ON THE RELATIONSHIP BETWEEN PREVENTIVE MAINTENANCE AND MANUFACTURING SYSTEM PERFORMANCE

By

Diwakar Gupta, Yavuz Günalay, and Mandyam M. Srinivasan

Michael G. DeGroote School of Business
McMaster University
Hamilton, Ontario

Working Paper # 437
June, 1999
ON THE RELATIONSHIP BETWEEN PREVENTIVE MAINTENANCE AND MANUFACTURING SYSTEM PERFORMANCE

By
Diwakar Gupta, Yavuz Günalay, and Mandyam M. Srinivasan

Michael G. DeGroote School of Business
McMaster University
Hamilton, Ontario

Working Paper # 437
June, 1999
On the Relationship Between Preventive Maintenance and Manufacturing System Performance

ABSTRACT

A common lament of the preventive maintenance (PM) crusaders is that production supervisors are often unwilling to lose valuable machine time when there are job waiting to be processed and do not assign high enough priority to PM. Maintenance activities that depend dynamically on system state are too complicated to implement and their overall impact on system performance, measured in terms of average tardiness or work-in-process (WIP) inventory, is difficult to predict. In this article, we present some easy to implement state-dependent PM policies that are consistent with the realities of production environment. We also develop polling models based analyses that could be used to obtain system performance metrics when such policies are implemented. We show that there are situations in which increased PM activity can lower total expected WIP (and overall tardiness) on its own, i.e., without accounting for the lower unplanned downtime. We also include examples that explain the interaction between duration of PM activity and switchover times. We identify cases in which a simple state-independent PM policy outperforms the more sophisticated state-dependent policies.

Keywords: Preventive-maintenance, polling systems, queueing models, stochastic production models.
1 Introduction

It is widely accepted that preventive maintenance (PM) reduces random breakdowns, and therefore unplanned downtime of facilities. Consequently, there is a large body of literature that deals with the scheduling and optimization of preventive maintenance operations. Comprehensive surveys of models for preventive maintenance systems may be found, for instance, in Sherif and Smith (1981), and Valdez-Flores and Feldman (1989). The models analyzed range from complex stochastic models that use system state information dynamically to schedule PM, to relatively simple deterministic models that set PM schedule using aggregate static information. An example of the latter is a PM schedule that requires maintenance activity after a fixed number of machine cycles.

Despite a relatively large body of literature on this topic, analysis of dynamic PM schedules, and their effect on the performance of the system, remains an open problem. Even when static PM schedules are used, the complex interaction between the uncertain lengths of maintenance activity, and the processing and changeover times presents a challenge to analysis. The models we analyze in this article represent a “middle ground” between the dynamic and static models, in the sense that PM activity is triggered when the system reaches a particular state. These trigger states are different in different models and are chosen for their intuitive and practical appeal. We consider a facility (which could represent a single machine, a manufacturing cell, or an integrated manufacturing plant, depending upon the situation) that processes many different types of jobs. Jobs arrive for processing at this facility according to a Poisson process, and the facility takes a random (arbitrary in distribution) amount of time to service each job, depending on its type.

The manufacturing operations are analyzed using a polling model in which the facility is represented by a single server that visits the various job types in a cyclic sequence. Jobs are modeled as customers, with each customer type held in a separate logical queue. The server administers service in an exhaustive manner. That is, it services type $i$ customers until there are no more customers of that type waiting to be processed, and then switches to service customer type $i + 1$. The facility is also subject to preventive maintenance.
PM could be triggered by one of the following events (depending upon the model): whenever the server becomes idle, before each new switchover to a different customer type, or else once in each complete cycle of service (i.e., after serving all customer types once). The PM activity is modeled by a vacation of random length. At the end of each vacation, the server either returns to a predetermined “base” queue, or returns to the same queue from where the vacation began. We assume that the vacation lengths include any tear down time and any set up time necessary to go from the queue where vacation begins to the base station. The latter is indexed as station 1, without loss of generality. If the server takes only one vacation during each uninterrupted idle period of the facility, we call it the single-vacation (SV) model. We also consider the multiple-vacation (MV) model in which the server takes one or more vacations, continuing in this mode until it finds at least one customer, somewhere in the system, at the end of a vacation.

In a more general sense, our models can be characterized as polling models with state-dependent server vacations. In addition to PM activity, vacations can represent server training, or some low contribution-margin activities that the server performs only when it is idle. Another situation in which such models arise is when the server is switched off upon becoming idle in order to conserve energy/cost, however once turned off, it is able to monitor system status only periodically. Lately, there has been considerable interest in polling models with forced server idling (see, for example, Duenyas, 1994, Duenyas and Van Oyen, 1995, and Cooper, Niu and Srinivasan, 1998) as such idling practices can sometimes improve overall system performance. Our models can also be described as polling systems with state-dependent and forced idling. Finally, the models described here are analogs of the $T$-policy for $M/G/1$ queues, albeit in the polling systems context.

Our primary intent in this paper is to determine the effect of the alternate PM strategies described above on the average waiting time experienced by jobs in the system. We do not model random breakdowns/repairs and the impact of PM activity on the likelihood of such events. In this sense, our models provide lower bounds on the change in system performance that results from any particular PM schedule. Incorporating random breakdown durations into our models is particularly easy in one specific instance, which is described next. If the failure rate is affected by PM frequency and duration, but once failure rate has been ascertained, the probability of failure is same in any two intervals of same length (i.e., time to failure is exponentially
distributed), and furthermore, service is resumed from where it was interrupted, unplanned downtimes can be modeled by appropriately inflating service time durations.

In addition to characterizing system performance for a given set of parameters, we are also interested in the system behavior as its parameters are varied. In particular, we identify situations in which the system exhibits counterintuitive behavior, wherein system performance degrades upon using dynamic state-dependent PM policies, or upon reducing the length of PM activities. Identifying conditions under which such paradoxical behavior occurs is important since it highlights the fact that efforts at performance improvement when using heuristic PM policies may actually result in the exact opposite effect. Such counterintuitive behavior has been the subject of considerable discussion recently, in the context of polling models without vacations (see, for instance, Sarkar and Zangwill (1989), Gerchak and Zhang (1994), Gupta and Srinivasan (1996), and Cooper, Niu and Srinivasan (1998), and references therein). To help explain such phenomenon better, we show an example first.

The Preventive Maintenance Conundrum

Tom Taylor is the Maintenance Superintendent of Tools-R-Us, a large press shop. He is interested in increasing the up-time of the $6 million Neptune SXM press which is a critical resource for the plant. Currently the press only processes a single type of job and it is scheduled for a PM activity every time it becomes idle (single vacation model in our terminology introduced earlier). Each job takes 3 minutes to complete (constant service time). Preventive maintenance takes the form of either a simple 5 minute lubrication or a 45 minute press overhaul. Which of the two is necessary can only be determined after shutting down the press and setting up for PM. At present, Tom thinks each one of these maintenance operations occurs with equal probability. Tom installs an automatic lubrication system which reduces the time for this operation down to zero. He also works on reducing the duration of press overhauls and succeeds in bringing it down to 44 minutes.

On the face of it, Tom has done something any maintenance manager would be proud of. He has effectively eliminated one of the preventive maintenance operations without, at the same time, affecting the quality of
the finished product. Unfortunately, he finds, to his chagrin, that the expected waiting time for a job has
gone up after the improvements are put in place.

To see how this can happen, we recognize that this system can be modeled as a single server queueing
systems with a single *vacation* at the end of each busy period. If $W$ denotes the expected time a job waits
in queue, $V$ the random vacation length, $B$ the service time, and $\lambda$ the arrival rate of jobs, then it is a
well-known fact that (see, for example, Fuhrmann and Cooper, 1985):

$$E[W] = E[W^{M/G/1}] + \frac{\lambda E[V^2]}{2(\vartheta_0 + \lambda E[V])}. \quad (1)$$

$E[W^{M/G/1}] = \frac{\lambda E[B^2]}{2(1-\lambda E[B])}$ is the expected waiting time in a standard $M/G/1$ queueing system without server
vacations and $\vartheta_0$ is the probability that the system is empty at the end of a vacation. This quantity is
easily calculated as $\vartheta_0 = E[e^{-\lambda V}]$. A seemingly paradoxical behavior will be observed if, whenever $E[V]$ is
reduced, the quantity $\lambda E[V^2]/2(\vartheta_0 + \lambda E[V])$ goes up because $E[V^2]$ is not reduced by a sufficient amount.

In the above example, let $V_1$ and $V_2$ respectively denote the preventive maintenance (vacation) duration
before and after the automatic lubrication system is installed. Thus Tom Taylor’s efforts have resulted in
the average vacation duration going down from $E[V_1] = 25$ minutes down to $E[V_2] = 22$ minutes, with the
second moment going down from $E[V_1^2] = 1025$ down to $E[V_2^2] = 968$. Let $W_1$ and $W_2$, respectively, denote
the expected waiting times before and after the automatic lubrication system is installed. From (1), we
obtain $E[W_1] = 33.70$ minutes and $E[W_2] = 33.95$ minutes. (Note that $\vartheta_0 = E[e^{-\lambda V}]$.)

Just as a reduction in preventive maintenance produces an effective increase in mean waiting times
(and therefore the average WIP), similarly the addition of preventive maintenance activity could, in some
cases, *decrease* the average work-in-process! Of course, knowing when this happens is non-trivial in more
complicated models such as the ones we propose. We also present analysis to explore when it makes sense
to do PM in a state-independent fashion.

The remainder of this paper is organized as follows. We present a formal description of our models in
section 2. The detailed analysis of MV model (which provides a skeleton for other models as well) is presented
in section 3. In the same section the results for other models are also provided. Section 4 summarizes results obtained from a number of numerical examples. Following that, in section 5, we also articulate managerial implications and develop some guidelines for manufacturing managers considering the use of dynamic PM scheduling rules.

2 The Models

We consider a polling model with $N$ stations, a strictly cyclic server routing and the exhaustive service protocol. Customers arrive at queue/station $j$ according to an independent Poisson process with rate $\lambda_j$, and we let $\Lambda = \sum_{j=1}^{N} \lambda_j$ denote the total arrival rate. Station buffers are assumed to be infinitely large. The service time for a customer at station $j$ is an independent random variable $B_j$, and the time taken by the server to switch to station $j$, from station $j-1$, is an independent random variable, $R_j$. The traffic intensity at station $j$ is denoted by $\rho_j = \lambda_j E[B_j]$, and $\rho = \sum_{j=1}^{N} \rho_j$ denotes the server utilization. For the polling system to be stable, $\rho$ must be less than 1, and this is assumed to be the case. The busy period at station $j$ is denoted by $\Theta_j$. The waiting time for a customer, which is the time it spends in the system before its service begins, is denoted by $W_j$.

We adopt the convention that an empty product equals 1, and that an empty sum equals 0. For any random variable $A$, the first two moments are denoted by $E[A]$ and $E[A^2]$. The Laplace Stieltjes Transform (LST) of a random variable $A$ is denoted by $A^*(\cdot)$. We will use $\mathbf{1}$ to denote the $1 \times N$ vector $(1, \cdots, 1)$ and $\mathbf{0}$ to denote the $1 \times N$ vector $(0, \cdots, 0)$. The description of models in the next three paragraphs applies primarily to SV and MV models. State-independent PM scheduling models are simpler and they are described in the last paragraph of this section.

The stations are visited in a strictly cyclic sequence. Without loss of generality, let the order in which the stations are visited be $1, 2, \cdots, N, 1, \cdots$. When the server arrives at a station it registers a polling epoch, and customers present at the station, if any, are served on an exhaustive basis. When the server completes service on all the customers at the station, it immediately registers a station-completion epoch. As long as the entire system is not empty, the server continues to register polling and station completion epochs, even
if no service is rendered at a particular station. Also, as long as the entire system is not empty, the server registers a *switch point* at the same time as a station-completion epoch, and begins to switch to the next station in the polling sequence. On the other hand, if the system is empty at a station-completion epoch at, say, station $j$, then the server immediately takes a vacation of length $V_j$, without registering a switch point. Parameter $V_j$ thus represents the length of PM activity when it is started from facility state $j$. Note that whereas a station completion epoch is registered at every station that is visited, a switch point is not registered if the system is empty at the service-completion instant. Hence, the number of switch points is a proper subset of the number of station completion epochs.

At the end of a vacation the server always registers a polling epoch at station 1. We assume that a vacation includes any setup that is necessary to start processing customers at station 1. We use the convention that the start of a vacation is not a switch point, since it is not the same as the server moving one station down in cyclic order. In general, the server never polls a station when the system is empty. The only exception to this rule is when the server returns (to station 1), following a vacation, and finds the system still empty, since the server will register a polling epoch at this instant.

In the multiple vacations model (MV), if the server finds the system empty when it returns from a vacation, it immediately takes another vacation, and continues in the vacation mode until it finds the system is populated by at least one customer at the end of a vacation. In this case, all vacations after the first one, start from station 1 and are, therefore, of duration $V_1$. (For this reason, we assume that $V_1$ is strictly positive.) In the single vacation (SV) model, the server idles at station 1 if it finds the system empty upon returning from a vacation; in this case the server is reactivated as soon as the first customer arrives (to an empty system). When the system reactivates, following a vacation (for both MV and SV models), the server moves to the first station at which customers are waiting, in cyclic order. Customers encountered at any intermediate stations are served en route.

The two models of state-independent PM scheduling policies are also called "continuously roving server" models, or CR and CR1 in short, on account of the fact that the server never stops switching among stations. In the CR model, the server begins PM at each station completion epoch, i.e., before switching over to a
new job type, whereas in the CR1 model, PM activity takes place only at the end of a station completion
epoch at station 1. In short, both server movement and scheduling of PM activity is determined apriori and
does not depend on the number of waiting jobs in the system.

3 The Analysis

Let \( X_i \) denote the number of customers at station \( i \), at the instant that station is polled. Let \( F_i(z_1, \ldots, z_N) \)
denote the PGF of the number of customers found at the various stations when (i.e., \textit{given that}) the server
has polled station \( i \). Note that the PGF of \( X_i \) is \( F_i(1, \ldots, 1, z_i, 1, \ldots, 1) \). We use the descendant sets (DS)
approach to find \( F_i(1, \ldots, 1, z_i, 1, \ldots, 1) \). Consider an arbitrary polling instance at station \( i \), and call it the
type-\( i \) reference point. Clearly, \( X_i \) represents the queue length at station \( i \) at the type-\( i \) reference point. The
DS approach determines \( F_i(1, \ldots, 1, z_i, 1, \ldots, 1) \) using a recursive method.

Define \( L_{j,c}(z_i) \), for \( j = 1, \ldots, N \), and \( c \geq -1 \), as follows.

\[
L_{j,c}(z_i) = \Theta_j^* \left( \sum_{k=j+1}^{N} [\lambda_k - \lambda_k L_{k,c}(z_i)] + \sum_{k=1}^{j-1} [\lambda_k - \lambda_k L_{k,c-1}(z_i)] \right).
\] (2)

The subscript \( i \) in \( z_i \) is used to signify the fact that we are counting contributions to \( X_i \) and \( c \) denotes the
number of cycles prior to reference point, with the latter indexed as the start of \( c = -1 \). Note that customers
that are served after the reference point do not contribute to \( X_i \). Hence, the above recursive expression is
initialized as follows:

\[
L_{j,-1}(z_i) = \begin{cases} 
  z_i & \text{if } j = i, \\
  1 & \text{if } j > i. 
\end{cases}
\] (3)

We also have \( L_{j,c}(z_i) = 1 \) for all \( c \leq -2 \) and all \( j \), since customers served in cycle \(-2\), or any other cycle
thereafter, do not contribute to \( X_i \). Similarly, for \( j = 1, \ldots, N \), and \( c \geq -1 \), let

\[
R_{j,c}(z_i) = R_j^* \left( \sum_{k=j+1}^{N} [\lambda_k - \lambda_k L_{k,c}(z_i)] + \sum_{k=1}^{j-1} [\lambda_k - \lambda_k L_{k,c-1}(z_i)] \right),
\] (4)

and note that \( R_{j,-1}(z_i) = 1 \) for all \( j > i \). Finally, define \( V_{j,c}(z_i) \) in an analogous manner, as follows:

\[
V_{j,c}(z_i) = V_j^* \left( \sum_{k=1}^{N} [\lambda_k - \lambda_k L_{k,c}(z_i)] \right).
\] (5)
3.1 The Multiple Vacations (MV) Model

The MV model assumes that the server goes through a PM process, every time the system becomes empty, and if the system is still empty at the end of a vacation (i.e., PM period), the process repeats itself. MV model is analyzed in detail and all the derivation steps will be reported. First we carefully analyze station 1, and then extend our analysis to any station i.

3.1.1 Station 1

When dealing with station 1, we simplify notation by using $z$ to denote $z_1$. The boundary conditions for the $L_{j,c}(z)$ recursion now become as follows: $L_{j,-1}(z) = z$ if $j = 1$, and $L_{j,c} = 1$ for all $c \leq -1$ and $j > 1$.

1. There is no net change in contribution to $X_1$ between a station polling and station-completion instant from the same station that occur in the same cycle. This makes sense since all new arrivals, if any, are offsprings of customers already present in the system and their contributions are already accounted for. Thus,

$$g_{j,c}(z) = f_{j,c}(z), \quad \forall \ j \text{ and } c.$$  \hfill (6)

Also, in the long run, we have an equal number of polling and station-completion instants, i.e., $f_j(I) = g_j(I)$, $j = 1, \ldots, N$.

2. Switch point from station $j$ coincides with a station-completion epoch from the same station, except when the entire system is empty at the station-completion instant. In that case, the server takes a vacation without registering a switch point.

$$h_{j,c}(z) = g_{j,c}(z) - g_j(\bar{0}), \quad \forall \ j \text{ and } c.$$  \hfill (7)

3. At a station $j$, $j \neq 1$, a polling instant always follows the switch point from the previous station. Since a type $j$ switchover (setup) time separates the two, we have

$$f_{j,c}(z) = h_{j-1,c}(z)R_{j,c}(z) \quad \forall \ j \neq 1.$$  \hfill (8)
4. A polling instant at station 1 follows either a switch point from station \( N \), or it coincides with the end of a vacation. We can write this relationship as follows:

\[
f_{1,c}(z) = \sum_{j=1}^{N} g_j(\bar{0}) V_{j,c}(z) + h_{N,c+1}(z) R_{1,c}(z). \tag{9}
\]

Notice that \( f_1(\bar{0}) \neq 0 \), but \( f_j(\bar{0}) = 0, \forall \, j \neq 1 \). In fact, \( f_1(\bar{0}) = \sum_{j=1}^{N} g_j(\bar{0}) V_j^*(\Lambda) \). However, \( h_j(\bar{0}) = 0, \forall \, j \).

From above relationships, we can write

\[
h_{j,c}(z) = h_{j-1,c}(z) R_{j,c}(z) - g_j(\bar{0}), \quad \forall \, j \neq 1, \quad \text{and} \tag{10}
\]

\[
h_{1,c}(z) = h_{N,c+1}(z) R_{1,c}(z) + \sum_{j=1}^{N} g_j(\bar{0}) V_{j,c}(z) - g_1(\bar{0}). \tag{11}
\]

We start with \( j = N \) and \( c = 0 \), express \( h_{N,0} \) in terms of \( h_{N-1,0} \), which in turn is expressed in terms of \( h_{N-2,0} \) and continue in this fashion. After \( N \) such iterations we obtain,

\[
h_{N,0}(z) = h_{N,1}(z) \prod_{j=1}^{N} R_{j,0}(z) - g_1(\bar{0}) \prod_{j=2}^{N} R_{j,0}(z)[1 - V_{1,0}(z)] - \sum_{j=2}^{N} g_j(\bar{0}) \{ \prod_{k=j+1}^{N} R_{k,0}(z) - V_{j,0}(z) \prod_{\ell=2}^{N} R_{\ell,0}(z) \}. \tag{12}
\]

Next, we express \( h_{N,1} \) in terms of \( h_{N,2} \), and so on infinitely many times. Using the fact that \( h_{N,\infty}(z) \to h_N(1) \), we get

\[
h_{N,0}(z) = h_N(1) \prod_{c=0}^{\infty} \prod_{j=1}^{N} R_{j,c}(z) - g_1(\bar{0}) \prod_{c=0}^{\infty} \prod_{j=2}^{N} R_{j,c}(z)[1 - V_{1,c}(z)] \prod_{p=0}^{c-1} \prod_{k=1}^{N} R_{k,p}(z)
\]

\[
- \sum_{j=2}^{N} g_j(\bar{0}) \{ \prod_{c=0}^{\infty} \prod_{k=j+1}^{N} R_{k,c}(z) - V_{j,c}(z) \prod_{\ell=2}^{N} R_{\ell,c}(z) \} \prod_{p=0}^{c-1} \prod_{k=1}^{N} R_{k,p}(z). \tag{13}
\]

Since

\[
f_1(z, 1, \cdots, 1) = f_{1,-1}(z) = h_{N,0}(z) R_{1,-1}(z) + \sum_{j=1}^{N} g_j(\bar{0}) V_{j,-1}(z), \tag{14}
\]

the above can be simplified to yield

\[
f_1(z, 1, \cdots, 1) = h_N(1) \prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(z) - g_1(\bar{0}) \sum_{c=-1}^{\infty} [1 - V_{1,c}(z)] K_{1,c}(z)
\]

\[
- \sum_{j=2}^{N} g_j(\bar{0}) \{ K_{j,c}(z) - V_{j,c}(z) K_{1,c}(z) \} + \sum_{j=1}^{N} g_j(\bar{0}), \tag{15}
\]

10
where the term
\[ K_{j,c}(z) = \prod_{\ell=j+1}^{N} R_{\ell,c}(z) \prod_{p=1}^{c-1} \prod_{k=1}^{N} R_{k,p}(z) \] (16)
has been used for expository clarity. In fact, the second and the third term on the right hand side can be
collapsed into one, yielding the following compact form:
\[ f_1(z,1,\cdots,1) = h_N(\bar{1}) \prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(z) - \sum_{j=1}^{N} g_j(\bar{0}) \sum_{c=-1}^{\infty} \left\{ K_{j,c}(z) - V_{j,c}(z)K_{1,c}(z) \right\} + \sum_{j=1}^{N} g_j(\bar{0}). \] (17)

Setting \( z = 1 \) in (17) we see that
\[ f_1(\bar{1}) = h_N(\bar{1}) + \sum_{j=1}^{N} g_j(\bar{0}) \] (18)
Define
\[ \vartheta_j = g_j(\bar{0})/h_N(\bar{1}), \] (19)
for each \( j = 1, 2, \cdots, N \). Next, divide both sides of (17) by \( f_1(\bar{1}) \) and simplify the right hand side with the
help of (18) and (19) to obtain:
\[ F_1(z,1,\cdots,1) = \left( \frac{1}{1 + \sum_{j=1}^{N} \vartheta_j} \right) \left( \prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(z) - \sum_{j=1}^{N} \vartheta_j \sum_{c=-1}^{\infty} \left\{ K_{j,c}(z) - V_{j,c}(z)K_{1,c}(z) \right\} + \sum_{j=1}^{N} \vartheta_j \right). \] (20)

### 3.1.2 Station \( i \)

Note that the basic relationships shown in equations 6 through 11 do not change*. Also note that the
relationship between \( h_{N,0}(z_i) \) and \( h_N(\bar{1}) \) in equation 13 does not change. From equation 8, 7, and then 6,
we can write
\[ f_{i,-1}(z_i) = h_{i,-1,-1}(z_i)R_{i,-1}(z_i) \]
\[ = [g_{i,-1,-1}(z_i) - g_{i,-1}(\bar{0})]R_{i,-1}(z_i) \]
\[ = f_{i,-1,-1}(z_i)R_{i,-1}(z_i) - g_{i,-1}(\bar{0})R_{i,-1}(z_i). \] (21)
Continuing in this fashion, we can obtain,
\[ f_{i,-1}(z_i) = f_{i,-1}(z_i) \prod_{j=2}^{i} R_{j,-1}(z_i) - \sum_{j=1}^{i-1} g_j(\bar{0}) \prod_{k=j+1}^{i} R_{k,-1}(z_i). \] (22)

*For every polling instant at station 1 when the system is empty, we do register a station completion instant and then start
either another vacation (MV) or server idle period (SV) until a new arrival occurs.
Next, \( f_{i-1}(z_i) \) is written in terms of \( h_{N,0}(z_i) \) (from equation 9), and we obtain

\[
f_{i-1}(z_i) = h_{N,0}(z_i) \prod_{j=1}^{i} R_{j-1}(z_i) + \sum_{j=1}^{N} g_j(\bar{0}) V_{j-1}(z_i) \prod_{k=2}^{i} R_{k-1}(z_i) - \sum_{j=1}^{i-1} g_j(\bar{0}) \prod_{k=j+1}^{i} R_{k-1}(z_i). \tag{23}
\]

Notice that \( R_{j-1}(z_i) = 1 \) for all \( j > i \), and therefore the upper limit of the products in the first, second and third terms on the RHS of the above expression can be changed from \( i \) to \( N \). This gives,

\[
f_{i-1}(z_i) = h_{N,0}(z_i) \prod_{j=1}^{N} R_{j-1}(z_i) + \sum_{j=1}^{N} g_j(\bar{0}) V_{j-1}(z_i) \prod_{k=2}^{N} R_{k-1}(z_i) - \sum_{j=1}^{i-1} g_j(\bar{0}) \prod_{k=j+1}^{N} R_{k-1}(z_i). \tag{24}
\]

Substituting from equation 13, we obtain the following expression after simplification

\[
f_{i-1}(z_i) = h_N(\bar{1}) \prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(z_i) - \sum_{j=1}^{N} g_j(\bar{0}) \sum_{c=0}^{\infty} \left\{ \prod_{k=j+1}^{N} R_{k,c}(z_i) - V_{j,c}(z_i) \prod_{\ell=2}^{N} R_{\ell,c}(z_i) \right\} \prod_{p=-1}^{\infty} \prod_{m=1}^{N} R_{m,p}(z_i)
  + \sum_{j=1}^{N} g_j(\bar{0}) V_{j-1}(z_i) \prod_{k=2}^{N} R_{k-1}(z_i) - \sum_{j=1}^{i-1} g_j(\bar{0}) \prod_{k=j+1}^{N} R_{k-1}(z_i). \tag{25}
\]

Substituting \( K_{j,c}(z_i) \) from its definition in 16, and noting that \( K_{j,-1}(z_i) = \prod_{\ell=j+1}^{N} R_{\ell,-1}(z_i) \), we can rewrite the above equation as follows:

\[
f_{i-1}(z_i) = h_N(\bar{1}) \prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(z_i) - \sum_{j=1}^{N} g_j(\bar{0}) \sum_{c=0}^{\infty} \left\{ K_{j,c}(z_i) - V_{j,c}(z_i)K_{1,c}(z_i) \right\} + \sum_{j=1}^{N} g_j(\bar{0}) K_{j,-1}(z_i). \tag{26}
\]

But, \( K_{j,-1}(z_i) = 1 \) for all \( j \geq i \). Thus, the following simplified form of the above equation can now be obtained.

\[
f_i(1, \cdots, z_i, \cdots, 1) = f_{i-1}(z_i) = h_N(\bar{1}) \prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(z_i) - \sum_{j=1}^{N} g_j(\bar{0}) \sum_{c=0}^{\infty} \left\{ K_{j,c}(z_i) - V_{j,c}(z_i)K_{1,c}(z_i) \right\} + \sum_{j=1}^{N} g_j(\bar{0}). \tag{27}
\]

If we set \( z_i = 1 \) in 27, we will obtain

\[
f_i(1) = h_N(\bar{1}) + \sum_{j=1}^{N} g_j(\bar{0}). \tag{28}
\]

Like before we divide both sides of 27 by \( f_i(1) \) and simplify using relationships 19 and 28. This yields the following relationship:

\[
F_i(1, \cdots, z_i, 1, \cdots, 1) = \left( \frac{1}{1 + \sum_{j=1}^{N} \vartheta_j} \right) \left( \prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(z_i) - \sum_{j=1}^{N} \vartheta_j \sum_{c=0}^{\infty} \left\{ K_{j,c}(z_i) - V_{j,c}(z_i)K_{1,c}(z_i) \right\} + \sum_{j=1}^{N} \vartheta_j \right). \tag{29}
\]
The only unknowns in expressions 20 and 29 are the $\vartheta_j$ terms. In the next section we concentrate on how to find $\vartheta_j$'s.

### 3.1.3 Computing $\vartheta_j$

Like in S&G (1996), we first express equations 29 in terms of vector arguments. Next, we notice that $F_i(\bar{0}) = 0$ for all $i > 1$. This comes from the fact that the server does not poll a station other than station 1, if the entire polling system is empty. On the other hand, the server may poll station 1 at the end of a vacation even when the polling system is empty. In fact, we know from 8 that $F_1(\bar{0}) = \sum_{j=1}^{N} \vartheta_j V_j^*(\lambda)/(1 + \sum_{j=1}^{N} \vartheta_j)$. These relationships give us $N$ equations in $N$ unknowns, as described below.

For station 1:

$$\prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(\bar{0}_1) - \sum_{j=1}^{\infty} \vartheta_j \sum_{c=-1}^{\infty} \{ K_{j,c}(\bar{0}_1) - V_{j,c}(\bar{0}_1)K_{1,c}(\bar{0}_1) \} + \sum_{j=1}^{N} \vartheta_j [1 - V_j^*(\lambda)] = 0 \tag{30}$$

For station $i$, $i \neq 1$:

$$\prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(\bar{0}_i) - \sum_{j=1}^{\infty} \vartheta_j \sum_{c=-1}^{\infty} \{ K_{j,c}(\bar{0}_i) - V_{j,c}(\bar{0}_i)K_{1,c}(\bar{0}_i) \} + \sum_{j=1}^{N} \vartheta_j = 0 \tag{31}$$

A subscript attached to vector $\bar{0}$ in equations 30 and 31 to emphasize the starting point for the $L_{j,c}(\cdot)$ recursion (see equations 2 and 3 for details).

### 3.1.4 Waiting Times: Distribution

WLOG we show first how waiting times are derived for station 1. Later these expressions will be generalized for any station $i$. Notice that the Fuhrmann-Cooper decomposition applies and that the following equation tells us how we would find the waiting time distribution in terms of the PGF of queue lengths at station polling instants.

$$W^*_i(\lambda_i - \lambda_1) = \left[ \frac{f_1(\bar{1})}{f_i^*(\bar{1})} \right] \frac{1 - \rho_1}{B^*_i(\lambda_i - \lambda_1z) - z} \tag{32}$$

where the second term on the right hand side is the well-known Pollaczek-Khintchine transform for the waiting-time distribution in the standard $M/G/1$ queue. Dividing both the numerator and denominator of the first term on the right hand side by $f_i(\bar{1})$ and upon simplification we obtain the following form of the
LST of waiting time distribution.

\[
W_l^*(\lambda_1 - \lambda_1 z) = \left[1 - \frac{F_l(z, 1, \cdots, 1)}{F_l^*(1)}\right] \frac{1 - \rho_l}{B_l^*(\lambda_1 - \lambda_1 z) - z},
\]  

(33)

The above expression remains valid for stations other than 1 upon changing all station indices appropriately.

3.1.5 Mean Waiting Time: Station 1

Substituting from 20 into equation 33, we get the exploded form of the LST of the waiting time distribution. This can now be differentiated with respect to \(z\) to obtain the mean waiting time.

Like before, we will work out the analysis for station 1 first. In this model, mean waiting times for stations other than 1 cannot simply be obtained by a renumbering of station indices. However, the method described below remains intact.

\[
E[W_1] = \frac{F_l''(1)}{2\lambda_1 F_l^*(1)} + \frac{\lambda_1 E[B_l^2]}{2(1 - \rho_l)}
\]

(34)

Our main task is to determine \(F_l''(1)\) and \(F_l'(1)\). In order to do that, we first define some additional notation and obtain results similar to S&G (1996). The results are presented without proof.

We define \(\gamma_{j,c} = (\lambda_j/\lambda_1)E[L_{j,c}]\) for all \(j\), when \(c \geq 0\), and \(\gamma_{1,-1} = 1\) and \(\gamma_{j,-1} = 0\) for all \(j > 1\). Then, the following results can be obtained:

\[
\gamma_{k,c} = \rho_k \left[\sum_{j \geq k} \gamma_{j,c} + \sum_{j < k} \gamma_{j,c-1}\right].
\]

(35)

\[
\sum_{c=0}^{\infty} \sum_{j=1}^{N} \gamma_{j,c} = (\rho - \rho_1)/(1 - \rho).
\]

(36)

\[
\sum_{c=-1}^{\infty} \sum_{j=1}^{N} \gamma_{j,c} = (1 - \rho_1)/(1 - \rho).
\]

(37)

\[
\sum_{c=0}^{\infty} \gamma_{j,c} = \begin{cases} \frac{\rho_l(\rho - \rho_1)}{(1 - \rho)} & \text{when } j = 1, \\ \frac{\rho_l(1 - \rho_1)}{(1 - \rho)} & \text{otherwise.} \end{cases}
\]

(38)

\[
\sum_{c=-1}^{\infty} \sum_{j=1}^{N} R_{j,c}(1) = \frac{\lambda_1(1 - \rho_1)E[R_T]}{(1 - \rho)},
\]

(39)

where \(E[R_T] = \sum_{j=1}^{N} E[R_j]\) is the sum of expected switchover times.
Next, if we define \( t_{j,c} = \frac{1}{\lambda_1} K_{j,c}(1) \), for all \( j \) and \( c \geq 0 \), then differentiating equation 16, we obtain:

\[
t_{j,c} = E[R_1] + \sum_{m=j+1}^{N} \left( \gamma_{m,c} \rho_m \right) E[R_m] + \sum_{p=0}^{c-1} \sum_{k=1}^{N} \left( \gamma_{k,p} \rho_k \right) E[R_k].
\]

(40)

Notice that \( K_{j,c}(z) = 1 \), whenever \( c \leq -1 \), and therefore we define \( t_{j,c} = 0 \) for all \( j \) and \( c \leq -1 \).

Differentiating equation 20, and subsequently setting \( z = 1 \), we obtain:

\[
F'(1) = \left( \frac{1}{1 + \sum_{j=1}^{N} \theta_j} \right) \left( \sum_{c=-1}^{N} \sum_{j=1}^{N} R'_{j,c}(1) - \sum_{j=1}^{N} \sum_{c=-1}^{N} \theta_j \{ \lambda_1(t_{j,c} - t_{1,c}) - V'_{j,c}(1) \} \right).
\]

(41)

Finally, substituting from 35 - 40, we obtain:

\[
F'(1) = \left( \frac{\lambda_1(1 - \rho_1)}{(1 - \rho)(1 + \sum_{j=1}^{N} \theta_j)} \right) \left( E[R_T] + \sum_{j=1}^{N} \theta_j E[V_j] + \sum_{j=2}^{N} \sum_{k=2}^{j} E[R_k] \right).
\]

(42)

Let \( C_1 \) denote the length of a server cycle at station 1. It is defined as the time elapsed between two consecutive polling instants at station 1. Then, \( F'(1) \) is the average number of type 1 customers found waiting at the start of a type-1 cycle. The total number of type 1 customers served in this cycle is therefore \( F'(1)/(1 - \rho_1) \). We also know that the average number of type 1 customers served per cycle equals \( \lambda_1 E[C_1] \).

Upon equating these two quantities we obtain:

\[
E[C_1] = \frac{E[R_T] + \sum_{j=1}^{N} \theta_j E[V_j] + \sum_{j=2}^{N} \sum_{k=2}^{j} E[R_k]}{(1 - \rho)(1 + \sum_{j=1}^{N} \theta_j)}.
\]

(43)

Twice differentiating 20 and setting \( z = 1 \) we obtain

\[
(1 + \sum_{j=1}^{N} \theta_j) F''(1) = \left( \sum_{c=-1}^{N} \sum_{j=1}^{N} (R''_{j,c}(1) - (R''_{j,c}(1))^2) + \left( \sum_{c=-1}^{N} \sum_{j=1}^{N} R''_{j,c}(1) \right)^2 \right)
\]

\[
- \sum_{j=1}^{N} \theta_j \sum_{c=-1}^{N} \{ K''_{j,c}(1) - K''_{1,c}(1) - V''_{j,c}(1) - 2V'_{j,c}(1)K'_{1,c}(1) \}.
\]

(44)

We define \( \gamma^{(2)}_{j,c} = (\lambda_j/\lambda_1)E[L^2_{j,c}] \) for all \( j \) and \( c \geq 0 \). When \( c \leq -1 \), we define \( \gamma^{(2)}_{j,c} = 0 \) for all \( j \). Like S&G (1996), we also define \( \Gamma_j = \sum_{c=0}^{\infty} \gamma^{2}_{j,c} \) for all \( j \). Then, it can be confirmed that

\[
\sum_{j=1}^{\infty} \sum_{c=0}^{N} \gamma^{(2)}_{j,c} = \frac{\lambda_1}{(1 - \rho)} \sum_{j=1}^{N} \frac{\Gamma_j}{\rho_j^2} \lambda_j E[B^2_j]
\]

(45)

\[
\sum_{c=-1}^{N} \sum_{j=1}^{N} (R''_{j,c}(1) - (R''_{j,c}(1))^2) = \lambda_1 \sum_{j=1}^{N} \left\{ \frac{\Gamma_j}{\rho_j^2} \left( Var(R_j) + \frac{\lambda_j E[B^2_j] E[R_T]}{(1 - \rho)} \right) + Var(R_1) \right\}.
\]

(46)

\[
\sum_{c=-1}^{N} (K''_{j,c}(1) - K''_{1,c}(1)) = -\lambda_1 \sum_{m=2}^{N} \frac{\Gamma_m}{\rho_m^2} Var(R_m) + \frac{E[R_m]}{(1 - \rho)} \sum_{k=1}^{N} \frac{\Gamma_k}{\rho_k} \lambda_k E[B^2_k] + \lambda_1 \sum_{c=0}^{\infty} (t^2_{j,c} - t^2_{1,c}).
\]

(47)
In simplifying the above expression, we have used the fact that \( t_{1,c} - t_{j,c} = \sum_{m=2}^{\infty} \gamma_{m,c} E[R_m]/\rho_m \), for \( c \geq 0 \) which follows from equation 40.

\[
\sum_{c=-1}^{\infty} V''_{j,c}(1) = \frac{\lambda^2 E[V_j]}{(1-\rho)} \sum_{k=1}^{N} \left( \frac{\Gamma_k}{\rho_k^2} + \lambda_k E[B_k^2] \right) \left[ 1 + \sum_{c=0}^{\infty} \left( \frac{\gamma_{1,c}}{\rho_1} \right)^2 \right],
\]

and

\[
\sum_{c=-1}^{\infty} V''_{j,c}(1) K'_{1,c}(1) = \frac{\lambda^2 E[V_j]}{\rho_1} \sum_{c=0}^{\infty} t_{1,c} \gamma_{1,c}.
\]

Substituting from above into equation 44, simplifying and collecting terms, we obtain the following expression after some effort:

\[
F{(1)} \quad = \left( \frac{\lambda^2}{1 + \sum_{j=1}^{N} \theta_j} \right) \left( \frac{(1-\rho_1)^2 E[R_T]^2}{(1-\rho)^2} + Var(R_1) + \sum_{j=1}^{N} \left( \frac{\Gamma_j}{\rho_j^2} \right) Var(R_j) \right)
+ \sum_{j=1}^{N} \theta_j \left( \sum_{k=2}^{j} Var(R_k) + \sum_{c=0}^{\infty} \left( t_{1,c}^2 - t_{j,c}^2 \right) + 2E[V_j] \sum_{c=0}^{\infty} \left( \frac{t_{1,c} \gamma_{1,c}}{\rho_1} \right) + E[V_j^2] (1 + \Gamma_1/\rho_1^2) \right)
+ \left( 1 + \sum_{j=1}^{N} \theta_j \right) E[C_1] \sum_{j=1}^{N} \left( \frac{\Gamma_j}{\rho_j^2} \right) \lambda_j E[B_j^2].
\]

After substituting equations 42, 43 and 50 into the equation 34 we get

\[
E[W_1] = \frac{(1-\rho_1)E[R_T]^2}{2(1-\rho)^2 E[C_1]} + \frac{\lambda_1 E[B_1^2]}{2(1-\rho_1)} + \frac{\sum_{j=1}^{N} (\Gamma_j/\rho_j)^2 \lambda_j E[B_j^2]}{2(1-\rho_1)} + \frac{Var(R_1) + \sum_{j=1}^{N} (\Gamma_j/\rho_j^2) Var(R_j)}{2(1-\rho_1)(1 + \sum_{j=1}^{N} \theta_j) E[C_1]}
+ \frac{\sum_{j=1}^{N} \theta_j \left( \sum_{k=2}^{j} (\Gamma_k/\rho_k^2) Var(R_k) + E[V_j^2] (1 + \Gamma_1/\rho_1) + \sum_{c=0}^{\infty} \left( 2E[V_j] \left( \frac{t_{1,c} \gamma_{1,c}}{\rho_1} \right) + (t_{j,c}^2 - t_{j,c}^2) \right) \right)}{2(1-\rho_1)(1 + \sum_{j=1}^{N} \theta_j) E[C_1]}.
\]

**Observation:** Setting \( N = 1 \) leads to the mean waiting time formula for an M/G/1 queue with T-policy.

Note that \( N = 1 \) in our model sets all switchover times to zero, i.e., \( R_j = 0, j = 1, \ldots, N \), and only \( V_1 > 0 \).

Also, \( N = 1 \) leads \( \theta_1 = 1 \), since there exists only one station. Then, equation 51 becomes

\[
E[V_1^2] \quad = \quad \frac{\lambda_1 E[B_1^2]}{2E[V_1]} + \frac{\lambda_1 E[B_1^2]}{2(1-\rho_1)},
\]

which identical to the mean waiting time formula of the T-policy M/G/1 queueing models.
3.1.6 Mean Waiting Time: Station $i$, $i \neq 1$

Using similar arguments to those in section 3.1.4 the waiting time distribution of type-$i$ customers can be calculated from the following relationship

$$W_i^*(\lambda_i - \lambda_i z_i) = \left[ 1 - F_i(1, \ldots, z_i, \ldots, 1) \right] \frac{1 - \rho_i}{B_i^*(\lambda_i - \lambda_i z_i) - z_i}.$$  \hfill (53)

Then by differentiating relation 53 with respect to $z_i$ and setting $z_i = 1$ we get the mean waiting time for type-$i$ customers,

$$E[W_i] = \frac{F''_i(1)}{2\lambda_i F'_i(1)} + \frac{\lambda_i E[B_i^2]}{2(1 - \rho_i)}.$$  \hfill (54)

We need to calculate $F''_i(1)$ and $F'_i(1)$. First, we generalize some of the definitions given for station 1, by introducing the parameter $i$ to indicate that the variable will be used to calculate the factorial moments of the queue length at station $i$ polling instants. Next, we obtain similar results to relations 35 – 40.

Then for any $i = 1, \ldots, N$ we define

$$\gamma_{j,c}(i) = \left( \frac{\lambda_j}{\lambda_i} \right) \frac{\partial}{\partial z_i} L_{j,c}(z_i)$$

\hfill (55)

$$\gamma^{(2)}_{j,c}(i) = \left( \frac{\lambda_j}{\lambda_i} \right) \frac{\partial^2}{\partial z_i^2} L_{j,c}(z_i) \quad j = 1, \ldots, N, \ c \geq -1.$$  \hfill (56)

Note that, due to the boundary conditions we have $\gamma_{i,-1}(i) = 1$, and $\gamma_{j,-1}(i) = \gamma^{(2)}_{k,-1}(i) = 0$, for $j > i$ and $k \geq i$. Other $\gamma_{j,c}(i)$ and $\gamma^{(2)}_{j,c}(i)$ variables with indices $j = 1, \ldots, N$, $c \geq 0$ and $j = 1, \ldots, i - 1$, $c = -1$, can be calculated using the definition of $L_{j,c}(z_i)$,

$$\gamma_{j,c}(i) = \rho_j \left[ \sum_{k \geq j} \gamma_{k,c}(i) + \sum_{k < j} \gamma_{k,c-1}(i) \right],$$

\hfill (57)

$$\gamma^{(2)}_{j,c}(i) = \rho_j \left[ \sum_{k \geq j} \gamma^{(2)}_{k,c}(i) + \sum_{k < j} \gamma^{(2)}_{k,c-1}(i) \right] + \lambda_i \lambda_j E[B_j^2] \left[ \sum_{k \geq j} \gamma_{k,c}(i) + \sum_{k < j} \gamma_{k,c-1}(i) \right]^2,$$

\hfill (58)

Then the sum of all calculated first factorial moments of the contributions to queue $i$ is,

$$\sum_{c=0}^{\infty} \sum_{j=1}^{i-1} \gamma_{j,c}(i) + \gamma_{j,-1}(i) = \sum_{c=0}^{\infty} \rho_j \left[ \sum_{k \geq j} \gamma_{k,c}(i) + \sum_{k < j} \gamma_{k,c-1}(i) \right].$$
\[ + \sum_{j=1}^{i-1} \rho_j \left[ \sum_{k \geq j} \gamma_{k-1}(i) + \sum_{k < j} \gamma_{k-2}(i) \right], \quad (60) \]
\[ = \sum_{j=1}^{N} \rho_j \left( \sum_{c=0}^{\infty} \sum_{k=1}^{N} \gamma_{k,c}(i) + \sum_{k=1}^{j-1} \gamma_{k,c}(i) \right) \]
\[ + \sum_{j=1}^{i-1} \rho_j \left( \sum_{k=1}^{N} \gamma_{k,c}(i) \right), \quad (61) \]
\[ = \sum_{j=1}^{N} \rho_j \left( \sum_{c=0}^{\infty} \sum_{k=1}^{N} \gamma_{k,c}(i) \right) + \sum_{j=1}^{i-1} \rho_j \sum_{k=1}^{N} \gamma_{k-1}(i) \]
\[ + \sum_{j=1}^{i-1} \rho_j \left( \sum_{k=1}^{N} \gamma_{k,c}(i) \right). \quad (62) \]

In relationship 62 the last term is zero. Also, using the fact that \( \gamma_{j-1}(i) = 0 \) for all \( j > i \), we get
\[
\sum_{c=0}^{\infty} \sum_{j=1}^{N} \gamma_{j,c}(i) + \sum_{j=1}^{i-1} \gamma_{j-1}(i) = \sum_{j=1}^{N} \rho_j \left( \sum_{c=0}^{\infty} \sum_{k=1}^{N} \gamma_{k,c}(i) \right) + \sum_{j=1}^{i-1} \rho_j \gamma_{k-1}(i) + \rho_i \sum_{k=1}^{N} \gamma_{k-1}(i) \]
\[ + \sum_{j=1}^{i-1} \rho_j \left( \sum_{k=1}^{N} \gamma_{k,c}(i) \right), \quad (63) \]
\[ = \sum_{j=1}^{N} \rho_j \left( \sum_{c=0}^{\infty} \sum_{k=1}^{N} \gamma_{k,c}(i) \right) + \sum_{j=1}^{i-1} \rho_j \gamma_{k-1}(i) \]
\[ + \rho_i \gamma_{k-1}(i). \quad (64) \]

Therefore,
\[
\sum_{c=0}^{\infty} \sum_{j=1}^{N} \gamma_{j,c}(i) + \sum_{j=1}^{i-1} \gamma_{j-1}(i) = \left( \frac{\rho - \rho_i}{1 - \rho} \right), \quad (65) \]
and
\[ \sum_{c=-1}^{\infty} \sum_{j=1}^{N} \gamma_{j,c}(i) = \left( \frac{1 - \rho_i}{1 - \rho} \right), \quad (66) \]

Also by using relation 57 in above equations we obtain
\[
\sum_{c=-1}^{\infty} \gamma_{j,c}(i) = \begin{cases} \frac{\rho_i(1 - \rho)}{(1 - \rho)} & \text{when } j = i, \\ \frac{\rho_i}{(1 - \rho)} & \text{otherwise.} \end{cases} \quad (67) \]

By differentiating equation 29 once and twice, and setting \( z_i = 1 \), we obtain
\[
F'_i(\tilde{t}_i) = \left( \frac{1}{1 + \sum_{j=i}^{N} \theta_j} \right) \left( \sum_{c=-1}^{\infty} \sum_{j=1}^{N} R_{j,c}^i(1_i) - \sum_{j=1}^{N} \sum_{c=-1}^{\infty} \theta_j \{ \lambda_i(t_{j,c}(i)) - t_{1,c}(i) \} - V'_{j,c}(1_i) \right), \quad (68) \]
\[
F''_i(\tilde{t}_i) = \left( \frac{1}{1 + \sum_{j=i}^{N} \theta_j} \right) \left( \sum_{c=-1}^{\infty} \sum_{j=1}^{N} \{ R_{j,c}^i(1_i) - (R_{j,c}^i(1_i))^2 \} + \left( \sum_{c=-1}^{\infty} \sum_{j=1}^{N} R_{j,c}^i(1_i) \right)^2 \right)^2 \]

18
\[- \sum_{j=1}^{N} \varphi_j \sum_{c=-1}^{\infty} \{ K_{j,c}''(1_t) - K_{1,c}''(1_t) - V_{j,c}''(1_t) - 2V_{j,c}'(1_t)K_{1,c}'(1_t) \} \] (69)

Note that the subscript \( i \) in the parameters of terms \( R_j,c, V_j,c \) and \( K_j,c \) are used to indicate that the partial derivative of the term is with respect to \( z_i \).

Using equation 67 we get

\[
\sum_{c=-1}^{\infty} \sum_{j=1}^{N} R_j,c(1_t) = \frac{\lambda_i(1 - \rho_i)E[R_T]}{(1 - \rho)} \tag{70}
\]

\[
\sum_{c=-1}^{\infty} V_j,c(1_t) = \frac{\lambda_i(1 - \rho_i)E[V_j]}{(1 - \rho)} \tag{71}
\]

\[
\sum_{c=-1}^{\infty} t_{j,c}(i) - t_{1,c}(i) = \frac{-1}{\lambda_i} \sum_{c=-1}^{\infty} \sum_{k=2}^{j} R_k,c(1_t),
\]

\[
= \frac{-1}{\lambda_i} \sum_{c=-1}^{j} \lambda_i E[R_k] \sum_{c=-1}^{\infty} \left[ \sum_{\ell=k}^{N} \gamma_{\ell,c}(i) + \sum_{\ell=1}^{k-1} \gamma_{\ell,c-1}(i) \right],
\]

\[
= \frac{-1}{\lambda_i} \sum_{k=2}^{j} \lambda_i E[R_k] \left( \sum_{c=-1}^{\infty} \sum_{\ell=1}^{N} \gamma_{\ell,c}(i) + \sum_{\ell=1}^{k-2} \gamma_{\ell,c-2}(i) \right),
\]

\[
= \frac{(1 - \rho_i) \sum_{k=2}^{j} E[R_k]}{1 - \rho} \tag{72}
\]

Finally, substituting from 70 - 72, we obtain:

\[
F_i'[\bar{1}_i] = \left( \frac{\lambda_i(1 - \rho_i)}{(1 - \rho)(1 + \sum_{j=1}^{N} \varphi_j)} \right) \left( E[R_T] + \sum_{j=1}^{N} \varphi_j E[V_j] + \sum_{j=2}^{N} \varphi_j \sum_{k=2}^{j} E[R_k] \right) \tag{73}
\]

Similar to the variable \( C_1 \), let \( C_i \) denote the length of a server cycle at station \( i \), which is defined as the time elapsed between two consecutive polling instants at station \( i \). Then,

\[
E[C_i] = \frac{E[R_T] + \sum_{j=1}^{N} \varphi_j E[V_j] + \sum_{j=2}^{N} \varphi_j \sum_{k=2}^{j} E[R_k]}{(1 - \rho)(1 + \sum_{j=1}^{N} \varphi_j)} \tag{74}
\]

From equation 59 we can calculate the sum of the second factorial moments of contributions to the queue length at station \( i \) polling instants.

\[
\sum_{c=-1}^{\infty} \sum_{j=1}^{N} \gamma_{j,c}^{(2)}(i) = \sum_{c=0}^{\infty} \sum_{j=1}^{N} \gamma_{j,c}^{(2)}(i) + \sum_{j=1}^{i-1} \gamma_{j,c}^{(2)}(i),
\]

\[
= \sum_{j=1}^{N} \varphi_j \sum_{c=0}^{\infty} \left[ \sum_{k \geq j} \gamma_{k,c}^{(2)}(i) + \sum_{k < j} \gamma_{k,c-1}^{(2)}(i) \right] + \lambda_i \sum_{j=1}^{N} \varphi_j [B_j^2] \sum_{c=0}^{\infty} \left( \frac{\gamma_{j,c}(i)}{\rho_j} \right)^2
\]

19
Note that $\gamma_{2,i-1}(i) = 0$ for $j \geq i$. Also, define

$$
\Gamma_j(i) = \begin{cases} 
\sum_{c=-1}^{\infty} \gamma_{2,c}(i), & j < i \\
\sum_{c=0}^{\infty} \gamma_{2,c}(i), & j \geq i
\end{cases}
$$

Then,

$$
\sum_{c=-1}^{\infty} \sum_{j=1}^{N} \gamma_{j,c}^{(2)}(i) = \left( \frac{\lambda_i}{1-\rho} \right) \sum_{i=1}^{\infty} \left( \frac{\Gamma_j(i)}{\rho_j^2} \right) \lambda_j E[B_j^2].
$$

Using the relations 66 and 78 we can obtain the following results:

$$
\sum_{c=-1}^{\infty} \sum_{j=1}^{N} R_{j,c}(1_i) - R_{j,c}(1_i)^2 = \lambda_i^2 \left[ \sum_{c=-1}^{\infty} \sum_{j=1}^{N} (E[R_j^2] - E[R_j])^2 \left( \sum_{k \geq j} \gamma_{k,c}(i) + \sum_{k < j} \gamma_{j,c-1}(i) \right)^2 \right] + \lambda_i \sum_{c=-1}^{\infty} \sum_{j=1}^{N} E[R_j] \left( \sum_{k \geq j} \gamma_{k,c}(i) + \sum_{k < j} \gamma_{j,c-1}(i) \right) + \lambda_i \sum_{j=i+1}^{\infty} \sum_{c=1}^{\infty} \gamma_{j,c}^{(2)}(i) \rho_j^2 + E[R_{T_j}] \left( \sum_{c=-1}^{\infty} \sum_{j=1}^{N} \gamma_{j,c}^{(2)}(i) + \sum_{k=1}^{j-1} \gamma_{j,c-2}(i) \right).
$$

Using the definition of $\Gamma_j(i)$ from equation 77 we get

$$
\sum_{c=-1}^{\infty} \sum_{j=1}^{N} R_{j,c}(1_i) - R_{j,c}(1_i)^2 = \lambda_i^2 \left[ \sum_{j=1}^{\infty} \left( \frac{\Gamma_j(i)}{\rho_j^2} \right) \left( Var(R_j) + \frac{\lambda_j E[B_j^2] E[R_{T_j}]}{(1-\rho)} + Var(R_i) \right) \right] + Var(R_i).
$$

$$
\sum_{c=0}^{\infty} K_{j,c}^{(2)}(1_i) - K_{j,c}^{(2)}(1_i) = -\lambda_i^2 \sum_{c=0}^{\infty} \sum_{m=2}^{j} \left[ \frac{\gamma_{m,c}^{(2)}(i)}{\rho_m^2} \right] Var(R_m) + \frac{E[R_m]}{\lambda_i} \left( \sum_{k \geq m} \gamma_{k,c}(i) \right) + \left( t_{1,c}^{2}(i) - t_{j,c}^{2}(i) \right) + \left( t_{k,c}^{2}(i) - t_{k-1,c}^{2}(i) \right).
$$

$$
\sum_{c=0}^{\infty} K_{j,c}^{(2)}(1_i) - K_{j,c}^{(2)}(1_i) = -\lambda_i^2 \sum_{c=0}^{\infty} \sum_{m=2}^{j} \left[ \frac{\gamma_{m,c}^{(2)}(i)}{\rho_m^2} \right] Var(R_m) + \frac{E[R_m]}{\lambda_i} \left( \sum_{k \geq m} \gamma_{k,c}(i) \right) + \left( t_{1,c}^{2}(i) - t_{j,c}^{2}(i) \right) + \left( t_{k,c}^{2}(i) - t_{k-1,c}^{2}(i) \right).
$$
For \( c = -1 \) we have three cases:

(i) \( i = 1 \), then \( K_{j,-1}(z_1) = 1 \), for all \( j \), and hence \( K''_{j,-1}(1_i) = 0 \) for all \( j = 1, \ldots, N, i = 1 \). Therefore, cycle index start from \( c = 0 \) for \( i = 1 \).

(ii) \( 1 \leq j < i \), then

\[
K''_{j,-1}(1_i) - K''_{1,-1}(1_i) = -\lambda_i^2 \sum_{m=2}^{j} \left( \text{Var}(R_m) \frac{\gamma_{m,-1}(i)}{\rho_m^2} + \frac{E[R_m]}{\lambda_i} \left( \sum_{k=m}^{N} \gamma_{k,-1}(i) + \sum_{k=1}^{m-1} \gamma_{k,-2}(i) \right) \right) - \lambda_i^2 (t_{i,-1}(i) - t_{j,-1}(i)).
\]

(iii) \( 1 < i \leq j \), then

\[
K''_{j,-1}(1_i) - K''_{1,-1}(1_i) = -\lambda_i^2 \sum_{m=2}^{i-1} \left( \text{Var}(R_m) \frac{\gamma_{m,-1}(i)}{\rho_m^2} + \frac{E[R_m]}{\lambda_i} \left( \sum_{k=m}^{N} \gamma_{k,-1}(i) + \sum_{k=1}^{m-1} \gamma_{k,-2}(i) \right) \right) - \lambda_i^2 (t_{i,-1}(i) + t_{j,-1}(i)).
\]

Then, for \( \sum_{c=-1}^{\infty} K''_{j,c}(1_i) - K''_{1,c}(1_i) \) we have two situations (of course for \( i > 1 \)). Note that the \( \gamma_{k,-2}(i) \) terms are ignored, since they are all zero, and for notational simplicity in some summation terms the upper limit is extended from \( i \) to \( N \) for terms \( \gamma_{k,-1}(i) \), where \( k \geq i \).

\[ j < i : \]

\[
\sum_{c=-1}^{\infty} K''_{j,c}(1_i) - K''_{1,c}(1_i) = -\lambda_i^2 \left( \sum_{m=2}^{j} \left[ \sum_{c=1}^{\infty} \left( \frac{\gamma_{m,c}(i)}{\rho_m^2} \right) \text{Var}(R_m) + \frac{E[R_m]}{\lambda_i} \left( \sum_{k=0}^{N} \gamma_{k,c}(i) \right) \right] + \sum_{k=1}^{m-1} \gamma_{k,-1}(i) + \sum_{k=1}^{N} \gamma_{k,-1}(i) \right) + \sum_{c=-1}^{\infty} \left( t_{i,c}(i) - t_{j,c}(i) \right),
\]

\[ j \geq i : \]

\[
\sum_{c=-1}^{\infty} K''_{j,c}(1_i) - K''_{1,c}(1_i) = -\lambda_i^2 \left( \sum_{m=2}^{j-1} \left[ \sum_{c=1}^{\infty} \left( \frac{\gamma_{m,c}(i)}{\rho_m^2} \right) \text{Var}(R_m) + \frac{E[R_m]}{\lambda_i} \left( \sum_{k=0}^{N} \gamma_{k,c}(i) \right) \right] + \sum_{k=1}^{m-1} \gamma_{k,-1}(i) + \sum_{k=1}^{N} \gamma_{k,-1}(i) \right) + \sum_{c=-1}^{\infty} \left( t_{i,c}(i) - t_{j,c}(i) \right) + \text{Var}(R_i).
\]
Using the definition of $\Gamma_j(i)$ we can combine these two cases into a single equation. Also, we must take care of the case $i = 1$, in which case we do not have a $-\lambda_1 \text{Var}(R_1)$ term, since there exist no switchovers (setups) at station 1. Then we get
\[
\sum_{c=-1}^{\infty} K''_{j,c}(1_i) - K''_{1,c}(1_i) = -\lambda_i^2 \left( \sum_{m=2}^{j} \left( \frac{\Gamma_m(i)}{\rho_m^2} \right) \text{Var}(R_m) + \frac{E[R_m]}{1 - \rho} \sum_{k=1}^{N} \left( \frac{\Gamma_k(i)}{\rho_k^2} \right) \lambda_k E[B_k^2] \right) \\
+ \text{Var}(R_i) \left[ (1 - \delta_1(i))(1 - \sum_{k=1}^{i-1} \delta_j(k)) \right] + \sum_{c=-1}^{\infty} \left( t_{1,c}(i) - t_{j,c}(i) \right),
\]
(87)
where
\[
\delta_i(j) = \begin{cases} 
1, & \text{if } j = i \\
0, & \text{otherwise}.
\end{cases}
\]

Similarly,\n\[
\sum_{c=-1}^{\infty} V''_{j,c}(1_i) = \lambda_i^2 \sum_{c=-1}^{\infty} E[V_j^2] \left( \sum_{k=1}^{N} \gamma_{k,c}(i) \right)^2 + \lambda_i \sum_{c=-1}^{\infty} E[V_j] \sum_{k=1}^{N} \gamma_{k,c}(i)^2, 
\]
(88)
\[
= \lambda_i^2 \left[ E[V_j^2] \sum_{c=0}^{\infty} \left( \frac{\gamma_{c}(i)}{\rho_1} \right)^2 + \delta_1(i) E[V_j^2] + (1 - \delta_1(i)) E[V_j^2] \left( \frac{\gamma_{1-1}(i)}{\rho_1} \right)^2 \\
+ \frac{E[V_j]}{\lambda_i} \sum_{c=-1}^{\infty} \sum_{k=1}^{N} \gamma_{k,c}(i)^2 \right],
\]
(89)

Using the definition of $\Gamma_j(i)$ for $j = 1, \ldots, N$, we get
\[
\sum_{c=-1}^{\infty} V''_{j,c}(1_i) = \lambda_i^2 E[V_j] \left( \frac{\Gamma_j(i)}{1 - \rho} \right) \sum_{k=1}^{N} \left( \frac{\Gamma_k(i)}{\rho_k^2} \right) \lambda_k E[B_k^2] + \lambda_i^2 E[V_j^2] \left[ \delta_1(i) + \frac{\Gamma_1(i)}{\rho_1^2} \right],
\]
(90)
\[
\sum_{c=-1}^{\infty} V'_{j,c}(1_i) K'_{1,c}(1_i) = \lambda_i^2 E[V_j] \sum_{c=-1}^{\infty} \frac{t_{1,c}(i) \gamma_{1,c}(i)}{\rho_1}.
\]
(91)

Note that, for $i = 1$, $t_{1,-1}(i) = 0$, and the above relationship matches with equation 49.

Substituting from above into equation 44, simplifying and collecting terms, we obtain the following expression after some effort:
\[
F''_i(1_i) = \left( \frac{\lambda_i^2}{1 + \sum_{j=1}^{N} \theta_j} \right) \left( \frac{(1 - \rho_i)^2 E[R_i]^2}{(1 - \rho)^2} \right) + \text{Var}(R_i) + \sum_{j=1}^{N} \left( \frac{\Gamma_j(i)}{\rho_j^2} \right) \text{Var}(R_j) \\
+ (1 + \sum_{j=1}^{N} \theta_j) E[C_i] \sum_{j=1}^{N} \left( \frac{\Gamma_j(i)}{\rho_j^2} \right) \lambda_j E[B_j^2] + \sum_{j=1}^{N} \theta_j \left[ \sum_{k=2}^{j} \left( \frac{\Gamma_k(i)}{\rho_k^2} \right) \text{Var}(R_k) + \sum_{c=-1}^{\infty} \left( t_{1,c}(i) - t_{j,c}(i) \right) \right] \\
+ \text{Var}(R_i) \left[ (1 - \delta_1(i))(1 - \sum_{k=1}^{i-1} \delta_j(k)) \right] + 2E[V_j] \sum_{c=-1}^{\infty} \left( \frac{t_{1,c}(i) \gamma_{1,c}(i)}{\rho_1} \right) + E[V_j^2] \left( \delta_1(i) + \frac{\Gamma_1(i)}{\rho_1^2} \right)
\]
(92)
After substituting equations 73, 74 and 92 into the equation 54 we get

\[
E[W_i] = \frac{(1 - \rho_i)E[R_T]^2}{2(1 - \rho)^2 E[C_i](1 + \sum_{j=1}^N \theta_j)} + \frac{\lambda_i E[B_i^2]}{2(1 - \rho_i)} + \frac{\sum_{j=1}^N \left( \begin{array}{c} \Gamma_j(i) \rho_j \end{array} \right) \lambda_j E[B_j^2]}{2(1 - \rho_i) + \sum_{j=1}^N \left( \begin{array}{c} \Gamma_j(i) \rho_j \end{array} \right) \text{Var}(R_j)} + \frac{\text{Var}(R_i) + \sum_{j=1}^N \left( \begin{array}{c} \Gamma_j(i) \rho_j \end{array} \right) \text{Var}(R_j)}{2(1 - \rho_i)(1 + \sum_{j=1}^N \theta_j) E[C_i]}
\]

\[+ \sum_{j=1}^N \theta_j \left\{ \sum_{k=2}^N \left( \begin{array}{c} \Gamma_k(i) \rho_k \end{array} \right) \text{Var}(R_k) + E[V_j^2](\delta_1(i) + \frac{\Gamma_1(i)}{\rho_1}) + \text{Var}(R_i) \left[ (1 - \delta_1(i))(1 - \sum_{k=1}^{i-1} \delta_j(k)) \right] \right\}
\]

\[+ \sum_{j=1}^N \theta_j \sum_{c=-1}^{\infty} \left[ 2E[V_j] \left( \frac{t_{1,c}(i)}{\rho_1} \right) + (t_{1,c}(i) - t_{2,c}(i)) \right]
\]

\[
\frac{1}{2(1 - \rho_i)(1 + \sum_{j=1}^N \theta_j) E[C_i]}
\]

(93)

3.1.7 The Fraction of Facility Time Spent on PM

Using the key renewal theorem, the fraction of facility (or server) time spent on PM (or vacations) is the ratio of the average amount of time spent on vacations in a renewal cycle divided by the average length of the cycle (see, for example, Ross, 1996, pp. 141). A natural choice for system renewal epochs is the set of time epochs at which the server registers a switch point from station \( N \). These epochs correspond to the moments at which the server ends a complete cycle, i.e., in which all stations are visited once. We shall define the time interval between two such consecutive renewal epochs as a R-cycle and denote it by \( RC \).

The expected number of station 1 polling instants that occur during a R-cycle are \( 1 + \sum_{j=1}^N \theta_j \). (See Remark 1.) Since \( E[C_1] \) represents the expected time between two successive polling epochs at station 1, we obtain \( E[RC] = \left( 1 + \sum_{j=1}^N \theta_j \right) E[C_1] \) as

\[
E[RC] = \frac{E[R_T] + \sum_{j=1}^N \theta_j E[V_j] + \sum_{j=2}^N \theta_j \sum_{k=2}^j E[R_k]}{(1 - \rho)}.
\]

(94)

Next, using the fact that the average amount of time spent on vacations during a R-cycle is equal to \( \sum_{j=1}^N \theta_j E[V_j] \), the average proportion of time server spends on PM equals

\[
t_v = \frac{E[R_T] + \sum_{j=1}^N \theta_j E[V_j] + \sum_{j=2}^N \theta_j \sum_{k=2}^j E[R_k]}{(1 - \rho)}.
\]

(95)

3.2 The SV Model

The basic difference in the single vacation model is the behavior of the server at the end of a vacation. The server waits until the first arrival in case it finds the system empty at the end of a vacation, which is a polling
instant of station 1. Therefore, all the relationships among the station polling and completion instants and switch points (i.e., relations 6 – 8) are valid, except the equation 9. Before presenting the new equation for \( f_{1,c}(z_i) \), we describe the system behavior when station 1 is polled. If a station 1 polling instant has occurred due to a vacation completion, then the server behavior depends on the system state (i.e., patient server behavior). Therefore, we need to differentiate these time epochs from the usual polling instants. We define the time epochs that the server reached to station 1 as station 1 beginning instants. At a station 1 beginning instant if the system is not empty, simultaneously a station 1 polling instant is marked and the server resumes its roving; otherwise the server becomes patient (no type-1 polling instant is marked yet) and the station 1 polling instant is marked by the first arrival to the system, which reactivates the server for roving.

Let \( b_1(n_1, \ldots, n_N) \) to denote the probability of the system state at a station 1 beginning epoch. Then, \( b_1(0) \neq 0 \), but \( f_j(0) = 0 \), for all \( j = 1, \ldots, N \). Also, let the PGF of the contributions to queue \( i \) from a station 1 beginning instant, \( c \) cycles prior to the reference point is denoted by \( b_{1,c}(z_i) \), then

\[
b_{1,c}(z_i) = h_{N,c+1}(z)R_{1,c}(z) + \sum_{j=1}^{N} g_j(0)V_{j,c}(z_i)
\]

Using the PS behavior, now we can write the modified equation for \( f_{1,c}(z_i) \),

\[
f_{1,c}(z_i) = b_{1,c}(z_i) - b_1(0)(1 - \sum_{j=1}^{N} p_{j}L_{j,c}(z_i)).
\]

Note that \( b_1(0) = \sum_{j=1}^{N} g_j(0)V_j^*(\lambda) \). Also, since we mark the station 1 polling instant (following a vacation completion with empty system) at the arrival epoch of the first customer, we still have \( g_j(1) = f_j(1) \) identity. (Therefore, \( F_i(\bar{z}) \) is still a conditional PGF and \( \sum_{i=1}^{N} f_i(1) = 1 \).)

In our analysis, using equation 97 instead of equation 9 leads the equation 13 to become

\[
h_{N,0}(z) = h_{N,1}(z) \prod_{i=1}^{N} R_{i,0}(z) - \sum_{j=1}^{N} g_j(0)\{ \prod_{k=j+1}^{N} R_{k,0}(z) - [V_{j,0}(z) - J_{j,c}(z)] \prod_{k=2}^{N} R_{k,0}(z) \},
\]

where \( J_{j,c}(z) = (V_j^*(\lambda)/\lambda) \sum_{k=1}^{N} (\lambda_k - \lambda_k L_{k,c}(z)) \).
Similarly equation 29 becomes

\[
F_l(1, \ldots, z_i, \ldots, 1) = \prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(z) - \sum_{j=1}^{N} \sum_{c=-1}^{\infty} \theta_j \sum_{j=1}^{\infty} \left\{ K_{j,c}(z) - [V_{j,c}(z) - J_{j,c}(z)]K_{1,c}(z) \right\} + \sum_{j=1}^{N} \theta_j \\
1 + \sum_{j=1}^{N} \theta_j
\]  

(99)

Again, the only terms to be evaluated are \( \theta_j \)'s, and they can be evaluated by solving \( N \) linear equations of the form:

For station \( i = 1, \ldots, N \):

\[
\prod_{c=-1}^{\infty} \prod_{j=1}^{N} R_{j,c}(\bar{0}_i) - \sum_{j=1}^{N} \theta_j \sum_{c=-1}^{\infty} \left\{ K_{j,c}(\bar{0}_i) - [V_{j,c}(\bar{0}_i) - J_{j,c}(\bar{0}_i)]K_{1,c}(\bar{0}_i) \right\} + \sum_{j=1}^{N} \theta_j = 0
\]  

(100)

Now in order to calculate \( E[W_i] \), again we need to calculate the first and second factorial moments of the queue length at station beginning instants which are not totally covered by the polling instants, any more. Therefore, \( E[W_i] \) term has to be driven.

\[
J'_{j,c}(1_i) = -p_v V'_j(\Lambda) \sum_{k=1}^{N} \gamma_{k,c}(i),
\]

(101)

\[
J''_{j,c}(1_i) = -p_v V''_j(\Lambda) \sum_{k=1}^{N} \gamma_{k,c}^{(2)}(i),
\]

(102)

\[
J'_j(1_i)K'_j(1_i) = -\frac{\lambda^2}{\Lambda} V'_j(\Lambda) t_{j,c}(i) \sum_{k=1}^{N} \gamma_{k,c}(i).
\]

(103)

Then, summing these variables over all \( c \) we obtain

\[
\sum_{c=-1}^{\infty} J'_{j,c}(1_i) = -p_v V'_j(\Lambda) \frac{(1 - p_v)}{(1 - \rho)},
\]

(104)

\[
\sum_{c=-1}^{\infty} J''_{j,c}(1_i) = -\frac{\lambda^2 V''_j(\Lambda)}{\Lambda(1 - \rho)} \sum_{k=1}^{N} \left( \frac{\Gamma_k(i)}{\rho_k^2} \right) \lambda_k E[B_{k}''],
\]

(105)

\[
\sum_{c=-1}^{\infty} J'_j(1_i)K'_j(1_i) = -\frac{\lambda^2 V'_j(\Lambda)}{\Lambda} \sum_{c=-1}^{\infty} t_{j,c}(i) \gamma_{j,c}(i).
\]

(106)

By differentiating equation 99 once and twice with respect to \( z_i \) and then setting \( z_i = 1 \) we obtain

\[
(1 + \sum_{j=1}^{N} \theta_j)F'_l(\bar{1}_i) = \sum_{c=-1}^{\infty} \sum_{j=1}^{N} R'_{j,c}(1_i) - \sum_{j=1}^{N} \theta_j \sum_{c=-1}^{\infty} \left\{ K'_{j,c}(1_i) - K'_1(1_i) - V'_{j,c}(1_i) + J'_j(1_i) \right\}
\]

(107)
\( (1 + \sum_{j=1}^{N} \phi_j) F''_p(\bar{\lambda}_i) = \sum_{c=-1}^{\infty} \sum_{j=1}^{N} \{ R''_{j,c}(1) - (R'_{j,c}(1))^2 \} + \left( \sum_{c=-1}^{\infty} \sum_{j=1}^{N} R'_{j,c}(1) \right)^2 - \sum_{j=1}^{N} \phi_j \sum_{c=-1}^{\infty} \{ K''_{j,c}(1) - K'_{j,c}(1) - V''_{j,c}(1) + J''_{j,c}(1) - 2(V'_{j,c}(1) - J'_{j,c}(1))K'_{1,c}(1) \}. \) (108)

After simplification of equations (107) and (108) we obtain

\[
F'_p(\bar{\lambda}_i) = \left( \frac{\lambda_i (1 - \rho_i)}{(1 - \rho)(1 + \sum_{j=1}^{N} \phi_j)} \right) \left( E[R_T] + \sum_{j=1}^{N} \phi_j [E[V_j] + (V_j^*(\Lambda)/\Lambda)] + \sum_{j=2}^{N} \phi_j \sum_{k=2}^{\infty} E[R_k] \right), \quad (109)
\]

and

\[
F''_p(\bar{\lambda}_i) = \left( \frac{\lambda_i^2}{1 + \sum_{j=1}^{N} \phi_j} \right) \left( \frac{(1 - \rho_i)^2 E[R_T]^2}{(1 - \rho)^2} + \text{Var}(R_i) + \sum_{j=1}^{N} \left( \frac{\Gamma_j(i)}{\rho_i^2} \right) \text{Var}(R_j) \right) + (1 + \sum_{j=1}^{N} \phi_j) E[C_i] \sum_{j=1}^{N} \left( \frac{\Gamma_j(i)}{\rho_i^2} \right) \lambda_j E[B_j^2] + \sum_{j=1}^{N} \phi_j \sum_{k=2}^{\infty} \left( \frac{\Gamma_k(i)}{\rho_k^2} \right) \text{Var}(R_k) + \sum_{c=-1}^{\infty} \left( t_{1,c}(i) - t_{j,c}(i) \right) + \text{Var}(R_i) \left[ (1 - \delta_i(i))(1 - \sum_{k=1}^{i-1} \delta_j(k)) \right] + 2 \{ E[V_j] + (V_j^*(\Lambda)/\Lambda) \} \sum_{c=-1}^{\infty} \left( t_{1,c}(i) t_{1,c}(i) / \rho_1 \right) \right), \quad (110)
\]

where \( E[C_i] \) becomes

\[
E[C_i] = \frac{E[R_T] + \sum_{j=1}^{N} \phi_j [E[V_j] + (V_j^*(\Lambda)/\Lambda)] + \sum_{j=2}^{N} \phi_j \sum_{k=2}^{\infty} E[R_k]}{(1 - \rho)(1 + \sum_{j=1}^{N} \phi_j)}. \quad (111)
\]

After substituting equations (109) – (111) into the equation 54 we get

\[
E[W_i] = \frac{(1 - \rho_i) E[R_T]^2}{2(1 - \rho)(1 + \sum_{j=1}^{N} \phi_j) E[C_i]} + \sum_{j=1}^{N} \left( \frac{\Gamma_j(i)}{\rho_j^2} \right) \text{Var}(R_j) + \text{Var}(R_i) \delta_i(i) + \sum_{j=1}^{N} \phi_j E[B_j^2] \left[ \delta_i(i) + \Gamma_i(i)/\rho_i^2 \right] + \sum_{j=1}^{N} \phi_j \left\{ \sum_{k=2}^{\infty} \left( \frac{\Gamma_k(i)}{\rho_k^2} \right) \text{Var}(R_k) + \sum_{c=-1}^{\infty} \left[ 2 \{ E[V_j] + (V_j^*(\Lambda)/\Lambda) \} \left( t_{1,c}(i) t_{1,c}(i) / \rho_1 \right) + (t_{1,c}(i) - t_{j,c}(i)) \right] \right\} \right) \] \( + \frac{\lambda_i E[B_i^2]}{2(1 - \rho)(1 + \sum_{j=1}^{N} \phi_j) E[C_i]} \) + \( \frac{\lambda_i E[B_i^2]}{2(1 - \rho)} \) \( + \frac{(1 - \rho_i) E[R_i]^2}{2(1 - \rho_i)(1 + \sum_{j=1}^{N} \phi_j) E[C_i]} \) \( + \frac{(1 - \rho_i) E[R_i]^2}{2(1 - \rho)} \). \quad (112)
\]

Note that,

### 3.3 Continuously Roving Server Models: The CR and CR1 Models

In these models, we assume the server takes a vacation independent of system state at that time. In the CR model, we assume that the server takes a vacation after each station completion epoch, whereas in the CR1 model, vacations occur only at the end of station completion epochs at station 1.
Thus, the CR model is just the standard polling model with a continuously roving server, but with the switchover time at each station inflated by the vacation time. Hence its waiting time can be calculated easily by replacing \( E[R_j] \) with \( E[R_j] + E[V_{j-1}] \) and \( \text{Var}(R_j) \) with \( \text{Var}(R_j) + \text{Var}(V_{j-1}) \), in the mean waiting time formula of Konheim et al. (1994):

\[
E[W_i] = \frac{(1 - \rho_i)}{2E[C]} \left( \frac{E[R_T] + E[V_T]}{1 - \rho} \right)^2 + \frac{\lambda_i E[B_i^2]}{2(1 - \rho_i)} + \frac{\lambda_i [\text{Var}(R_i) + \text{Var}(V_{i-1})]/E[C]}{2(1 - \rho_i)} + \sum_{j=1}^{N} \frac{\Gamma_j}{\rho_j^2} \lambda_j E[B_j^2] + \frac{\text{Var}(R_j) + \text{Var}(V_{j-1})]/E[C]}{2(1 - \rho_i)}
\]

(113)

where \( E[V_T] = \sum_{j=1}^{N} E[V_j] \). The CR1 model is obtained as a special case with \( V_j = 0 \), for all \( j \neq 1 \), i.e., by having only \( V_1 > 0 \) in the above procedure. Note that a renewal cycle under both CR and CR1 policies is the ordinary server cycle observed at station 1.

4 Numerical Examples

The presence of the \( \vartheta_j \) terms makes it difficult to analytically compare the mean waiting times in MV and SV models. As seen above, the mean waiting time expressions are complex, highly non-linear functions of the input parameters. However, our approach makes it possible to quickly compute mean waiting times for a given system configuration. Consequently, we performed a number of numerical experiments for different system configurations. The results of these experiments are summarized in Figures 1 through 7 below. However, we begin by demonstrating an example first, similar in spirit to the PM conundrum presented in section 1.

We consider a manufacturing facility with 5 distinct job types. In the equivalent polling model, this results in 5 stations. We assume that all these jobs have identical characteristics, i.e., \( \lambda_i = \Lambda/5 \) and \( B_i = B \) for all \( i \). Service times are assumed to be exponentially distributed and switchover times constant with the following parameters: \( E[B_i] = 1.0 \), \( \lambda_i = 0.08 \), \( \rho = 0.4 \), and \( R_j = 0.01 \). The PM activity has a 2 point distribution, consisting of a typical value with a high probability, and an unusual value with a low probability, as shown below (measured in seconds):
The unusual value represents major overhaul which becomes necessary occasionally. It is also implicit in our model that the necessity of a major overhaul becomes known only after the PM process has been started.

Next, we used our models and calculated the weighted sum of mean waiting times or \( \sum_{i=1}^{N} \rho_i E[W_i] \). We use this quantity as an overall measure of system performance. It is also called the pseudo-conservation law (or PCL for short) in polling systems literature. When \( E[B_i] = E[B] \) for all \( i \), it follows from Little’s law that \( \sum_{i=1}^{N} \rho_i E[W_i] \) is equal to average total work-in-process multiplied by \( E[B] \). In all our examples below, we assume \( E[B_i] = 1 \), and thus PCL is a measure of both weighted average tardiness and total average WIP.

The following values of PCL were obtained for the example described above:

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Typical Value</th>
<th>Prob.</th>
<th>Unusual Value</th>
<th>Prob.</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.99</td>
<td>19,981</td>
<td>0.01</td>
<td>200.80</td>
<td>1,987.98</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0.99</td>
<td>19,985</td>
<td>0.01</td>
<td>204.80</td>
<td>1,987.98</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>0.99</td>
<td>20,000</td>
<td>0.01</td>
<td>219.80</td>
<td>1,987.98</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>0.99</td>
<td>20,030</td>
<td>0.01</td>
<td>249.80</td>
<td>1,987.98</td>
</tr>
<tr>
<td>5</td>
<td>500</td>
<td>0.99</td>
<td>20,480</td>
<td>0.01</td>
<td>699.80</td>
<td>1,987.98</td>
</tr>
<tr>
<td>6</td>
<td>5000</td>
<td>0.99</td>
<td>24,980</td>
<td>0.01</td>
<td>5,199.80</td>
<td>1,987.98</td>
</tr>
</tbody>
</table>

Clearly, increasing PM duration can have the unexpected benefit of lowering PCL (i.e., weighted average waiting times as well as average total WIP) while at the same time lowering unplanned downtime!

Figures 1, 2 and 3 show the effect of increasing expected duration of PM activity on various performance measures in all four models. The basic data is closely related to our example above, i.e., we have 5 identical stations, service times are exponential with rate 1, arrival rate is 0.08 for each job type, and overall facility utilization is 0.4. In figures 1 and 2, both switchover and PM activity times follow two point distributions as described below: \( R_j = 1 \) with probability (w.p.) 0.99 and 20 w.p. 0.01, \( V_j = x_1 \) w.p. 0.99 and \( x_2 \) w.p. 0.01, where \( x_1 \) and \( x_2 \) are chosen such that \( \text{Var}(V) = 49.9 \), and \( E[V] \) varies from 1.7 to 20.7. In figure 3, the switchover time distribution is the same as it was in figures 1 and 2, but the PM activity time is Erlang distributed with \( \text{Var}(V_j) = 50 \) while its parameters are varied to produce different values of \( E[V_j] \).
From figure 1 we see that for a large range of $E[V]$ values, SV and MV are indistinguishable in terms of PCL. This happens because once PM activity times become large enough, the opportunity to have multiple such activities during the same idle period of the server becomes negligible. Notice that SV consistently does as well as or better than the other three policies and that there is a range of $E[V]$ values for which PCL actually decreases under MV policy upon increasing the expected duration of PM activity. These observations, however, are not true in general; making it difficult to predict the impact of PM activity without the aid of models like ours. Focusing on figure 2, we notice that there are decreasing marginal returns in terms of the fraction of time the facility is available for PM activity when $E[V]$ is increased. Clearly, saturation point is reached earliest under the CR policy. For the MV policy, we notice that an increase in $E[V]$ can, in a certain range, lead to a drop in $t_o$. Combining these figures, it is apparent that manufacturing managers should attempt to control PM activity duration if it goes beyond the saturation point. Past saturation point, there is little marginal benefit (in terms of increased $t_o$) of higher $E[V]$ whereas both PCL and $E[RC]$ increase rapidly.

Figures 4, 5 and 6 show the effect of changing the variance of $V$ on PCL, $t_o$ and $E[RC]$ respectively. Figure 4 shows that when $\text{Var}(V)$ is small, the simple CR policy of performing PM at each switch from one job type to another outperforms the other three. Our experiments have revealed that this occurs only when both $V_j$ and $R_j$ are variable. Here too, for sufficiently large $\text{Var}(V)$, SV outperforms the remaining three PM policies. Figures 5 and 6 bring out the non-monotonic behavior of $t_o$ and $E[RC]$ as functions of $\text{Var}(V)$. We observe that both $t_o$ and $E[RC]$ are increasing in $\text{Var}(V)$ for the MV model whereas these metrics are nearly at their saturation points for the remaining three models. If $\text{Var}(V)$ is increased even further, $t_o$ does eventually reach an asymptotic maximum even for the MV model, similar to figure 2.

The interaction between $R_j$ and $V_j$, when they are both random can be seen in figure 7. Notice that as switchover times become more random (as compared to the variability of PM activity), CR becomes more desirable. In other experiments we performed (not reported), we also noticed that CR never minimizes PCL when $E[V]$ is made very large, which is consistent with our observation in figure 1.
5 Managerial Implications

Based on the numerical results, the following practical guidelines for manufacturing managers can be formulated.

- The State dependent protocols (MV, SV) are suitable only when $p$ is not too high. When $p > 0.8$, vacations are rare and state-dependent PM strategy becomes ineffective. Also, in such cases, PCL increases sharply with increasing PM activity duration under CR and CR1 regimes.

- Increasing PM duration can have the unexpected benefit of lowering PCL while at the same time lowering unplanned downtime! This phenomenon is, however, hard to predict. We have numerical evidence, as well as analogous analytical results in symmetrical two-station polling models without vacations (see, for example, Cooper, Niu and Srinivasan, 1998), which indicate that this counter-intuitive phenomenon cannot occur when both $R_j$ and $V_j$ are constants for all $j$.

- The fraction of time spent on vacations, namely, $t_v$, is increasing (though not monotonically) in $E[V]$ and it eventually exhibits decreasing marginal returns. To determine how much PM activity is appropriate in any particular case requires detailed modeling. Specifically, it depends on where the saturation point lies, which in turn depends in a complicated manner on system parameters. However, it is clear that increasing PM activity duration beyond the saturation point yields little benefits and only causes system performance to deteriorate. In such cases, it would only be justified by a substantial reduction in unplanned downtime due to failures. Generally, CR and CR1 policies have shorter saturation points and therefore these might be appropriate where benefits of PM come from frequent but short PM activity. On the other hand, if it is more beneficial to have longer PM activities (albeit less frequent), SV and MV policies make more sense.

- As either $N$, $E[R_j]$, or $E[V_j]$ become large, and we keep $p$ constant by adjusting either $A$ or $E[B_j]$, SV will eventually dominate other policies (irrespective of variabilities of $R_j$ and $V_j$) on account of the high overhead associated with switchovers for the remaining three policies.
• When switchover times are highly variable, and PM activity is not very time consuming, then a state-independent PM schedule results in best overall system performance, i.e., PCL is smallest under either CR or CR1.

• Whether SV, MV, or CR is superior in the PCL sense depends on the relative magnitudes and variabilities of $R_j$ and $V_j$. For example,
  
  - CR is superior when $R_j$'s are highly variable, and $N$ and $E[V_j]$ are not too large.
  - SV is superior when either $V_j$'s are highly variable or $E[V_j]$ are large.
  - SV/MV is usually superior under a wide range of parameter values, excluding extremes mentioned above.

Overall, we see that MV and SV give very close results. This suggests that when there are no jobs waiting to be processed at the end of a PM activity, it may be worth setting the facility aside for an additional level of preventive maintenance!

Acknowledgement

This research was supported, in part, by the Natural Sciences and Engineering Research Council of Canada through a research grant to DG (research grant #45904).

References


Figure 3
WORKING PAPERS - RECENT RELEASES


419. Robert F. Love and Halit Uster, "Comparison of the Properties and the Performance of the Criteria Used to Evaluate the Accuracy of Distance Predicting Functions", November, 1996.


