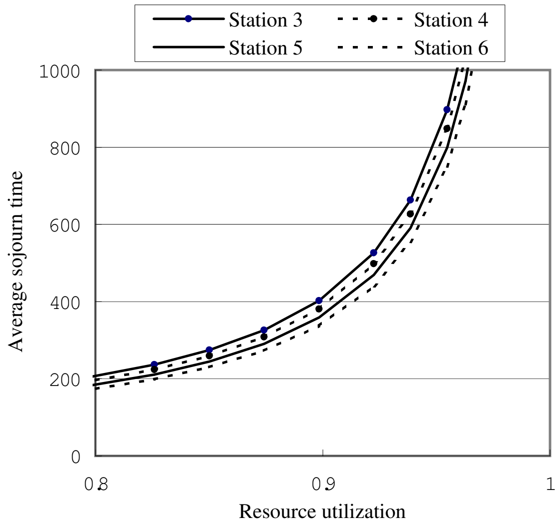

Figure 4.2: Priority 2 (all stations)

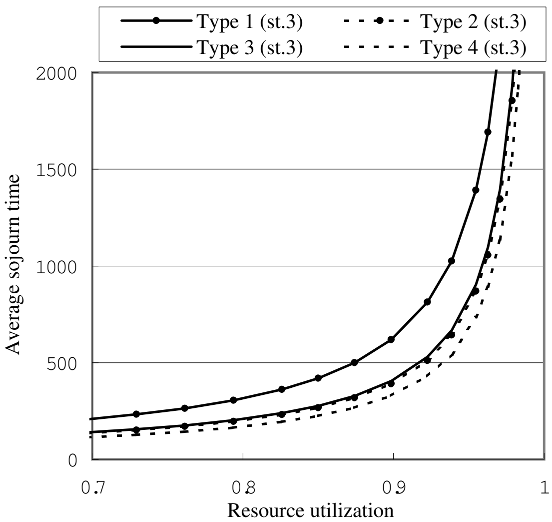

Figure 5.1: Priority 1 (station 3)

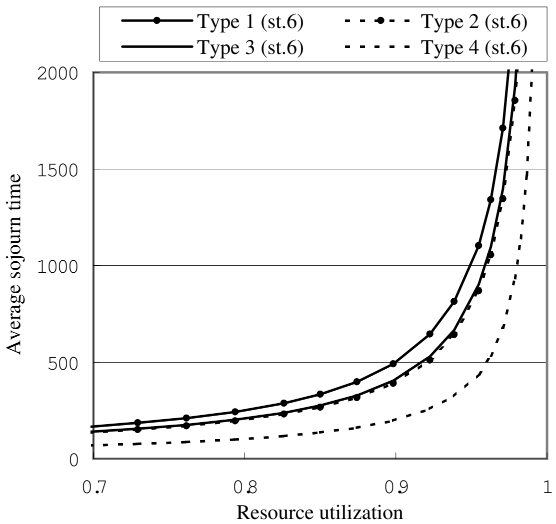

Figure 5.2: Priority 1 (station 6)

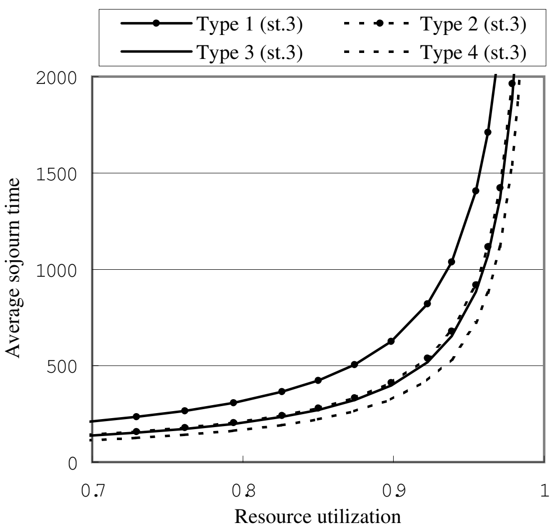

Figure 6.1: Priority 2 (station 3)

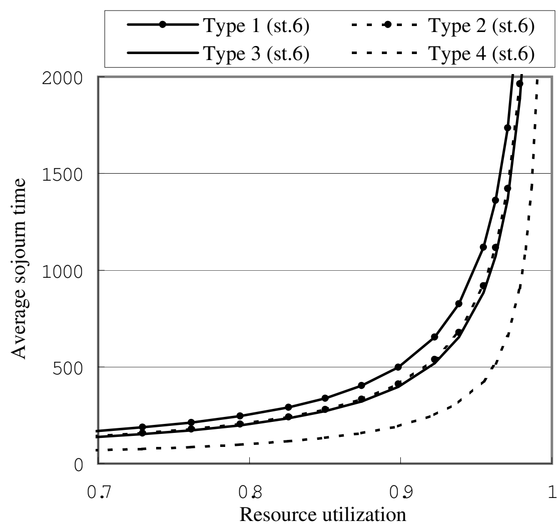

Figure 6.2: Priority 2 (station 6)

For $i = k \in \mathcal{H}_g$:

$$\begin{aligned}
\tilde{n}_k^k &= \frac{1}{E[C]}\left\{f_k^*(k)\frac{\lambda_k \overline{s_k^2}}{2} + f_k^*(k,k)\frac{E[S_k]}{2}\left(\lambda_k E[S_k] + p_{k,k}\right)\right\};\\
\tilde{n}_k^{k^s} &= \frac{1}{E[C]}\left\{f_k^*(k)\left(\lambda_k E[S_k] + p_{k,k}\right)\overline{s_k^o} + \frac{\lambda_k \overline{s_k^{o2}}}{2}\right\};\\
\tilde{g}_k^k &= \frac{1}{E[C]}\left\{f_k^*(k,k)\frac{E[S_k]}{2}\right\}.
\end{aligned}$$

For $i = k \in \mathcal{H}_e$:

$$\begin{aligned}
\tilde{n}_k^k &= 0;\\
\tilde{n}_k^{k^s} &= \frac{1}{E[C]}\left\{\frac{\lambda_k \overline{s_k^{o2}}}{2}\right\};\\
\tilde{g}_k^k &= \frac{1}{E[C]}\left\{f_k^*(k)\frac{\lambda_k \hat{b}_k^{(2)}}{2(1-\lambda_k \hat{b}_k)^2} + f_k^*(k,k)\frac{\hat{b}_k}{2(1-\lambda_k \hat{b}_k)}\right\}.
\end{aligned}$$

Note In the above expressions, C is the (steady state) cycle time of the system and they can be obtained by evaluating the integrals of the system states on a cycle C. For example,

$$\tilde{n}_i^k = \frac{E\left[\int_0^C n_i(s)\mathbf{1}\{\kappa(s) = k\}ds\right]}{E[C]}.$$

Similar expressions can be found in [25] (equations (3.12), (3.13), (4.10) and (4.11)). ■

7. Conclusion

We have been concerned with multiclass polling systems with feedback, and the method developed in [11, 14, 15] has been extended. The average sojourn time $\bar{w}_{i,\alpha}(j,\beta)$ of (i,α)-customers arriving from outside the system as (j,β)-customers and a pseudo-conservation law have been obtained. We have compared our method with the buffer occupancy method.

We conclude the paper by summarizing the key features and the advantages of our method. The first feature is that we can analyze composite scheduling algorithms.

- The scheduling algorithms include mixtures of the exhaustive and the gated rules, and the (local) priority and the FCFS orders.
- We can easily calculate mean path times for many types of customers which are mean amounts of times spent by customers traversing specific paths of services, by setting the system parameters appropriately.

The second feature is the following analytical advantages.

- By solving the feedback equations, we can reduce the complicated derivation of the overall average sojourn times to the derivation of the expected sojourn times in each stage of service and the expected numbers of customers at each feedback epoch.
- By using the linear functional expressions, the performance measures for all scheduling algorithms considered have the same expressions. Furthermore, similar expressions are found in the priority systems [11, 14] and the Markovian polling system [12].
- We overcome the difficulties in analyzing the FCFS orders. (See the expressions and the explanations in Section 4 of $F_{j,\beta}^I(\cdot)$ and the expected numbers of customers at feedback epochs, which explicitly depend on the (local) service orders).

- We scarcely require any second moment in our analysis. The second moments appear only in the steady state average values through the average remaining time $\tilde{r}^{\kappa,a}$.

The method developed in this paper has a lot in common with the method for priority queueing systems [11, 14] and can also be applied to a Markovian polling system [12]. The advantage of our method is its wide applicability to the analysis of mean sojourn times in many types of M/G/1 multiclass queueing systems.

Acknowledgment
The author would like to thank the referees for their valuable comments.

References

[1] J.L. van den Berg, O.J. Boxma and W.P. Groenendijk: Sojourn times in the M/G/1 queue with deterministic feedback. *Communications in Statistics-Stochastic Models*, **5** (1989), 115-129.

[2] O.J. Boxma: Workloads and waiting times in single-server systems with multiple customer classes. *Queueing Systems*, **5** (1989), 185-214.

[3] O.J. Boxma and W.P. Groenendijk: Pseudo-conservation laws in cyclic-service systems. *Journal of Applied Probability*, **24** (1987), 949-964.

[4] R.B. Cooper: Queues served in cyclic order: waiting times. *The Bell System Technical Journal*, **49** (1970), 399-413.

[5] R.B. Cooper and G. Murray: Queues served in cyclic order. *The Bell System Technical Journal*, **48** (1969), 675-689.

[6] J.N. Daigle and C.E. Houstis: Analysis of a task oriented multipriority queueing system. *IEEE Transactions on Communications*, **29** (1981), 1669-1677.

[7] R.L. Disney: A note on sojourn times in M/G/1 queues with instantaneous, Bernoulli feedback. *Naval Research Logistics Quarterly*, **28** (1981), 679-684.

[8] B.T. Doshi and J.S. Kaufman: Sojourn time in an M/G/1 queue with Bernoulli feedback. In O.J. Boxma and R. Syski (eds.): *Queueing Theory and its Applications* (Elsevier, North-Holland, Amsterdam, 1988), 207-233.

[9] M. Eisenberg: Queues with periodic service and changeover time. *Operations Research*, **20** (1972), 440-451.

[10] M.J. Ferguson and Y.J. Aminetzah: Exact results for nonsymmetric token ring systems. *IEEE Transactions on Communications*, **33** (1985), 223-231.

[11] T. Hirayama: Analysis of multiclass M/G/1 queues with a mixture of 1-limited disciplines and gated disciplines. *Journal of the Operations Research Society of Japan*, **42** (1999), 237-255.

[12] T. Hirayama: Mean waiting times in Markovian polling systems. In H. Sakasegawa (ed.): *Proceedings of the Symposium on Queueing Theory and its Applications* (Nagoya, Japan, 2002), 58-65.

[13] T. Hirayama: A multiclass polling system with feedback. In J. Li, K. Kato and H. Kameda (eds.): *Proceedings of the IASTED International Conference on Networks, Parallel and Distributed Processing, and Applications* (ACTA Press, Anaheim, 2002), 154-159.

[14] T. Hirayama: Mean sojourn times in multiclass feedback queues with gated disciplines. *Naval Research Logistics*, **50** (2003), 719-741.

[15] T. Hirayama, S.J. Hong and M.M. Krunz: A new approach to analysis of polling systems. *Queueing Systems*, **48** (2004), 135-158.

[16] T. Katayama: Mean sojourn times in a multi-stage tandem queue served by a single server. *Journal of the Operations Research Society of Japan*, **31** (1988), 233-251.

[17] T. Katayama: Performance analysis and optimization of a cyclic-service tandem queueing system with multi-class customers. *Computers & Mathematics with Applications*, **24** (1992), 25-33.
[18] T. Katayama: Analysis of an alternating-service tandem queue with server vacations and conversion relationships between the performance measures. *IEICE Transactions on Communications*, **E78-B** (1995), 1075-1079.
[19] A.G. Konheim, H. Levy and M.M. Srinivasan: Descendant set: an efficient approach for the analysis of polling systems. *IEEE Transactions on Communications*, **42** (1994), 1245-1253.
[20] H. Levy: Delay computation and dynamic behavior of non-symmetric polling systems. *Performance Evaluation*, **10** (1989), 35-51.
[21] H. Levy and M. Sidi: Polling systems: applications, modeling, and optimization. *IEEE Transactions on Communications*, **38** (1990), 1750-1760.
[22] H. Levy and M. Sidi: Polling systems with simultaneous arrivals. *IEEE Transactions on Communications*, **39** (1991), 823-827.
[23] M. Paterok and O. Fischer: Feedback queues with preemption-distance priorities. *Performance Evaluation Review*, **17** (1989), 136-145.
[24] S. Shimogawa and Y. Takahashi: A note on the pseudo-conservation law for a multi-queue with local priority. *Queueing Systems*, **11** (1992), 145-151.
[25] M. Sidi, H. Levy and S.W. Fuhrmann: A queueing network with a single cyclically roving server. *Queueing Systems*, **11** (1992), 121-144.
[26] B. Simon: Priority queues with feedback. *Journal of the Association for Computing Machinery*, **31** (1984), 134-149.
[27] H. Takagi: Analysis and applications of a multiqueue cyclic service system with feedback. *IEEE Transactions on Communications*, **35** (1987), 248-250.
[28] H. Takagi: Analysis of polling systems with a mixture of exhaustive and gated service disciplines. *Journal of the Operations Research Society of Japan*, **32** (1989), 450-461.
[29] H. Takagi: Analysis and application of polling models. In G. Haring, C. Lindemann and M. Reiser (eds.): *Performance Evaluation: Origins and Directions; Lecture Notes in Computer Science*, **1769** (Springer, Berlin, 2000), 423-442.
[30] Y. Takahashi and B.K. Kumar: Pseudo-conservation law for a priority polling system with mixed service strategies. *Performance Evaluation*, **23** (1995), 107-120.
[31] T. Takine and T. Hasegawa: Average waiting time of a symmetric polling system under Bernoulli scheduling. *Operations Research Letters*, **10** (1991), 535-539.
[32] T. Takine, H. Takagi and T. Hasegawa: Sojourn times in vacation and polling systems with Bernoulli feedback. *Journal of Applied Probability*, **28** (1991), 422-432.
[33] W. Whitt: A review of $L = \lambda W$ and extensions. *Queueing Systems*, **9** (1991), 235-268.
[34] R.W. Wolff: *Stochastic Modeling and the Theory of Queues* (Prentice-Hall, Englewood Cliffs, NJ, 1989).

Appendix A: Summary of Notation

The following is a list of the notations frequently used in this paper.

System parameters and related quantities (I)

J, L_i, J_c	: number of groups, number of classes in group i, and $J_c = \sum_{i=1}^{J} L_i$;
$\lambda_{i,\alpha}, \lambda$	: Poisson arrival rate of (i,α)-customers and $\lambda \equiv \sum_{i=1}^{J}\sum_{\alpha=1}^{L_i} \lambda_{i,\alpha}$;
$E[S_{i,\alpha}], \overline{s_{i,\alpha}^2}$	: mean and second moment of a service time of an (i,α)-customer;
$p_{i,\alpha,j,\beta}, \boldsymbol{P}$	: feedback probability from class (i,α) to class (j,β) and its matrix;
$\overline{s_i^o}, \overline{s_i^{o2}}$	: mean and second moment of a switchover time from station i;
$\boldsymbol{c}^e, \sigma_0^e$	: the e^{th} arriving customer from outside the system and its arrival epoch;
σ_k^e	: the k^{th} service stage completion epoch of $\boldsymbol{c}^e$;
M^e	: total number of $\boldsymbol{c}^e$'s service stages during its stay in the system.

System parameters and related quantities (II)

$T_{i,\alpha}, \overline{T}_{i,\alpha}$	: total amount of service times received by a customer arriving as an (i,α)-customer until it departs from the system, and its expectation;
ρ	: resource utilization $(= \sum_{i=1}^{J} \sum_{\alpha=1}^{L_i} \lambda_{i,\alpha} \overline{T}_{i,\alpha} < 1)$;
$N_{i,\alpha,k,\gamma}, \overline{N}_{i,\alpha,k,\gamma}, \overline{N}_{i,\alpha,k,\gamma}(r)$	: number of (k,γ)-customers who arrive during a service period of station i starting with an (i,α)-customer, and who still stay at station k at its completion epoch, and its expectations (eqs. (3.4), (3.5));
$T^{\delta}_{i,\alpha}, \overline{T}^{\delta}_{i,\alpha}, \overline{T}^{\delta}_{i,\alpha}(r)$	: total amount of service times received by a customer who is initially an (i,α)-customer until it departs from the set of classes $(i,1),\ldots,(i,\delta)$ at station i for the first time after at least receiving a service, and its expectations (eq. (4.1));
$\varrho^{+}_{i,\delta}$	: utilization of the classes $(i,1),\ldots,(i,\delta)$ $(= \sum_{\alpha=1}^{\delta} \lambda_{i,\alpha} \overline{T}^{\delta}_{i,\alpha})$;
$N^{\delta}_{j,\alpha,k,\gamma}, \overline{N}^{\delta}_{j,\alpha,k,\gamma}, \overline{N}^{\delta}_{j,\alpha,k,\gamma}(r)$	: number of (k,γ)-customers who arrive (from outside or by feedback) during a (j,δ)-busy period starting with a (j,α)-customer, and who still stay at station k at the end of the period, and its expectations (eqs. (4.18), (4.19));
$D(\boldsymbol{Y})$	: expected remaining length of the current service period (eq. (4.3));
$\delta(i,\alpha)$	: expected length of a service period i starting with an (i,α)-customer (eq. (4.4)).

Sets of the groups and the periods, and related quantities

i^s	: switchover period from station i to station $i+1$;
$\mathcal{S}$	: set of all customer classes $(= \{(i,\alpha) : i = 1,\ldots,J \text{ and } \alpha = 1,\ldots,L_i\})$;
$\mathcal{S}_A$	: set of all server status $(= \mathcal{S} \cup \{(1^s,0),(2^s,0),\ldots,(J^s,0)\})$;
Π	: set of indices of the service periods $(= \{1,\ldots,J\})$;
Π^s	: set of indices of the switchover periods $(= \{1^s,\ldots,J^s\})$;
$\Pi_{k,j}, \Pi^s_{k,j}$	: subsets of Π and Π^s defined in equations (2.2) and (2.3), resp.;
k^-	: period just before the period k (eq. (2.4));
$\mathcal{H}_g, \mathcal{H}_e$	: sets of the gated groups and the exhaustive groups, resp.;
$\mathcal{H}_{gF}, \mathcal{H}_{eF}$	: sets of the gated FCFS groups and the exhaustive FCFS groups, resp.;
$\mathcal{H}_{gP}, \mathcal{H}_{eP}$	: sets of the gated priority groups and the exhaustive priority groups, resp.

State of the system and related quantities

$(\kappa(t), a(t))$	: status of the server at time t: the server serves a $(\kappa(t), a(t))$-customer if $\kappa(t) \in \Pi$, or it is in the switchover period $\kappa(t)$ if $\kappa(t) \in \Pi^s$;
$r(t)$	: remaining service time of a customer being served at time t if $\kappa(t) \in \Pi$, or remaining length of a switchover period if $\kappa(t) \in \Pi^s$;
$g_{i,\alpha}(t)$	: number of (i,α)-customers in the service facility (not being served) at time t;
$n_{i,\alpha}(t)$	: number of (i,α)-customers in the waiting room at time t;
$\boldsymbol{g}(t), \boldsymbol{n}(t)$	: vectors $(g_{i,\alpha}(t) : (i,\alpha) \in \mathcal{S})$ and $(n_{i,\alpha}(t) : (i,\alpha) \in \mathcal{S})$, resp.;
$X(t), \Gamma(t)$	: station and class, resp., of a customer arriving at the last transition epoch;
$L(t)$	: information of the system at time t;
$\boldsymbol{Y}(t)$	: system state at time t $(= (X(t), \Gamma(t), \kappa(t), a(t), r(t), \boldsymbol{g}(t), \boldsymbol{n}(t), L(t)))$;
$\mathcal{Q}, \mathcal{E}$	: the stochastic process $(\mathcal{Q} = \{\boldsymbol{Y}(t) : t \geq 0\})$ and its state space.

System performance measures and related quantities

$W^e_{i,\alpha}$	: W.T. of $\boldsymbol{c}^e$ spent in the waiting room as an (i,α)-customer (eq. (2.6));
$W_{i,\alpha}(\boldsymbol{Y}, e, l)$	: C.E.W.T. of $\boldsymbol{c}^e$ related to $W^e_{i,\alpha}$ (eq. (2.7));
$W^I_{i,\alpha}(\boldsymbol{Y}, e, l)$	: C.E.W.T. of $\boldsymbol{c}^e$ per service stage related to $W^e_{i,\alpha}$ (eq. (2.8));

$H^e_{i,\alpha}(k)$: W.T. of $\boldsymbol{c}^e$ spent in the waiting room as an (i,α)-customer while the system is in period k (eq. (2.10));
$H_{i,\alpha}(\boldsymbol{Y},e,l,k)$: C.E.W.T. of $\boldsymbol{c}^e$ related to $H^e_{i,\alpha}(k)$ (eq. (2.11));
$H^I_{i,\alpha}(\boldsymbol{Y},e,l,k)$: C.E.W.T. of $\boldsymbol{c}^e$ per service stage related to $H^e_{i,\alpha}(k)$ (eq. (2.12));
$F^e_{i,\alpha}$: S.T. of $\boldsymbol{c}^e$ spent in the service facility as an (i,α)-customer (eq. (2.14));
$F_{i,\alpha}(\boldsymbol{Y},e,l)$: C.E.S.T. of $\boldsymbol{c}^e$ related to $F^e_{i,\alpha}$ (eq. (2.15));
$F^I_{i,\alpha}(\boldsymbol{Y},e,l)$: C.E.S.T. of $\boldsymbol{c}^e$ per service stage related to $F^e_{i,\alpha}$ (eq. (2.16)).
(W.T.: Waiting Time; C.E.W.T.: Conditional Expected Waiting Time)
(S.T.: Sojourn Time; C.E.S.T.: Conditional Expected Sojourn Time)

State of the system at polling instants and related quantities

$\nu^i_{k,\gamma}(\tau)$: number of (k,γ)-customers at the first polling instant of station i after any arrival epoch τ;
$\overline{\nu}^i_{k,\gamma}(\boldsymbol{Y})$: $E[\nu^i_{k,\gamma}(\tau)|\boldsymbol{Y}(\tau)=\boldsymbol{Y}]$;
$\overline{\boldsymbol{\nu}}^i(\boldsymbol{Y}),\overline{\boldsymbol{\nu}}(\boldsymbol{Y})$: vectors $(\overline{\nu}^i_{k,\gamma}(\boldsymbol{Y}):(k,\gamma)\in\mathcal{S})$ and $(\overline{\nu}^k_{k,\gamma}(\boldsymbol{Y}):(k,\gamma)\in\mathcal{S})$, resp.

Coefficients for the expressions of the system performance measures

$\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,k),\boldsymbol{w}^{j,\beta}(\kappa_0,k),w^{j,\beta}(\kappa_0,k)$: coefficients for $W^I_{i,\alpha}(\cdot)$ and $H^I_{i,\alpha}(\cdot)$;
$\boldsymbol{\eta}^{j,\beta}(\kappa_0,a_0),\boldsymbol{f}^{j,\beta}(\kappa_0),f^{j,\beta}(\kappa_0)$: coefficients for $F^I_{i,\alpha}(\cdot)$;
$\boldsymbol{\upsilon}^{j,\beta}(\kappa_0,a_0),\boldsymbol{U}^{j,\beta}(\kappa_0),\boldsymbol{u}^{j,\beta}(\kappa_0)$: coefficients for the vector of the conditional expected numbers of customers at σ^e_{l+1};
$\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,k),\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,k),w_{i,\alpha}(j,\beta,\kappa_0,k)$: coefficients for $W_{i,\alpha}(\cdot)$ and $H_{i,\alpha}(\cdot)$;
$\boldsymbol{\eta}_{i,\alpha}(j,\beta,\kappa_0,a_0),\boldsymbol{f}_{i,\alpha}(j,\beta,\kappa_0),f_{i,\alpha}(j,\beta,\kappa_0)$: coefficients for $F_{i,\alpha}(\cdot)$.

Steady state values

$\bar{w}_{i,\alpha}(j,\beta)$: average sojourn time of (i,α)-customers arriving from outside the system as (j,β)-customers;
$\tilde{\boldsymbol{Y}}^{k,\gamma},\tilde{\boldsymbol{Y}}^k$: time average values of the system state defined by (6.5) and (6.6), resp.;
$\tilde{\boldsymbol{g}}^{k,\gamma},\tilde{\boldsymbol{g}}^k$: components of $\tilde{\boldsymbol{Y}}^{k,\gamma}$ and $\tilde{\boldsymbol{Y}}^k$, resp., related to the numbers of customers in the service facility;
$\tilde{\boldsymbol{n}}^{k,\gamma},\tilde{\boldsymbol{n}}^k$: components of $\tilde{\boldsymbol{Y}}^{k,\gamma}$ and $\tilde{\boldsymbol{Y}}^k$, resp., related to the numbers of customers in the waiting room;
$\tilde{q}^{\kappa_0,a_0}$: long-run fraction of time that the server's status is (κ_0,a_0) (eq. (6.8));
$\tilde{r}^{\kappa_0,a_0}$: average value of the remaining time (eq. (6.9));
$\Lambda_{i,\alpha}$: composite arrival rate of (i,α)-customers (eq. (6.7)).

Other quantities

$\mathcal{R},\mathcal{R}_+,\mathcal{I}_+$: sets of real numbers, nonnegative real numbers, and nonnegative integers;
$\boldsymbol{1}\{\mathcal{K}\}$: 1 if the event $\mathcal{K}$ is true, or 0 otherwise;
$\boldsymbol{1}_{i,\alpha}(j,\beta)$: 1 if $(j,\beta)=(i,\alpha)$, or 0 otherwise;
$\boldsymbol{1}(r)$: 1 if $r>0$, or 0 if $r=0$.

Appendix B: Proof of Proposition 3

Let $e(=1,2,\ldots)$ be a customer number and $l(=0,1,2,\ldots)$ be the number of feedbacks of $\boldsymbol{c}^e$. For any $\boldsymbol{Y}=(j,\beta,\kappa_0,a_0,r,\boldsymbol{g},\boldsymbol{n},L)\in\mathcal{E}$, let $\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}$ be the state of the system at σ^e_l. Then it can be easily seen that $\kappa(\sigma^e_{l+1})=j$ and $r(\sigma^e_{l+1})=0$. We prove them for $W_{i,\alpha}(\cdot)=\hat{W}_{i,\alpha}(\cdot)$, and the other proofs are similar.

First we can show from equations (5.1), (5.3), (5.6) and (5.7) that

$$\begin{array}{l}\sum_{m=1}^J\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\boldsymbol{w}_{i,\alpha}(m,\delta,j,k)=\hat{\boldsymbol{w}}_{i,\alpha}(j,\beta,k);\\ \sum_{m=1}^J\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}w_{i,\alpha}(m,\delta,j,k)=\hat{w}_{i,\alpha}(j,\beta,k);\end{array}\quad\left(\begin{array}{l}(i,\alpha),(j,\beta)\in\mathcal{S},\\ k\in\Pi\cup\Pi^s\cup\{0\}\end{array}\right).\tag{7.1}$$

For $(i,\alpha)=(j,\beta)\in\mathcal{S}$, we have

$$
\begin{aligned}
&W^I_{j,\beta}(\boldsymbol{Y},e,l)+E[\hat{W}_{j,\beta}(\boldsymbol{Y}(\sigma^e_{l+1}),e,l+1)|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}]\\
&=(r,\mathbf{1}(r))\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,0)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{w}^{j,\beta}(\kappa_0,0)+w^{j,\beta}(\kappa_0,0)\\
&\qquad+E[(\boldsymbol{g}(\sigma^e_{l+1}),\boldsymbol{n}(\sigma^e_{l+1}))\boldsymbol{w}_{j,\beta}(X(\sigma^e_{l+1}),\Gamma(\sigma^e_{l+1}),\kappa(\sigma^e_{l+1}),0)\\
&\qquad\qquad+w_{j,\beta}(X(\sigma^e_{l+1}),\Gamma(\sigma^e_{l+1}),\kappa(\sigma^e_{l+1}),0)|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}]\\
&=(r,\mathbf{1}(r))\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,0)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{w}^{j,\beta}(\kappa_0,0)+w^{j,\beta}(\kappa_0,0)\\
&\quad+\sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\left\{E[(\boldsymbol{g}(\sigma^e_{l+1}),\boldsymbol{n}(\sigma^e_{l+1}))|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}]\boldsymbol{w}_{j,\beta}(m,\delta,j,0)+w_{j,\beta}(m,\delta,j,0)\right\}\\
&=(r,\mathbf{1}(r))\boldsymbol{\varphi}^{j,\beta}(\kappa_0,a_0,0)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{w}^{j,\beta}(\kappa_0,0)+w^{j,\beta}(\kappa_0,0)\\
&\quad+\sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\left\{\left((r,\mathbf{1}(r))\boldsymbol{v}^{j,\beta}(\kappa_0,a_0)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{U}^{j,\beta}(\kappa_0)+\boldsymbol{u}^{j,\beta}(\kappa_0)\right)\boldsymbol{w}_{j,\beta}(m,\delta,j,0)\right.\\
&\qquad\qquad\qquad\left.+w_{j,\beta}(m,\delta,j,0)\right\}\\
&=(r,\mathbf{1}(r))\boldsymbol{\varphi}_{j,\beta}(j,\beta,\kappa_0,a_0,0)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{w}_{j,\beta}(j,\beta,\kappa_0,0)+w_{j,\beta}(j,\beta,\kappa_0,0).
\end{aligned}
$$

The first equality comes from (4.22) and (5.11), the third equality comes from (4.26), and the last equality comes from (5.5), (5.6), (5.7) and (7.1).

For $(i,\alpha)\neq(j,\beta)\in\mathcal{S}$, we have

$$
\begin{aligned}
&W^I_{i,\alpha}(\boldsymbol{Y},e,l)+E[\hat{W}_{i,\alpha}(\boldsymbol{Y}(\sigma^e_{l+1}),e,l+1)|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}]\\
&=E[(\boldsymbol{g}(\sigma^e_{l+1}),\boldsymbol{n}(\sigma^e_{l+1}))\boldsymbol{w}_{i,\alpha}(X(\sigma^e_{l+1}),\Gamma(\sigma^e_{l+1}),\kappa(\sigma^e_{l+1}),0)\\
&\qquad+w_{i,\alpha}(X(\sigma^e_{l+1}),\Gamma(\sigma^e_{l+1}),\kappa(\sigma^e_{l+1}),0)|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}]\\
&=\sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\left\{E[(\boldsymbol{g}(\sigma^e_{l+1}),\boldsymbol{n}(\sigma^e_{l+1}))|\boldsymbol{Y}(\sigma^e_l)=\boldsymbol{Y}]\boldsymbol{w}_{i,\alpha}(m,\delta,j,0)+w_{i,\alpha}(m,\delta,j,0)\right\}\\
&=\sum_{m=1}^{J}\sum_{\delta=1}^{L_m}p_{j,\beta,m,\delta}\left\{\left((r,\mathbf{1}(r))\boldsymbol{v}^{j,\beta}(\kappa_0,a_0)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{U}^{j,\beta}(\kappa_0)+\boldsymbol{u}^{j,\beta}(\kappa_0)\right)\boldsymbol{w}_{i,\alpha}(m,\delta,j,0)\right.\\
&\qquad\qquad\qquad\left.+w_{i,\alpha}(m,\delta,j,0)\right\}\\
&=(r,\mathbf{1}(r))\boldsymbol{\varphi}_{i,\alpha}(j,\beta,\kappa_0,a_0,0)+(\boldsymbol{g},\boldsymbol{n})\boldsymbol{w}_{i,\alpha}(j,\beta,\kappa_0,0)+w_{i,\alpha}(j,\beta,\kappa_0,0).
\end{aligned}
$$

The first equality comes from (4.25) and (5.11), the third equality comes from (4.26), and the last equality comes from (5.5), (5.6), (5.7) and (7.1). ■

Tetsuji Hirayama
Graduate School of Systems and Information Engineering
University of Tsukuba
1-1-1 Tennoudai, Tsukuba-shi
Ibaraki 305-8573, Japan
E-mail: `hirayama@cs.tsukuba.ac.jp`