Closed Polling Models with Failing Nodes

HOVAV DROR and URI YECHIALI

Department of Statistics and Operations Research
School of Mathematical Sciences
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University, Tel Aviv 69978, Israel
email: hovav@hotmail.com and uriy@math.tau.ac.il
fax: +972-3-6409357

Abstract

Closed polling systems with station breakdowns, under the Gated, Exhaustive or Globally Gated services regimes, are studied and analyzed. Multidimensional sets of probability generating functions of the system's state are derived. They are further utilized to obtain an approximate solution for the mean number of jobs residing in the system's various queues at polling instants. The analysis is then concentrated on the case of Cyclic Bernoulli Polling. Explicit formulae for the mean number of jobs, as well as for the expected cycle duration and system utilization, are derived. Comparison of the throughputs of the three regimes concludes the paper.

Keywords: Polling, Closed Networks, Failing Nodes, Breakdowns.

1 Introduction

Polling systems are queueing networks where a single server attends N stations, visiting them according to some 'routing table' (most commonly, cyclic). Due to their wide range of application, those systems have been studied extensively in the literature (see, for example, Takagi [1986], Browne and Yechiali [1989], Levy and Sidi [1990], Yechiali [1993] and references there). However, most investigated models did not consider the important phenomenon of breaking, or failing, nodes (as is the case in real systems). Only recently Eliazar and Yechiali [1998] analyzed Randomly Timed Gated regimes, that can be considered as node failing problems, and Kofman and Yechiali [1996],[1997] explicitly introduced and analyzed polling systems with station breakdowns. Furthermore, most of the work in the literature dealt with 'open' systems where a stream of *transient* jobs flow into the system, get served and leave never to return. Boxma and Cohen [1991] analyzed a *single-node,* $M/G/1$ configuration, with regular (transient) jobs and with a fixed number of permanent jobs, where each such job immediately returns to the end of the (overall) queue whenever it completes a service. Altman and Yechiali [1994] analyzed fully 'closed' polling systems in which a fixed number of *permanent* jobs always reside in the system in a way that each job, after being served in one station, is immediately routed (randomly) to another station where it waits to be served again by the moving server. Armony and Yechiali [1999] further extended the investigation by studying polling systems with both permanent and transient jobs.

In this work we study fully 'closed' polling systems that suffer station breakdowns under the Gated, Exhaustive, or Globally-Gated service regimes. We assume that there are two modes of station breakdowns: node-dependent, and general. Extended investigation is made on Cyclic Bernoulli polling (CBP), where only the general breakdown process is possible (as was done by Altman and Yechiali [1993] for 'open' systems).

We begin section 2 with a description of the assumptions leading to the model and introduce the required notation. Section 3 deals with the Gated service regime. We derive a set of multi-dimensional Probability Generating Functions (PGFs) from which we obtain an approximated solution for the mean number of jobs residing in the system's various queues at polling instants. Section 4 is devoted to the analysis of Gated Cyclic Bernoulli polling. We derive the probability generating functions at polling instants and at arbitrary moments, and obtain an explicit formulae for the mean number of jobs residing in the system's various queues at polling instants. We also calculate the mean cycle time, sojourn times, throughputs and the system's utilization. Sections 5 and 6 are devoted to the analysis of the Exhaustive service regime (regular and CBP). Sections 7 and 8 conclude our analyses by studying the system state under the Globally Gated regime and comparing the throughputs of the three regimes.

2 Model and Notation

A closed network of N channels (stations, queues, nodes), denoted by Q_1, Q_2, ..., Q_N, is attended by a single server in a cyclic manner. The system is 'closed' in the sense that there are M permanent jobs that always reside in the system and are routed from one channel to another.

Service time of a job (customer, message) served in Q_i is a random variable B_i with mean b_i and Laplace Stieltjes Transform (LST) $\tilde{B}_i(s) \equiv E[e^{-sB_i}]$. After being served in Q_i, the job is immediately routed to Q_j ($j = 1, 2, \ldots, N$) with probability P_{ij}, $\sum_{j=1}^{N} P_{ij} = 1$. The routed job then waits in Q_j to be served again by the cycling server (with service time duration B_j) and rerouted further to another queue, etc.

When moving from channel i to channel $i+1$, the server incurs a random switch-over time D_i, having mean d_i. Total switching time during a full cycle is $D = \sum_{i=1}^{N} D_i$ with mean $d = \sum_{i=1}^{N} d_i$.

Stations (nodes) may fail from time to time, and we assume that there are two modes of station breakdowns: node-dependent, and general. If the server enters a non-empty station and starts serving jobs present there, the station may fail due to a node-dependent breakdown process. We assume that a breakdown (if occurs) is observed only at the end of a service (e.g. transmission of a message). The server then switches on to the next station, while the job (message) 'suffering' the breakdown will have to be processed (transmitted) again during the next server's visit - with a new, resampled, service time. We further assume that due to an overall general 'noise' in the system at large, whenever the server switches to station i, the station is found to be 'up' with probability p_i, and 'down' with probability $(1 - p_i)$. These probabilities are independent of whether the station was up or down when last visited. If the server finds that the station is down, it immediately switches on to the next node.

We assume independence between channels, arrival processes, service times, switch-over times and breakdowns.

Notation

We define the following:

- N = number of channels, indexed $i = 1, 2, \ldots, N$.
- Polling instant = the moment when the server 'enters' a channel.
- X_i^j = number of jobs at Q_j when Q_i is polled. $f_i^j = E[X_i^j]$
- $\underline{X_i} = (X_i^1, X_i^2, \ldots, X_i^j, \ldots, X_i^N)$ = system state at a polling instant of channel i (in steady state).
- $G_i(z_1, z_2, \ldots, z_N) = G_i(\underline{z}) \equiv E\left[\prod_{j=1}^{N} z_j^{X_i^j}\right]$ = PGF of the system state at a polling instant of node i.
- $\hat{W}(z) \equiv E\left[z^W\right]$ = PGF of a non-negative, integer-valued, random variable W.
- $\tilde{S}(\omega) = E\left[e^{-\omega S}\right]$ = LST of a non-negative continuous random variable S.
- C = cycle time.
- Y_i = number of jobs that are *successfully* served in queue i during a visit of the server, given that node i is up at the polling instant.
- Y_i^* = number of service attempts (successful and unsuccessful) made by the server each time it visits an 'up' queue i.
- $K_i^j(y)$ = number of jobs, out of y successful services, routed from Q_i to Q_j while the server is visiting channel i.
- T_i = time to breakdown of channel i. T_i is exponentially distributed with parameter γ_i. This channel-dependent process is regenerated at the beginning of every new visit of the server.
- Throughout the paper, all summations are cyclic. i.e. for any $i, j = 1, \ldots, N,\ j < i$ and for any function $g(k)$, $\sum_{k=i}^{j} g(k) = \sum_{k=i}^{N} g(k) + \sum_{k=1}^{j} g(k)$

3 The Gated Regime

In this section we analyze the Gated service discipline by which, in every visit to a queue, the server may serve only the jobs that are present at the polling instant. A job present at the polling instant will receive service if and only if the channel does not suffer a failure before its turn to be processed. Jobs rerouted back into the served queue (upon service completion) will wait for the next server's visit, before receiving service again..

3.1 Evolution Equations

The evolution equations for the system state $\underline{X}$ at polling instants depend on whether a node is *up* or *down* when polled.

If node i is *down* when polled, the server immediately moves to the next station, and $X_{i+1}^j = X_i^j$. If node i is *up* when polled, then

$$X_{i+1}^j = \begin{cases} X_i^j + K_i^j(Y_i) & j \neq i \\ X_i^i - Y_i + K_i^i(Y_i) & j = i \end{cases} \tag{1}$$

If node i is *up* but $X_i^i = 0$, the server switches over to the next station.

An interesting deduction from Eq. (1) is that the X_i^j's are independent of the service durations B_i or the switch over times D_i.

3.2 The Joint PGF of $\left\{K_i^j(Y_i)\right\}_{j=1}^N$

For the development of the PGFs of $\underline{X}_i$'s we need the joint PGF of $\left\{K_i^j(\bullet)\right\}_{j=1}^N$. If Q_i is up when polled, there will be Y_i successful completions during the server sojourn time, of which $K_i^j(Y_i)$ will be routed to Q_j. Thus,

$$\begin{aligned} \hat{K}_i(\underline{z}) &\equiv E\left[\prod_{j=1}^N z_j^{K_i^j(Y_i)}\right] \\ &= E\left[E\left[\prod_{j=1}^N z_j^{K_i^j(Y_i)} \middle| Y_i\right]\right] \end{aligned} \tag{2}$$

$$= E\left[\sum_{\sum_{j=1}^{N} m_j = Y_i} \prod_{j=1}^{N} z_j^{m_j} P\left[K_i^j(Y_i) = m_j, j = 1, 2, \ldots, N \middle| Y_i\right]\right]$$

Using Newton's multinomial equation,

$$\begin{aligned}
\hat{K}_i(\underline{z}) &= E\left[\sum_{\sum_{j=1}^{N} m_j = Y_i} \frac{Y_i!}{\prod_{j=1}^{N} m_j!} \prod_{j=1}^{N} (P_{ij} z_j)^{m_j}\right] \\
&= E\left[\left(\sum_{j=1}^{N} P_{ij} z_j\right)^{Y_i}\right] \qquad (3)
\end{aligned}$$

Defining $s_i(\underline{z}) \equiv \sum_{j=1}^{N} P_{ij} z_j$, Eq. (3) can be written as

$$\hat{K}_i(\underline{z}) = \hat{Y}_i(s_i(\underline{z})) \qquad (4)$$

Define $a_i \equiv P[T_i > B_i] = E[P[T_i > B_i | B_i]] = E\left[e^{-\gamma_i B_i}\right] = \tilde{B}_i(\gamma_i)$. Using the memoryless property of the breakdown process we have:

$$P\left[Y_i = j | X_i^i\right] = \begin{cases} a_i^j (1 - a_i) & j = 0, 1, \ldots, X_i^i - 1 \\ a_i^{X_i^i} & j = X_i^i \end{cases} \qquad (5)$$

$$\begin{aligned}
\hat{Y}_i\left(z | X_i^i\right) &\equiv E\left[z^{Y_i} \middle| X_i^i\right] \\
&= (1 - a_i) \frac{1 - (a_i z)^{X_i^i}}{1 - a_i z} + (a_i z)^{X_i^i} \qquad (6)
\end{aligned}$$

Combining (6) and (4) we conclude:

$$\begin{aligned}
\hat{K}_i(\underline{z}) &= \hat{Y}_i(s_i(\underline{z})) = E\left[\hat{Y}_i\left(s_i(\underline{z}) | X_i^i\right)\right] \\
&= E\left[(1 - a_i) \frac{1 - (a_i s_i(\underline{z}))^{X_i^i}}{1 - a_i s_i(\underline{z})} + (a_i s_i(\underline{z}))^{X_i^i}\right] \qquad (7)
\end{aligned}$$

Using (5) we can further deduct that

$$E\left[Y_i\right] = \frac{a_i}{1-a_i}\left(1-\hat{X}_i^i\left(a_i\right)\right) \tag{8}$$

and

$$E\left[K_i^j\left(Y_i\right)\right] = P_{ij}E\left[Y_i\right] = \frac{a_iP_{ij}}{1-a_i}\left(1-\hat{X}_i^i\left(a_i\right)\right) \tag{9}$$

Also,

$$P\left[Y_i^* = j \mid X_i^i\right] = \begin{cases} a_i^{j-1}(1-a_i) & j = 1, 2, \ldots, X_i^i - 1 \\ a_i^{X_i^i-1} & j = X_i^i \end{cases}$$

This readily implies that, for $j = 1, 2, \ldots, X_i^i$, $P\left[Y_i^* = j \mid X_i^i\right] = \frac{1}{a_i}P\left[Y_i = j \mid X_i^i\right]$. Thus, $E\left[Y_i^*\right] = \frac{1}{a_i}E\left[Y_i\right]$. Clearly, the expected number of service attempts until a successful one is $1/a_i$.

3.3 Generating Functions at Polling Instants

Let $G_i\left(\underline{z}\right) = G_i\left(z_1, z_2, \ldots z_N\right) = E\left[\prod_{j=1}^{N} z_j^{X_i^j}\right]$ be the PGF of the number of customers at the various queues at polling instants of queue i. Then, assuming node i is *up* when polled,

$$\begin{aligned} G_{i+1}\left(\underline{z}|up\right) &= E\left[\left(\prod_{j=1}^{N} z_j^{X_i^j+K_i^j(Y_i)}\right) z_i^{-Y_i}\right] \\ &= E\left[E\left[E\left[\prod_{j=1}^{N} z_j^{X_i^j+K_i^j(Y_i)} z_i^{-Y_i} \middle| Y_i, X_i^i\right] \middle| X_i^i\right]\right] \end{aligned} \tag{10}$$

From (4), $E\left[\prod_{j=1}^{N} z_j^{K_i^j(Y_i)}\right] = E\left[s_i\left(\underline{z}\right)^{Y_i}\right]$

Therefore

$$G_{i+1}\left(\underline{z}|up\right) = E\left[E\left[\left(\frac{s_i\left(\underline{z}\right)}{z_i}\right)^{Y_i} \prod_{j=1}^{N} z_j^{X_i^j} \middle| X_i^i\right]\right] \tag{11}$$

Using (6) we conclude that if node i is up when polled, then

$$G_{i+1}\left(\underline{z}|up\right) = E\left[\hat{Y}_i\left(\left.\frac{s_i\left(\underline{z}\right)}{z_i}\right| X_i^i\right)\prod_{j=1}^{N} z_j^{X_i^j}\right] \tag{12}$$

$$= E\left[\left((1-a_i)\frac{1-\left(a_i\frac{s_i(\underline{z})}{z_i}\right)^{X_i^i}}{1-a_i\frac{s_i(\underline{z})}{z_i}} + \left(a_i\frac{s_i\left(\underline{z}\right)}{z_i}\right)^{X_i^i}\right)\prod_{j=1}^{N} z_j^{X_i^j}\right]$$

Now,

$$E\left[(1-a_i)\frac{1-\left(a_i\frac{s_i(\underline{z})}{z_i}\right)^{X_i^i}}{1-a_i\frac{s_i(\underline{z})}{z_i}}\prod_{j=1}^{N} z_j^{X_i^j}\right]$$

$$= \frac{1-a_i}{1-\frac{a_i s_i(\underline{z})}{z_i}} E\left[\prod_{\substack{j=1\\ j\neq i}}^{N} z_j^{X_i^j}\left(z_i^{X_i^i} - \left(a_i s_i\left(\underline{z}\right)\right)^{X_i^i}\right)\right] \tag{13}$$

$$= \frac{1-a_i}{1-\frac{a_i s_i(\underline{z})}{z_i}}\left(G_i\left(\underline{z}\right) - G_i\left(z_1, z_2, \ldots, z_{i-1}, a_i s_i\left(\underline{z}\right), z_{i+1}, \ldots, z_N\right)\right)$$

Similarly,

$$E\left[\left(a_i\frac{s_i\left(\underline{z}\right)}{z_i}\right)^{X_i^i}\prod_{j=1}^{N} z_j^{X_i^j}\right]$$

$$= G_i\left(z_1, z_2, \ldots, z_{i-1}, a_i s_i\left(\underline{z}\right), z_{i+1}, \ldots, z_N\right) \tag{14}$$

Therefore (12) is simplified into

$$G_{i+1}(\underline{z}|up) = \frac{1-a_i}{1-\frac{a_i s_i(\underline{z})}{z_i}} G_i\left(\underline{z}\right) \tag{15}$$

$$+\frac{a_i\left(1-\frac{s_i(\underline{z})}{z_i}\right)}{1-\frac{a_i s_i(\underline{z})}{z_i}} G_i\left(z_1, z_2, \ldots, z_{i-1}, a_i s_i\left(\underline{z}\right), z_{i+1}, \ldots, z_N\right)$$

Finally,

$$
\begin{aligned}
G_{i+1}\left(\underline{z}\right) &= p_i \frac{1-a_i}{1-\frac{a_i s_i(\underline{z})}{z_i}} G_i\left(\underline{z}\right) \\
&+ p_i \frac{a_i\left(1-\frac{s_i(\underline{z})}{z_i}\right)}{1-\frac{a_i s_i(\underline{z})}{z_i}} G_i\left(z_1, z_2, \ldots, z_{i-1}, a_i s_i\left(\underline{z}\right), z_{i+1}, \ldots, z_N\right) \\
&+ (1-p_i) G_i\left(\underline{z}\right) \qquad (16)
\end{aligned}
$$

Remark 1 The PGF's $\{G_i(\cdot)\}$ are well-defined and exist whenever the invariant distribution (or the corresponding Cezaro limits) of the routing matrix P exist. Clearly, when P is irreducible the above holds.

3.4 The Busy Period and Expected Number of Jobs Left at the End of a Visit

Let $\theta^i_{X^i_i}$ be the duration of a busy period in station i, under the Gated regime (and with possible failing node). If the server is up when station i is polled, then $\theta^i_{X^i_i} = \sum_{m=1}^{\min\left(Y_i+1, X^i_i\right)} B_{im}$, where $\min\left(Y_i+1, X^i_i\right) = Y^*_i$. Otherwise $\theta^i_{X^i_i} = 0$. That is, the length of a busy period is comprised of either X^i_i consecutive successful services, or of $Y_i (< X^i_i)$ consecutive successful services followed by an unsuccessful service.

Let S_i $[\underline{S}_i]$ be the duration of a successful [unsuccessful] service attempt at channel i. Clearly, $S_i \sim B_i|_{B_i<T_i}$ $[\underline{S}_i \sim B_i|_{B_i \geq T_i}]$. It has been shown by Kofman and Yechiali [1996] that:

$$
E\left[S_i\right] = \frac{E\left[B_i e^{-\gamma_i B_i}\right]}{\tilde{B}_i\left(\gamma_i\right)} \qquad (17)
$$

$$
E\left[\underline{S}_i\right] = \frac{b_i - E\left[B_i e^{-\gamma_i B_i}\right]}{1-\tilde{B}_i\left(\gamma_i\right)} \qquad (18)
$$

Let ρ^i_k denote the probability of failing during the service of the k-th customer in station i. Clearly,

$$
\rho^i_k = a_i^{k-1}(1-a_i) \qquad\qquad k = 1, \ldots, X^i_i \qquad (19)
$$

Then,

$$
\begin{aligned}
E\left[\theta^i_{X^i_i}\right] &= p_i E\left[\sum_{k=1}^{X^i_i} \rho^i_k[(k-1)\,E\left[S_i\right] + E\left[\underline{S}_i\right]] + (1-\sum_{k=1}^{X^i_i}\rho^i_k)X^i_i E\left[S_i\right]\right] \\
&= p_i E\left[\frac{1-a_i^{X^i_i}}{1-a_i} a_i E\left[S_i\right] + (1-a_i^{X^i_i})E\left[\underline{S}_i\right]\right] \qquad (20)
\end{aligned}
$$

Combining (20) with (17) and (18) we conclude:

$$
E\left[\theta^i_{X^i_i}\right] = \frac{p_i b_i}{1-a_i}\left(1-\hat{X}^i_i\left(a_i\right)\right) = p_i b_i E\left[Y^*_i\left(X^i_i\right)\right] \qquad (21)
$$

That is, the mean duration of a busy period equals the mean number of service attempts, $E\left[Y^*_i\right]$, multiplied by the mean duration of a service attempt, b_i, and further factored by the probability p_i that the station is up.

Let R^i_r be the number of *remaining* jobs at the end of a sojourn period of the server at station i, initiated with r jobs.

$$
E\left[R^i_{X^i_i}\right] = E\left[X^i_i\right] - p_i E\left[Y_i\right](1-P_{ii}) \qquad (22)
$$

Using (22) and (8) we get:

$$
E\left[R^i_{X^i_i}\right] = E\left[X^i_i\right] - p_i a_i \frac{1-\hat{X}^i_i\left(a_i\right)}{1-a_i}\left(1-P_{ii}\right)p_i \qquad (23)
$$

3.5 First-Order Moments

Let $f^j_i = E\left[X^j_i\right]$. It can be evaluated either by differentiating the PGF with respect to z_j and evaluating at $\underline{z} = \underline{1}$ or by using the Evolution Equations. Using the latter method we obtain:

$$
f^j_{i+1} = \begin{cases} f^j_i + p_i P_{ij} E\left[Y_i\right] & j \neq i \\ f^i_i - \left(1-P_{ii}\right)p_i E\left[Y_i\right] & j = i \end{cases} \qquad (24)
$$

We can simplify (24) by defining $x_i \equiv p_i E\left[Y_i\right]$, where $E\left[Y_i\right]$ is given by (8). x_i is the expected number of *successful* service completion per visit of channel i.

$$
f^j_{i+1} = \begin{cases} f^j_i + x_i P_{ij} & j \neq i \\ f^i_i - x_i\left(1-P_{ii}\right) & j = i \end{cases} \qquad (25)
$$

we now show:
Proposition 1

$$\underline{x}^T = \underline{x}^T P \tag{26}$$

Proof
Since $\sum_{i=1}^{N} f_{i+1}^j = \sum_{i=1}^{N} f_i^j$, we obtain, for every j,

$$0 = \sum_{\substack{i=1 \\ i \neq j}}^{N} x_i P_{ij} - (1 - P_{jj}) x_j \tag{27}$$

or $\underline{x}^T A = \underline{0}$ where

$$A = \begin{bmatrix} 1 - P_{1,1} & -P_{1,2} & -P_{1,3} & \cdots & -P_{1,N} \\ -P_{2,1} & 1 - P_{2,2} & -P_{2,3} & \cdots & -P_{2,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -P_{N,1} & -P_{N,2} & -P_{N,3} & \cdots & 1 - P_{N,N} \end{bmatrix}$$

That is, $A = I - P$, and proposition 1 follows.

Thus, $\underline{x}$ is an eigenvector of P^T with an eigenvalue of 1, and for some $c > 0$ there exists a solution $\underline{x} = c\underline{\pi}$, where $\underline{\pi}$ is the steady state probability distribution of P. Moreover, since $\sum_i \pi_i = 1$, $c = \sum_{i=1}^{N} x_i$. That is, c expresses the expected total number of successful service completions during a full cycle.

It follows that, for all i, j:

$$\frac{x_i}{x_j} = \frac{\pi_i}{\pi_j} = \frac{p_i E\left[Y_i\right]}{p_j E\left[Y_j\right]} \tag{28}$$

This gives us $N - 1$ linearly independent equations for x_i:

$$x_i = \frac{\pi_i}{\pi_1} x_1, \quad i = 2, 3, \ldots, N \tag{29}$$

Still, without the knowledge of the constant c, the x_i's are not uniquely determined.

3.5.1 Finding c

We first observe that, for every i,

$$\sum_{j=1}^{N} f_i^j = M \tag{30}$$

Using equations (25)

$$\begin{aligned} f_1^j &= f_N^j + P_{Nj} x_N \\ &= (f_{N-1}^j + P_{N-1,j} x_{N-1}) + P_{Nj} x_N \\ &= \cdots = f_{j+1}^j + \sum_{k=j+1}^{N} P_{kj} x_k = f_j^j - x_j + \sum_{k=j}^{N} P_{kj} x_k \end{aligned} \tag{31}$$

Since $x_i = c\pi_i$, we write (31) as

$$f_1^j = f_j^j - c\pi_j + c \sum_{k=j}^{N} P_{kj} \pi_k \tag{32}$$

Combining (30) for $i = 1$ $\left(\sum_{j=1}^{N} f_1^j = M \right)$ with (32) we have

$$M = \sum_{j=1}^{N} f_j^j + c \sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj} \pi_k - \pi_j \right] \tag{33}$$

Defining K as the constant: $K \equiv \sum_{j=1}^{N} f_j^j$ we conclude that

$$c = \frac{M - K}{\sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj} \pi_k - \pi_j \right]} = \frac{K - M}{\sum_{j=1}^{N} \left[\sum_{k=1}^{j-1} P_{k_j} \pi_k \right]} \tag{34}$$

However, the value of K is unknown. Yet, since c must be non-negative and $\pi_j = \sum_{k=1}^{N} P_{kj} \pi_k \geq \sum_{k=j}^{N} P_{kj} \pi_k$, we conclude that

$$K \equiv \sum_{i=1}^{N} f_i^i \geq M \tag{35}$$

Since the information on the x_i's may be obtained only through their ratios (Eq.(29)) we cannot obtain explicitly their values. Furthermore, given the nature of equations (25), even the knowledge of x_i would not be sufficient to determine the values of the f_i^j's. We therefore turn to find an approximated solution.

3.6 An Approximated Solution

Using Taylor series expansion for $\hat{X}_i^i(a_i)$ we have

$$\begin{aligned} \hat{X}_i^i(a_i) &= E\left[\sum_{k=0}^{\infty} \frac{(\ln a_i)^k}{k!} \left(X_i^i\right)^k\right] \\ &\approx E\left[1 + (\ln a_i)\ X_i^i\right] \\ &= 1 + (\ln a_i)\ E\left[X_i^i\right] \end{aligned} \tag{36}$$

Combining (36) with the x_i $\left(x_i \equiv p_i E[Y_i] = p_i \frac{a_i}{1-a_i}\left(1 - \hat{X}_i^i(a_i)\right)\right)$ we write

$$x_i \approx \left(-p_i \frac{a_i}{1-a_i} \ln a_i\right) f_i^i \tag{37}$$

defining $e_i \equiv -p_i \frac{a_i}{1-a_i} \ln a_i$, we can write (37) as

$$x_i \approx e_i f_i^i \tag{38}$$

Going back to equation (33),

$$M = \sum_{j=1}^{N} f_j^j + c \sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj}\pi_k - \pi_j\right]$$

and substituting $\frac{x_j}{e_j}$ for f_j^j we get,

$$\begin{aligned} M &\approx \sum_{j=1}^{N} \frac{x_j}{e_j} + c \sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj}\pi_k - \pi_j\right] \\ &= c \sum_{j=1}^{N} \frac{\pi_j}{e_j} + c \sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj}\pi_k - \pi_j\right] \end{aligned} \tag{39}$$

Therefore,

$$c \approx \frac{M}{\sum\limits_{j=1}^{N} \frac{\pi_j}{e_j} + \sum\limits_{j=1}^{N} \left[\sum\limits_{k=j}^{N} P_{kj}\pi_k - \pi_j \right]} \tag{40}$$

Hence,

$$\begin{aligned} f_i^i &\approx \frac{x_i}{e_i} = \frac{c\pi_i}{e_i} \\ &\approx \frac{M\pi_i}{-p_i \frac{a_i}{1-a_i} \ln a_i \left(\sum\limits_{j=1}^{N} \frac{\pi_j}{e_j} + \sum\limits_{j=1}^{N} \left[\sum\limits_{k=j}^{N} P_{kj}\pi_k - \pi_j \right] \right)} \end{aligned} \tag{41}$$

To complete the solution for all f_i^j, we use (25) in a similar manner used for the derivation of (31):

$$\begin{aligned} f_i^j &= f_{i-1}^j + x_{i-1}P_{i-1,j} = \cdots = f_{j+1}^j + \sum_{k=j+1}^{i-1} x_k P_{kj} \\ &= f_j^j - x_j\left(1 - P_{jj}\right) + \sum_{k=j+1}^{i-1} x_k P_{kj} \end{aligned} \tag{42}$$

Thus for $j \neq i$,

$$f_i^j \approx x_j \left(\frac{1}{e_j} - 1 \right) + \sum_{k=j}^{i-1} x_k P_{kj} \tag{43}$$

Therefore

$$f_i^j \approx \begin{cases} x_j \left(\frac{1}{e_j} - 1 \right) + \sum\limits_{k=j}^{i-1} x_k P_{kj} & j \neq i \\ \frac{x_j}{e_j} & j = i \end{cases} \tag{44}$$

3.7 How Good is the Approximation?

For ease of notation we set in this discussion $X_i^i = X$, $a_i = a$. We're interested in situations where $a < 1$, implying that $\ell_n a < 0$.

In Section 3.6 we use the relation $\hat{X}(a) = E\left[a^X\right] = E\left[e^{(\ell_n a)X}\right]$. Now, for real $t > 0$, it can be shown that

$$1 - t < e^{-t} < 1 - t + \frac{t^2}{2}$$

Hence,

$$1 + (\ell_n a)\, E[X] < \hat{X}(a) < 1 + (\ell_n a)\, E[X] + \frac{(\ell_n a)^2}{2} E\left[X^2\right]$$

It follows that the error of the approximation is bounded as follows:

$$\text{error } \equiv \hat{X}(a) - (1 + (\ell_n a)\, E[X]) < \frac{(\ell_n a)^2}{2} E\left[X^2\right]$$

In modern (communication) systems the probability of successful service completion should not be far from 1. Thus, considering situations where $a \geq 0.95$ (say), then $(\ell_n a)^2 / 2 \leq 0.0013155$. Consequently, if we wish an error ≤ 0.1, our approximation is good for all cases where $E\left[X^2\right] < 0.1/0.0013155 = 76$. (The corresponding value for $a = 0.9$ is $E\left[X^2\right] < 0.1/0.00555 = 18$). Thus, when $E\left[X^2\right]$ is not too large, our approximation is good enough.

4 Gated, Cyclic Bernoulli Polling

A special case of polling in a closed network with station breakdowns is Cyclic Bernoulli Polling. As in the general gated polling regime, there are breakdown occurrences. At polling instants, after leaving channel $i-1$, the server enters channel i only if it is "up", which has a probability of p_i, $0 < p_i \leq 1$. If channel i is "down" the server moves immediately to the next channel. However, in Cyclic Bernoulli Polling, the server can not have a breakdown while serving a channel. That is, for every i, $a_i = 1$.

The class of open Bernoulli random polling systems was analyzed extensively in Altman and Yechiali [1993]. We shall now analyze the gated regime for a closed system.

4.1 PGF and First-Order Moments

Letting $a_i = 1$ in (16) we readily have:

$$
\begin{aligned}
G_{i+1}(\underline{z}) &= p_i G_i\left(z_1, z_2, \ldots, z_{i-1}, s_i(\underline{z}), z_{i+1}, \ldots, z_N\right) \\
&\quad +(1-p_i)G_i(\underline{z})
\end{aligned} \tag{45}
$$

Since $E[Y_i] = E[X_i^i] = f_i^i$ equations (24) reduce to:

$$
f_{i+1}^j = \begin{cases} f_i^j + p_i P_{ij} f_i^i & j \neq i \\ (1-p_i) f_i^i + p_i P_{ii} f_i^i & j = i \end{cases} \tag{46}
$$

and

$$
x_i = p_i E[Y_i] = p_i f_i^i \tag{47}
$$

Using (33) we obtain

$$
\begin{aligned}
M &= \sum_{j=1}^{N} f_j^j + c \sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj}\pi_k - \pi_j \right] \\
&= c \sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj}\pi_k + \pi_j \left(\frac{1}{p_j} - 1 \right) \right]
\end{aligned}
$$

from which we conclude that the constant c is:

$$
c = \frac{M}{\sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj}\pi_k + \frac{\pi_j}{p_j}(1-p_j) \right]} \tag{48}
$$

It is easy to see that when setting $p_i = 1$ for all i, we have the same results as in Altman and Yechiali [1994].

Knowing c and the f_i^i values, we can determine the values of any f_i^j, using (42):

$$
\begin{aligned}
f_i^j &= f_j^j - x_j(1 - P_{jj}) + \sum_{k=j+1}^{i-1} x_k P_{kj} \\
&= c \left(\frac{1-p_j}{p_j} \pi_j + \sum_{k=j}^{i-1} P_{kj}\pi_k \right)
\end{aligned} \tag{49}
$$

4.2 Cycle Duration

Define the cycle duration C as the time, in steady state, between two consecutive arrival instants of the server to some given queue (say queue 1). Then,

$$E[C] = \sum_{i=1}^{N} p_i E[Y_i] b_i + \sum_{i=1}^{N} d_i$$
$$= \sum_{i=1}^{N} x_i b_i + \sum_{i=1}^{N} d_i \tag{50}$$

As the mean total switching time during a full cycle is $d = \sum_{i=1}^{N} d_i$,

$$E[C] = c \sum_{i=1}^{N} \pi_i b_i + d \tag{51}$$

It will be shown in the sequel that $E[C] = \frac{d}{1-\rho}$, as in many polling configurations (see e.g. Watson [1984]).

4.3 PGF of Number of Customers at an Arbitrary Moment

Let $F^*(\underline{z})$ be the PGF of the number of customers at different queues at an arbitrary moment (in stationary regime). Similar to Altman and Yechiali [1994], $F^*(\underline{z})$ can be expressed as

$$F^*(\underline{z}) = \frac{1}{E[C]} \sum_{i=1}^{N} p_i E[Y_i] b_i F^*(\underline{z}|\,\text{service period } i)$$
$$+ \frac{1}{E[C]} \sum_{i=1}^{N} d_i F^*(\underline{z}|\,\text{switch} - \text{over period } i)$$
$$= \frac{1}{E[C]} \sum_{i=1}^{N} x_i b_i F^*(\underline{z}|\,\text{service period } i)$$
$$+ \frac{1}{E[C]} \sum_{i=1}^{N} d_i F^*(\underline{z}|\,\text{switch} - \text{over period } i) \tag{52}$$

To obtain $F^*(\underline{z})$ we introduce the following PGFs of the number of customers at:

1. an arbitrary switch-over instant at queue i as $\bar{F}_i(\underline{z})$
2. an arbitrary service beginning instant at queue i as $V_i(\underline{z})$
3. an arbitrary service completion instant at queue i as $\bar{V}_i(\underline{z})$

Adapting Eisenberg's [1972] result (see also Fuhrmann, Levy and Sidi [1992] and Altman and Yechiali [1994]) we have

$$G_i(\underline{z}) + p_i E[Y_i]\bar{V}_i(\underline{z}) = \bar{F}_i(\underline{z}) + p_i E[Y_i] V_i(\underline{z}) \tag{53}$$

$$\bar{V}_i(\underline{z}) = V_i(\underline{z}) \sum_{j=1}^{N} P_{ij} \frac{z_j}{z_i} \tag{54}$$

$$G_i(\underline{z}) = \bar{F}_{i-1}(\underline{z}) \tag{55}$$

Eq. (53) states that against each event of entering a station or a job service ending there is one event of station leaving or a job service beginning. Eq. (54) states that at the end of a service there is one customer less in the queue being served but one more in the queue to which this customer was routed. Eq. (55) follows because there are no arrivals to any queue during a switch-over period. For the same reason:

$$F^*(\underline{z} | \,\mathrm{switch-over\ period}\ i) = G_{i+1}(\underline{z}) \tag{56}$$

Since there are no arrivals during a service period

$$F^*(\underline{z} | \,\mathrm{service\ period}\ i) = V_i(\underline{z}) \tag{57}$$

Combining (53), (54) and (55) we have

$$\begin{aligned}
\bar{V}_i(\underline{z}) - V_i(\underline{z}) &= V_i(\underline{z}) \left(\sum_{j=1}^{N} P_{ij} \frac{z_j}{z_i} - 1 \right) \\
&= \frac{\bar{F}_i(\underline{z}) - G_i(\underline{z})}{p_i E[Y_i]} \\
&= \frac{G_{i+1}(\underline{z}) - G_i(\underline{z})}{p_i E[Y_i]}
\end{aligned} \tag{58}$$

Substituting into (57) yields

$$
\begin{aligned}
F^*\left(\underline{z}\right|\text{service period } i) &= V_i\left(\underline{z}\right) \\
&= \left(\frac{G_{i+1}\left(\underline{z}\right) - G_i\left(\underline{z}\right)}{\sum\limits_{j=1}^{N} P_{ij} z_j - z_i}\right) \frac{z_i}{p_i E\left[Y_i\right]}
\end{aligned}
\tag{59}
$$

Thus,

$$
\begin{aligned}
F^*\left(\underline{z}\right) &= \frac{1}{E\left[C\right]} \sum_{i=1}^{N} \frac{G_{i+1}\left(\underline{z}\right) - G_i\left(\underline{z}\right)}{\sum\limits_{j=1}^{N} P_{ij} z_j - z_i} b_i z_i \\
&\quad + \frac{1}{E\left[C\right]} \sum_{i=1}^{N} d_i G_{i+1}\left(\underline{z}\right)
\end{aligned}
\tag{60}
$$

The first moments of the queue's length at an arbitrary moment can now be obtained by differentiating $F^*\left(\underline{z}\right)$.

4.4 Sojourn Times, Throughputs and System's Utilization

Denote by Λ the throughput of the system, which is the rate at which customers are successfully served (or rerouted). It is given by

$$
\Lambda = \frac{\sum\limits_{i=1}^{N} p_i E\left[Y_i\right]}{E\left[C\right]} = \frac{c}{E\left[C\right]}
\tag{61}
$$

This follows from the fact (indicated before) that the expected number of successfully served customers during a cycle is $\sum\limits_{i=1}^{N} p_i E\left[Y_i\right] = \sum\limits_{i=1}^{N} x_i = c \sum\limits_{i=1}^{N} \pi_i = c$.

Let λ_i be the throughput of queue i, i.e. the rate at which customers leave that queue. It is equal to the expected total number of customers leaving queue i during a cycle divided by the expected cycle duration:

$$
\lambda_i = \frac{p_i E\left[Y_i\right]}{E\left[C\right]} = \frac{c\pi_i}{E\left[C\right]} = \Lambda \pi_i
\tag{62}
$$

Let ℓ_i be the expected length of queue i at an arbitrary moment at the stationary regime (it may be obtained by differentiating Eq. (60)). Let w_i be the expected sojourn time of an arbitrary customer at queue i (total time from arrival to departure). By Little's Law we have

$$w_i = \frac{\ell_i}{\lambda_i} = \frac{\ell_i}{\Lambda \pi_i} \tag{63}$$

The utilization of the system is given by

$$\begin{aligned} \rho &\equiv \frac{\sum_{i=1}^{N} p_i E\left[Y_i\right] b_i}{E\left[C\right]} = \frac{c}{E\left[C\right]} \sum_{i=1}^{N} \pi_i b_i \\ &= \Lambda \sum_{i=1}^{N} \pi_i b_i \end{aligned} \tag{64}$$

where ρ is the proportion of time that the server is working effectively, namely serving 'successful' customers. Dividing both sides of (51) by $E\left[C\right]$ we get

$$\begin{aligned} 1 &= \frac{c}{E\left[C\right]} \sum_{i=1}^{N} \pi_i b_i + \frac{d}{E\left[C\right]} \\ &= \rho + \frac{d}{E\left[C\right]} \end{aligned} \tag{65}$$

which yields, as can be expected from the general balance argument, that

$$E\left[C\right] = \frac{d}{1-\rho} \tag{66}$$

The results for $E\left[C\right]$, $F^*\left(\underline{z}\right)$, Λ and λ_i are identical in appearance to their parallel in Altman and Yechiali [1994]. The *difference* between the problems is hidden in the value of the constant c.

5 The Exhaustive Regime

Under the Exhaustive service discipline, if channel i is up when polled, the server starts serving the jobs present there (if any), and continues to do so until either all work accumulated is done (including service of jobs rerouted back into channel i) or the channel fails, whichever occurs first. The server then moves to node $i+1$. As before, we assume that a failure is discovered only at the end of a service period of a job.

The evolution laws of $\underline{X}_i$ have the same form as in the gated regime:

if node i is down upon polling then for all j

$$X_{i+1}^j = X_i^j \tag{67}$$

if node i is up upon polling, then

$$X_{i+1}^j = \begin{cases} X_i^j + K_i^j(Y_i) & j \neq i \\ X_i^i - Y_i(X_i^i) + K_i^i(Y_i) & j = i \end{cases} \tag{68}$$

However, Y_i and $K_i^j(Y_i)$, differ from their gated counterparts. Note that under the exhaustive regime Y_i may exceed X_i^i.

5.1 Number of Rerouted Jobs and Successful Services

Denoting $y_i(k) = E[Y_i | X_i^i]$ we can write the following recursive equation

$$\begin{aligned} y_i(k+1) &= a_i(1 + (1-P_{ii})y_i(k) + P_{ii}y_i(k+1)) \\ y_i(0) &= 0 \end{aligned} \tag{69}$$

or

$$y_i(k+1) = \frac{a_i(1 + (1-P_{ii})y_i(k))}{1 - a_i P_{ii}} \tag{70}$$

The solution for (70) is

$$y_i(k) = \frac{a_i}{1-a_i}\left(1 - \left(a_i \frac{1-P_{ii}}{1-a_i P_{ii}}\right)^k\right) \tag{71}$$

Hence,

$$\begin{aligned} E[Y_i] &= \frac{a_i}{1-a_i}\left(1 - E\left[\left(a_i \frac{1-P_{ii}}{1-a_i P_{ii}}\right)^{X_i^i}\right]\right) \\ &= \frac{a_i}{1-a_i}\left(1 - \hat{X}_i^i\left(a_i \frac{1-P_{ii}}{1-a_i P_{ii}}\right)\right) \end{aligned} \tag{72}$$

and

$$E\left[K_i^j\left(Y_i\right)\right]=P_{ij}E\left[Y_i\right] \tag{73}$$

5.2 Generating Functions at Polling Instants

Let $G_i\left(\underline{z}\right)=G_i\left(z_1,z_2,\ldots z_N\right)=E\left[\prod_{j=1}^{N} z_j^{X_i^j}\right]$ be the PGF of the number of customers at the various queues at polling instants of queue i. Then, assuming node i is *up* when polled,

$$G_{i+1}\left(\underline{z}|up\right)=E\left[\left(\prod_{j=1}^{N} z_j^{X_i^j+K_i^j(Y_i)}\right)z_i^{-Y_i}\right] \tag{74}$$

Eq. (3), $E\left[\prod_{j=1}^{N} z_j^{K_i^j(Y_i)}\right]=E\left[s_i\left(\underline{z}\right)^{Y_i}\right]$, holds for the exhaustive case, too. Hence

$$G_{i+1}\left(\underline{z}|up\right)=E\left[\hat{Y}_i\left(\left.\frac{s_i\left(\underline{z}\right)}{z_i}\right|X_i^i\right)\prod_{j=1}^{N} z_j^{X_i^j}\right] \tag{75}$$

Therefore

$$G_{i+1}\left(\underline{z}\right)=p_iE\left[\hat{Y}_i\left(\left.\frac{s_i\left(\underline{z}\right)}{z_i}\right|X_i^i\right)\prod_{j=1}^{N} z_j^{X_i^j}\right]+\left(1-p_i\right)G_i\left(\underline{z}\right) \tag{76}$$

In order to evaluate (76) we need an expression for $\hat{Y}_i\left(\left.\frac{s_i(\underline{z})}{z_i}\right|X_i^i\right)$. It can be resolved using a recursive equation for $P\left[Y_i=k|X_i^i=n\right]$. Denoting $t_i\left(k,n\right)\equiv P\left[Y_i=k|X_i^i=n\right]$, the recursive equation is:

$$t_i\left(k,n\right)=\begin{cases} a_i\left(\left(1-P_{ii}\right)t_i\left(k-1,n-1\right)+P_{ii}t_i\left(k-1,n\right)\right) & n>0,\ k>0 \\ 1-a_i & n>0,\ k=0 \\ 0 & n=0,\ k>0 \\ 1 & n=0,\ k=0 \end{cases} \tag{77}$$

5.3 First Order Moments

Having the same form for the evolution equations (though with different values of $E[Y_i]$), we immediately have the Gated regime's expressions:

$$f_{i+1}^j = \begin{cases} f_i^j + p_i P_{ij} E[Y_i] & j \neq i \\ f_i^i - (1 - p_i) P_{ii} E[Y_i] & j = i \end{cases} \tag{78}$$

$$x_i \equiv p_i E[Y_i] = c\pi_i$$

$$f_{i+1}^j = \begin{cases} f_i^j + x_i P_{ij} & j \neq i \\ f_i^i - x_i (1 - P_{ii}) & j = i \end{cases} \tag{79}$$

$$c = \frac{M - \sum_{j=1}^{N} f_j^j}{\sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj}\pi_k - \pi_j \right]} \tag{80}$$

Since $E[Y_i] = \frac{a_i}{1-a_i}\left(1 - \hat{X}_i^i\left(a_i \frac{1-P_{ii}}{1-a_i P_{ii}}\right)\right)$, an approximated solution will have a slightly different form.

5.4 An Approximated Solution

Using Taylor series expansion for $\hat{X}_i^i\left(a_i \frac{1-P_{ii}}{1-a_i P_{ii}}\right)$ we have

$$\hat{X}_i^i\left(a_i \frac{1-P_{ii}}{1-a_i P_{ii}}\right) \approx 1 + \ln\left(a_i \frac{1-P_{ii}}{1-a_i P_{ii}}\right) E\left[X_i^i\right] \tag{81}$$

Hence,

$$x_i \approx -\frac{p_i a_i}{1-a_i} \ln\left(a_i \frac{1-P_{ii}}{1-a_i P_{ii}}\right) f_i^i \tag{82}$$

Defining

$$e_i \equiv -\frac{p_i a_i}{1-a_i} \ln\left(a_i \frac{1-P_{ii}}{1-a_i P_{ii}}\right) \tag{83}$$

we can write (82) as

$$x_i \approx e_i f_i^i \tag{84}$$

Thus, using results from the gated regime's approximated solution (see equations (40) and (44))

$$c \approx \frac{M}{\sum_{j=1}^{N} \frac{\pi_j}{e_j} + \sum_{j=1}^{N} \left[\sum_{k=j}^{N} P_{kj} \pi_k - \pi_j \right]} \tag{85}$$

$$f_i^j \approx \begin{cases} x_j \left(\frac{1}{e_j} - 1 \right) + \sum_{k=j}^{i-1} x_k P_{kj} & j \neq i \\ \frac{x_j}{e_j} & j = i \end{cases} \tag{86}$$

6 Exhaustive, Cyclic Bernoulli Polling

Limiting ourselves again into cases where $a_i = 1 \; \forall i$, Eq. (70) takes the form

$$y_i (k+1) = \frac{1 + (1 - P_{ii}) y_i (k)}{1 - P_{ii}} \tag{87}$$

to which the solution is

$$y_i (k) = \frac{k}{1 - P_{ii}} \tag{88}$$

Hence, for exhaustive cyclic Bernoulli polling

$$E [Y_i] = \frac{E [X_i^i]}{1 - P_{ii}} \tag{89}$$

$$E \left[K_i^j (Y_i) \right] = \frac{P_{ij}}{1 - P_{ii}} E \left[X_i^i \right] \tag{90}$$

6.1 First-Order Moments

For CBP under the exhaustive regime

$$x_i = p_i E [Y_i] = \frac{p_i f_i^i}{1 - P_{ii}} = c \pi_i \tag{91}$$

hence,

$$f_i^i = \frac{1 - P_{ii}}{p_i} x_i \tag{92}$$

Using Eq. (33), which is valid for the exhaustive regime, too,

$$\begin{aligned} M &= \sum_{j=1}^{N} f_j^j + c\sum_{j=1}^{N}\left[\sum_{k=j}^{N} P_{kj}\pi_k - \pi_j\right] \\ &= c\sum_{j=1}^{N}\left[\sum_{k=j}^{N} P_{kj}\pi_k + \frac{1-P_{jj}-p_j}{p_j}\pi_j\right] \end{aligned} \tag{93}$$

Thus,

$$c = \frac{M}{\sum_{j=1}^{N}\left[\sum_{k=j}^{N} P_{kj}\pi_k + \frac{\pi_j}{p_j}\left(1-p_j-P_{jj}\right)\right]} \tag{94}$$

and, similar to (42),

$$\begin{aligned} f_i^j &= f_j^j - x_j\left(1-P_{jj}\right) + \sum_{k=j+1}^{i-1} x_k P_{kj} \\ &= c\left(\left(\frac{1}{p_j}-1\right)\left(1-P_{jj}\right)\pi_j + \sum_{k=j+1}^{i-1}\pi_k P_{kj}\right) \end{aligned} \tag{95}$$

The expected cycle duration, the PGF of number of customers at an arbitrary moment, the sojourn times, the throughputs and system's utilization are developed in the same way as for the gated regime. Their value differ only by the value of the constant c.

7 The Globally Gated Regime

The (cyclic) Globally Gated service regime brings the polling system closer to the (fair) first-come first-served discipline (in comparison to the regular Gated and Exhaustive disciplines). Under this service policy *all* gates are closed simultaneously (globally) at the moment the server polls queue 1. At that instant the system state is $\underline{X}_1$, and during the coming cycle only those X_1^j customers "captured" (present) at Q_j at the start of the cycle can be served at Q_j.

7.1 Evolution Equations

The evolution equations keep the form of the other two regimes:

If node i is down when polled, the server immediately moves to the next station, and hence for every j

$$X_{i+1}^{j} = X_i^j \tag{96}$$

If node i is up when polled, then

$$X_{i+1}^{j} = \begin{cases} X_i^j + K_i^j(Y_i) & j \neq i \\ X_i^i - Y_i + K_i^i(Y_i) & j = i \end{cases} \tag{97}$$

But, the value of Y_i depends on X_1^i, unlike X_i^i previously. By replacing X_i^i with X_1^i in equations (20) and (21), we learn that the expected duration of a busy period in station i, $\theta^i_{X_1^i}$ is

$$E\left[\theta^i_{X_1^i}\right] = \frac{p_i b_i}{1-a_i}\left(1 - \hat{X}_1^i(a_i)\right) \tag{98}$$

Eq. (8) turns into

$$E[Y_i] = \frac{a_i}{1-a_i}\left(1 - \hat{X}_1^i(a_i)\right) \tag{99}$$

and Eq. (9) into

$$E\left[K_i^j(Y_i)\right] = \frac{a_i P_{ij}}{1-a_i}\left(1 - \hat{X}_1^i(a_i)\right) \tag{100}$$

7.2 Generating Functions at Polling Instants

Let $\underline{W}$=$\underline{X}_1$ denote the state vector of the number of jobs at the various queues at the *start* of a cycle (clearly, $W_j = X_1^j$, $j = 1, 2, \ldots, N$). We extend the system's state vector to include both $\underline{W}$ and $\underline{X}_i$, such that at a polling instant of station i, the state is given by

$(\underline{W}; \underline{X}_i) = \left(W_1, W_2, \ldots, W_N; X_i^1, X_i^2, \ldots, X_i^N\right)$.

The PGF at that instant is given by

$$G_i(\underline{w}; \underline{z}) = E\left[\prod_{j=1}^{N} w_j^{W_j} \prod_{j=1}^{N} z_j^{X_i^j}\right] \tag{101}$$

Using the evolution equations (97), when node i is up

$$
\begin{aligned}
G_{i+1}\left(\underline{w};\underline{z}|up\right) &= E\left[\prod_{j=1}^{N} w_j^{W_j}\prod_{j=1}^{N} z_j^{X_i^j+K_i^j(Y_i)} z_i^{-Y_i}\right] \\
&= E\left[\prod_{j=1}^{N} w_j^{W_j}\prod_{j=1}^{N} z_j^{X_i^j+K_i^j(Y_i)} z_i^{-Y_i}\right] \qquad (102)\\
&= E\left[E\left[E\left[\prod_{j=1}^{N} w_j^{W_j}\prod_{j=1}^{N} z_j^{X_i^j+K_i^j(Y_i)} z_i^{-Y_i}\middle| Y_i, W_i\right]\middle| W_i\right]\right]
\end{aligned}
$$

It is easy to see that, similar to equations (4) and (6) in the regular gated regime, using $W_i = X_1^i$,

$$
\hat{K}_i(\underline{z}) = \hat{Y}_i\left(s_i\left(\underline{z}\right)\right) \qquad (103)
$$

and

$$
\hat{Y}_i\left(z\right) = E\left[(1-a_i)\frac{1-\left(a_i z\right)^{W_i}}{1-a_i z} + \left(a_i z\right)^{W_i}\right] \qquad (104)
$$

Therefore,

$$
\begin{aligned}
G_{i+1}\left(\underline{w};\underline{z}|up\right) &= E\left[E\left[E\left[\prod_{j=1}^{N} w_j^{W_j}\prod_{j=1}^{N} z_j^{X_i^j+K_i^j(Y_i)} z_i^{-Y_i}\middle| Y_i, W_i\right]\middle| W_i\right]\right] \\
&= E\left[E\left[\prod_{j=1}^{N} w_j^{W_j}\prod_{j=1}^{N} z_j^{X_i^j}\hat{Y}_i\left(W_i\right)\left(\frac{s_i\left(\underline{z}\right)}{z_i}\right)\middle| W_i\right]\right] \qquad (105)\\
&= E\left[\prod_{j=1}^{N} w_j^{W_j}\prod_{j=1}^{N} z_j^{X_i^j}\left((1-a_i)\frac{1-\left(a_i\frac{s_i(\underline{z})}{z_i}\right)^{W_i}}{1-a_i\frac{s_i(\underline{z})}{z_i}} + (a_i\frac{s_i\left(\underline{z}\right)}{z_i})^{W_i}\right)\right]
\end{aligned}
$$

Dividing the RHS of (105) into two, we can simplify the first part to

$$
E\left[\prod_{j=1}^{N} w_j^{W_j}\prod_{j=1}^{N} z_j^{X_i^j}(1-a_i)\frac{1-\left(a_i\frac{s_i(\underline{z})}{z_i}\right)^{W_i}}{1-a_i\frac{s_i(\underline{z})}{z_i}}\right]
$$

$$= \frac{1-a_i}{1-a_i\frac{s_i(\underline{z})}{z_i}} E\left[\prod_{j=1}^{N} z_j^{X_i^j} \prod_{\substack{j=1\\ j\neq i}}^{N} w_j^{W_j}\left(w_i^{W_i} - \left(a_i\frac{s_i(\underline{z})}{z_i}w_i\right)^{W_i}\right)\right] \tag{106}$$

$$= \frac{1-a_i}{1-a_i\frac{s_i(\underline{z})}{z_i}}\left(G_i(\underline{w};\underline{z}) - G_i\left(w_1, w_{2,} \ldots, w_{i-1}, a_i\frac{s_i(\underline{z})}{z_i}w_i, w_{i+1}, \ldots, w_N; \underline{z}\right)\right)$$

Similarly, we can simplify the second part into

$$E\left[\prod_{j=1}^{N} w_j^{W_j} \prod_{j=1}^{N} z_j^{X_i^j}(a_i\frac{s_i(\underline{z})}{z_i})^{W_i}\right]$$

$$= G_i\left(w_1, w_{2,} \ldots, w_{i-1}, a_i\frac{s_i(\underline{z})}{z_i}w_i, w_{i+1}, \ldots, w_N; \underline{z}\right) \tag{107}$$

Hence, Eq. (105) is given by

$$G_{i+1}(\underline{w};\underline{z}|up) = \frac{1-a_i}{1-a_i\frac{s_i(\underline{z})}{z_i}}G_i(\underline{w};\underline{z}) \tag{108}$$

$$+\frac{a_i\left(1-\frac{s_i(\underline{z})}{z_i}\right)}{1-a_i\frac{s_i(\underline{z})}{z_i}}G_i\left(w_1, w_{2,} \ldots, w_{i-1}, a_i\frac{s_i(\underline{z})}{z_i}w_i, w_{i+1}, \ldots, w_N; \underline{z}\right)$$

Finally,

$$\begin{aligned} G_{i+1}(\underline{w};\underline{z}) &= p_i\frac{1-a_i}{1-a_i\frac{s_i(\underline{z})}{z_i}}G_i(\underline{w};\underline{z}) \\ &+p_i\frac{a_i\left(1-\frac{s_i(\underline{z})}{z_i}\right)}{1-a_i\frac{s_i(\underline{z})}{z_i}}G_i\left(w_1, w_{2,} \ldots, w_{i-1}, a_i\frac{s_i(\underline{z})}{z_i}w_i, w_{i+1}, \ldots, w_N; \underline{z}\right) \\ &+(1-p_i)\,G_i(\underline{w};\underline{z}) \end{aligned} \tag{109}$$

7.3 First Order Moments

Since the evolution equations keep their form from the gated regime, we can use the results already obtained there:

$$f_{i+1}^j = \begin{cases} f_i^j + p_i P_{ij} E[Y_i] & j \neq i \\ f_i^i - (1-p_i)\,P_{ii} E[Y_i] & j = i \end{cases} \tag{110}$$

$$x_i \equiv p_i E\left[Y_i\right] = c\pi_i \tag{111}$$

$$f_{i+1}^j = \begin{cases} f_i^j + x_i P_{ij} & j \neq i \\ f_i^i - x_i \left(1 - P_{ii}\right) & j = i \end{cases} \tag{112}$$

$$c = \frac{M - \sum\limits_{j=1}^{N} f_j^j}{\sum\limits_{j=1}^{N} \left[\sum\limits_{k=j}^{N} P_{kj}\pi_k - \pi_j\right]} \tag{113}$$

Yet, for the Globally Gated regime we can find an approximated solution more easily.

7.4 An Approximated Solution

Using Taylor series expansion for $\hat{X}_1^i\left(a_i\right)$ we have, similarly to Eq. (36)

$$\hat{X}_1^i\left(a_i\right) \approx 1 + \left(\ln a_i\right)\, E\left[X_1^i\right] \tag{114}$$

Combining (114) with the definition of x_i and. (99),

$$x_i \approx -p_i \frac{a_i}{1-a_i}\left(\ln a_i\right)\, E\left[X_1^i\right] \tag{115}$$

or

$$x_i \approx \left(-p_i \frac{a_i}{1-a_i} \ln a_i\right)\, f_1^i \tag{116}$$

defining, as before, $e_i \equiv -p_i \frac{a_i}{1-a_i} \ln a_i$, we write (116) as

$$x_i \approx e_i f_1^i \tag{117}$$

Since $\sum\limits_{j=1}^{N} f_1^j = M$,

$$\sum_{j=1}^{N} \frac{x_j}{e_j} = c \sum_{j=1}^{N} \frac{\pi_j}{e_j} \approx M \tag{118}$$

Thus,

$$c \approx \frac{M}{\sum_{j=1}^{N} \frac{\pi_j}{e_j}} \tag{119}$$

Note the difference between this result and the approximated value in the regular Gated regime, Eq. (40).

Hence,

$$x_i = c\pi_i \approx \frac{M\pi_i}{\sum_{j=1}^{N} \frac{\pi_j}{e_j}} \tag{120}$$

When applied back to (117) this result is an approximation for the f_1^i's themselves:

$$f_1^i \approx \frac{x_i}{e_i} \approx \frac{M\pi_i}{-p_i \frac{a_i}{1-a_i} \ln a_i \sum_{j=1}^{N} \frac{\pi_j}{e_j}} \tag{121}$$

To complete the solution for all f_i^j, we use (112) which, together with (121), gives f_i^j's approximation for all j and all $i > 1$ almost instantly:

$$\begin{aligned} f_i^j &= f_{i-1}^j + P_{i-1,j}x_{i-1} - \delta_{i-1,j}x_j \\ &= \cdots = \begin{cases} f_1^j + \sum_{k=1}^{i-1} P_{kj}x_k & j \geq i \\ f_1^j + \sum_{k=1}^{i-1} P_{kj}x_k - x_j & j < i \end{cases} \\ &\approx \begin{cases} f_1^j + \sum_{k=1}^{i-1} P_{kj}e_k f_1^k & j \geq i \\ f_1^j + \sum_{k=1}^{i-1} P_{kj}e_k f_1^k - e_j f_1^j & j < i \end{cases} \end{aligned} \tag{122}$$

Finally,

$$f_i^j \approx \begin{cases} \frac{x_j}{e_j} & i = 1 \\ \frac{x_j}{e_j} + \sum_{k=1}^{i-1} P_{kj}x_k & i > 1,\, j \geq i \\ \frac{x_j}{e_j} + \sum_{k=1}^{i-1} P_{kj}x_k - x_j & i > 1,\, j < i \end{cases} \tag{123}$$

8 Globally Gated, Cyclic Bernoulli Polling

Limiting ourselves into cases where $a_i = 1\ \forall i$, equations (99) and (100) reduce to

$$E\left[Y_i\right] = E\left[X_1^i\right] = f_1^i \tag{124}$$

$$E\left[K_i^j\left(Y_i\right)\right] = P_{ij} E\left[X_1^i\right] \tag{125}$$

and equations (110) through (112) reduce to

$$f_{i+1}^j = \begin{cases} f_i^j + p_i P_{ij} f_1^i & j \neq i \\ f_i^i - p_i\left(1 - P_{ii}\right) f_1^i & j = i \end{cases} \tag{126}$$

$$x_i \equiv p_i f_1^i \tag{127}$$

$$f_{i+1}^j = \begin{cases} f_i^j + P_{ij} x_i & j \neq i \\ f_i^i - \left(1 - P_{ii}\right) x_i & j = i \end{cases} \tag{128}$$

8.1 First Order Moments

The first order moments for Cyclic Bernoulli polling are found the same way as in the general approximated solution, replacing e_i by p_i.

Using $\sum\limits_{j=1}^{N} f_1^j = M$,

$$\sum_{j=1}^{N} \frac{x_j}{p_j} = c \sum_{j=1}^{N} \frac{\pi_j}{p_j} = M \tag{129}$$

Thus

$$c = \frac{M}{\sum\limits_{j=1}^{N} \frac{\pi_j}{p_j}} \tag{130}$$

and

$$x_i = c\pi_i = \frac{M\pi_i}{\sum\limits_{j=1}^{N} \frac{\pi_j}{p_j}} \tag{131}$$

applying back to (127)

$$f_1^i = \frac{x_i}{p_i} = \frac{M\pi_i}{p_i \sum_{j=1}^{N} \frac{\pi_j}{p_j}} \tag{132}$$

To complete the solution for all f_i^j, we use (126) which, together with (132), gives f_i^j for all j and all $i > 1$ almost instantly:

$$\begin{aligned} f_i^j &= f_{i-1}^j + p_{i-1} P_{i-1,j} f_1^{i-1} - \delta_{i-1,j} p_j f_1^j \\ &= \cdots = \begin{cases} f_1^j + \sum_{k=1}^{i-1} P_{kj} p_k f_1^k & j \geq i \\ f_1^j + \sum_{k=1}^{i-1} P_{kj} p_k f_1^k - p_j f_1^j & j < i \end{cases} \end{aligned} \tag{133}$$

Finally,

$$f_i^j = \begin{cases} \frac{x_j}{p_j} & i = 1 \\ \frac{x_j}{p_j} + \sum_{k=1}^{i-1} P_{kj} x_k & i > 1,\, j \geq i \\ \frac{x_j}{p_j} + \sum_{k=1}^{i-1} P_{kj} x_k - x_j & i > 1,\, j < i \end{cases} \tag{134}$$

The expected cycle duration, the PGF of number of customers at an arbitrary moment, the sojourn times, the throughputs and system's utilization keep the exact form as in the gated and exhaustive regimes; yet, their actual values are different, depending on the varying value of the constant c in each regime.

We now compare the throughput values of the three regimes(i.e. Gated, Exhaustive and Globally Gated), and show:

Theorem 1

The throughput for the CBP Exhaustive regime, denoted by Λ_E, is larger than the throughput for the CBP Gated regime, Λ_G, which, in turn, is larger than Λ_{GG}, the throughput for the CBP Globally Gated regime.

Proof

According to Eq. (51) and (61), which hold for all regimes:

$$
\begin{aligned}
E[C] &= c\sum_{i=1}^{N}\pi_i b_i + d \\
\Lambda &= \frac{c}{E[C]}
\end{aligned}
$$

Thus,

$$\Lambda = \frac{c}{c\sum_{i=1}^{N}\pi_i b_i + d} = \frac{1}{\sum_{i=1}^{N}\pi_i b_i + \frac{d}{c}} \tag{135}$$

Therefore, in order to show $\Lambda_E \geq \Lambda_G \geq \Lambda_{GG}$ for CBP, it is enough to show that for Cyclic Bernoulli Polling,

$$c_E \geq c_G \geq c_{GG} \tag{136}$$

Indeed, comparing Eq. (94),

$$c_E = \frac{M}{\sum\limits_{j=1}^{N}\left[\sum\limits_{k=j}^{N} P_{kj}\pi_k + \frac{\pi_j}{p_j}\left(1 - p_j - P_{jj}\right)\right]}$$

to Eq. (48),

$$c_G = \frac{M}{\sum\limits_{j=1}^{N}\left[\sum\limits_{k=j}^{N} P_{kj}\pi_k + \frac{\pi_j}{p_j}\left(1 - p_j\right)\right]}$$

we immediately see that $c_E \geq c_G$ since $\sum\limits_{j=1}^{N}\frac{\pi_j}{p_j}P_{jj} \geq 0$

Comparing the latter equation to Eq. (130),

$$c_{GG} = \frac{M}{\sum\limits_{j=1}^{N}\frac{\pi_j}{p_j}}$$

we see that $c_G \geq c_{GG}$ as $\sum\limits_{j=1}^{N}\left[\sum\limits_{k=j}^{N} P_{kj}\pi_k - \pi_j\right] \leq 0$. This completes the proof.

Acknowledgement

We thank Dr. Eitan Altman for very helpful discussions.

References

1. Altman, E. and Yechiali, U. [1993]. "Cyclic Bernoulli Polling", *ZOR - Methods and Models of Operations Research,* **38**, pp 55-76.

2. Altman, E. and Yechiali, U. [1994]. "Polling in a Closed Network", *Probability in the Engineering and Informational Sciences,* **8**, pp 327-343.

3. Armony, R. and Yechiali, U. [1999]. "Polling Systems with Permanent and Transient Customers", *Stochastic Models,***15**, pp 395-427.

4. Boxma, O.J. and Cohen, J.W. [1991]. "The $M/G/1$ Queue with Permanent Customers", *IEEE Transactions on Communications*, **9**, pp. 179-184.

5. Browne, S. and Yechiali, U. [1989]. "Dynamic Priority rules for Cyclic-Type Queues", *Advances in Applied Probability*, **21**, pp. 432-450.

6. Eisenberg, M. [1972]. "Queues with Periodic Service and Changeover Time", *Operations Research*, **20**, pp 440-451.

7. Eliazar, I. and Yechiali, U. [1998]. "Polling under the Randomly Timed Gated Regime", *Stochastic Models,* **14**, pp. 79-93.

8. Fuhrmann, S.W., Levy, H. and Sidi, M. [1992]. "A Queueing Network with a Single Cyclically Roving Server", *Queuing Systems,* **11**, pp. 121-144.

9. Levy, H. and Sidi, M. [1990]. "Polling Systems: Applications, Modeling and Optimization", *IEEE Transactions on Communications*, **8**, pp. 1750-1760.

10. Kofman, D. and Yechiali, U. [1996]. "Polling Systems with Station Breakdowns". *Performance Evaluation,* **27** & **28**, pp. 647-672.

11. Kofman, D. and Yechiali, U. [1997]. "Queueing Networks with Station Breakdowns and Globally Gated Service Regime", *Telletraffic Contribution for the Information Age,* (V. Ramaswamy and P.E. Wirth, Eds.) North Holland, pp. 285-296.

12. Takagi, H. [1986]. *Analysis of Polling Systems*, MIT Press.

13. Watson, K.S. [1984]. "Performance Evaluation of Cyclic Service Strategies - A Survey", *Performance '84*, (E. Gelenbe, Ed.) North Holland, pp. 521-533.

14. Yechiali, U. [1993]. "Analysis and Control of Polling Systems". *Performance Evaluation of Computer and Communications Systems* (L. Donatiello and R. Nelson, Eds.), Springer Verlag, pp. 630-650.