PERFORMANCE EVALUATION OF FLEXIBLE MANUFACTURING SYSTEMS WITH STATION BREAKDOWNS, MATERIAL HANDLING SYSTEM DELAY, AND GENERAL PROCESSING TIMES

By

LONG-GENG ZHAO

Master of Business Administration
Oklahoma State University
Stillwater, Oklahoma
1991

Master of Computer Science
Fudan University
Shanghai, China
1985

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 1992

PERFORMANCE EVALUATION OF FLEXIBLE MANUFACTURING

SYSTEMS WITH STATION BREAKDOWNS, MATERIAL

HANDLING SYSTEM DELAY, AND GENERAL

PROCESSING TIMES

Thesis Approved:

Thesis Adviser

Marilyn G. Kletke

D. E. Hedrick

Thomas C. Collins

Dean of the Graduate College

PREFACE

This study is concerned with the modeling and performance evaluation of Flexible Manufacturing Systems (FMS's) with station breakdowns, Material Handling System (MHS) delay, and general repair and processing times. The primary objective is to determine the system's steady state output rate, sojourn time, and station utilizations. The open queueing network model is used to analyze the FMS. A tranformation algorithm is developed to deal with station breakdowns and MHS delay, and then an iterative procedure is devised to derive the system's performance measures.

The author wishes to express his appreciation to his major adviser, Dr. Hon-Shiang Lau, for his guidance and assistance throughout this study. Appreciation is also expressed to the other committee members, Dr. G.E. Hedrick, Dr. David C. Ho, Dr. M.G. Kletke, and Dr. R. Wilson, for their invaluable assistance.

Finally, special gratitude is expressed to my wife, Bifen Wu, our daugher, Congzhi, and my parents Yan Zhao and Yueming Zhang, for their understanding, encouragement, and many sacrifices.

TABLE OF CONTENTS

Chapter Page

Chapter | Page

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

a_i	e_i/δ_i
b_i	Probability that a job is blocked on arriving at station i
e_i	Expected number of times a job will visit station i
e_{ir}	Expected number of times a class-r job will visit station i
f_i	See f_{ir}, when there is only one class of jobs, or r is understood
f_{ir}	Probability that a class-r job leaves the system after processing at station i
g_i	Probability that a job visits another station after processing at station i, i.e., $(1 - f_i)$
k	Mean time between breakdowns
pdf	Probability density function
$g(\cdot)$	Processing time S's pdf
$h(\cdot)$	Time between breakdowns T's pdf
$m(\cdot)$	General processing time GS's pdf
n_i	Station i's buffer capacity
$r(\cdot)$	Repair time R's pdf
q_{ir}	Probability that a class-r job has its first operation on machine i
C	Maximum number of jobs allowed in a system
C_k^i	$k!/[(k-i)!i!]$
B	Number of classes of jobs
CDF	Cumulative density function

$G(\cdot)$	Processing time S's CDF
$H(\cdot)$	Time between breakdowns T's CDF
$R(\cdot)$	Repair time R's CDF
D_i	Transition time, or MHS delay, of station i
$E[Z]$	Expected value (mean) of the random variable Z
GS_i	General processing time of station i, i.e., $S_i + D_i$
$H(\cdot)$	T's CDF (see T)
L	The number of jobs in the station, i.e, station size
N_r	Number of class-r jobs in the ststem
M	Number of stations in a system
P	see $\mathbf{P}_r$, when there is only one class of jobs, or r is or r is understood
P_{ij}	see $P_{ij,r}$, when there is only one class of jobs, or r is understood
$P_{ij,r}$	Probability that a class-r job will go to station j after processing at statin i
$\mathbf{P}_r$	Transition matrix $(P_{ij,r})$ for class-r jobs
PC	Production capacity, i.e., maximum possible output rate
Q	The number of jobs in the queue, i.e, queue length
R	Repair time
R_i	Repair time of station i
S	Processing time, or service time
S_i	Processing time, or service time, of station i
T	Time between two consecutive breakdowns
T_i	Time between two consecutive breakdowns of station i
TH	Throughput rate, i.e., output rate
U_i	The utilization of the machine (station) i

W	The total time a job stays in the queue, i.e, waiting time
X	The total time a job stays in the station, i.e, station time
Y	Effective processing time
Y_i	Effective processing time of staion i
$\alpha_1(Z)$	$\mu_3(Z)/\sigma_Z^3$, i.e., Z's skewness measure
$\alpha_2(Z)$	$\mu_4(Z)/\sigma_Z^4$, i.e., Z's kurtosis measure
δ_i	Mean processing rate of machine i
λ	Mean arrival rate (to the system)
k_j	$\int_0^\infty t^j H(t)m(t)dt$
λ_i	Mean arrival rate (to the station i)
μ	$1/\delta_i$ (for a specified station i)
μ_Z	Expected value (mean) of the random variable Z
$\mu_n(Z)$	The n-th central moment of Z ($n>1$)
$\nu_n(Z)$	The n-th moment (about zero) of Z ($n\geq 1$)
σ_Z	Standard deviation of the random variable Z

CHAPTER I

INTRODUCTION

This chapter outlines a dissertation that investigates flexible manufacturing systems (FMS's) with station breakdowns, material handling system (MHS) delay, and general processing times. In this investigation, a powerful approach is developed to theoretically transform an FMS with station breakdowns and MHS delay into an equivalent system without station breakdowns and MHS delay. This transformation facilitates performance evaluations of individual stations, as well as the whole FMS with general processing times. A recursive algorithm is devised to calculate stations' performance measures, such as the number of jobs in the station or in its queue, and the total time a job stays in the station. An iterative procedure is then developed to obtain the FMS's performance measures, such as the output rate, sojourn time, and station utilizations.

This chapter begins with a brief introduction of the research problem, followed by a summary of research methodologies and findings. Then theoretical and practical implications and importance of the dissertation research will be described.

The Research Problem

An FMS is usually modeled as a queueing network in which the customers are the jobs to be processed by the system and the servers are the CNC (Computer Numerical Control) machines. Each CNC machine has a local control computer linked by a communication network. The model is schematically depicted in Figures 4 and 5 on pages 25 and 26. The basic assumptions of the queueing network model are as follow:

(1) The total number of jobs in the system could be a fixed constant N, or various.

(2) All stations use the FCFS (First-Come-First-Served) queue discipline and all job classes have the same service rate at a station. The stations' service times could be exponential or non-exponential.

(3) There may be a central storage to accommodate all jobs in the system. Each station may have a local buffer with infinite or limited capacity.

(4) Machines are always available for processing jobs, (i.e., no machine breakdowns,) and any set-up/tool changing time is included in the service (processing) times.

(5) No MHS delay.

The queueing network models have been successfully applied to study various planning and control aspects of FMS's, largely because the first three basic assumptions are so flexible that they impose few restrictions. However, the last two basic assumptions seem to be very restrictive and unrealistic. In practice, assumptions (4) and (5) are often violated, because machines do break down from time to time; and the MHS needs some time to move workpieces among stations and the central storage. Therefore, it is desired to develop new models and new methodologies which allow station breakdowns and MHS delay, and can incorporate general processing times.

This dissertation will present a powerful and convenient approach to transform a system with MHS delay and breakdowns into an equivalent system without MHS delay and breakdowns. This transformation is based upon the method of moments. After this transformation, the earlier techniques for the case of no MHS delay and breakdowns can be applied to analyze the system. Furthermore, to take the advantage of this moments-oriented transformation, this dissertation develops new iterative procedures to obtain performance measures for individual stations as well as the entire FMS's.

Methodologies and Findings

In this dissertation, a two-step transformation is conducted to transform a system with MHS delay and breakdowns into an equivalent system without MHS delay and breakdowns. In the first step, the repair time of each machine will be treated as a part of the machine's processing time. This is called the "effective processing time," which includes the machine's normal processing time and repair time. When a machine breaks down while processing a job, the effective processing time will be the sum of the processing time and the machine's repair time. Of course, if the machine does not break down, the effective processing time is equivalent to the processing time. Let Y be the effective processing time, which is a random variable. If the distributions, or only the moments, of the processing time, repair time, and time between breakdowns are given, then Y's moments can be determined.

The second step is to address the MHS delay issue. This step, much more complicated than the first, is directly related to the storage system. If there is only a central storage with infinite capacity, one can model the central storage and the MHS as a virtual station with the effective processing time that includes the queue waiting time and MHS delay. If there are also local moving buffers (say, conveyors) with limited capacity, MHS delay, denoted by D, can be treated as a part of stations' "generalized processing times." It is more difficult to handle the traditional local fixed buffers. This dissertation provides a hybrid model to absorb MHS delay.

After absorbing the repair times and MHS delay into the effective processing time Y's or generalized processing time GS's, the approaches of Hahn and Shapiro (1968), Kendall and Stuart (1969), and Kottas and Lau (1979, 1980) can be used to fit Y's or GS's first k (k=3 or 4) moments to a k-parameter distribution function. Thus the system with the effective (or generalized) processing times is equivalent to the original system, but without breakdowns and MHS delay. Fitting the first k moments to a k-parameter distribution is a convenient method that provides good approximations.

Moreover, when the processing time distributions are known, the moments of Y's and/or GS's can be calculated analytically. These moments are all that are needed to evaluate each station's performance, such as the total time in the station (station time X, which is a use-ful byproduct to determine the sojourn time) or in its queue (waiting time W), the number of jobs in the station (station size L) or in the queue (queue length Q), and the station output rate (TH). A step-by-step recursive algorithm (Procedure I) is shown to calculate the moments of X, W, L, and Q.

To employ this unique moments-oriented feature, Yao and Buzacott (1985b)'s open queueing network model, with necessary modifications, is adopted to deal with station breakdowns and MHS delay, and an iterative procedure (Procedure II) is developed to evaluate the whole FMS's performance. This procedure employs a different, natural iteration scheme, and a simple method to compute each station's blocking rate, or the probability that the station will reject

2a coming job. The blocking rates are critical to derive the system output rate, sojourn time, and machine utilization.

Numerical examples are used to show how to transform an FMS with breakdowns and MHS delay into an equivalent system without breakdown and MHS delay. After the transformation, the effective (or genera-lized) processing times are never exponential, no matter whether the original processing times are exponential or not. Then, iterative Procedure II can be used to analyze the system's performance. Computer simulations are conducted to verify the analytical results.

Theoretical and Practical Implications of the Dissertation Research

This dissertation presents a powerful and convenient approach to model FMS's with machine breakdowns, MHS delay and stochastic repair and processing times. By transforming the system with breakdowns and MHS delay into an equivalent system without breakdown and MHS delay, many well-established methodologies can be applied to evaluate the system's performance. This approach concentrates on the effective processing time (or the generalized processing time) Y. One can analytically calculate Y's (central) moments necessary in fitting any n-parameter distribution for Y. Thus, the approach provides a solid base for the further analysis of the whole queueing network.

Usually, one fits Y's first three or four moments to a three or four parameter distribution. The major advantage of the approach in this paper is that there is no need to know the repair time R and/or MHS delay D's probability density functions (pdf's). Moreover, if the processing time S's pdf is unknown, but S's first three or four moments are known, one can fit them to a three or four parameter distribution as S's pdf. So in practice, only the mean breakdown rate and the first three or four moments of R, S and D are needed, which are easy to obtain from collected data. This feature makes the proposed method very attractive to both researchers and practitioners, as determining those probability density functions is not necessary.

After the transformation, one can evaluate each station's perfor-mance, that is necessary to evaluate the whole system's performance. By assuming a Poisson arrival process, a recursive procedure is devised to use Y's moments to calculate the moments of the station time, wait-ing time (in the queue), and station size for each station. A new algorithm is proposed to calculate the moments of the length of queue. It is relatively easy to obtain each station's output rate and utiliza-tion. Here the knowledge of the station time is very important, since that makes it possible to obtain the sojourn time. While early models concentrate on the derivation of systems' output rates, it is believed that the sojourn time is more relevant to customers, especially for make-to-order productions. This dissertation develops an iterative procedure (Procedure II) to calculate the FMS's output rate, machine utilization, and <u>sojourn time</u>. This procedure is very efficient, and easy to program.

Queuing network simulation models are also easy to be established to verify the analytical conclusions.

Outline of the Dissertation

This chapter (Chapter I) is an introduction. Chapter II gives a literature review. In chapter III, the necessary, basic assumptions and definitions are introduced. Chapter IV deals with the performance evaluations of queueing networks without MHS delay and breakdown. Chapter V discusses machine breakdowns and the calculations of the effective processing time Y's moments, and then in chapter VI several numerical examples are given to show how to transform an FMS with breakdowns into an equivalent system with no breakdowns. Chapter VII discusses MHS delay and how to deal with MHS delay. In fact, MHS delay can be absorbed in the queue waiting time or in the effective processing time Y. Chapter VIII shows how to use Y's moments to compute the waiting time, queue length, time in the station, and number of jobs in the station (their moments and distributions). In chapter IX, an iterative procedure is proposed to derive the FMS's performance measures. Chapter X gives conclusions and discussions.

CHAPTER II

A GENERAL LITERATURE REVIEW AND EARLY RELATED RESEARCH

Definition of an FMS

§1.1 b

A flexible manufacturing system is a computer-controlled configuration, consisting of a group of processing stations, each containing a set of computer numerical control (CNC) machine tools, interconnected by means of an automated material handling system (MHS) and storage systems (Groover 1987). It is capable of processing a variety of different types of parts simultaneously at the various stations. The FMS is designed to combine the mass-production efficiency of transfer lines and the flexibility of job shops to handle batch production at medium volume and medium product variety. FMS also combines the existing technology of NC (numerical control) manufacturing, automated material handling, and computer hardware and software to create an integrated system for the automatic random processing of palletized parts across various work stations in the system. Figure 1 gives two examples of FMS layout (Goetsch 1988).

An FMS has four essential physical components:

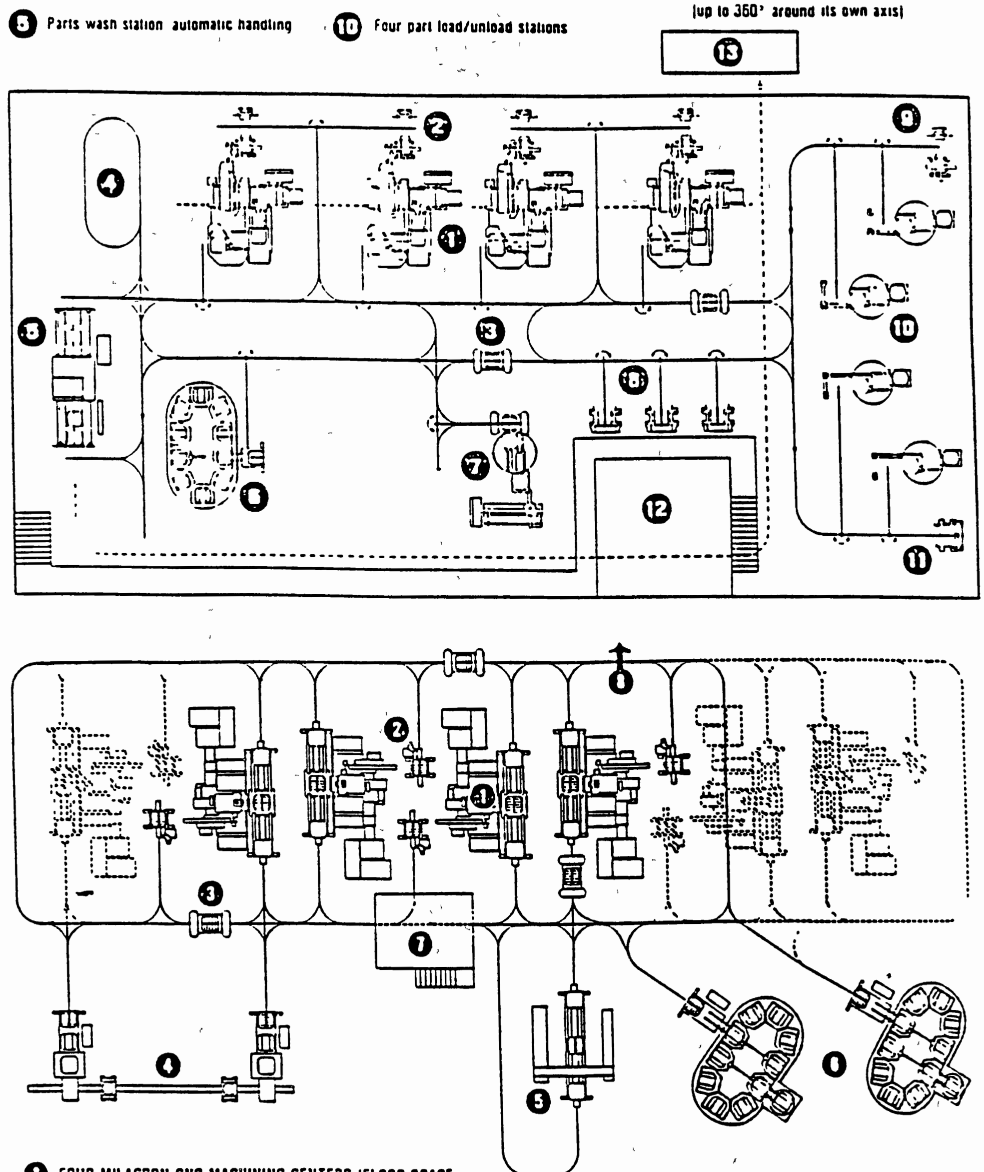
1. CNC machine tools;

2. An MHS to move parts and sometimes tools among machines and fixturing stations (so machines are linked by the MHS);

3. An overall computer control network that coordinates the

1. Four Milacron T 30 CNC Machining Centers
2. Four tool interchange stations one per machine for tool storage chain delivery via computer controlled cart
3. Three computer controlled carts with wire guided path
4. Cart maintenance station
5. Parts wash station automatic handling
6. Automatic Workchanger (10 pallets) for online pallet queue
7. One inspection module — horizontal type coordinate measuring machine
8. Three queue stations for tool delivery chains
9. Tool delivery chain load/unload station
10. Four part load/unload stations
11. Pallet/fixture build station
12. Control center computer room (elevated)
13. Centralized chip/coolant collection/recovery system (----- flume path)

Cart turnaround station (up to 360° around its own axis)



1. FOUR MILACRON CNC MACHINING CENTERS (FLOOR SPACE RESERVED FOR ADDING THREE MORE)
2. FOUR TOOL INTERCHANGE SYSTEMS (ONE PER MACHINE). COMPUTER CONTROLLED TOOL DELIVERY VIA CART
3. THREE REMOTELY CONTROLLED CARTS WITH WIRE GUIDED PATH
4. TWO LOAD/UNLOAD CLEAN/ORIENT STATIONS WITH COOLANT/CHIP HANDLING
5. ONE INSPECTION MODULE (COORDINATE MEASURING MACHINE)
6. TWO AUTOMATIC WORKCHANGERS (10 PALLETS EACH) FOR PART OVERFLOW AND QUEUE
7. RAISED OFFICE (CART PATH UNDER)
8. CART MAINTENANCE STATION

Source: D. L. Goetsch, Fundamentals of CIM Technology, (1988), Delmar Publisher Inc.

Figure 1. Two Examples of FMS's

machine tools, the parts-moving elements, and the workpieces.

4. A storage system that is needed to store raw materials, work-in-process (WIP), and finished products.

A storage system is necessary for the smooth operation of an FMS. Different machines in an FMS have different production and utilization rate. Therefore, when one machine finishes processing a job, the subsequent machine for the job may not be free to process the job immediately. If there is no available space to which the job can be stored, the first machine will be "blocked;" if the job can be moved away, but there is no job waiting for the first machine, the machine will be "starved." Machine breakdowns are extreme cases of short-run imbalance of production rates. If a machine breaks down and the system has no storage capacity, other machines that process jobs either going to or coming from the failed machine may be forced down. Therefore, storage spaces are needed as buffers to decouple the machines. Generally, the local buffers before each machine have limited capacities, whereas a central storage can be much larger, and its capacity can often be assumed to be infinite. However, significant "transit times" (MHS delay) are often incurred in moving WIP among the stations and the central storage. To reduce these transit times, many FMS's employ a mixed storage system consisting of a central storage and some local buffers.

FMS Models

An FMS is usually modeled as a queueing network in which the customers are the jobs to be processed by the system and the servers are the machines. Each CNC machine has a local control computer linked by a communication network.

The computer control network in the FMS is in fact a computer communication network, which can also be modeled as a queueing network, in which the customers are packets of data, and servers are local control computers and the central control computer.

Closed Queueing Network Models

An FMS can be modeled as a closed queueing network (CQN, Solberg 1977) following the approach of Gordon and Newell (1967), and Posner and Bernholtz (1968). The model is schematically depicted in Figure 2 (Co and Li, 1989). The basic assumptions of this model are as follows (Buzacott and Yao 1986):

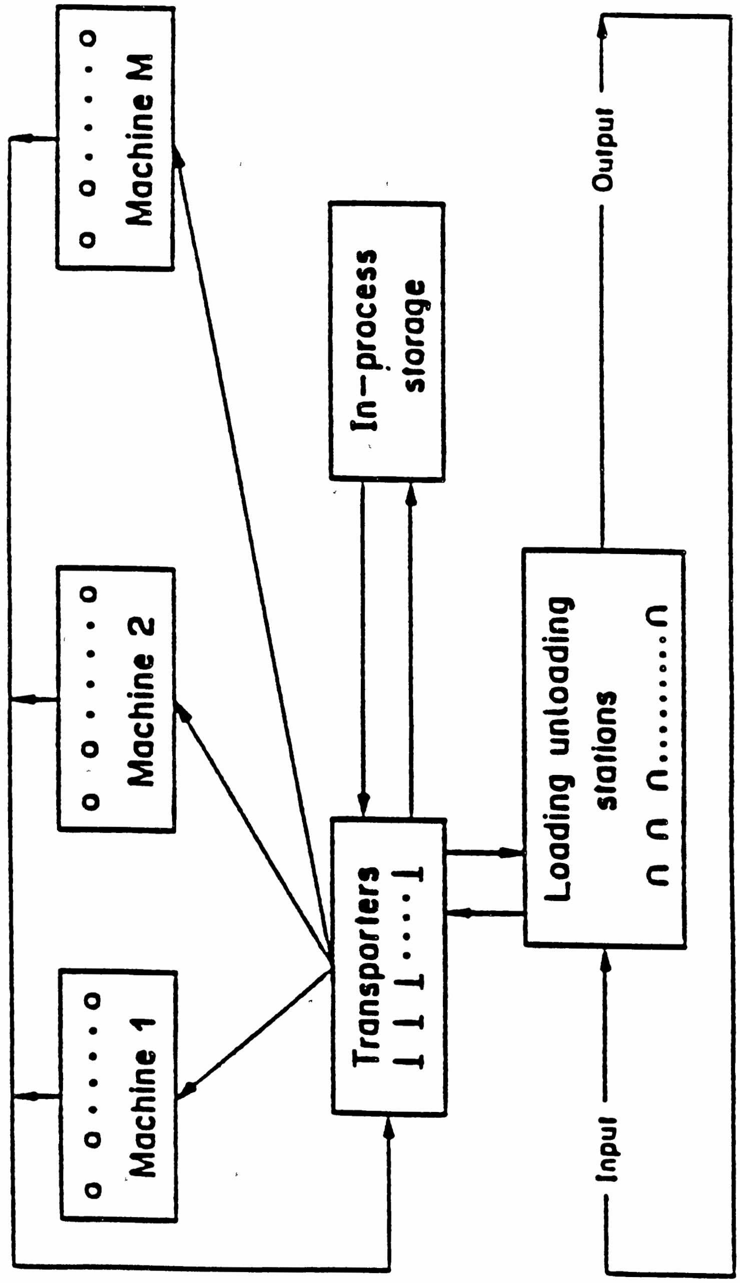
(1) The total number of jobs in the system is a fixed constant N, that implies that when a finished job leaves the system, a new job enters the system immediately (as if that leaving job re-enters the system).

(2) All stations use the FCFS queue discipline and have exponential service time distributions. All job classes have the same service rate at a station.

(3) All stations have a local storage large enough to accommodate all N jobs in the system; i.e., stations will never be blocked.

(4) Machines are always available for processing jobs, i.e., no breakdown, and any set-up/tool changing time is included in the service (processing) times.

(5) No MHS delay.



Source: H. C. Co, and G. Li, "A mean value analysis model for job shops and job shop-like systems," Computers ind. Engng, Vol.16, No.1, (1989), 9-18.

Figure 2. Conceptual Illustration of the Closed Network Model

The CQN model has been successfully applied to study various planning and control aspects of FMS's (Stecke and Solberg 1985, Stecke and Morin 1985, Yao and Buzacott 1986, Shanthikumar and Stecke 1986, and Dallery and Stecke 1990, among others). However, the model's five basic assumptions seem to be very restrictive, and in many cases one or more of them are violated.

Open Queueing Network Models

An FMS can also be modeled as an open queueing network, in which the total number of jobs varies through out the operation (Jackson 1963, Buzacott and Shanthikumar 1980), implying that the first assumption in the last section is dropped. This model is based on the works of Jackson (1963), Schweitzer (1977), and Baskett et al. (1975). This model is discussed in more detail in chapters III and IV.

Although tremendous effort has been made to relax the second and third assumptions with some success (Marie 1979, Shanthikumar and Buzacott 1980 and 1981, Whitt 1982, Altiok 1985, Yao and Buzacott 1985a and 1985b, among others), the fourth and fifth assumptions are retained, i.e., there will be no breakdowns and MHS delay in the FMS's.

Related Research on Machine Breakdowns

Machine breakdowns have been investigated by many authors in the transfer line area. A transfer line (Figure 3) is a number of automatic machines, in series, integrated into one system by a common transfer mechanism and a common control system (Buzacott 1967). Transfer lines can be considered a "simplified" FMS.

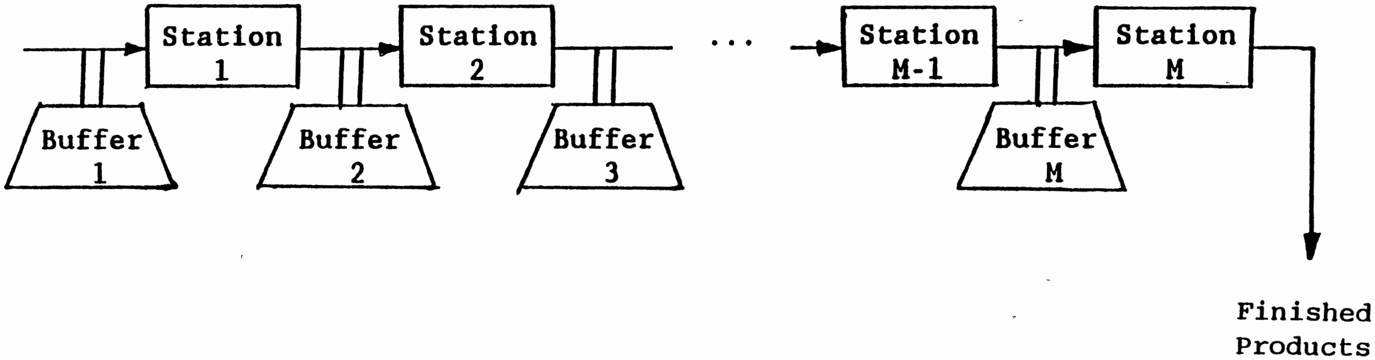


Figure 3. A Transfer Line with Buffers

Buzacott (1967, 1972), and Barlor and Proschan (1975) derived several basic formulas for the line's output rate R. They assume each station has one machine, and use discrete time units (cycles). They also assume all operation times (processing times, times between breakdowns, and times to repair) are geometrically distributed, which allows them to apply Markov process methods to analyze the two-station case and obtain the exact solution for R and other system performance measures. Due to the computational difficulties of their methods, many authors have, in turn, proposed reasonably good approximations. Among them are Buzacott (1967), Ingnall and Silver (1977), who also considered the case that each station has two or more of the same kind of machines, and the two stations have different processing rates, and Wijngaard (1979), who considered two single-machine station lines with different processing rates while assuming all operation times are exponentially distributed. However, when the number of stations is greater than two, these approaches are not feasible. Therefore, many other authors have proposed and tested other approximations (Murphy 1975, Sheshkin 1976, Gershwin and Berman 1981, Jafari and Shanthikumar 1987a, and Liu and Buzacott 1990).

Notably, all the above research assumes that operation times are either exponentially (for continuous times) or geometrically (for discrete times) distributed. This assumption is often too restrictive, since in practice the processing and repair times could follow any distribution.

Lau and Martin (1987) have investigated how the processing time's distribution forms affect system performance. They found that the stations' steady-state output rates are sensitive to the third and fourth (central) moments of processing times, which suggests that the ordinary one- or two-parameter distributions are not satisfactory to describe the real random processing times. Hence, they suggest the use of four-parameter distribution families to describe the processing times, which are more accurate and reliable.

Related Research on MHS Delay

p.116 b

In the CQN model, one can define a special station (say, station 0) to model the MHS. Assuming all WIP stays in station 0 before and/ or after processing at each normal station, then MHS delay in essence becomes the station 0's processing time. Since the CQN model assumes the total number of jobs in the system is a fixed constant N, and all stations have a local storage large enough to accommodate all N jobs in the system, one can use station 0 to replace all the local buffers. Posner and Bernholtz (1968) solved this system and obtained the steady-state solution (the distribution of the number of customers at each station).

Unfortunately, this method does not work in the open queueing network model, because local buffer capacities are usually limited, while the total number of jobs in the system is various.

Recently, MHS delay in transfer lines was considered by Commault and Semery (1990). They investigated to what extent this delay influences the line's performance and, in particular, the output rate R. They show in a two machine example how R is affected when this delay parameter varies, and then propose to define an "equivalent line," with unchanged machine characteristics, no MHS delay, but reduced local buffer capacity (Figure 8).

The main advantage of this approach is that, after this equivalent transformation, performance may be evaluated using an existing method. Their simulation results show that this capacity reduction technique provides very good approximations for the output rate, the mean buffer level and the mean time in system. But they only verify it for a two machine line, and, as the authors put it, "the really appealing problem is how to deal with longer lines." Another shortcoming of this model is that it cannot handle systems with central storage.

Related Research on Systems Without Breakdowns and MHS Delay

On the other hand, there are many well established methods to analyze the transfer lines with general processing times, but without breakdowns (i.e., no need of repair; Altiok and Ranjan 1987, Brandwajn and Jow 1988, Gershwin 1987, Jun and Perros 1990, etc.). For FMS's, several authors have considered systems with general processing times, but no breakdown and MHS delay. Many approaches for FMS's are suggested by Shanthikumar and Buzacott (1981), Whitt (1982), Yao and Buzacott (1985a, 1985b, 1986), among others.

For the open queueing network model, when processing times are exponentially distributed and local buffers' capacity is infinite, Jackson (1963) showed that the stations can be decomposed and analyzed separately, and the joint probability distribution of queue lengths can be expressed as a product form.

Shanthikumar and Buzacott (1981) pointed out that when processing times are not exponential, Jackson decomposition is not exact; they developed an approximate decomposition approach to analyze FMS's with non-exponential processing times.

Recent Efforts, Conclusions and Implications §1.7 6

It seems that there is a gap between the groups of research in the last three sections. The gap is the consideration of MHS delay and machine breakdowns. Attempts have been made to fill the gap.

Federgruen and Green (1986) have studied queues with service interruptions. The service times and service interruptions can be properly interpreted as times between breakdowns (on-time, or up-time, periods) and repair times (off-time, or down-time, periods), respectively. Their model has been generalized by Sengupta (1990), who has considered the situations in which both the arrival and service rates of the customers who arrive during the up-time periods could be different from that they would be during the down-time periods. Unfortunately, their models are computationally impractical.

Therefore, Federgruen and Green (1986) have suggested several approximations of completion times, waiting times, and the number of customers in the system. Their approximations are accurate only if the expected up- and down-times are short compared to the expected processing times. Sengupta (1990) has provided an approximation of waiting times that works well only if the processing times are very long. However, in manufacturing environments, the expected processing time is usually very <u>short</u> compared to the down-time (repair time) and, especially, the up-time (time between breakdowns). Another problem is that their methods don't provide vehicles to aggregate a group of queues; therefore, their methods are not suitable for analyzing transfer lines and FMS's, which consist of a group of stations, that are subject to starvation and blocking.

Jafari and Shanthikumar (1987b), and Yeralan and Muth (1987) have considered a two-station transfer line with general up-time and down-time distributions. The former assumes the processing times are constant, while the latter assumes the processing times could be stochastic. However, they fail to show how to extend their models to deal with longer lines or larger systems.

Generally speaking, any machine will break down sooner or later. Machine breakdowns interrupt smooth production, dramatically affecting the systems' performance. Therefore, by no means can breakdowns be ignored. For MHS delay, as Posner and Bernholtz (1968) pointed, "The assumption ..., namely that a unit takes zero time to move from one station to the next, is in general incorrect, and in many instances may be a poor approximation to the real situation."

Thus, developing new models to cover station breakdowns and MHS delay, and new methodologies to handle breakdowns, MHS delay, and inevitable general processing times is highly desired. This dissertation is a formal attempt to solve these problems.

CHAPTER III

BASIC ASSUMPTIONS AND DEFINITIONS

Assumptions and Definitions Pertaining to Stations

In this study, it is assumed that each station has one machine and a queue before the machine (Figure 6). For a particular station, its processing time S is a random variable. Let $m(\cdot)$, $M(\cdot)$, μ_S, and $\mu_n(S)$ be S's probability density function (pdf), cumulative density function (CDF), mean, and n-th central moment for $n>1$, respectively. When the station breaks down, it can be repaired in R time units. R is also a random variable with mean μ_R, and n-th central moment $\mu_n(R)$ for $n>1$. S and R are independent of each other.

When a station ceases to run due to an inability within itself, one says it "breaks down." When the station finishes processing a unit, if there is no available space to which the unit can be moved, the station will be blocked; if the unit can be moved away, but there is no unit waiting in the queue of the station, the station will be starved. A station (machine) is idle if it is not working. This could occur when the station is under repair, blocked, starved, or simply shut down. The station is busy if it is not idle. It is assumed repairmen are always available whenever breakdowns occur.

The time between breakdowns, T, is also a random variable. T, as well as basic assumptions pertaining MHS delay in chapter V, will be discussed in the next section, because different storage systems will impose special problems. Only after that can one define equivalency of stations and systems (also in chapter V).

This study makes three assumptions about breakdowns.

<u>Assumption 1: No-Failure-While-Idle</u>

When stations are idle, no breakdown occurs. This is a logical assumption. In some rare cases, machines do break down while idle, but usually people are not aware of it until they are going to work on the next job. Many authors adopt some similar assumption (Ignall and Silver 1977, Buzacott and Hanifin 1978, among others).

<u>Assumption 2: Exponential-Up-Time</u>

The time between breakdowns, T, is a r.v. with pdf $h(t) = \omega e^{-\omega t}$. Suppose a breakdown occurs at the clock time t_1, and the next breakdown occurs at the clock time t_2. Then the up-time between these two breakdowns is $(t_2 - t_1)$ - (all idle time periods during $t_2 - t_1$).

While this exponential distribution assumption seems restrictive, it is used frequently in the literature (see references in Buzacott and Hanifin 1978, Gershwin and Berman 1981, and Liu and Buzacott, among others). Additionally, it is a recent phenomenon that the frequency of wear-out caused breakdowns is decreasing steadily, so that other causes of breakdowns, such as mis-operation, electronic parts/circuits failure (different from mechanical wear-out), and communication network jam (Ghosh and Wysk 1989), are becoming relatively more significant. This makes the exponential distribution assumption more plausible. Later, one will see that this assumption is critical to analyze the system with breakdowns to obtain analytical results.

Assumption 3: No-Double-Fault

During one job's processing time, at most one breakdown may occur. It is easy to justify this assumption, because usually during a relatively short time period, a machine should not break down twice or more. If the machine does break down frequently, the firm should replace it, or enforce some kind of preventive maintenance measure. Another way to state this assumption is that when a breakdown occurs, the station can always complete the job on hand after repair, and the total processing time is the same (any extra set-up time is thought to be a part of the repair time). Many authors simply assume that processing times are deterministic, and breakdowns may only occur right after completing a job (see the references in Buzacott and Hanifin 1978, Gershwin and Berman 1981, among others). To reconcile these assumptions with those used in this paper, consider that the repair time will be charged to the last job before breakdown.

Some Basic Probability and Statistics Concepts and Definitions

In this study, theories of probability and statistics play a very important role. Hence, all necessary and basic probability and statistics concepts and definitions are presented in this section.

The mean of a r.v. X with pdf f(.) is

$$E[X] = \mu_X = \mu_1(X) = \int_{-\infty}^{\infty} xf(x)dx.$$

For $n>1$, X's n-th central moment is

$$\mu_n(X) = E[(X-E[X])^n] = \int_{-\infty}^{\infty} (x - E[X])^n f(x)dx,$$

and X's n-th moment is

$$\nu_n(X) = \int_{-\infty}^{\infty} x^n f(x)dx, \quad \text{for } n \geq 0.$$

Specifically, $\nu_o(X) = \mu_o(X) = 1$, $\mu_1(X) = 0$, and $\nu_1(X) = \mu_X$. Usually, X's variance is $\sigma_X{}^2 = \mu_2(X)$, and X's coefficient of variation is $CV(X) = \sigma_X/\mu_X$. X's skewness measure is $\alpha_1 = \mu_3(X)/\sigma_X{}^3$, and X's kurtosis measure is $\alpha_2 = \mu_4(X)/\sigma_X{}^4$.

For simplicity, ν_n and μ_n are used to denote $\nu_n(X)$ and $\mu_n(X)$, respectively, whenever the r.v. X is understood. $\nu_k(X)$'s ($k\geq1$) are called X's first k moments, and μ_X and $\mu_k(X)$ ($k\geq2$) are called X's first k central moments. μ_n's can be expressed in terms of ν_n's: (Wilks 1962, Kendall and Stuart 1969)

$$\left.\begin{aligned} \mu_2 &= \nu_2 - \nu_1{}^2, \\ \mu_3 &= \nu_3 - 3\nu_2\nu_1 + 2\nu_1{}^3, \\ \mu_4 &= \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1{}^2 - 3\nu_1{}^4, \end{aligned}\right\} \quad (3.1)$$

Generally, $\mu_k = \sum_{i=0}^{k} (C_k^i)\nu_{k-i}(-\nu_1)^i$, where $C_k^i = k!/[(k - i)!i!]$.

ν_n's can also be expressed in terms of μ_n's and μ_X:

$$\left.\begin{aligned} \nu_2 &= \mu_2 + \mu_X{}^2, \\ \nu_3 &= \mu_3 + 3\mu_2\mu_X + \mu_X{}^3, \\ \nu_4 &= \mu_4 + 4\mu_3\mu_X + 6\mu_2\mu_X{}^2 + \mu_X{}^4, \end{aligned}\right\} \quad (3.2)$$

Generally, $\nu_k = \sum_{i=0}^{k} (C_k^i)\mu_{k-i}\mu_X{}^i$.

In this study, Kendall and Stuart (1969)'s notations for probability terms will be followed. For instance, Prob[E] means the probability that an event E occurs.

Chapter VIII will introduce necessary queueing theory notations and related concepts.

Assumptions and Definitions Pertaining to the Entire FMS

The Basic System Model

The basic system model is an open queueing network, consisting of M stations and an initial queue (Figure 4). Jobs or customers waiting to enter the system are kept in the initial queue. Each station consists of a server (or a machine) and a queue before it (Figure 5). This paper will consider "customer" and "job", as well as "server" and "machine," interchangeable terms throughout. The general assumptions are (Schweitzer 1977, Buzacott and Shanthikumar 1980) :

1. Each machine can only process one job at a time.

2. Jobs arrive at the system and join the initial queue according to a Poisson process with parameter λ.

3. There are B classes of jobs. The probability that an arriving job A is class r and has its first operation on machine i is q_{ir}, i.e., q_{ir} = Prob(A $\in$ class r AND A's first operation is on i).

4. FCFS (first-come-first-served) is assumed for all machines.

5. Within each class the routing of jobs is determined by the transition matrix $\mathbf{P}_r = (P_{ij,r})$ for class r, where $P_{ij,r}$ is the transition probability that a class-r job (i.e., the job class is known) will go to station j after processing at station i. When i=j, $P_{ij,r}=0$.

Let $f_{ir} = 1 - \sum_{j=1}^{M} P_{ij,r}$ be the probability that a class-r job

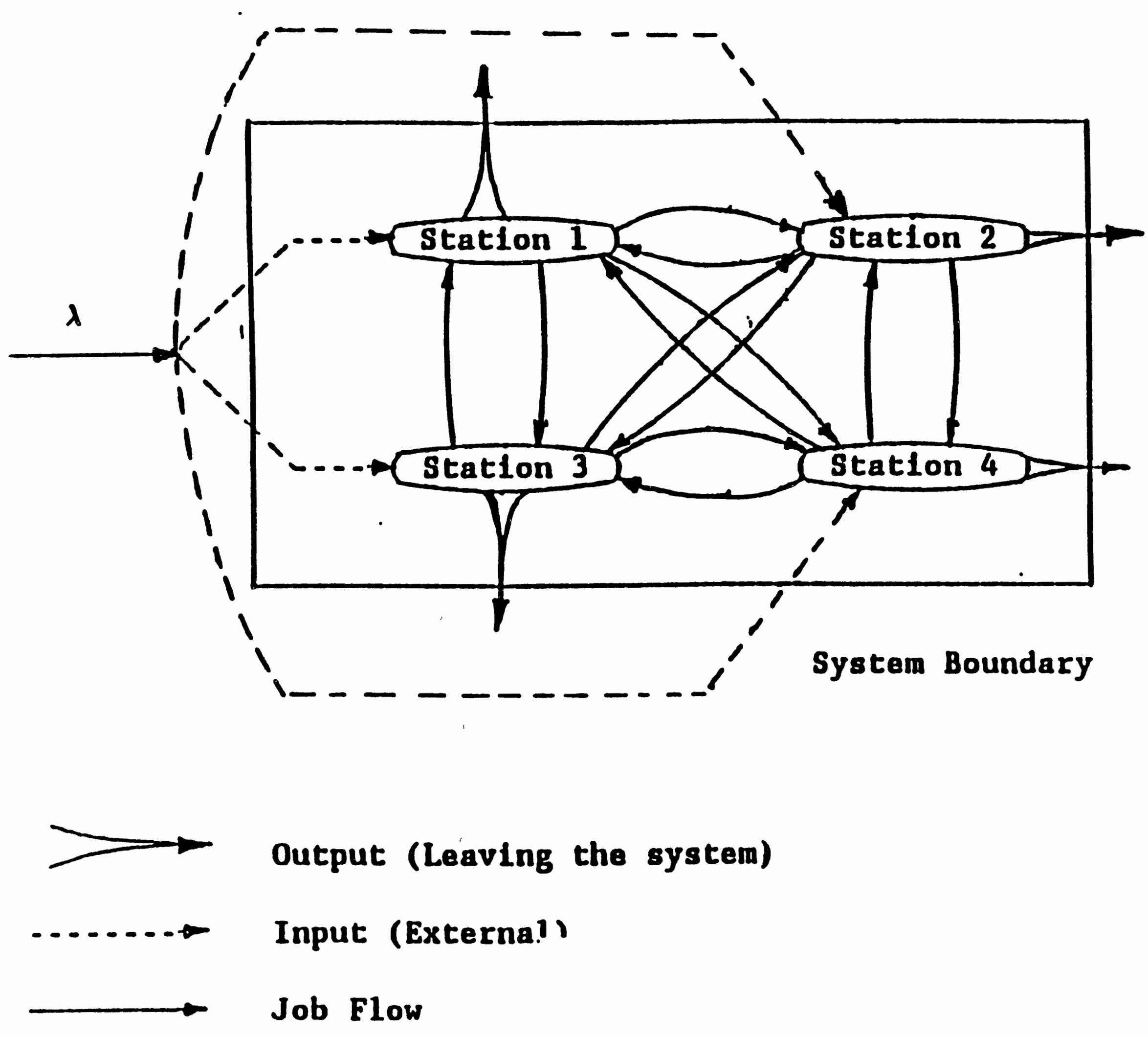
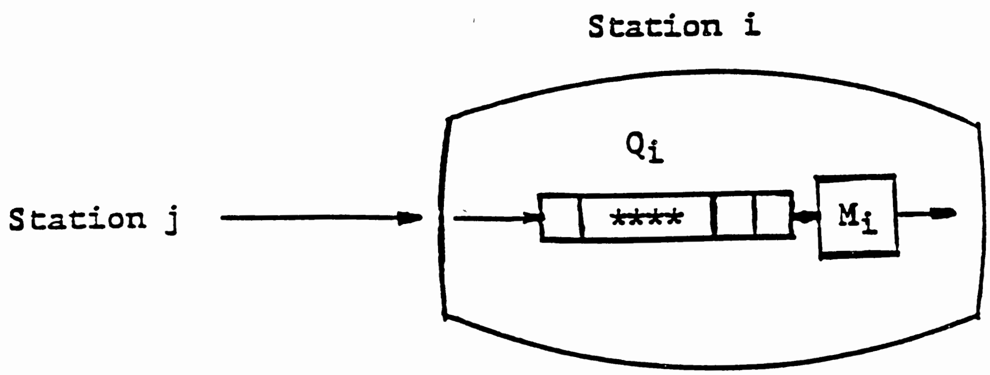


Figure 4. Diagram of an FMS



Q --- Queue M --- Machine

n_i --- Q_i's capacity

S_i --- M_i's processing time with pdf m(.)

R_i --- M_i's repair time

k_i --- mean time between M_i's breakdowns

Figure 5. Diagram of Station i

leaves the system after processing at machine j. When there is only one class of jobs ($B = 1$), the subscript r could be dropped.

6. The processing times are independent of job class. In this research, we will consider the processing times to be general, i.e., they could follow any distribution, and $1/\delta_i$ is the mean processing time at machine i (i.e., machine i's processing rate is δ_i).

7. At station i there is buffer space n_i, <u>including the space for the job on the machine</u> ($n_i>1$).

8. In the FMS, there is a central storage space for N_0 jobs. Usually N_0 is very large such that it can be treated as infinite.

<u>The Expected Number of Times a Job Visits Machine i</u>

It is desired to know an important parameter, e_{ir}, the expected number of times an arriving class-r job A will visit machine i.

Let q_i^r be Prob(A goes to i first | A $\in$ class r), q_{ir} be Prob(A $\in$ class r AND A goes to i first), and h^r be Prob(A $\in$ class r). Apparently, $q_{ir} = h^r * q_i^r$ $(i = 1, \ldots, M)$.

Let e_i^r be the expected number of times that A will visit the machine i, given A $\in$ class r, and e_i be the expected number of times that A will visit i. Then

$$e_i = e_i^1 * h^1 + e_i^2 * h^2 + \ldots + e_i^B * h^B$$

$$= \sum_{r=1}^{B} e_i^r * h^r. \qquad (3.3)$$

Note that $e_{ir} = e_i^r * h^r$, (3.3) can be rewritten as $e_i = \sum_{r=1}^{B} e_{ir}$.

Now notice that $e_i^r = q_i^r + \sum_{r=1}^{B} P_{ij,r} * e_j^r$ $(i = 1, \ldots M)$. $\qquad (3.4)$

Using matrix notations, one has

$\mathbf{e}^r = \mathbf{q}_i{}^r + \mathbf{P}_r{}^T\mathbf{e}^r$, where $\mathbf{e}^r = (e_1{}^r, e_2{}^r, \ldots, e_M{}^r)T$, and $\mathbf{q}^r = (q_1{}^r, q_2{}^r, \ldots, q_M{}^r)^T$. Or, $\mathbf{e}^r = (\mathbf{I} - \mathbf{P}_r{}^T)^{-1}\mathbf{q}^r$, where $\mathbf{I}$ is the MxM identity matrix.

Because of (2.4), $h^r{*}e_i{}^r = h^r{*}q_i{}^r + \sum_{j=1}^{M} P_{ij,r}(h^r{*}e_j{}^r)$, therefore

$$e_{ir} = q_{ir} + \sum_{j=1}^{M} P_{ji,r}{*}e_{jr} \quad (i = 1, \ldots, M), \tag{3.5}$$

or $\mathbf{E}_r = \mathbf{Q}_r + \mathbf{P}_r{}^T\mathbf{E}_r$, where $\mathbf{E}_r = (e_{1r}, \ldots, e_{Mr})^T$, and $\mathbf{Q}_r = (q_{1r}, \ldots, q_{Mr})^T$. Again,

$$\mathbf{E}_r = (\mathbf{I} - \mathbf{P}_r{}^T)^{-1}\mathbf{Q}_r. \tag{3.6}$$

Then one can obtain e_{ir} by solving the simultaneous equations (3.5) or (3.6). Here it is assumed that a unique solution for the e_{ir}'s exist, and all $e_{ir} \geq 0$. e_{ir} can be interpreted as the arrival rate of class r jobs to station i. All q_{ir}'s are non-negative. If at least one $q_{ir} >$ 0, the network is open, and

$$\sum_{r=1}^{M} \sum_{i=1}^{M} q_{ir} = 1.$$

If all $q_{ir} = 0$, it is a closed network, and $\sum_{i=1}^{M} e_{ir} = N_r$, where N_r is the number of class-r jobs in the system. Of course $\sum_{r=1}^{B} N_r = N$. Then one should solve following equations to get e_{ir}'s:

$$\sum_{i=1}^{M} e_{ir} = N_r, \text{ and } (\mathbf{I} - \mathbf{P}_r{}^T)\mathbf{E}_r = \mathbf{0}.$$

<u>System States and Steady State</u>

Suppose at a specific time t there are v_i jobs in station i

($1 \le i \le M$). Then it is said the system is at state $\underline{v} = (v_1, \ldots v_i \ldots v_M)$. $\underline{v}$ is called the system's state vector, because it describes the system's current state.

In the last subsection it is assumed that,

(i) A unique non-negative solution for the e_{ir}'s exist;

(ii) All transition probabilities $P_{ij,r}$ are constant, regardless of which successive time periods are considered (i.e., any $P_{ij,r}$ does not change over time).

For performance evaluation purpose, only the systems' behaviors at steady state or equilibrium will be considered. It is said that the system is at steady state if at that time the system already reaches the state such that (i) and (ii) hold asymptotically.

Blocking Rate

Because the capacity of local buffers are limited, it is possible that when a job A arrives at station i, i's queue is full and A is not allowed to enter i. The blocking rate b_i can be defined as the probability that a coming job will be rejected by the station i. When the station is in the steady state, its output rate is equal to its input rate, and both input and output rates are denoted by λ_i.

In chapter IX more terminology pertaining to the queueing networks will be introduced.

CHAPTER IV

PERFORMANCE EVALUATION OF QUEUEING NETWORKS WITHOUT MHS DELAY AND BREAKDOWNS

Some major objectives of FMS research and development are to maximize the FMS's production capacity or output rate, to increase machine utilization, and to reduce WIP inventory. In other words, production capacity (hereafter PC), output rate (or throughput rate, hereafter TH), machine utilization, and WIP inventory level are important measures in evaluating the FMS's performance. These measures can be used to compare alternative designs, and hopefully, to obtain optimal system configuration. Among these measures, the simplest is PC, which is the maximal possible output rate.

System Output Rate When Processing Times Are Exponential

Now assume no MHS delay and station breakdowns, and the input rate for the system is λ. Schweitzer (1977) shows that, if all processing times are exponentially distributed, and at most C customers are allowed to be in the system at any time, then the FMS's output rate TH is a function depending on C and λ:

$$TH(\lambda, C) = \lambda\left[\sum_{m=1}^{C-1} \lambda^m q(m)\right] / \left[\sum_{m=1}^{C} \lambda^m q(m)\right] \tag{4.1}$$

where $q(K) = \sum_{n_1+n_2+\ldots+n_M=K} \left[\prod_{i=1}^{M} a_i^{n_i}\right]$, $0 \le K \le C$, and $a_i = \sum_{r=1}^{B} e_{ir}/\delta_i$.

Note that here the summation is over all possible combinations of $(n_1, \ldots, n_i, \ldots, n_M)$ such that $n_1+n_2+\ldots+n_M=K$, and for each K, q(K) is independent of λ and C.

Let A(C) be $\sum_{M=1}^{C} \lambda^m q(m)$. Because $A(C) \sim \lambda^C q(C)$ for large λ, (4.1) indicates

$$\lim_{\lambda\to\infty} TH(\lambda, C) = q(C-1)/q(C). \tag{4.2}$$

When B=1, i.e. only one class of customers, $a_i = e_i/\delta_i$. Let $a_{max} = \max\{a_i\}$, Schweitzer (1977) shows $PC = \lim_{C\to\infty} q(C-1)/q(C) = 1/a_{max}$, and

$$\left.\begin{aligned} \lim_{C\to\infty} PC(\lambda, C) = PC(\lambda, \infty) &= \lambda && \text{if } \lambda < 1/a_{max} \\ &= 1/a_{max} && \text{if } \lambda \geq 1/a_{max}. \end{aligned}\right\} \tag{4.3}$$

Here $\lambda\to\infty$ is equivalent to the situation where there are infinite customers staying in the initial queue, and $C\to\infty$ means there may be infinite customers in the system. Hence (4.3) gives the maximum possible output rate of the system (Buzacott and Shanthikumar 1980). However, when λ and/or C are finite, maybe small, there is no easy way to calculate the throughput rate.

If all stations have a local buffer with infinite capacity, Jackson (1963) shows the system can be decomposed and the equilibrium joint probability distribution of queue lengths can be expressed as a product form as follows.

Recall that $\underline{v} = (v_1, \ldots, v_i, \ldots, v_M)$ is the system's state vector, where v_i $(1 \leq i \leq M)$ is the number of jobs in station i.
Let $w(\underline{v}) = \prod_{m=1}^{M} [e_m/\delta_m]^{v_m}$. For any given $\underline{v}$,

$T(K) = \Sigma\, w(\underline{v})$ summed over all $\underline{v}$'s with $S(\underline{v}) = [\sum_{i=1}^{M} v_i] = K$, and $$\pi = 1/\sum_{K=0}^{\infty} \{[\prod_{i=0}^{K-1} \lambda_i]\ T(K)\}.$$

Then the unique equilibrium state probability distribution exists, and is given by

$$p(\underline{v}) = \pi w(\underline{v})[\prod_{i=0}^{S(\underline{v})-1} \lambda_i]. \tag{4.4}$$

Note here $p(\underline{v})$ is the probability that the system state is $\underline{v}$. It is easy to see that, for a particular station i, its station size L_i's distribution is

$$\text{Prob}[L_i=j] = \sum_{\forall(\underline{v}|v_i=j)} p(\underline{v}). \tag{4.5}$$

Since the central storage's capacity is infinite and the transit time (MHS delay) is zero, that is equivalent to that every local buffer's capacity is infinite, so Jackson's model can be applied.

System Output Rate When Processing Times Are General

(4.1) to (4.3) require that the processing times must be exponential, and the storage capacity is infinite. This may be not realistic.

As Yao and Buzacott (1985b) pointed out, most real systems have carefully designed procedures to ensure no blocking. One example is the Caterpillar Omniline's "deliver-and-pick-up" scheme (Hutchinson 1979); another example is Toyota's "return conveyors" (Hatvany 1983), which continuously takes away finished jobs from machines.

Here it can be shown that, when the processing times follow any distribution, as long as there is no blocking, the FMS's output rate is still λ. Appendix A provides the derivation. This leads to the following proposition.

PROPOSITION 1.

For an FMS with general processing times and no blocking, IF $\lambda < 1/a_{max}$, (or $\lambda_i < 1$ for $i = 1, \ldots, M$) the throughput rate is λ. ##

If blocking does occur (due to limited storage capacity, for example), several authors (Gelenbe 1975, Yao and Buzacott 1985b, among others) provide iterative procedures to calculate blocking probabilities b_i for $i = 1, \ldots, M$. In chapter IX, I will develop an easier procedure to calculate the b_i's.

Other Measures

It is shown (Buzacott and Shanthikumar 1980) that U_i, the utilization of machine i, is given by

$$U_i = \lambda e_i/\delta_i; \tag{4.6}$$

and, if only a maximum of C jobs are allowed in the system, (4.1) suggests

$$PC = \lim_{\lambda\to\infty} TH(\lambda, C) = q(C-1)/q(C). \tag{4.7}$$

When all e_i/δ_i are the same, the system is called a balanced system with PC equal to $C/([a(C+M-1)]$, where $a=e_i/\delta_i$.

Although Buzacott and Shanthikumar (1980) assume all processing times are exponential, it will be shown in chapter IX that (4.6) is valid for any processing time distribution.

It is shown in chapter III that a job (say A) will visit station i on average e_i times. Let $E[W_i]$ be the mean waiting time in station i's queue, and $E[X_i]$ be the mean time in the station i. Each time A is expected to spend $E[W_i]$ of time in the queue i, and then to be served in $1/\delta_i$ of time. Apparently, (see, say, Ross 1989)

$$E[X_i] = 1/\delta_i + E[W_i]. \tag{4.8}$$

So the sojourn time SJ, i.e., the expected time a job stays in the system, is given by

$$SJ = \sum_{i=1}^{M} e_i E[X_i] = \sum_{i=1}^{M} e_i (1/\delta_i + E[W_i]). \tag{4.9}$$

Using Little's Law, one has

$$\left.\begin{aligned} E[W_i] &= E[Q_i]/\lambda_i, \text{ and} \\ E[X_i] &= E[L_i]/\lambda_i, \end{aligned}\right\} \tag{4.10}$$

where λ_i is station i's input rate, $E[Q_i]$ is the expected length of queue i, and $E[L_i]$ is the expected station size. Since $E[L_i]$ can be obtained from (4.4) and (4.5),

$$E[L_i] = \sum_{j=0}^{\infty} j\text{Prob}[L_i=j], \tag{4.11}$$

Now $E[X_i]$ can be derived from (4.10). Then, using (4.8), one can determine $E[W_i]$; and from (4.10) one can determine $E[W_i]$. Finally, it is straightforward to calculate SJ from (4.9).

The only difficulty is that (4.11) requires the summation of infinite terms. In practice, one has to assume $\text{Prob}[L_i \geq J]=0$ for some large integer J, so that (4.11) can be approximated by

$$E[L_i] = \sum_{j=0}^{J} j\text{Prob}[L_i=j]. \tag{4.12}$$

It should be pointed out that, the mean WIP inventory level of station i, in terms of units, denoted by WIP_i, is just $E[L_i]$; i.e.,

$$WIP_i = E[L_i] \quad \text{for } 1 \le i \le M.$$

If $B>1$, i.e., there are two or more classes of jobs in the system, jobs in different classes should have different values (or costs). Furthermore, the same job may have different values (or costs) when it is going through different stations. So (4.12) should be revised accordingly to reflect these considerations. Since this heavily depends on the cost structures of these products, which are beyond the scope of this dissertation, it will not be discussed here. However, this is a topic worthy of future research.

CHAPTER V

TECHNIQUES TO COPE WITH BREAKDOWNS

In the FMS literature, it is always assumed that machines never break down. This is an unrealistic assumption. This study will investigate machine breakdowns and present a powerful and convenient approach to deal with them. This approach transforms the line with breakdowns into an equivalent line without breakdowns, and then many well-established methodologies can be applied to evaluate the system's performance.

The Effective Processing Time Y and Its Mean E[Y]

Consider that, when a unit A arrives at a station, which is ready to work on A, it will take S time units to process A if no breakdown occurs during the S time units. If the station breaks down during processing A, because of the No-Double-Fault assumption, A will stay in the station for S + R time units. Formally, the effective processing time is

$$Y = \begin{cases} S, & \text{if no breakdown during } S; \\ S + R, & \text{if breakdown during } S. \end{cases} \tag{5.1}$$

According to the Exponential-Up-Time and No-Double-Fault assumptions, the probability that the machine breaks down during a given busy time duration t is

$$H(t) = \int_0^t h(x)dx$$

$$= \int_o^t \omega e^{-\omega x} dx$$

$$= 1 - e^{-\omega t},$$

and the probability of no breakdown during t is then

$$1 - H(t) = e^{-\omega t},$$

where the exponential distribution's memoryless property is exploited.

Now E[Y] can be calculated by conditioning (Ross 1989). Let E[X|W] be the function of the random variable W whose value at W=w is E[X|W=w]. Note that E[X|W] itself is a random variable. It is shown (Ross 1989) that, for all random variables X and W,

$$E[X] = E[E[X|W]]. \tag{5.2}$$

Therefore, E[Y|S] is a function of S, whose value at S = t is E[Y|S=t], and

$$\begin{aligned} E[Y|S=t] &= t[1 - H(t)] + (t + \mu_R)H(t) \\ &= t + \mu_R H(t). \end{aligned} \tag{5.3}$$

Note the probability that S = t, i.e., Prob[S=t], is m(t)dt; according to (5.2),

$$\begin{aligned} E[Y] &= \int_o^\infty E[Y|S=t]m(t)dt \\ &= \int_o^\infty [t + \mu_R H(t)]m(t)dt \\ &= \mu_S + \mu_R \int_o^\infty H(t)m(t)dt. \end{aligned} \tag{5.4}$$

For convenience, let k_i denote $\int_o^\infty t^i H(t)m(t)dt$ for $i \geq 0$. Then (5.4) can be rewritten as

$$E[Y] = \mu_S + \mu_R k_0. \tag{5.4'}$$

It is clear that, according to (5.4'), E[Y] depends only on μ_S, μ_R, and S's and T's distributions, and the repair time distribution is irrelevant here. Buzacott (1967) guessed that "certain results hold irrespective of the repair time distribution." Equation (5.4') proves that is right.

Calculation of Y's Moments and Central Moments

In the last section, it is known that $E[Y] = \mu_Y = \nu_1$. Now, consider the calculation of Y's higher moments ν_i and central moments μi for i = 2, 3, and 4. In fact, one can calculate ν_i and μ_i for all i>1. First, one calculates ν_i for i>1, and then, with the help of (3.1), one can easily obtain μ_i for i>1. Instead of computing each ν_i individually, it is appropriate to introduce the following theorem to handle ν_i's for all i>0.

THEOREM 1. Assume two r.v.'s X and W have a joint pdf f(x,w), and their marginal pdf's are $f_X(x)$ and $f_W(w)$, respectively. Then, for i>0,

$$E[X^i] = \int_{-\infty}^{\infty} E[X^i|W=w]f_W(w)dw. \tag{5.5}$$

Proof. See Appendix B. ##

It is known (Kendall and Stuart 1969, Ross 1989) that

$$E[X^i|W=w] = \int_{-\infty}^{\infty} x^i f_{X|W}(x|w)dx. \tag{5.6}$$

The beauty of Theorem 1 is that one does not need to know f(x,w) if (5.6) is available. Now it is ready to compute ν_i's.

$$\text{Let } \bar{Y}^i = \begin{cases} t^i, & \text{if no breakdown during t;} \\ (t + R)^i, & \text{if breakdown during t;} \end{cases}$$

$$\text{and } Q = \begin{cases} 0, & \text{if no breakdown during t;} \\ 1, & \text{if breakdown during t.} \end{cases}$$

It is easy to see $\bar{Y}^i = [Y^i|S=t] = t^i + (t + R)^iQ$, and

$$E[Y^i|S=t] = E[\bar{Y}^i] = t^i[1 - H(t)] + E[(t + R)^i]H(t) \tag{5.7}$$

Therefore, according to Theorem 1,

$$\begin{aligned}
\nu_i(Y) &= E[Y^i] \\
&= \int_0^\infty E[Y^i|S=t]m(t)dt \\
&= \int_0^\infty \{t^i[1 - H(t)] + E[(t + R)^i]H(t)\}m(t)dt \\
&= \nu_i(S) - k_i + \int_0^\infty \sum_{j=0}^{i} C_i^j \nu_j(R)t^{i-j}]H(t)m(t)dt \\
&= \nu_i(S) - k_i + \sum_{j=0}^{i} [C_i^j \nu_j(R)k_{i-j}] \\
&= \nu_i(S) + \sum_{j=1}^{i} (C_i^j)\nu_j(R)k_{i-j} \qquad (5.8)
\end{aligned}$$

Since k_j $(j<i)$ can be numerically calculated, and R's and S's moments are either given or easy to compute, $\nu_i(Y)$ can be obtained from (5.8).

As a summary, the effective processing time Y's first four moments and central moments are explicitly identified.

$$\left.\begin{aligned}
\nu_1(Y) &= E[Y] = \nu_1(S) + \nu_1(R)k_0 = \mu_S + \mu_R k_0. \\
\nu_2(Y) &= \nu_2(S) + 2\nu_1(R)k_1 + \nu_2(R)k_0; \\
\nu_3(Y) &= \nu_3(S) + 3\nu_1(R)k_2 + 3\nu_2(R)k_1 + \nu_3(R)k_0; \\
\nu_4(Y) &= \nu_4(S) + 4\nu_1(R)k_3 + 6\nu_2(R)k_2 + 4\nu_3(R)k_1 + \\
&\qquad \nu_4(R)k_0.
\end{aligned}\right\} \tag{5.9}$$

Now applying (3.1) one obtains $\mu_i(Y)$'s for i = 2, 3, and 4. Namely,

$$\left.\begin{aligned}
\mu_1(Y) &= \nu_1(Y), \\
\mu_2(Y) &= \nu_2(Y) - [\nu_1(Y)]^2, \\
\mu_3(Y) &= \nu_3(Y) - 3\nu_2(Y)\nu_1(Y) + 2[\nu_1(Y)]^3, \\
\mu_4(Y) &= \nu_4(Y) - 4\nu_3(Y)\nu_1(Y) + 6\nu_2(Y)[\nu_1(Y)]^2 - 3[\nu_1(Y)]^4.
\end{aligned}\right\} \tag{5.10}$$

Some Discussions

Notably it is assumed that machine breakdowns do not affect processing times. If breakdowns affect processing times, the equation (5.1) should be rewritten as

$$Y = \begin{cases} S, & \text{if no breakdown;} \\ S_1 + R + S_2, & \text{if breakdown occurs,} \end{cases} \qquad (5.1')$$

where S_1 is the processing time <u>before breakdown</u> and S_2 is the remaining processing time <u>after breakdown</u> (Residual Life Time, Figure 6). Essentially in this study it is assumed $S_1 + S_2 = S$, that means the interruption does not affect the processing time.

```
No breakdown:
                              S
     ---------------|-----|-------

With breakdown:
                             S1       R       S2
     ---------------|----|---------|--|-------
```

FIGURE 6. The effects of breakdown

Since it is unlikely that $S_1 + S_2 < S$, the only other possible case is $S_1 + S_2 > S$. While it is intractable when S_1 and S_2 are general, several important special cases can be handled accordingly.

Case A: $S_1 + S_2 = S + c^*$, where c^* is a random variable.

This will happen when a breakdown only causes some extra time to set up and restart. Let $S' = S + c^*$. If S and c^*'s first four moments are known, then the first four moments of S' can be calculated by using formulas (7.2) and (7.3) (see chapter VII).

Case B: $S_1 + S_2 = wS$, where w is a constant multiplier.

Since $E[wS] = wE[S]$, $Var[wS] = w^2Var[S]$, etc., the first four moments of wS can be obtained easily.

A Special Case: S and S_2 are identically exponential.

The No-Double-Fault assumption is useful to derive a series of results in this study. However, this assumption is apparently invalid when the processing time is exponential. In this case, when a breakdown occurs, the interrupted job will be reprocessed after repair. Due to the exponential distribution, the remained processing time is still S. In other words, the interrupted job will be treated as a new job, and its history is forgot. Federgruen and Green (1986) considered and solved this problem.

Summary

Here a powerful and convenient approach is presented to handle station breakdowns, stochastic processing times and repair times. By transforming the line with breakdowns into an equivalent line without breakdown, many well-established methodologies can be applied to evaluate the system's performance. This approach concentrates on the effective processing time Y. It can analytically calculate Y's moments and central moments, that facilitate fitting any n-parameter distribution for Y, and therefore provides a solid base for further analyses of the whole queueing network.

Usually one fits Y's first three or four moments to a three or four parameter distribution, respectively. To calculate Y's moments, it's necessary to known S and R's moments, and the following integral,

$$k_i = \int_0^\infty t^i H(t)m(t)dt$$

$$= \int_0^\infty t^i(1 - e^{-kt})m(t)dt,$$

which only depends on the mean breakdown rate k and m(.). In other words, one doesn't need to know the repair time R's and MHS delay D's pdf's! Moreover, if one doesn't know the processing time S's pdf, but know S's first three or four moments, he/she can fit a three or four parameter distribution as S's pdf m(.). So in practice, one only needs to know k and the first three or four moments of S, and R, and that are easy to obtain from collected data. This feature makes this method very attractive to researchers as well as practitioners, for they don't need to figure out those pdf's of S and R.

The numerical examples in chapter VI will clearly show that the processing time S's third and fourth (central) moments have significant impact on the effective processing time Y, that indicates a four-parameter distribution is more appropriate to be used to fit S's first four moments.

CHAPTER VI

NUMERICAL EXAMPLES

Here several numerical examples are given to show how to transform an FMS with breakdowns into an equivalent system with no breakdowns.

The program and results for example 1 are in Appendix C.

The Effects of Uncertainty and R's Distribution

Example 1.

Consider the following data.

Repair time R:

$\mu_R=5$, $\sigma_R{}^2=9$, $\mu_3(R)=100$, and $\mu_4(R)=525$. So $CV_R=0.6$.

Processing time S:

$\mu_G=1.5$, $\sigma_T{}^2=0.81$, $\mu_3(S)=0.8748$, and $\mu_4(S)=3.9366$. So $CV_S=0.6$.

Time between breakdowns T:

$\omega=1/100$, or the mean time between breakdowns is 100.

From (3.1),

$$\left.\begin{array}{l} \nu_1(R)=5,\ \nu_2(R)=34,\ \nu_3(R)=360,\ \text{and } \nu_4(R)=4500. \\ \nu_1(S)=1.5,\ \nu_2(S)=3.06,\ \nu_3(S)=7.8948,\ \nu_4(S)=25.1829. \end{array}\right\} \quad (6.1)$$

To calculate k_i's $(0\leq i\leq 3)$, one needs to determine m(t). Here one can fit $\nu_i(S)$'s $(1\leq i\leq 4)$ to an S-D distribution (Schmeiser and Deutsch 1977), which has four parameters, a, b, c, and d.

First, one can obtain α_1 and α_2 as follows.

$\alpha_1 = \mu_3(S)/\sigma_{GS}{}^3 = 0.8748/(0.81)^{1.5} = 1.2,$

$$\alpha_2 = \mu_4(S)/\sigma_{GS}{}^4 = 3.9366/(0.81)^2 = 6.$$

Then, one can look up the graph in Schmeiser and Deutsch (1977) to find that, for $(\alpha_1, \alpha_2) = (1.2, 6)$, c=4.63836, and d=0.47117. (An accurate table can be obtained from the authors.) Then b and a can be calculated as follows.

$$b = \sigma_S \left\{ \frac{(2c+1)(c+1)^2}{(c+1)^2[d^{2c+1}+(1-d)^{2c+1}]-(2c+1)[(1-d)^{c+1}]^2} \right\}^{1/2} = 45.1361$$

$$a = \mu_S - \left[\frac{(1-d)^{c+1} - d^{c+1}}{c+1} \right] b = 1.3945.$$

Note that

$$\begin{aligned} k_i &= \int_o^\infty t^i(1 - e^{-t/100})m(t)dt \\ &= \lim_{x\to\infty} \int_o^X t^i(1 - e^{-t/100})dM(t). \end{aligned} \tag{6.2}$$

Let BL be $a - bd^c$, and BR be $a + b(1 - d)^c$. The S-D distribution has a very good property: Its CDF is given as the following closed-form.

$$M(t) = \begin{cases} 0, & t < BL; \\ d - [(a-t)/b]^{1/c}, & BL \le t \le a; \\ d + [(t-a)/b]^{1/c}, & a \le t \le BR; \\ 1, & t > BR. \end{cases}$$

Hence, (6.2) can be rewritten as $k_i = \int_{BL}^{BR} t^i(1 - e^{-t/100})dM(t)$. Let $\delta = (BR-BL)/N$ for some big integer N, $\Delta M(j\delta) = M(j\delta) - M((j-1)\delta)$, and presume $BL \ge 0$. Then

$$k_i = \lim_{N\to\infty} \sum_{j=1}^{N} (j\delta)^i[1 - e^{-(j\delta/100)}]\Delta M(j\delta). \tag{6.3}$$

Equation (6.3) can be solved numerically, and the results are

$$k_0 = 0.05160,\ k_1 = 0.16336,\ k_2 = 0.56675,\ k_3 = 2.04813. \tag{6.4}$$

Now, from (5.9), (6.1), and (6.4), one knows

$\nu_1(Y) = 1.758$, $\nu_2(Y) = 6.448$, $\nu_3(Y) = 51.636$, $\nu_4(Y) = 649.211$.

Using (5.10), it is easy to see that $\mu_Y = \nu_1(Y) = 1.758$, $\sigma_Y^2 = 3.357$, and $CV_Y = 1.042$. By the way, $\mu_3(Y) = 28.495$ and $\mu_4(Y) = 377.021$.

Example 2.

Consider the following data.

Repair time R:

$\mu_R=5$, $\sigma_R^2=32$, $\mu_3(R)=45$, and $\mu_4(R)=875$. So $CV_R=1.13$.

Processing time S:

$\mu_S=1.5$, $\sigma_S^2=0.81$, $\mu_3(S)=0.8748$, and $\mu_4(S)=3.9366$. So $CV_S=0.6$.

Time between breakdowns T: $\omega=1/100$.

From (3.1),

$\nu_1(R)=5$, $\nu_2(R)=57$, $\nu_3(R)=650$, and $\nu_4(R)=7200$.

$\nu_1(S)=1.5$, $\nu_2(S)=3.06$, $\nu_3(S)=7.8948$, and $\nu_4(S)=25.1829$.

Similarly, one can know $\alpha_1 = 1.2$, $\alpha_2 = 6$, and all a, b, c, d, and k_i's are the same as those in example 1. Therefore, one obtains that $\nu_1(Y)=1.7580$, $\nu_2(Y)=7.6352$, $\nu_3(Y)=77.8790$, and $\nu_4(Y)=1056.3809$. Then $\mu_Y = 1.7580$ and $CV_Y = 1.2126$.

Discussions about the First Two Examples

In example 1, assume the time unit is an hour. If all times are deterministic, the station processes 66 jobs in 1.5*66=99 hours without interruption. When it is processing the 67-th job, the machine breaks down, and that will take 5 hours to repair it. In total it needs 99+1.5+5=105.5 hours to process these 67 jobs. So the mean effective processing time E[Y] is about 1.575 hours. Now in fact all times are stochastic, E[Y] is 1.758, about 11.62% greater than 1.575, that clearly shows the negative effect of uncertainty. It is not surprising, because equation (5.3) indicates that both T's and S's distributions affect E[Y].

The Effects of the Variance, Skewness, and Kurtosis of Processing Times

Example 3.

Assume $\mu_R=5$, $\sigma_R^2=9$, $\mu_3(R)=100$, and $\mu_4(R)=525$. So $CV_R=0.6$. For T, $\omega=1/100$. Now one may be interested in how σ_S^2, $\alpha_1(S)$, and $\alpha_2(S)$ affect the effective processing time Y. Because the skewness measure $\alpha_1(S)=\mu_3(S)/\sigma_S^3$, the kurtosis measure $\alpha_2(S)=\mu_4(S)/\sigma_S^4$, and $\sigma_S^2= \mu_2(S)$, the effects of S's second, third, and fourth central moments are being investigated. Let $\mu_S=1.5$. Note $CV_S=\sigma_S/\mu_S$ and is set at $CV_S=0.4$, 0.8 and 1.2. Table I shows the results.

In example 2, only R's second, third, and 4-th moments are changed. As equations (5.3) and (5.8) suggest, E[Y] is still 1.758, but CV_Y, $\nu_3(Y)$, and $\nu_4(Y)$ increase dramatically.

Here considering α_1 and α_2 jointly (denoted by $[\alpha_1, \alpha_2]$), Table I reveals that both μ_Y and CV_Y are monotonously, positively correlated with CV_S and $[\alpha_1(S), \alpha_2(S)]$. It seems that CV_S has stronger impact on Y than $[\alpha_1(S), \alpha_2(S)]$. So in section 3 it will be shown if $\alpha_1(S)$ and $\alpha_2(S)$, separately, have significant impact on Y.

TABLE I

THE JOINT EFFECTS OF SKEWNESS AND KURTOSIS

S:	$CV_S = 0.4$		$CV_S = 0.8$		$CV_S = 1.2$	
$[\alpha_1, \alpha_2]$	μ_Y	CV_Y	μ_Y	CV_Y	μ_Y	CV_Y
0.8 2.0	1.685	0.788	1.721	1.070	1.757	1.380
1.2 3.0	1.692	0.805	1.735	1.097	1.778	1.410
1.6 4.0	1.696	0.814	1.744	1.112	1.790	1.428
1.6 5.0	1.715	0.860	1.781	1.186	1.823*	1.522*
2.0 6.0	1.721	0.873	1.792	1.207	1.861	1.542

* The original S-D distribution has a small left tail stretching into the negative area, so the distribution is adjusted. Hereafter, the number with star is obtained after adjustment.

Example 4.

In this example, only ω value is changed to 1/200. All others are the same as those in example 3. Note that ω is the breakdown rate, and $1/\omega$ is the mean time between breakdowns.

TABLE II

EFFECTS OF THE BREAKDOWN RATE ω

S:	$CV_S = 0.4$		$CV_S = 0.8$		$CV_S = 1.2$	
$[\alpha_1, \alpha_2]$	μ_Y	CV_Y	μ_Y	CV_Y	μ_Y	CV_Y
0.8 2.0	1.593	0.650	1.611	0.969	1.630	1.314
1.2 3.0	1.597	0.663	1.619	0.987	1.641	1.335
1.6 4.0	1.599	0.670	1.623	0.998	1.647	1.348
1.6 5.0	1.608	0.705	1.642	1.052	1.659*	1.382*
2.0 6.0	1.611	0.715	1.647	1.068	1.683	1.432

Once again, one sees clearly that both μ_Y and CV_Y are positively correlated with CV_S and $[\alpha_1(S), \alpha_2(S)]$. But, comparing Table II with Table I, one sees that μ_Y decreases by 5.5% to 9.6%, and CV_Y decreases by 7.2% to 17.5%. This indicates the importance of the parameter ω as expected.

Are S's Skewness and Kurtosis Important?

Example 5.

R's first four central moments and k are the same as those in example 4. μ_S is still 1.5. $CV_S = 0.4$. First, we set $\alpha_2(S)$ equal to 3, and change $\alpha_1(S)$ from -1.2 to 1.2, in step of 0.4. Then we set $\alpha_2(S)$ equal to 4, and change $\alpha_1(S)$ from -1.6 to 1.6, in step of 0.4. Note that $[\alpha_1=\pm1.6, \alpha_2=3]$ and $[\alpha_1=\pm2.0, \alpha_2=4]$ are impossible combinations (see Schmeiser and Deutsch 1977 for details). From Table III, it is evident that, except the marked rows, μ_Y and CV_Y are monotonously, positively correlated with $\alpha_2(S)$, and $\alpha_2(Y)$ is monotonously, negatively correlated with $\alpha_2(S)$. But at $[\alpha_1=1.2, \alpha_2=3]$ and $[\alpha_1=1.6, \alpha_2=4]$, μ_Y and CV_Y are decreasing, while $\alpha_2(Y)$ is increasing. Notice that $[\alpha_1=1.6, \alpha_2=3]$ and $[\alpha_1=2.0, \alpha_2=4]$ are impossible combinations, these anomalies might be attributed to the S-D distribution's boundary behavior.

The relation between $\alpha_1(Y)$ and $\alpha_1(S)$ is very interesting. When $\alpha_1(S)$ is increasing from -1.2 to 0, $\alpha_1(Y)$ <u>decreases</u> from 6.1550 to 6.0658; then when $\alpha_1(S)$ continues increasing to 1.2, $\alpha_1(Y)$ turns around and <u>increases</u> to 6.1900. It seems that while S skews to left <u>or</u> right, Y always skews further toward left. That does not mean $\alpha_1(Y)$ is always positive, but always increases while $\alpha_1(S)$ deviates further from the neutral point 0.

Table III shows $\alpha_1(S)$ has significant effects on Y. For $\alpha_2(S)=3$, while $\alpha_1(S)$ increases from -1.2 to 0.8, μ_Y increases by 1.94%, and CV_Y increases by 9.96%. For $\alpha_2(S)=4.0$, while $\alpha_1(S)$ increases from -1.6 to 1.2, μ_Y increases by 2.69%, and CV_Y increases by 14.0%.

TABLE III

EFFECTS OF SKEWNESS

α_1	α_2	μ_Y	CV_Y	$\alpha_1(Y)$	$\alpha_2(Y)$
-1.2	3.0	1.6673	0.7480	6.1550	64.4553
-0.8		1.6826	0.7824	6.0843	60.1234
-0.4		1.6898	0.7990	6.0671	58.2194
-0.0		1.6945	0.8101	6.0658	57.0227
0.4		1.6978	0.8179	6.0733	56.2022
0.8		1.6997	0.8225	6.0904	55.7254
@ 1.2		1.6923	0.8051	6.1900	57.5518
-1.6	4.0	1.6636	0.7401	6.1328	65.5955
-1.2		1.6820	0.7812	6.0486	60.3183
-0.8		1.6911	0.8022	6.0205	57.9150
-0.4		1.6976	0.8176	6.0067	56.2755
0.0		1.7021	0.8282	6.0061	55.2014
0.4		1.7056	0.8366	6.0101	54.3697
0.8		1.7074	0.8410	6.0265	53.9478
1.2		1.7084	0.8435	6.0483	53.7088
@ 1.6		1.6963	0.8144	6.1882	56.5933

<u>Example 6.</u>

R's first four central moments and ω are the same as those in example 4. μ_S is still 1.5. $CV_S = 0.4$. This time set $\alpha_1(S)$ equal to -0.4, and change $\alpha_2(S)$ from 2 to 8. Then set $\alpha_1(S)$ equal to 0.4, and change $\alpha_2(S)$ from 2 to 8 again.

From Table IV, it is evident that, when $\alpha_1(S)=0.4$, μ_Y and CV_Y are monotonously, positively correlated with $\alpha_2(S)$, and $\alpha_1(Y)$ and $\alpha_2(Y)$ are monotonously, negatively correlated with $\alpha_2(S)$. But when $\alpha_1(S)$ is -0.4, while μ_Y is still monotonously, positively correlated with $\alpha_2(S)$, once again, CV_Y, $\alpha_1(Y)$ and $\alpha_2(Y)$ show some anomalies that might be attributed to the S-D distribution's boundary behavior. Table IV also shows $\alpha_2(S)$ has significant effects on Y. For $\alpha_1(S)= -0.4$, while $\alpha_2(S)$ increases from 2 to 8, μ_Y increases by 2.31%, and CV_Y increases by 10.0%. For $\alpha_1(S)=0.4$, while $\alpha_1(S)$ increases from 2 to 8, μ_Y increases by 2.38%, and CV_Y increases by 12.2%. These observations indicate that S's fourth moment or central moment can not be ignored.

TABLE IV

EFFECTS OF KURTOSIS

α_1	α_2	μ_Y	CV_Y	$\alpha_1(Y)$	$\alpha_2(Y)$
-0.4	2.0	1.6794	0.7752	6.1492	60.9373
	3.0	1.6898	0.7990	6.0671	58.2194
	4.0	1.6976	0.8176	6.0067	56.2755
	5.0	1.7043	0.8335	5.9576	54.7128
	6.0	1.7105	0.8484	5.9133	53.3298
	7.0	1.7150	0.8531	5.9185	52.9702
	8.0	1.7182	0.8472	5.9743	53.6179
0.4	2.0	1.6888	0.7969	6.1484	58.4140
	3.0	1.6978	0.8179	6.0733	56.2022
	4.0	1.7056	0.8366	6.0101	54.3697
	5.0	1.7121	0.8524	5.9599	52.9231
	6.0	1.7183	0.8675	5.9139	51.6148
	7.0	1.7238	0.8811	5.8746	50.4997
	8.0	1.7290	0.8939	5.8390	49.4971

CHAPTER VII

DEALING WITH MATERIAL HANDLING SYSTEM DELAY

In the last two chapters a powerful and convenient approach is presented to handle breakdowns. By transforming the line with breakdowns into an equivalent line without breakdown, many well-established methodologies can be applied to evaluate the system's performance. In this chapter, the very same approach is employed to handle MHS delay. So far, MHS delay (or the transit time) is almost ignored in the FMS literature. Many FMS models simply assume transit times are zero. Others concluded that, in general, transit times have a negligible effect on the production capacity, if there are a very large number of jobs in the system and the ratio of mean transit time over mean processing time ($E[D]/E[S]$) is small (Posner and Bernholtz 1968, Buzacott and Shanthikumar 1980). However, recently it is evident that the number of jobs in the system could be small or medium, and the ratio $E[D]/E[S]$ could be as high as 80%, or even greater than 100% (Ghosh and Wysk 1989, Commault and Semery 1990).

A closer look reveals that the way to treat MHS delay is depend on storage system configurations. In practice there are three kinds of storage systems as discussed below.

Central Storage with Infinite Capacity

Assume that the central storage with infinite capacity is shared by all stations which have no local buffers.

Let D_i be the transit time for moving a job from station i (or the central storage) to the central storage (or station i). Posner and Bernholtz (1968) considered to treat the MHS and the central storage as an extra station (say, station 0) with the service time equal to the transit time. Then the system has one more station, but no MHS delay (Figure 7). Because any processed job has to either return to station 0 or leave the system, there will be no blocking,

and the probability $P_{i0}=1-f_i$ for all $i\geq 1$, and the probability $P_{ij}=0$ for all $i\geq 1$ and $j\geq 1$. It is easy to see that Proposition I can apply, i.e., the system's output rate is λ, and station i's utilization is $\lambda e_i/\delta_i$.

It should be pointed out that station 0's service time includes not only the transit time, but also the waiting time in the queue (hereafter, queue delay), because the central storage is virtually a collection of queues. Therefore, this model essentially aggregates all stations' MHS delay and queue delay. Station 0 is special not only because its server is the MHS with the central storage instead of a machine, but also because it can serve all other M stations simultaneously, that is equivalent to having M servers. Chapter IX will return to this point.

Local Buffers with Limited Capacity

Assume each station has a local buffer with limited capacity, but there is no central storage.

Commault and Semery (1990) described a kind of MHS's, in which MHS delay for station i is D_i, which is a r.v. (Figure 8), if MHS's

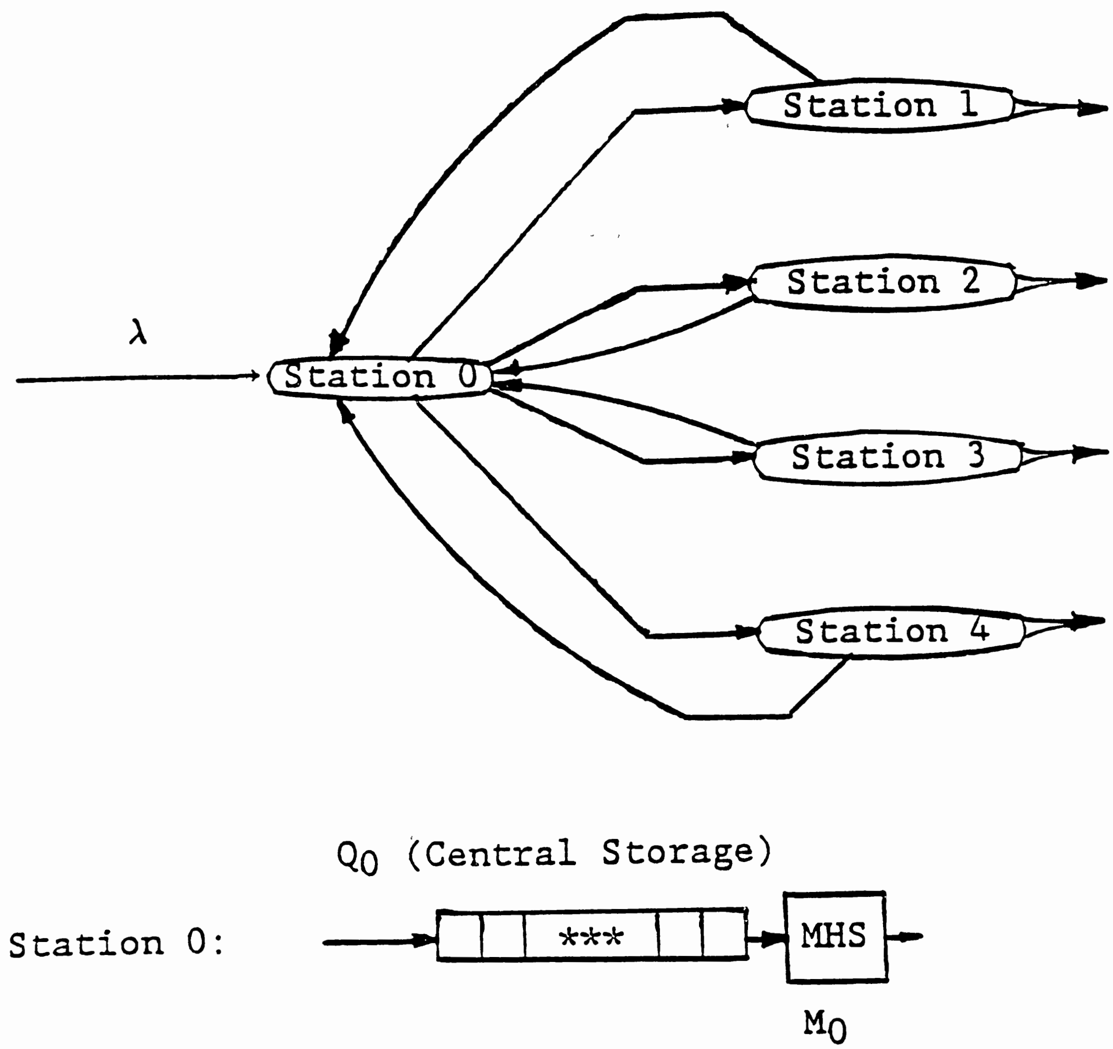
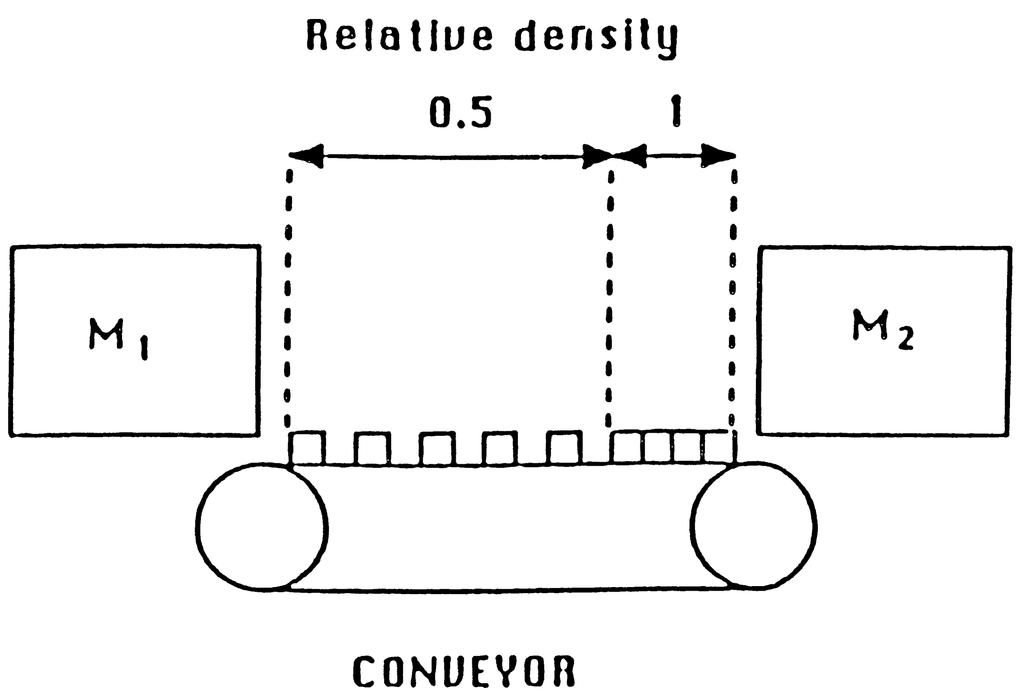


Figure 7. The FMS Model



Source: C. Commault, and A. Semery, "Taking into account delays in buffers for analytical performance evaluation of transfer lines," IIE Trans., Vol.22, No.2, (1990), 133-142.

Figure 8. A Two-Machine Transfer Line with a Conveyor

sends jobs to i immediately using a conveyor. For example, in Figure 8, when machine 2 (M2) finishes a job and sends the job out, it has to wait for the conveyor bringing a job. At that time, the nearest job (A) may be still **x** feet away. Assume the conveyor moves jobs at a constant speed of v feet per second, then M2 has to wait for $D_2 = x/v$ seconds to get A.

Commault and Semery (1990) suggest to transform the system into an "equivalent" system with unchanged station characteristics, no MHS delay, but reduced buffer capacities. As they point out, the main advantage of this approach is that, after a simple modification of the system input data, performance may be evaluated using existing method. However, they only consider a two-station case, and, as they put it, the really appealing problem is how to deal with larger systems.

Theoretically, they also suggest to aggregate MHS delay and queue delay, because the conveyor in Figure 8 is the buffer (station 2's queue), and modifying buffer capacities means modifying queue delay. We call this Backward Aggregation. A natural alternative is to aggregate MHS delay and effective processing times, or Forward Aggregation.

Let the general processing time of station i, GS_i, be the sum of its effective processing time Y_i and MHS delay D_i, i.e.,

$$GS_i = Y_i + D_i. \tag{7.1}$$

Assume Y_i and D_i are independent of each other, as they usually are; and D_i's pdf is $\theta(.)$ with moments $\mu_j(D_i)$ for j>0. If only one station is in question, the subscript i can be dropped.

For a particular station, let g(.) be GS's pdf, and y(.) be Y's pdf. If $\theta(.)$ and $y(.)$ are known, $g(.)$ is given as follows (Convolution Formula),

$$g(x) = \int_0^{\infty} y(z)\theta(x-z)dz. \tag{7.2}$$

Unfortunately, in most cases (7.2) is intractable even if $y(.)$ and $\theta(.)$ are given, not to mention that usually $y(.)$ and/or $\theta(.)$ are unknown. An alternative is the method of moments.

It is easy to see that GS's moments are (Wilks 1962, Kendall and Stuart 1969):

$$\left.\begin{aligned} &\mu_{GS} = \mu_Y + \mu_D; \\ &\mu_2(GS) = \mu_2(Y) + \mu_2(D); \\ &\mu_3(GS) = \mu_3(Y) + \mu_3(D); \\ &\mu_4(GS) = \mu_4(Y) + 6\mu_2(Y)\mu_2(D) + \mu_4(D). \end{aligned}\right\} \tag{7.3}$$

Given GS's first three or four moments, (7.3) could be fit to a three or four parameter distribution as GS's probability distribution. This distribution's density function is still denoted by $g(.)$.

Let's refer to Commault and Semery (1990)'s model the MB model (moving buffer model). Another popular model, which will be called FB model (fixed buffer model), describes fixed buffers. Fixed buffers impose some difficulties, and will be discussed in the next section.

Hybrid Storage System

Again, consider the central storage system. Let l_i be the transit time or MHS delay for moving a job from station i to the central storage, and L_i be the transit time for moving a job from the central storage to station i. When station i completes processing job A, the MHS will send A to the central storage within l_i time units, if no other station is waiting for A. Imagine that the central storage is a collection of M stations' queues, then l_i is only a part of A's queue waiting time. It is easy to see that, as long as no one needs A before a arrives at the central storage, l_i doesn't affect stations' processing times. But when station i finishes the job on hand, it has to spend time l_i to get the next job. Naturally, local buffers can be used to keep several job while stations are busy, because now the MHS can send jobs to the station (pre-load) whenever the buffer is not full. In other words, the local buffer can reduce MHS delay. That explains why so many FMS's employ both the central storage and local buffers, or hybrid storage systems.

If all buffers are MB's, the last two sections already provide a way to transform the system into an equivalent system without MHS delay. But how about FB's? The first section's method is recommended, i.e., using an extra station (0) to represent the MHS and the central storage. However, when local buffers are full, station 0 will be blocked, so Proposition I is no longer applicable. Chapter IX will return to this problem.

CHAPTER VIII

PERFORMANCE MEASURES OF STATIONS

So far the job arrival process is ignored, simply because it is irrelevant to the effective processing time Y or the general processing time GS. At this stage, the paper considers the equivalent system with (effective or general) processing time Y_i's, but without MHS delay and breakdowns; i.s., does not distinguish Y and GS.

Now it is desired to know the total time a job will stay in a station (X, the station time) or in its queue (W, the waiting time), the number of jobs in the station (L, the station size) or in the queue (Q, the length of the queue), and the station's output rate (TH). X, W, L, Q, and TH are all r.v.'s depending upon the job arrival process, as well as Y.

It is assumed that external jobs arrive at the system following a Poisson stream with rate λ. Then the arrival processes to stations are the merging/splitting of Poisson streams, which remain Poisson streams (Kelly 1979, Whitt 1982, Yao and Buzacott 1985b). Therefore, for a particular station i, it is assumed the arrival process is a Poisson stream with rate λ_i, and the effective processing time is Y_i with moments $\nu_n(Y_i)$, $n>0$; or simply λ and Y with moments ν_n, if no confusion.

Because the arrival process is Poisson, the service time (Y) is general, and there is one server (machine) in the station, it can be modeled as an M/G/1 queue. Hence the method of the imbedded Markov chain can be applied to find out the moments of L, Q, X, and W.

In this chapter, first several basic formulae for M/G/1 queues are introduced, and then a recursive procedure is developed to compute the moments of X, W, and L. The length of the queue, Q, has to be treated separately, because again we need the Theorem I to obtain Q's moments.

Basic M/G/1 Queue Formulae

Let $L(z) = \sum_{n=0}^{\infty} \text{Prob}[L=n]z^n$ be L's probability generating function (Gross and Harris 1985). Define the k-th factorial moment of L as

$$F_k(L) = E[L(L-1)(L-2)\ldots(L-K+1)],\ (k>0). \tag{8.1}$$

For example, $F_2(L) = E[L(L-1)] = E[L^2-L] = \nu_2(L) - \nu_1(L)$. One sees L(1) is 1, and

$$\begin{aligned} L'(1) &= dL(z)/dz|_{z=1} \\ &= \sum_{n=1}^{\infty} n*\text{Prob}[L = n] \\ &= \nu_1(L) \\ &= \mu_L. \end{aligned}$$

Generally, one has (Kleinrock 1975, Gross and Harris 1985)

$$L^{(k)}(1) = F_k(L), \tag{8.2}$$

i.e., one can obtain L's moments by calculating $L^{(k)}(1)$, the k-th derivative evaluated at z=1, which gives the k-th factorial moment of L. Since $F_k(L)$ can be expressed in terms of L's first k moments, one can recursively calculate L's moments. The only problem is, what is L(z)?

The Expression of L(z)

Let $Y^*(s)$ be the Laplace transform of Y's pdf m(t), i.e.,

$$Y^*(s) = \int_0^\infty e^{-st} m(t)dt.$$

Kleinrock (1975) shows that

$$L(z) = Y^*(\lambda-\lambda z) \frac{(1-\rho)(1-z)}{Y^*(\lambda-\lambda z) - z}, \tag{8.3}$$

where $\rho = \lambda\mu_Y$. Let μ be the station's mean processing rate. Since μ_Y is the mean processing time,

$$\mu = 1/\mu_Y, \text{ and } \rho = \lambda/\mu. \tag{8.4}$$

The equation (8.3) is the famous Pollaczek-Khinchin (P-K) transform equation, which yields the moments for the distribution of the number of jobs in the station. When one attempts to set z=1 in equations $L^{(k)}(z)$ $(k\geq 0)$, he obtains indeterminant forms and has to use the L'Hospital's rule. In carrying out this operation, it is necessary to evaluate

$$V^{(k)}(1) = \lim_{z\to 1} \frac{d^k Y^*(\lambda-\lambda z)}{dz^k}, \tag{8.5}$$

where $V(z) = Y^*(\lambda-\lambda z)$. Fortunately, it can be shown (Kleinrock 1975) that

$$V^{(k)}(1) = \lambda^k \nu_k(Y). \tag{8.6}$$

Specifically, $V'(1) = \lambda\nu_1(Y)$, and $V''(1) = \lambda^2\nu_2(Y)$, where V' and V" are V(z)'s first and second order derivatives, respectively.

Calculation of $\nu_1(Y)$

Now, as an example, let's calculate L'(1).

$$L'(z) = dL(z)/dz$$

$$= \frac{(1-\rho)(1-z)}{V-z} V' + (1-\rho)V(z) \frac{(V-z) + (1-z)(V'-1)}{(V-z)^2} .$$

And therefore,

$$L'(1) = \lim_{z\to 1} L'(z)$$

$$= \lim_{z\to 1} \frac{V'(1-\rho)}{V'-1} - (1-\rho) \lim_{z\to 1} \frac{V'(V-z)+(1-z)V'(V'-1)+(1-z)VV''}{2(V-z)(V'-1)}$$

$$= \rho - (1-\rho) \lim_{z\to 1} \left[\frac{V'}{2(V'-1)} + \frac{(1-z)V'}{2(V-z)} + \frac{(1-z)VV''}{2(V-z)(V'-1)}\right]$$

$$= \rho + (1-\rho)\left[\frac{\rho}{2(1-\rho)} - \lim_{z\to 1} \frac{(1-z)V''-V'}{2(V'-1)} + \right.$$

$$\left. \lim_{z\to 1} \frac{VV''-(1-z)V'V''-(1-z)VV^{(3)}}{2(V-z)V''+2(V'-1)^2}\right]$$

$$= \rho + (1-\rho)\frac{\lambda^2\nu_2(Y)}{2(1-\rho)^2}$$

$$= \rho + \frac{\lambda^2\nu_2(Y)}{2(1-\rho)} . \qquad (8.7)$$

Considering (8.4) and (3.1), one has

$$\nu_1(L) = L'(1)$$

$$= \rho + \frac{\lambda^2\nu_2(Y)}{2(1-\rho)}$$

$$= \rho + \frac{\lambda^2(\sigma_Y^2 + \mu_Y^2)}{2(1-\rho)}$$

$$= \rho + \frac{\rho^2(1 + CV_Y^2)}{2(1-\rho)} , \qquad (8.8)$$

which is the well-known P-K mean value formula.

One could go ahead to get L's second and higher moments. But it will be very troublesome and tedious to follow this way. In the next sections a recursive procedure is developed to calculate L's moments. In fact, it will calculate all the moments of L, X, and W.

A Recursive Procedure for Moments of X, W, and L

The Recursive Procedure

Let $A^*(s)$ be the Laplace transform of the r.v. A's pdf. According to Kleinrock (1975), we know that, for the station time X,

$$X^*(s) = Y^*(s)s(1-\rho)/[s-\lambda+\lambda Y^*(s)],$$

and for the waiting time W,

$$X^*(s) = Y^*(s)W^*(s),$$

which is from the well-known formula $X = Y + W$.

Then it is found that (Takacs 1962, Kleinrock 1975)

$$\nu_k(W) = \frac{\lambda}{1-\rho} \sum_{i=1}^{k} (C_k^i) \frac{\nu_{i+1}(Y)}{(i+1)} \nu_{k-i}(W), \tag{8.9}$$

$$\nu_k(X) = \sum_{i=0}^{k} (C_k^i) \nu_{k-i}(W)\nu_i(Y), \tag{8.10}$$

and $F_k(L) = \lambda^k \nu_k(X)$. (8.11)

Now one can calculate the first n moments of W, X, and L as follows. Q will be treated separately later.

PROCEDURE I.

Input: n, the input rate λ, and Y's first n moments.

Step 1. Let $i = 1$; Notice that $\nu_0(.) = 1$;

Step 2. Use (8.9) and $\nu_j(W)$'s $(j<i)$ to obtain $\nu_i(W)$;

Step 3. Use (8.10), $\nu_j(X)$'s ($j<i$) and $\nu_j(W)$'s ($j\le i$) to obtain $\nu_i(X)$;

Step 4. Use (8.11), $\nu_j(L)$'s ($j<i$) and $\nu_j(X)$'s ($j\le i$) to obtain $\nu_i(L)$;

Step 5. If $i<n$, increase i by one (i.e., $i \leftarrow i+1$), return to step 2;

If $i=n$, stop.

Output: The first n moments of W, X, and L. ##

Next, one can apply this procedure to calculate their first four moments.

The First Four Moments of W, X, and L

Following the above procedure, it is found that,

1. $$\nu_1(W) = \frac{\lambda}{1-\rho} * \frac{\nu_2(Y)}{2} = \frac{\lambda\nu_2(Y)}{2(1-\rho)} = \frac{\rho^2(1+CV_Y{}^2)}{2(1-\rho)},$$ which is also well-known.

$$\nu_1(X) = \nu_1(W) + \nu_1(Y) = \frac{\lambda\nu_2(Y)}{2(1-\rho)} + \mu_Y.$$

$$\nu_1(L) = \lambda\nu_1(X) = \frac{\lambda^2\nu_2(Y)}{2(1-\rho)} + \lambda\mu_Y = \rho + \frac{\rho^2(1 + CV_Y{}^2)}{2(1-\rho)},$$

which is reconciled with equations (8.7) and (8.8).

2. $$\nu_2(W) = \frac{\lambda}{1-\rho}\left[\frac{2\nu_2(Y)}{2}\nu_1(W) + \frac{\nu_3(Y)}{3}\right]$$

$$= 2[\nu_1(W)]^2 + \frac{\lambda\nu_3(Y)}{3(1-\rho)}.$$

$$\nu_2(X) = \nu_2(W) + 2\nu_1(W)\nu_1(Y) + \nu_2(Y).$$

$\nu_2(L) - \nu_1(L) = \lambda^2\nu_2(X)$, and therefore

$$\nu_2(L) = \lambda^2\nu_2(X) + \nu_1(L).$$

3. $$\nu_3(W) = \frac{\lambda}{1-\rho}\left[\frac{3\nu_2(Y)}{2}\nu_2(W) + \nu_3(Y)\nu_1(W) + \frac{\nu_4(Y)}{4}\right].$$

$$\nu_3(X) = \nu_3(W) + 3\nu_2(W)\nu_1(Y) + 3\nu_1(W)\nu_2(Y) + \nu_3(Y).$$

$\nu_3(L) - 3\nu_2(L) + 2\nu_1(L) = \lambda^3\nu_3(X)$, and therefore

$$\nu_3(L) = \lambda^3\nu_3(X) + 3\nu_2(L) - 2\nu_1(L).$$

4. $$\nu_4(W) = \frac{\lambda}{1-\rho}\left[2\nu_2(Y)\nu_3(W)+2\nu_3(Y)\nu_2(W)+\nu_4(Y)\nu_1(W)+ \frac{\nu_5(Y)}{5}\right].$$

$$\nu_4(X) = \nu_4(W) + 4\nu_3(W)\nu_1(Y) + 6\nu_2(W)\nu_2(Y) + 4\nu_1(W)\nu_3(Y) + \nu_4(Y).$$

$\nu_4(L) - 6\nu_3(L) + 10\nu_2(L) - 3\nu_1(L) = \lambda^3\nu_3(X)$, and therefore

$$\nu_4(L) = \lambda^4\nu_4(X) + 6\nu_3(L) - 10\nu_2(L) + 3\nu_1(L).$$

Moments of the Length of the Queue

Let Q represent the number of jobs in the queue, not counting the job, if any, in processing. Q can be expressed by L, the number of jobs in the station, including the job in processing:

$$Q = \begin{cases} L - 1, & \text{if } L > 0; \\ 0, & \text{if } L = 0. \end{cases} \tag{8.12}$$

By conditioning on L, (8.12) can be rewritten as

$$Q = \begin{cases} k - 1, & \text{if } L = k > 0; \\ 0, & \text{if } L = 0. \end{cases} \tag{8.13}$$

According to the Theorem I in chapter V, using "Σ" for discrete distributions, instead of "∫" for continuous distributions, one has

$$\nu_n(Q) = \sum_{k=1}^{\infty} (k-1)^n \mathrm{Prob}[L = k]$$

$$= \sum_{k=1}^{\infty} [\sum_{j=0}^{n} (C_n^j) k^{n-j} (-1)^j] * \mathrm{Prob}[L = k]$$

$$= \sum_{j=0}^{n} (-1)^j (C_n^j) \{ \sum_{k=1}^{\infty} k^{n-j} \mathrm{Prob}[L = k] \}$$

$$= \sum_{j=0}^{n-1} (-1)^j (C_n^j) \{ \sum_{k=0}^{\infty} k^{n-j} \mathrm{Prob}[L = k] \} + (-1)^n \sum_{k=1}^{\infty} \mathrm{Prob}[L = k]$$

$$= \sum_{j=0}^{n} (-1)^j (C_n^j) \nu_{n-j}(L) + (-1)^n \rho. \tag{8.14}$$

It is easy to see that $\nu_1(Q) = \nu_1(L) - \rho$, which is also a well-known formula in the queueing theory. Moreover,

$$\nu_2(Q) = \nu_2(L) - 2\nu_1(L) + \rho;$$

$$\nu_3(Q) = \nu_3(L) - 3\nu_2(L) + 3\nu_1(L) - \rho;$$

$$\nu_4(Q) = \nu_4(L) - 4\nu_3(L) + 6\nu_2(L) - 4\nu_1(L) + \rho.$$

Other Measures for the station

When the arrival rate λ is given, the station's steady state output rate TH = λ, and its utilization is ρ, whenever $\lambda < \mu$ (Gross and Harris 1985, Ross 1989, among others).

In the next chapter the blocking rate computation for each station will be considered. Blocking usually stems from limited buffer space.

Notice that, if the local buffer capacity is, say, two, then the probability that a coming job will be rejected by this station is simply $\text{Prob}[Q \geq 2]$, or $\text{Prob}[L \geq 3]$. Now the problem is, which one, $\text{Prob}[Q \geq 2]$ or $\text{Prob}[L \geq 3]$, is better? If both Q's and L's CDF's are known, there will be no difference. But now only their moments are known, and have to be fit to, say, a four parameter, continuous distribution, that will incur some error, especially at the boundary. Because

$$\text{Prob}[Q \geq 2] = 1 - \text{Prob}[Q = 1] - \text{Prob}[Q = 0],$$

$$\text{Prob}[L \geq 3] = 1 - \text{Prob}[L = 2] - \text{Prob}[L \leq 1], \text{ and}$$

$$\begin{aligned}\text{Prob}[L \leq 1] &= \text{Prob}[Q = 0] \\ &= \text{Prob}[L = 1] + \text{Prob}[L = 0],\end{aligned}$$

there will be more mass built up at the boundary point zero for Q than for L. So it is reasonable to choose working on L, rather than Q, not to mention the savings from avoiding the computation of (8.14).

Now let b_i be station i's blocking rate. Assume L's first three or four moments are fit to a three or four parameter CDF $F_i(.)$. Then

$$b_i = \text{Prob}[L \geq 3] = 1 - F_i(3). \tag{8.15}$$

Notably, when blocking occurs, the arrival process, therefore TH, could be changed. From station i's point of view (Figure 6), it may reject some coming jobs (in input flows), and its outputs could be rejected by other stations. In turn, the rejected jobs will affect the input flows. These system related problems will be discussed in the next chapter.

CHAPTER IX

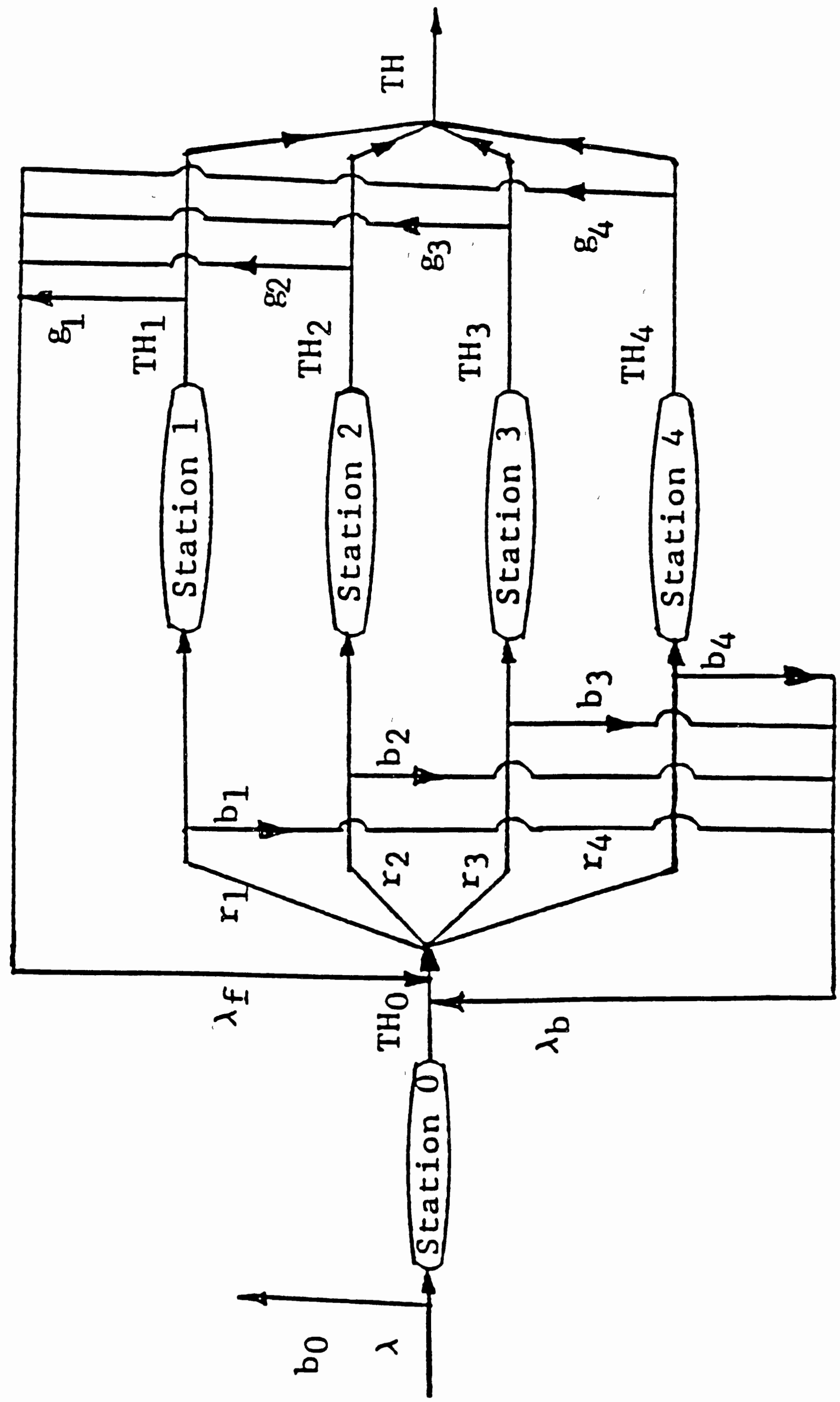
PERFORMANCE EVALUATION OF FLEXIBLE MANUFACTURING SYSTEMS

The Model of the System

Yao and Buzacott (1985b) described an open queueing network with a set of work stations, each having a local buffer with limited capacity and general processing times (Figure 9). The MHS is divided into two subsystems, the MHS(I) and the MHS(O). The MHS(I) consists of a set of carts (or input conveyors) to send jobs to the stations. The MHS(O) is a return conveyor at the output side of the stations.

Although the machines are never blocked in this model, the input jobs can be blocked if the local buffer at the destination station is fully occupied. The authors made an important assumption pertaining to the blocking mechanism: The blocked jobs will be recirculated ('block-and-recirculate'), instead of occupying the cart and waiting in front of the station ('block-and-hold', see Buzacott and Fanifin, 1978). They argued that, "Given the versatility of job routing and the variety of operations and operation sequences in an FMS environment, a block-and-hold mechanism just seems to be too restrictive." (Yao and Buzacott, 1985b)

This model is adopted here, with slight modifications, and described in the next two sections. Some of the assumptions previously stated in chapter III will be repeated and/or modified as appropriate.



Source: D. D. Yao, and J. A. Buzacott, "Modeling tne performance of flexible manufacturing systems," <u>Int. J. Prod. Res.</u>, Vol.23, No.5, (1985b), 945-959.
(Revised by the author Long-Geng Zhao)

Figure 9. The "Block-and-Recirculate" Model

Model Description

The FMS consists of M stations, each of which has one machine with a general processing time (e.g., Y_i for station i). The buffer capacity is n_i at station i. The MHS(I), which is treated as station 0, has c>0 carts to deliver jobs from the central storage to stations 1 to M. Let $r_i>0$ be the probability that MHS(I) will send a job to station i. Thus, the "distributing rates" r_i's must satisfy

$$\sum_{i=1}^{M} r_i = 1.$$

Jobs may be rejected by a station which queue is full, and rejected jobs will be recirculated, i.e., sent back to the central storage and get prepared for retrial. Note that it is assumed that rejected jobs follow the same probabilities r_i's when they join the central storage. In fact, r_i's are defined asymptotically, not individually. The service processes of carts are known, or at least its first k (k=4 or 5) moments, $\nu_k(D)$'s, are given. Here D represents one cart's service time and all carts are identical. A job leaving station i will either be fed back to the central storage with probability $g_i = 1 - f_i$, or leave the system with probability f_i. Both the feedback and the exit transits are handled by MHS(O), which is treated as station M+1, and modeled as an infinite-server queue with known parameters.

External jobs arrive at the system following a Poisson stream with rate λ. Whenever the total number of jobs at the central storage reaches N_0, external arrivals are turned away and lost. Therefore, the central storage should be able to contain a total of

$$N = N_0 + \sum_{i=1}^{M} n_i$$

jobs. The central storage imposes no limit on internal jobs. So the internal jobs have priority to occupy both the central storage and the MHS(I).

The blocking at the MHS(I) is modeled through an additional arrival stream, the blocking feedback. Because no real physical blocking may happen to the stations, the isolated stations can be analyzed using the techniques developed in chapters IV through VIII. The only problem is how to decide the arrival process for each station. This is discussed in the next section, adapted from Yao and Buzacott (1985b).

The Equivalent Arrival Process

Let the arrival rate to station i be λ_i. It is assumed that the arrival processes to stations can be approximated by renewal streams or the merging/splitting of renewal streams. TH_i is defined as the number of jobs completed in a unit of time. Let b_i ($0 \le i \le M$) denote the probability that a job is blocked (rejected) on arriving at station i. No job will be blocked at station M+1. These b_i's are unknown parameters to be derived.

The arrival flow to station 0 has three components: the external flow λ, the output feedback λ_f, and the blocking feedback λ_b. The arrival flow to station i ($1 \le i \le M$), λ_i, is a fraction r_i of the output of station 0 (TH_0). The arrival flow to station M+1 is the merging of the output (TH_i) from each station.

Yao and Buzacott (1985b) proved the following proposition valid for the model here.

PROPOSITION II.

The equivalent arrival rates to the stations can be formulated as follows:

$$\lambda_0 = \begin{cases} \lambda + \lambda_f + \lambda_b & \text{if } n<N_0 \\ \lambda_f + \lambda_b & \text{if } N_0 \le n \le N \end{cases} \tag{9.1}$$

$$\lambda_i = TH_0 r_i \quad (1 \le i \le M) \tag{9.2}$$

$$\lambda_{M+1} = TH_{M+1} = \sum_{i=1}^{M} TH_i \tag{9.3}$$

where

$$\lambda_f = \sum_{i=1}^{M} TH_i g_i, \quad \lambda_b = \sum_{i=1}^{M} \lambda_i b_i \tag{9.4}$$

and the outputs of the stations are as follows:

$$TH_0 = \frac{\lambda(1-b_0)}{1 - \sum_{i=1}^{M} r_i[b_i+(1-b_i)g_i]} \tag{9.5}$$

$$TH_i = TH_0 r_i (1-b_i) \quad (1 \le i \le M). \tag{9.6}$$

##

From (9.1), one can see that

$$\begin{aligned} \lambda_0 &= (\lambda + \lambda_f + \lambda_b)(1 - b_0) + (\lambda_f + \lambda_b) \\ &= \lambda(1 - b_0) + \lambda_f + \lambda_b. \end{aligned} \tag{9.1'}$$

9.1 to 9.6 are the flow balance equations. With the help of these equations, a general iterative procedure (called GIP here) can be employed to figure out the system's output rate TH. This GIP will be introduced in the next section.

Derivation of the Blocking Probabilities, Arrival Rates and System Output Rate

The formulas (9.1) to (9.6) are based on a set of unknown parameters, b_i's ($0 \leq i \leq M$), which, in turn, depend on λ_i's or TH_i's. This set of parameters can be derived through solving a fixed-point problem. The following procedure, Figure 10, is devised to find out b_i's, TH_i's and TH.

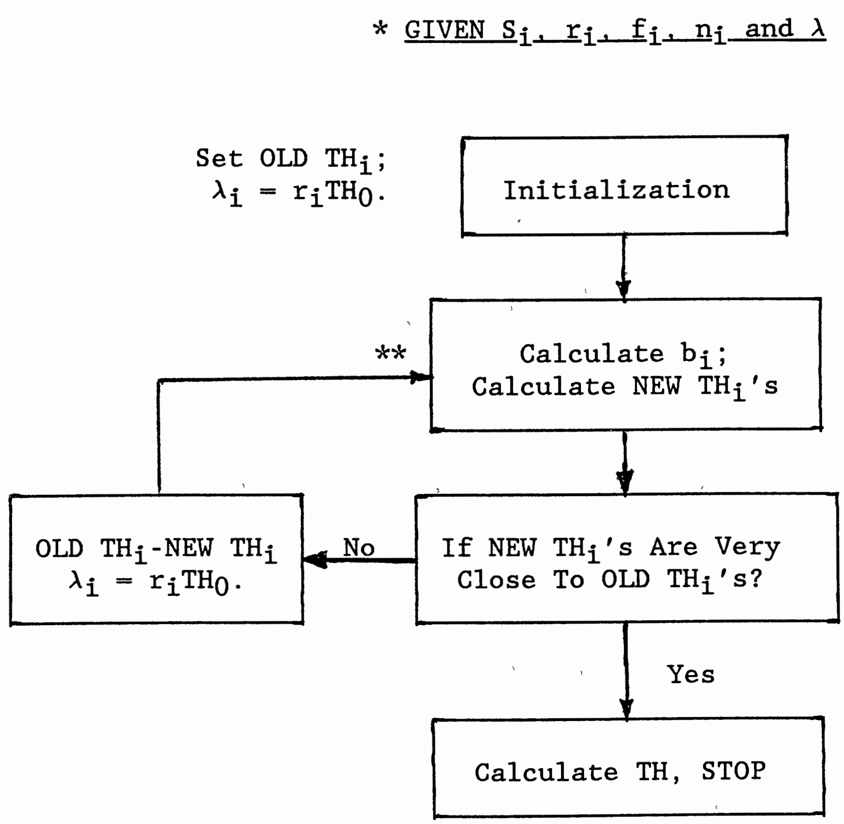


Figure 10. General Iterative Procedure (GIP)

Verbally, this procedure takes S_i, n_i (i=0,...,M), r_i, f_i (i=1,...,M) and λ as inputs. First, it arbitrarily assigns initial values to TH_i's (OLD TH_i, usually making a reasonable guess) and sets $\lambda_i = r_i TH_0$; then calculates b_i's (discussed later) and NEW TH_i's (applying the flow balance equations).

Next the NEW TH_i's are compared with the OLD TH_i's. If the NEW and OLD TH_i's are close enough, the fixed point is found and the NEW TH_i's are used to calculate TH; otherwise it sets the OLD TH_i's equal to the NEW TH_i's and recalculates λ_i's, b_i's and NEW TH_i's and goes to the next iteration.

Though there are several different versions of GIP in the literature, the underlying iterative structures are the same. Here the common assumptions are that S_i's (i.e., the distributions of S_i's) are known and there are no breakdowns. To cope with breakdowns, the input S_i should be replaced by Y_i, or the first four moments of S_i and R_i along with k_i (see Figure 11), and Procedure I in chapter VIII can be applied to calculate the first several moments of W_i, X_i and L_i; then according to (8.15), $b_i = \text{Prob}[L_i > n_i] = 1 - F_i(n_i)$, that leads to Figure 12.

Note that Figure 12 will replace the block with double asterisks (**) in Figure 10. In summary, Figure 10 should be transformed to Figure 13, which is the flow chart for Procedure II.

Procedure II is devised to find out TH, where the superscript (j) denotes the j-th iteration. Note that it sets $TH_0 = \lambda$, and consequently $\lambda_i = r_i TH_0$ and other TH_i's can be calculated by (9.5) and (9.6) if b_i's are known.

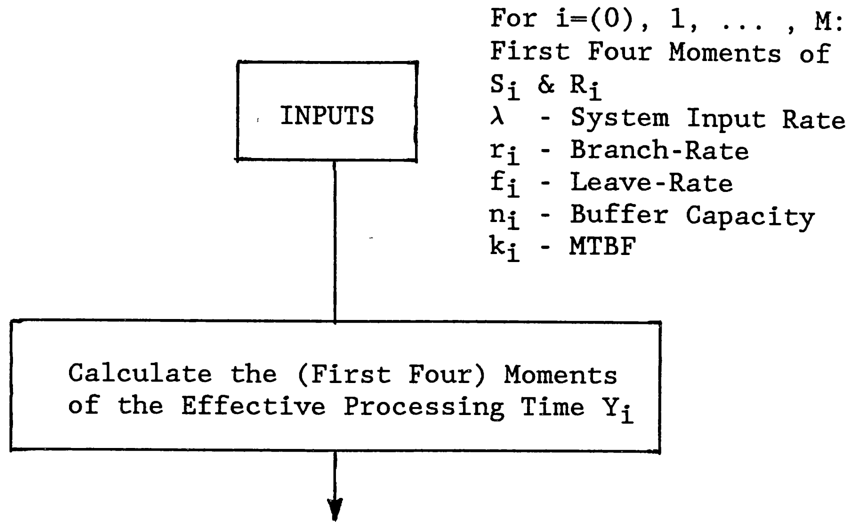


Figure 11. System Inputs Considering Breakdowns

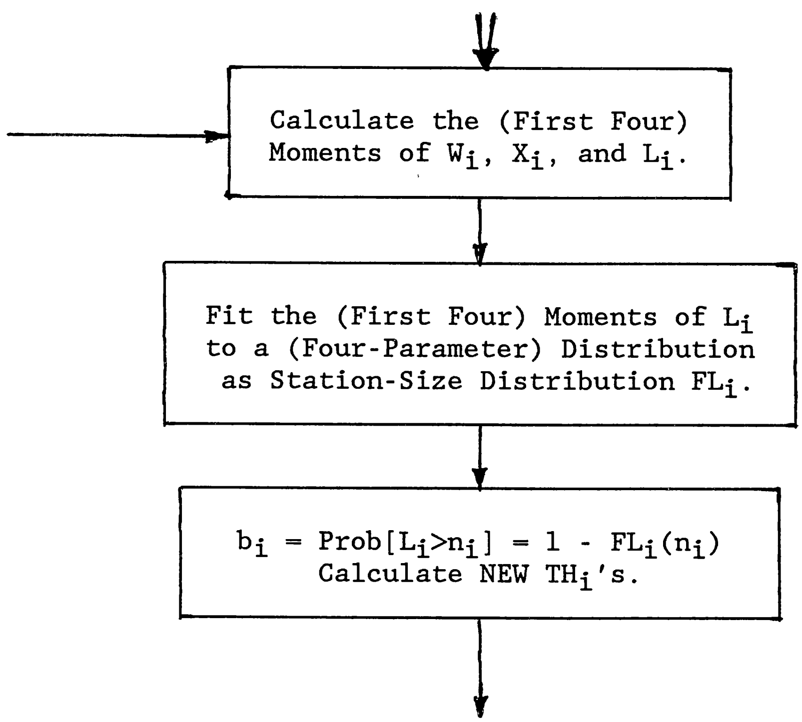


Figure 12. The New Method To Obtain b_i

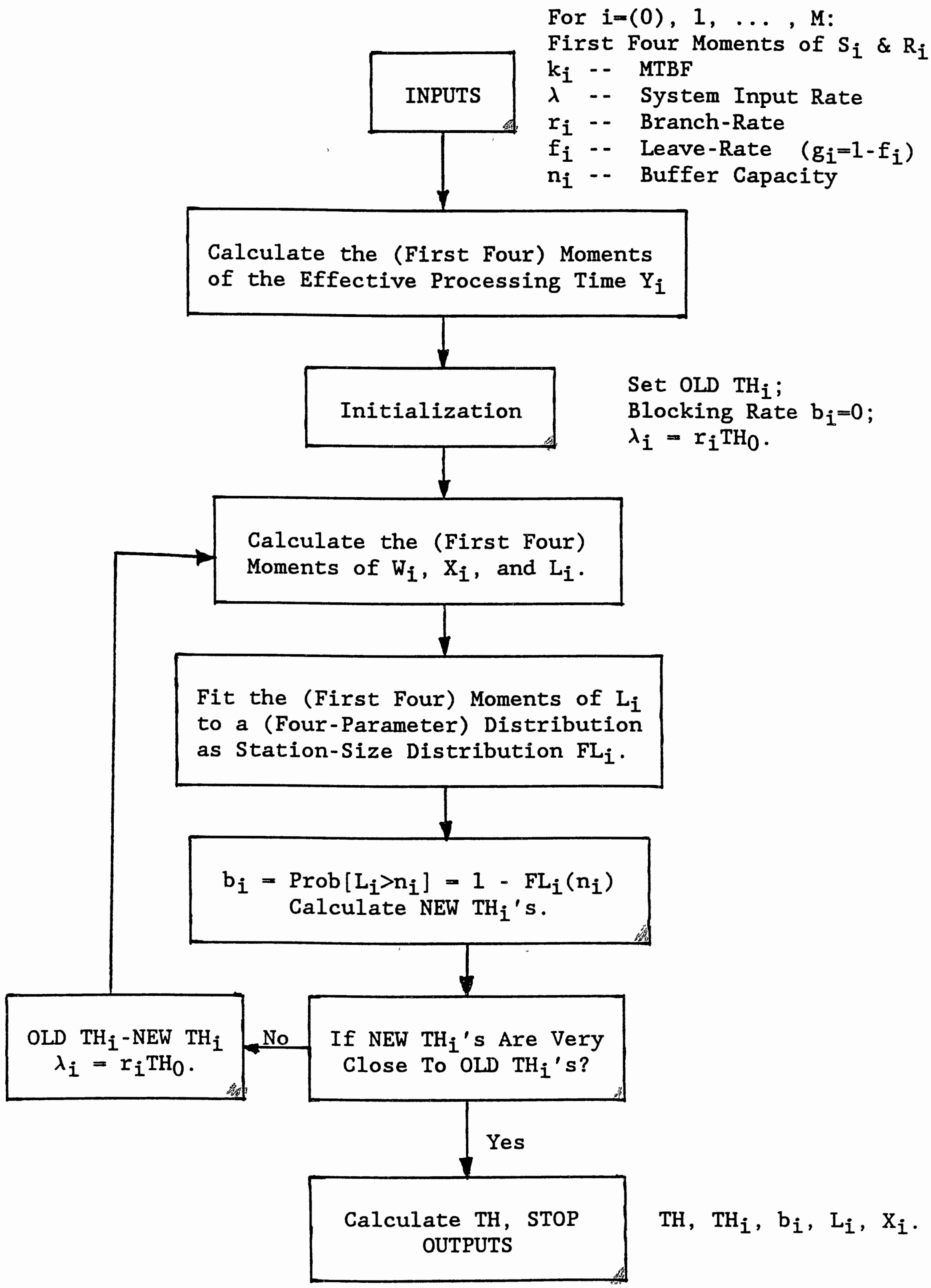


Figure 13. The Flow Chart For PROCEDURE II.

PROCEDURE II.

Input: λ, L_i's CDF ($1 \le i \le M$), and ϵ, which is a predetermined small positive number.

Step 1. Initially, set $j=0$, $b_0^{(j)}=0$, $\lambda_0^{(j)}=\lambda$, $TH_0^{(j)}=\lambda$, and $\lambda_i^{(j)}= r_i TH_0^{(j)}$ ($1 \le i \le M$).

Step 2. For all $i=1,\ldots,M$, derive $b_i^{(j)}= \text{Prob}[L_i > n_i]$ following the procedures introduced in chapter VIII.

Step 3. Solve (9.5), (9.6) and (9.2) to obtain $TH_i^{(j)}$ ($0 \le i \le M+1$).

Step 4. Use (9.4) to obtain $\lambda_f^{(j)}$ and $\lambda_b^{(j)}$, and use (9.1'), (9.2), and (9.3) to obtain $\lambda_i^{(j)}$ ($0 \le i \le M+1$).

Step 5. Set $j = j+1$;

Treat station 0 as a normal station with service time D and maximum buffer capacity N_0 for external inputs. Use the same method of Step 2 to Calculate $b_0^{(j)}$ and $b_i^{(j)}$ ($i = 1, \ldots, M$), and set $b_0^{(j)} = 1 - (TH_0 - \lambda_f - \lambda_b)/\lambda$;

Step 6. Solve (9.5), (9.6) and (9.2) to obtain $TH_i^{(j)}$ ($0 \le i \le M+1$).

If $\max_i |TH_i^{(j)} - TH_i^{(j-1)}| < \epsilon$, stop; else, go to step 4.

Output: b_i's, λ_i's, and TH_i's. ##

This procedure naturally imitates the system. At the beginning, the system is empty, and all capacity is free. So $b_0=0$, and $\lambda_0=\lambda$. Since the input into station 0 is a Poisson stream, and all carts are free, at least before the first time the number of jobs in station 0 reaches N_0, the output of station 0 is also a Poisson stream. Of course λ_f and λ_b may be taken in account, and later it will be shown that they are also thought of as Poisson.

Because the probabilistic splitting/merging of a set of (independent) Poisson streams preserve the Poisson property (Kelly 1979, Whitt 1982), all input flows of station i ($1 \leq i \leq M$) are Poisson streams. That is what Step 1 does. Since input flows are Poisson, Procedure I can be applied to calculate b_i's, which is Step 2. Yao and Buzacott (1985b) argued that, since the input into station M+1 is a superposition of the output processes from many machines <u>all the time</u>, it can be reasonably approximated by a Poisson stream; and since station M+1 is an infinite-server station, the Poisson input will yield a Poisson output, regardless of the processing time distribution of this station. That justifies (9.5) of Step 3 and (9.4) of Step 4. Following the same argument, λ_b can also be approximated by a Poisson stream, and so is λ_0. Let BF denote the blocking flow of station 0, which is the difference between the sum of in-flows ($\lambda_f + \lambda_b + \lambda$) and the out-flow TH_0, or $BF = (\lambda_f + \lambda_b + \lambda) - TH_0$. Also note that $b_0 = BF/\lambda$; so, combining the above equation, one has $b_0 = 1 - (TH_0 - \lambda_f - \lambda_b)/\lambda$, which justifies Step 5. Hillier and Boling (1967)'s "exit-oriented" approach (see Appendix I) justifies (9.6) of Step 3. Steps 4 to 6 compose an iteration procedure to find out the fixed points

$$\mathbf{TH} = (TH_0, TH_1, \ldots, TH_{M+1})^T,$$

and $$\mathbf{b} = (b_0, b_1, \ldots, b_M)^T.$$

When the real FMS can reach steady state, this simple iterative scheme should converge, although no formal proof exists.

System Performance Measures

It is clear that the system output rate is

$$TH = \sum_{i=1}^{M} f_i THi.$$

Station i's utilization is $U_i = \lambda_i THi \mu_1(Y_i)$, $1 \le i \le M$.

The sojourn time for a class-r job, T_s^r, is approximated as follows,

$$T_s^r = \sum_{i=0}^{M} e_{ir}[\rho_i + \mu_1(W_i)], \qquad \text{(see chapters III and IV.)} \qquad (9.7)$$

where $\rho_i = \lambda_i TH_i \mu(Y_i) < 1$. Note that ρ_i is station i's utilization. If MHS delay is already included in the effective processing time Y, there is no need to consider station 0 and M+1. When there is only one class of jobs, the subscript or superscript r can be dropped. The remaining problem is how to calculate $\mu_1(W_i)$ in (9.7).

Recall in step 2 of Procedure II, the above mentioned Procedure I is called, and that will calculate the moments of W_i, X_i, and Q_i for i=0, ... , M. Therefore, it is easy to determine each station's mean waiting time $\mu_1(W_i)$, mean length of queue $\mu_1(Q_i)$, and mean station time $\mu_1(X_i) = \rho_i + \mu_1(W_i)$. But due to the blocking mechanism, the real waiting time/queue length distributions' right tails will be truncated. For example, if station i's buffer capacity is n_i, then

$$E[Q_i] = \int_0^{n_i} xf(x)dx,$$

where f(x) is Q_i's probability density function, which can be fitted to Q_i's first three or four moments.

Numerical Examples and Simulation Verification

Example 9.1

Consider an FMS with four stations (station 1 to 4). Assume their service times are the same Y, which first four moments are 1.0, 0.36, 0.0864, and 0.2592, respectively. Also assume that $r_1 = r_2 = 0.3$, $r_3 = r_4 = 0.2$; $f_1 = f_2 = 0.35$, $f_3 = f_4 = 0.25$; Station 0's processing time (MHS delay) is uniformly distributed between 0.1 and 0.3; the buffer capacity is 3 for stations 1 to 4 and 200 for the central storage (N_0=200); and $\lambda = 0.3$.

Table V below shows the analytical results from Procedure II. Computer simulation is used to verify the analytical results. The computer program for Procedure II and the basic SLAM II simulation model (Pritsker 1986) is shown in Appendix D and E, respectively. It can be seen that almost all analytical results fall in the ranges of confidence intervals, set at 95% level, with a few exceptions (marked by an asterisk).

The sojourn time is determined in the following manner. First, treat the system as it is, i.e., only 4 stations. Let $\mathbf{Q} = (r_1, r_2, r_3, r_4)^T = (0.3, 0.3, 0.2, 0.2)^T$ and $\mathbf{E} = (e_1, e_2, e_3, e_4)^T$. The transition matrix $\mathbf{P}$ is

$$\begin{bmatrix} 0.3*0.65 & 0.3*0.65 & 0.2*0.65 & 0.2*0.65 \\ 0.3*0.65 & 0.3*0.65 & 0.2*0.65 & 0.2*0.65 \\ 0.3*0.75 & 0.3*0.75 & 0.2*0.75 & 0.2*0.75 \\ 0.3*0.75 & 0.3*0.75 & 0.2*0.75 & 0.2*0.75 \end{bmatrix}.$$

For example, when a job leaves station 2, it will stay in the system with the probability of 0.65, and if it stays in the system, it

TABLE V

EXAMPLE 9.1

Station Input Rate $r_i TH_0$:	STATION 1	2	3	4	IN_0@
Analytical Results	0.2903	0.2903	0.1935	0.1935	0.9677
Simulation Results	0.2851	0.2879	0.1943	0.1903	0.9569
Relative Error (%)	1.82%	0.83%	-0.41%	1.68%	1.13%
Confidence Interval (Confidence Level: 95%.)	±0.091	±0.088	±0.125	±0.130	±0.031

Station Output Rate:	TH_1	TH_2	TH_3	TH_4	System
Analytical Results	0.2903	0.2903	0.1935	0.1935	0.3000
Simulation Results	0.2843	0.2865	0.1940	0.1902	0.2980
Relative Error (%)	2.11%	1.33%	-0.26%	1.74%	0.67%
Confidence Interval	±0.090	±0.088	±0.124	±0.130	±0.066

Blocking Rate:	b_1	b_2	b_3	b_4	b_0
Analytical Results	0.000	0.000	0.000	0.000	0.000
Simulation Results	0.003	0.004	0.001	0.000	0.000
Confidence Interval	±0.018	±0.018	±0.006	±0.006	±0.006

(Since all blocking rates are very close to zero, the Relative Error are meaningless.)

Length of Queue:	1	2	3	4	5
Analytical:	0.081	0.081	0.032	0.032	0.000
Simulation:	0.085	0.096	0.035	0.039	0.000
Conf. Int.:	±0.007	±0.008*	±0.004	±0.005*	

Time in the Station:	1	2	3	4	0	Sojourn Time
Analytical:	1.278	1.278	1.163	1.163	0.026	4.704
Simulation:	1.303	1.356	1.198	1.239	0.000	4.733
Conf. Int.:	±0.040	±0.044*	±0.036	±0.042*	±0.09	±0.088

@ Note: IN_0 means station 0's input rate, which is determined by equations (9.1) and (9.1').

will go to station 3 with the probability of 0.2; hence P_{23}, the probability that it will go to station 3, is 0.2*0.65. Referring to Chapter III, one must solve the equations $\mathbf{E} = (\mathbf{I} - \mathbf{P}^T)^{-1}\mathbf{Q}$. It can be easily shown that $\mathbf{E} = (0.968, 0.968, 0.645, 0.645)^T$.

Now $e_0 = e_1 + e_2 + e_3 + e_4$, because any job must go to station 0 before it goes to any other station. Note that station 0's mean service time is (0.1 + 0.3)/2 = 0.2, and the waiting time in queue 0 is 0.026, then according to (9.7), the sojourn time is

$$
\begin{aligned}
T_s &= 1.278*0.968 + 1.278*0.968 + 1.163*0.645 + 1.163*0.645 \\
&\quad + (0.968 + 0.968 + 0.645 + 0.645)*(0.2 + 0.026) \\
&= 4.704
\end{aligned}
$$

which is very close to the simulation results.

More directly, taking station 0 in consideration, then one can see that $\mathbf{Q} = (1, 0, 0, 0, 0)^T$ and $\mathbf{E} = (e_0, e_1, e_2, e_3, e_4)^T$. This means a job first enters the central storage anyway. Consequently, the transition matrix $\mathbf{P}$ becomes

$$
\begin{array}{l}
\text{Station 0} \\
\text{Station 1} \\
\text{Station 2} \\
\text{Station 3} \\
\text{Station 4}
\end{array}
\qquad
\begin{bmatrix}
0 & 0.3 & 0.3 & 0.2 & 0.2 \\
0.65 & 0 & 0 & 0 & 0 \\
0.65 & 0 & 0 & 0 & 0 \\
0.75 & 0 & 0 & 0 & 0 \\
0.75 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

Again, equations $\mathbf{E} = (\mathbf{I} - \mathbf{P}^T)^{-1}\mathbf{Q}$ must be solved; doing so, one obtains $\mathbf{E} = (3.226, 0.968, 0.968, 0.645, 0.645)^T$ and therefore Ts = 4.704, verifying the above result.

Now it is straightforward to get Station i's utilization by $U_i = \lambda_i THi\mu_1(Y_i)$ for $1 \le i \le M$ (see section 9.3).

The sojourn time is an important measure in system analysis. As shown in Proposition I, as long as the system is stable and there is no physical blocking, its output rate is equal to the input rate; therefore, people will be more concerned about the production lead time, or the sojourn time, which tells how long a job will stay in the system. However, in the literature, it is overlooked.

Example 9.2: The Effects of λ

The system input rate λ has direct effects on the system performance. Here suppose λ changes from 0.1 to 0.7 in steps of 0.2, while all others are the same as those in the example 9.1. The table below shows the results. Since station 1 and station 2, as well as station 3 and station 4, are "identical" in terms of their parameters, station 2 and station 4 will not be shown hereafter.

The table below clearly shows that the system output rate is always equal to the system input rate λ, as Proposition I predicts. While λ increases, so do stations' input/output rates and b_i's.

An interesting observation is that b_0's are always close to zero. Can it be assumed that $b_0=0$? It appears that this assumption would be valid if every station (queue) is stable, i.e., $\rho_i<1$ (which is the standard assumption). However, it seems that a system could reach steady state when even one or more stations are not "stable," since whenever station 0's buffer is full, no more jobs will be accepted by the system. Unfortunately, there is no guarantee that the system can reach the steady state. therefore, only the systems' steady state performance is considered here, and it will be assumed that $\rho_i<1$ for $i=0, \ldots ,M$.

TABLE VI

THE EFFECTS OF λ

(St=Station; Ana=Analytical; Sim=Simulation; C.I=Confidence Interval; Sys.=System Output rate; IN_0=Station 0's Input Rate.)

λ	0.1			0.3			0.5			0.7		
St1	r_1TH_0	TH_1	b_1	r_1TH_0	TH_1	b_1	r_1TH_0	TH_1	b_1	r_1TH_0	TH_1	b_1
Ana	.097	.097	.00	.290	.290	.00	.500	.476	.047	.699	.668	.09
Sim	.096	.098	.00	.285	.284	.00	.511	.471	.079	.738	.652	.11
C.I	±.29	±.29		±.09	±.09		±.05	±.05	±.03	±.03	±.03	±.04
St3	r_3TH_0	TH_3	b_3	r_3TH_0	TH_3	b_3	r_3TH_0	TH_3	b_3	r_3TH_0	TH_3	b_3
Ana	.065	.065	.00	.194	.194	.00	.333	.333	.00	.466	.465	.045
Sim	.064	.064	.00	.194	.194	.00	.348	.338	.03	.496	.472	.046
C.I	±.42	±.42		±.12	±.12		±.07	±.07	±.04	±.05	±.03	±.01
	IN_0	Sys.	b_0	IN_0	Sys.	b_0	IN_0	Sys.	b_0	IN_0	Sys.	b_0
Ana	.323	.100	.00	.968	.300	.00	1.666	.500	.00	2.331	.700	.00
Sim	.319	.100	.00	.970	.298	.00	1.716	.497	.00	2.461	.695	.00
C.I	±.12	±.12		±.03	±.03		±.015	±.04		±.09	±.03	

The analytical procedure also shows that when λ approaches 0.8, station 1 and station 2's queues are fluctuating dramatically, and there is no evidence of convergence.

<u>The Effects of Local Buffer Capacity</u>

Following example 9.1, let us examine the effects of local buffer capacity changes (all others stay the same except that λ = 0.5). Let **QC** be the local buffer capacity vector. In example 9.1, **QC** = $(3, 3, 3, 3)^T$. In general, the i-th entry of **QC** gives station i's buffer capacity.

According to Tables V and VI, b3 and b_4 are always zero or very close to zero, that may suggest there is no need to increase stations 3 and 4's buffer capacity. Therefore, only stations 1 and 2's buffer capacity will be changed from 2 to 5 in steps of 1. Table VII below show the results.

Table VI clearly shows that <u>both</u> b_1 and b_3 decrease while **QC**(1) and **QC**(2) increase and **QC**(3) and **QC**(4) stay the same. When **QC** = (5, 5, 3, 3), b_1 and b_3 are close to zero, potentially suggesting that there is no need to further increase the buffer capacity.

TABLE VII

THE EFFECTS OF **QC**

QC		Station 1 Input	Station 1 TH_1	Station 1 b_1	Station 3 Input	Station 3 TH_2	Station 3 b_3	IN_0	System Output	b_0
(2, 2, 3, 3)	Ana	.525	.464	.161	.350	.350	.01	1.75	.500	.00
	Sim	.546	.451	.175	.377	.363	.03	1.86	.497	.00
	C.I	±.05	±.05	±.10	±.07	±.07	±.04	±.014	±.04	
(3, 3, 3, 3)	Ana	.500	.476	.047	.333	.333	.00	1.67	.500	.00
	Sim	.511	.471	.079	.348	.338	.02	1.72	.497	.00
	C.I	±.05	±.05	±.03	±.07	±.07	±.03	±.014	±.04	
(4, 4, 3, 3)	Ana	.484	.484	.00	.323	.323	.00	1.61	.500	.00
	Sim	.492	.476	.02	.333	.326	.02	1.66	.497	.00
	C.I	±.05	±.05	±.03	±.07	±.07	±.03	±.020	±.04	
(5, 5, 3, 3)	Ana	.484	.484	.00	.323	.323	.00	1.61	.500	.00
	Sim	.483	.476	.01	.329	.323	.00	1.63	.498	.00
	C.I	±.05	±.05	±.02	±.07	±.07	±.03	±.020	±.04	

The Effects of System Balance

In the above examples, r_i's and f_i's are different, therefore the system is not balanced (see chapter III, and Buzacott and Shanthikumar 1980). It is well accepted that a balanced system will perform better. Therefore, the study will examine what occurs if the system is balanced when, say, all r_i's are 1/4, all f_i's are 0.3, and all n_i's are 3. Also assume λ = 0.5 and 0.7. Table VIII shows the results. Since all four stations are the same, only station 1 and station 0 are listed.

Compared with Table VI, it is found that for the balanced system, all b_i's decrease. Since each station may have different settings, it is meaningless to compare each individual station's performance.

However, if all stations' blocking rates decrease, the entire system's performance improves.

TABLE VIII

BALANCED SYSTEMS

λ	0.5			0.7		
St1	r_1TH_0	TH_1	b_1	r_1TH_0	TH_1	b_1
Ana	.417	.417	.000	.629	.583	.073
Sim	.426	.421	.012	.638	.579	.092
C.I	±.035	±.034	±.021	±.034	±.034	±.037
	IN_0	Sys.	b_0	IN_0	Sys.	b_0
Ana	1.667	.500	.000	2.517	.700	.000
Sim	1.689	.498	.000	2.541	.695	.000
C.I	±.015	±.041		±.088	±.033	

Statistical Aspects of Simulation

Start-Up Policy

As Pritsker (1986) pointed out, the initial conditions for a simulation model may cause the values obtained from the model to be different from those obtained after a start-up period. When steady state performance is to be estimated, the initial responses influence the estimators of steady state performance. Start-up policies are used for setting the initial conditions for the simulation model and specifying a procedure for estimating a truncation point, say Tr, at which sample values should begin to be included in the estimators being computed. Considering the cost of simulation, Tr should be as small as possible, because all sample values collected up to the truncation point are discarded.

In this study, Schriber (1974)'s truncation rule is adopted to monitor the sojourn time. The Schriber truncation rule sets Tr whenever the batch means for the i most recent batches of size b all fall within an interval of length e. This study used i=20, b=20, and e=10%*x', where x' is the mean of the first 100 sample observations at the beginning, and will be updated after every 100 observations.

Stopping Rules

Determining the length of a simulation run as specified in terms of the number of sample observations is a complex problem. Assume n observations are taken to estimate the random variable X. Denote X_n' be the unbiased estimation of μ_X, and $\mathrm{Var}[X_n'] = \sigma_X^2/n$. Then n, the number contained in a prescribed interval can be decided by $\mathrm{Prob}[X_n' - \epsilon \le \mu_X \le X_n' + \epsilon] \ge 1-\alpha$, here ϵ is a prescribed half length for the confidence interval.

Let $Z = \sqrt{n}(X_n' - \mu_X)/\sigma_X$, then $\mathrm{Prob}[|Z| \le \epsilon\sqrt{n}/\sigma_X] \ge 1-\alpha$, and let n^* be the smallest value of n for which the above equation holds. Assuming n^* is large enough so that the central limit theorem applies, it is easy to see that $n^* = [(\sigma_X/\epsilon)Z_{\alpha/2}]^2$, where $Z_{\alpha/2}$ is such that

$$\frac{1}{\sqrt{2\pi}} \int_{Z_{\alpha/2}}^{\infty} (0.5e^{-y^2/2})dy = \alpha/2.$$

Usually ϵ is specified in relative terms of σ_X, that is, $\epsilon = c\sigma_X$ for c>0. Therefore, $n^* = (Z_{\alpha/2}/2)^2$. The following table displays n^* values with respect to the commonly used α-c values. In this study, c=0.02 and α=0.05 were utilized; therefore, the length of a simulation run is 9604, and the actual length is Tr + 9604, where Tr is determined by the Schriber truncation rule.

TABLE IX

NUMBER OF OBSERVATIONS

c \ α	0.02	0.05	0.10
0.01	54093	38416	27060
0.02	13698	9604	6765
0.05	2164	1536	1084
0.10	541	384	271
0.20	135	96	68

CHAPTER X.

CONCLUSIONS AND DISCUSSIONS

This dissertation has investigated flexible manufacturing systems with station breakdowns, material handling system delay, and general processing times.

In this investigation, this dissertation presents a powerful and convenient approach to transform a system with MHS delay and breakdowns into an equivalent system without MHS delay and breakdowns. This transformation is based upon the method of moments. After absorbing the repair times and MHS delay into the effective processing time Y's or generalized processing time GS's, the approaches of Hahn and Shapiro (1968), Kendall and Stuart (1969), and Kottas and Lau (1979, 1980) can be used to fit Y's or GS's first k (k=3 or 4) moments to a k-parameter distribution function. Thus the system with the effective (or generalized) processing times is equivalent to the original system, but without breakdowns and MHS delay. Fitting the first k moments to a k-parameter distribution is a convenient method that provides good approximations.

Moreover, when the processing time distributions are known, the moments of Y's and/or GS's can be calculated analytically. These moments are all that are needed to evaluate each station's performance, such as the total time in the station (station time X) or in its queue (waiting time W), the number of jobs in the station (station size L) or in the queue (queue length Q), and the station output rate (TH). A step-by-step recursive algorithm (Procedure I) is shown to calculate the moments of X, W, L, and Q. Therefore, this transformation facilitates performance evaluations of individual stations, as well as the whole FMS with general processing times. After this transformation, the previously used techniques for the case of no MHS delay and breakdowns can be applied to analyze the system.

Furthermore, to take the advantage of this moments-oriented transformation, this dissertation develops new iterative procedures to obtain performance measures for individual stations as well as the entire FMS's. Procedure I is devised to calculate stations' performance measures, such as the number of jobs in the station or in its queue, and the total time a job stays in the station, and Procedure II is then developed to obtain the FMS's performance measures.

Numerical examples are used to show how to transform an FMS with breakdowns and MHS delay into an equivalent system without breakdown and MHS delay. After the transformation, the effective or generalized processing times are never exponential, no matter whether the original processing times are exponential or not. Then, Procedure II can be used to analyze the system's performance. This procedure is very efficient, and easy to program. Computer simulations are conducted to verify the analytical results.

The analytical results show that when one or more stations are unstable, Procedure II can not converge, while in practice the system can continue operation anyway. Since it is of practical importance to evaluate those "unstable" FMS' performance, investigating this phenomenon is worthy of future research.

As mentioned in chapter IV, if there are two or more classes of jobs in the system, jobs in different classes could have different values (or costs). Furthermore, the same job may have different values (or costs) when it is going through different stations. So (4.12) should be revised accordingly to reflect these considerations. Since this heavily depends on the cost structures of these products, it is not discussed here. However, this is also a topic worthy of future research.

BIBLIOGRAPHY

Altiok, T., "On the phase-type approximations of general distributions," IIE Trans., Vol.17, No.2, (1985), 110-116.

Altiok, T., and Ranjan, R., "Analysis of production lines with general service times and finite buffers: A two-node decomposition approach," IE Report, (1987), Rutgers Univ.

Baskett, F., Chandy, K.M., Muntz, R.R, and Palacios, F.G., "Open, closed and mixed networks of queues with different classes of customers," JACM, Vol.22, No.2, (1975), 248-260.

Brandwajn, A., and Jow, Y.L. "An approximation method for tandem queues with blocking," Operations Research, Vol.36, No.1, (1988), 73-83.

√ Buzacott, J.A., "Automatic transfer lines with buffer stocks," Int. J. Prod. Res., Vol.5, No.3, (1967), 183-200.

------ , "The effects of station breakdowns and random processing times on the capacity of flow lines with in-process storage," AIIE Trans., Vol.4, No.4, (1972), 308-312.

------ , "The production capacity of job shops with limited storage space," Int. J. Prod. Res., Vol.14, No.5, (1976), 597-605.

√ Buzacott, J.A., and Hanifin, L.E. "Model of automatic transfer lines with inventory banks: A review and comparison," AIIE Trans., Vol.10, No.2, (1978), 197-207.

√ Buzacott, J.A., and Shanthikumar, J.G., " Models for understanding flexible manufacturing systems," AIIE Trans., Vol.12, No.4, (1980), 339-350.

√ Buzacott, J.A., and Yao, D.D., "Flexible manufacturing systems: A review of analytical models", Management Science, Vol.32, No.7, (1986), 890-905.

Co, H.C., and Li, G., "A mean value analysis model for job shops and job shop-like systems," Computers ind. Engng, Vol.16, No.1, (1989), 9-18.

√ Commault, C., and Semery, A., "Taking into account delays in buffers for analytical performance evaluation of transfer lines," IIE Trans., Vol.22, No.2, (1990), 133-142.

Dallery, Y., and Stecke, K.E., "On the optimal allocation of servers and workloads in closed queueing networks," Operations Research, Vol.38, No.4, (1990).

Federgruen, A., and Green, L., "Queueing Systems with Service Interruptions," Operations Research, Vol.34, No.5, (1986) 752-768.

Gelenbe, E., "On Approximate Computer System Models," JACM, Vol.22, No.2, (1975), 261-269.

Gershwin, S.B., "An efficient decomposition method for the approximate evaluation of tandem queues with finite storage space and blocking," Operations Research, Vol.35, No.2, (1987), 291-305.

Gershwin, S.B., and Berman, O., "Analysis of transfer lines consisting of two unreliable machines with random processing times and finite storage buffers," AIIE Trans., Vol.13, No.1, (1981), 2-11.

Ghosh, K.B., and Wysk, R.A., "Models for communication system performance in computer integrated manufacturing systems," Computers ind. Engng, Vol.16, No.1, (1989), 55-64.

Goetsch, D.L., Fundamentals of CIM Technology, (1988), Delmar Publisher Inc.

Gordon, W.J., and Newell, G.F., "Closed queueing networks with exponential servers," Operations Research, Vol.15, No.2, (1967), 254-265.

✓ Groover, M.P., Automation, Production Systems, and Computer Integrated Manufacturing, Prentice-Hall, Inc., (1987).

Gross, D., and Harris, G.M., Fundamentals of Queueing Theory, (1985), Wiley.

Hahn, G.J., and Shapiro, S.S., Statistical Methods in Engineering, (1968), Wiley.

Hatvany, J. (editor), World Survey on CAM, (1983), U.K. Butterworths.

Henneke, M.J., and Choi, R.H., "Evaluation of FMS parameters on overall system performance", Computers and ind. Engng, Vol.18, No.1, (1990), 105-110.

Hillier, F.S., and Boling, R.W., "The effect of some design factors on the efficiency of production lines with variable operation times," Journal of Industrial Engineering, Vol.17, No.11, (1966), 651-658.

---- and ---- , "Finite queues in series with exponential or Erlang services -- A numerical approach," Operations Research, Vol.15, No.2, (1967), 286-303.

Hutchinson, G.K., Flexible Manufacturing Systems in the United States, Management Research Center, The University of Wisconsin - Milwaukee, Milwaukee, (1979).

Ingnall, E., and Silver, P.E., "The output of a two-stage system with unreliable machines and limited storage," AIIE Trans. Vol.9, No.2, (1977), 183-188.

Jackson, J.R., "Jobshop-like queueing systems", Management Science, Vol.10, No.1, (1963), 131-142.

Jafari, M.A., and Shanthikumar, J.G. "Exact and approximate Solutions to two-stage transfer lines with general uptime and downtime distributions," IIE Trans., Vol.19, No.4, (1987a), 412-420.

------ and ------, "An approximate model of multistage automatic transfer lines with possible scrapping of workpieces," IIE Trans., Vol.19, No.3, (1987b), 252-265.

Jun, K.P., and Perros, H.G. "An approximate analysis of open tandem queueing networks with blocking and general service times," E.J.O.R., Vol.46, No.1, (1990), 123-135.

Kelly, F.P., Reversibility and Stochastic Networks, (1979), Wiley.

Kendall, M., and Stuart, A., The Advanced Theory of Statistics, (1969), Griffin.

Kleinrock, l., Queueing Systems, (1975), Wiley .

Kottas, J.F., and Lau, H.S., "The use of versatile distribution families in some stochastic inventory calculations," J. Opl. Res. Soc., Vol.31, No.5, (1980), 393-403.

------ and ------ , "A realistic approach for modeling stochastic lead time distributions," AIIE Trans., Vol. 11, No.1, (1979), 54-60.

Lau, H.S. and Martin, G.E., "The effects of skewness and kurtosis of processing times in unpaced lines," Int. J. Prod. Res., Vol.25, No.10, (1987), 1483-1492.

Marie, R.A., "An approximate analytical method for general queueing networks," IEEE, Vol.SE-5, No.5, (1979), 530-538.

Murphy, R.A., "The effect of surge on system availability," AIIE Trans., Vol.7, No.4, (1975), 439-443.

Posner, M., and Bernholtz, B., "Closed finite queuing networks with time lags," Operations Research, Vol.16, No.5, (1968), 962-976.

Pritsker, A.A.B., Introduction To Simulation and SLAM II, (1986), Wiley and Systems Publishing Corp.

Ross, S.M., Introduction to Probability Models, Academic Press, Inc., Harcourt Bruce Jovanovich, Publishers, 4th edition, (1989).

Schweitzer, P.J., "Maximum throughput in finite capacity open queueing networks with product-form solutions", Management Science, Vol.24, No.2, (1977), 217-223.

Sengupta, B., "A Queue with Service Interruptions in an Alternating Random Environment," Operations Research, Vol.38, No.2, (1990) 308-318.

Sheshkin, T.J., "Allocation of interstage storage along an automatic production line," AIIE Trans., Vol.8, No.1, (1976), 146-152.

Shanthikumar, J.G., and Stecke, K.E., "Reducing work-in-process inventory in certain classes of flexible manufacturing systems," E.J.O.R., Vol.26, No.2, (1986), 266-271.

Solberg, J.J., "A mathematical model of computerized manufacturing systems," Proceedings of the 4-th International Conference on Production Research, Tokyo, (1977).

Stecke, E.K., and Morin, T.L., "The optimality of balancing workloads in certain types of flexible manufacturing systems," E.J.O.R., Vol.20, No.1, (1985), 68-82

Stecke, E.K., and Solberg, J.J., "The optimality of unbalancing both workloads and machine group sizes in closed queueing networks of multiserver queues," Operations Research, Vol.33, No.4, (1985), 882-908.

Suri, R., "Robustness of queueing network formulas," JACM, Vol.30, No.3, (1987), 564-494.

Takacs, L., "A single-server queue with Poisson input," Operations Research, Vol.10, No.3, (1962), 388-394.

Whitt, W., "Approximating a point process by a renewal process, I: Two basic methods," Operations Research, Vol.30, No.1, (1982), 125-147.

Wijngraard, J., "The effect of interstage buffer storage on the output of two unreliable production units in series, with different production rates," AIIE Trans., Vol.11, No.1, (1979), 42-47.

Wilks, S.S., Mathematical Statistics, (1962), Wiley.

Yao, D.D., and Buzacott, J.A., "Queueing models for a flexible machining station. Part I: Diffusion approximations; Part II: The Method of Coxian phases," E.J.O.R., Vol.19, No.2, (1985a), 233-252.

------ and ------, "Modeling the performance of flexible manufacturing systems," Int. J. Prod. Res., Vol.23, No.5, (1985b), 945-959.

------ and ------ , "Models of flexible manufacturing systems with limited local buffers," Int. J. Prod. Res., Vol.24, No.1, (1986), 107-118.

Yeralan, S., and Muth, E.J., "A general model of a production line with intermediate buffer and station breakdown," IIE Trans., Vol.19, No.9, (1987), 130-139.

APPENDIX A

THE PROOF OF PROPOSITION I

THE PROOF OF PROPOSITION I

Consider a queueing system consisting of m stations in series. Hillier and Boling (1967) shows that the output rate

$$TH = \mu_m[1 - P_m(0)],$$

where μ_m is the last station m's service rate, $P_m(0)$ is the probability that station m is idle, or starving. Note that $\mu_m[1 - P_m(0)]$ represents the only exit station m's effective output rate, and m will never be blocked. In other words, one can concentrate on the exits and find out $P_i(0)$ so as to obtain TH as long as no blocking occurs. Now consider the open queueing network again. Every job, after being processed in station i, will leave the system with probability of f_i, when i is an exit. Let R_i be the effective output rate of station i, then $R_i = f_i * \delta_i[1 - P_i(0)]$, and

$$TH = \sum_{i=1}^{M} R_i = \sum_{i=1}^{M} \delta_i[1 - P_i(0)]f_i \qquad (11.1)$$

According to Baskett et al. (1975), $P_i(n_i) = (1 - \rho_i)\rho_i^{n_i}$, where n_i is the number of jobs in station i, and $\rho_i = \lambda(e_i/\delta_i)$, if $\rho_i < 1$ for i=1, ... , M for the equilibrium solution to exist. ($\rho_i = \lambda * a_i$ when B=1) Therefore

$$1 - P_i(0) = 1 - (1 - \rho_i) = \rho_i = \lambda(e_i/\delta_i). \qquad (11.2)$$

Note that $e_i = q_i + \sum_{i=1}^{M} e_j P_{ji}$ for i= 1, ... , M, and $\sum_{i=1}^{M} q_i = 1$, I have

$$\sum_{i=1}^{M} e_i = \sum_{i=1}^{M} q_i + \sum_{i=1}^{M} [\sum_{j=1}^{M} e_j P_{ji}]$$

$$= 1 + \sum_{j=1}^{M} [\sum_{i=1}^{M} e_j P_{ji}]$$

$$= 1 + \sum_{j=1}^{M} e_j [\sum_{i=1}^{M} P_{ji}]$$

Therefore,

$$\sum_{j=1}^{M} e_j - \sum_{j=1}^{M} e_j [\sum_{i=1}^{M} P_{ji}] = 1, \text{ or}$$

$$\sum_{j=1}^{M} e_j * f_j = 1. \tag{11.3}$$

The equation (11.3) simply says that any job, once entering the system, will eventually leave the system (with probability 1).

Now return to (11.1). Using (11.2) and (11.3), one can see that

$$TH = \sum_{i=1}^{M} \delta_i [1 - P_i(0)] f_i$$

$$= \sum_{i=1}^{M} \delta_i * \lambda (e_i/\delta_i) f_i \qquad \text{(from 11.2)}$$

$$= \lambda \sum_{i=1}^{M} e_i * f_i$$

$$= \lambda. \qquad \text{(from 11.3)}$$

Because a central storage with virtually infinite capacity can guarantee no-blocking, I get the following proposition.

PROPOSITION I*. For an FMS with general processing times and no blocking, when $\lambda < 1/a_{max}$, (or $\rho_i < 1$ for i = 1,... , M) the output rate is λ. ##

Proposition 1* can be easily extended to $B > 1$. Let f_{ir} be the probability that a class-r job will leave the system after processing at station i. Note that

$$R_i = \sum_{r=1}^{B} h^r f_{ir} \delta_i [1 - P_i(0)], \text{ so}$$

$$TH = \sum_{i=1}^{M} R_i$$

$$= \sum_{i=1}^{M} [\sum_{r=1}^{B} h^r f_{ir} \delta_i (\lambda e_i / \delta_i)]$$

$$= \lambda \sum_{r=1}^{B} [\sum_{i=1}^{M} h^r f_{ir} e_i]$$

$$= \lambda \sum_{r=1}^{B} h^r [\sum_{i=1}^{M} f_{ir} e_i]$$

$$= \lambda \sum_{r=1}^{B} h^r$$

$$= \lambda.$$

This proves proposition I.

APPENDIX B

THE PROOF OF THEOREM I

THE PROOF OF THEOREM 1.

THEOREM 1. Assume two r.v.'s X and W have a joint pdf $f(x,w)$, and their marginal pdf's are $f_X(x)$ and $f_W(w)$, respectively. Then, for $i>0$,

$$E[X^i] = \int_{-\infty}^{\infty} E[X^i|W=w]f_W(w)dw. \qquad (12.4) \quad [\text{i.e., } (5.5)]$$

Proof.

It is known (Kendall and Stuart 1969, Ross 1989) that

$$E[X^i|W=w] = \int_{-\infty}^{\infty} x^i f_{X|W}(x|w)dx, \qquad (12.5)$$

where $f_{X|W}(x|w) = f(x,w)/f_W(w)$ is the conditional pdf of X, given that $W=w$, and is defined for all values of w such that $f_W(w)>0$.

So the right-hand side of (12.4) is

$$\begin{aligned}
\int_{-\infty}^{\infty} E[X^i|W=w]f_W(w)dw &= \int_{-\infty}^{\infty}[\int_{-\infty}^{\infty} x^i f_{X|W}(x|w)dx]f_W(w)dw \\
&= \int_{-\infty}^{\infty}\{\int_{-\infty}^{\infty} x^i[f(x,w)/f_W(w)]dx\}f_W(w)dw \\
&= \int_{-\infty}^{\infty}[\int_{-\infty}^{\infty} x^i f(x,w)dx]dw \\
&= \int_{-\infty}^{\infty} x^i[\int_{-\infty}^{\infty} f(x,w)dw]dx \\
&= \int_{-\infty}^{\infty} x^i f_X(x)dx \\
&= E[X^i],
\end{aligned}$$

that is the left-hand side of (12.4). ##

APPENDIX C

COMPUTER PROGRAM AND PRINTOUT FOR NUMERICAL EXAMPLE 1

```
      REAL*8 SS
      DIMENSION FA(5),WA(5),FB(5),WB(5),CX(12,12),DX(12,12)
      COMMON D(25001),M,YY
      DIMENSION PV(10),DP(10),DQ(10),YM(5),RM(5),XP(5),WS(5)
C
      N3=6
      N4=7
      N3T=2*N3-1
      M3=N3-1
C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     CX AND DX ARE 7X11 TABLES CONTAINING S-D      C
C     DISTRIBUTION'S THIRD AND FOURTH PARAMETERS.   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      READ(5,910) ((CX(I,J),J=1,N4),I=N3,N3T)
      READ(5,910) ((DX(I,J),J=1,N4),I=N3,N3T)
      DO 10 J=1,N4
      DO 10 I=1,M3
      IT=N3T-I+1
      CX(I,J)=CX(IT,J)
   10 DX(I,J)=1.-DX(IT,J)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     XLEMDA IS THE MEAN TIME BETWEEN BREAKDOWNS.   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      XLEMDA=100
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     VECTOR RM CONTAINS THE FIRST FOUR MOMENTS (TO    C
C     ZERO) OF THE REPAIR TIME R.                      C
C     VECTOR FB CONTAINS THE FIRST FOUR MOMENTS        C
C     (CENTRAL) OF R.  HERE EQUATION 3.2 IS APPLIED.   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      RM(1)=5.
      FB(2)=9.
      FB(1)=RM(1)
      RM(2)=FB(2)+FB(1)**2
      IF (WB(2).LT.O.) GO TO 615
      FB(3)=100.
      FB(4)=525.
      RM(3)=FB(3)+3.*FB(1)*FB(2)+FB(1)**3
      RM(4)=FB(4)+4.*FB(1)*FB(3)+6.*FB(1)**2*FB(2)+FB(1)**4
      WRITE(6,66)
   66 FORMAT(1X,'MOMENTS OF R: ')
      WRITE(6,955) (RM(I),I=1,4)
      WRITE(6,70)
   70 FORMAT(1X,'CENTRAL MOMENTS OF R: ')
      WRITE(6,955) (FB(I),I=1,4)
      WRITE(6,69)
   69 FORMAT(1X,'LEMDA=')
      WRITE(6,955) XLEMDA
   68 FORMAT(F6.2,' ')
```

```
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      VECTOR XP CONTAINS THE FIRST FOUR MOMENTS OF   C
C      THE PROCESSING TIME P; CV IS THE COEFFICIENT   C
C      OF VARIATION OF P.                             C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       XP(1)=1.5
       CV=0.6
       SEG=CV*XP(1)
       XP(2)=SEG**2
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      A1 IS ALPHA 1, AND A2 IS ALPHA 2.   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       A2=6
       A1=1.2
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      IK=6 CORRESPONDS TO A1=0, IK=6+1 CORRESPONDS TO   C
C      A1=0+0.4, ETC., AND JK IS A2-1. (ACCORDING TO     C
C      THE STRUCTURES OF TABLEA CX AND DX.)              C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       JK=5
       IK=9
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      FIT A D-S DISTRIBUTION AS PROCESSING   C
C      TIME P'S PDF M(T) REQUIRED IN (5.4),   C
C      (5.5), AND (5.8).  P1, P2, P3 AND P4   C
C      ARE THE FOUR PARAMETERS OF THE D-S     C
C      DISTRIBUTION.                          C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       P3=CX(IK,JK)
       P4=DX(IK,JK)
       DO 400 II=1,4
       WA(II)=0.
   400 WB(II)=0.
       PA=P3+1
       PB=PA**2
       PC=2.*P3+1
       PD=(1.-P4)**PA-P4**PA
       P2=PB*(P4**PC+(1.-P4)**PC)-PC*PD**2
       P2=SEG*SQRT(PC*PB/P2)
       P1=XP(1)-PD*P2/PA
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      BL AND BR ARE THE LEFT ABD RIGHT BOUNDARIES   C
C      OF THE S-D DISTRUBUTION.                      C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       BL=P1-P2*P4**P3
       BR=P1+P2*(1.-P4)**P3
   866 Y1=0.
       WRITE (6,950)
       WRITE(6,951) CV,A1,A2
       XP(3)=XP(2)**1.5*A1
       XP(4)=XP(2)**2.*A2
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      HERE EQUATION 3.2 IS APPLIED.  C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       WS(1)=XP(1)
       WS(2)=XP(2)+XP(1)**2
```

```
      WS(3)=XP(3)+3*XP(2)*XP(1)+XP(1)**3
      WS(4)=XP(4)+4*XP(3)*XP(1)+6*XP(2)*XP(1)**2+XP(1)**4
      WRITE(6,72)
   72 FORMAT(1X,'MOMENTS OF S: ')
      WRITE(6,955) (WS(I),I=1,4)
      WRITE(6,71)
   71 FORMAT(1X,'CENTRAL MOMENTS OF S: ')
      WRITE(6,955) (XP(I),I=1,4)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     THE NUMERICAL INTEGRAL ALGORITHM TO CALCULATE    C
C     KO, K1, K2 AND K3 REQUIRED IN (5.9) AND (5.10).  C
C     SEE (6.2) AND (6.3) FOR THE ALGORITHM.           C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      M=5000
      T=ABS(BR-BL)
      SS=(BR-BL)/M
  343 CALL FDSD(SS,P1,P2,P3,P4,BR,BL)
  833 DO 234 I=1,4
      WA(I)=0.
      TIME=BL
       DO 238 J=1,M
          TIME=TIME+SS
          Z1=TIME**(I-1)*(1-EXP(-TIME/XLEMDA))*D(J)
  238     WA(I)=WA(I)+Z1
  234 CONTINUE
      WRITE (6,814) P1,P2,P3,P4,BR,BL,SS
  814 FORMAT(1X,'P ARE',4F10.5,1X,'BR/BL=',2F10.4,'  SS ',F8.5)
  829 WRITE(6,73)
   73 FORMAT(1X,'INTEGRAL KI: ')
      WRITE(6,955) (WA(I),I=1,4)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     APPLYING (5.9) TO OBTAIN Y'S MOMENTS.   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      YM(1)=WS(1)+RM(1)*WA(1)
      YM(2)=WS(2)+2*RM(1)*WA(2)+RM(2)*WA(1)
      YM(3)=WS(3)+3*RM(1)*WA(3)+3*RM(2)*WA(2)+RM(3)*WA(1)
      YM(4)=WS(4)+4*RM(1)*WA(4)+6*RM(2)*WA(3)+4*RM(3)*WA(2)+RM(4)*WA(1)
      WRITE(6,74)
   74 FORMAT(1X,'MOMENTS OF Y: ')
      WRITE(6,955) (YM(I),I=1,4)
  505 WB(1)=YM(1)
      WB(2)=YM(2)-YM(1)**2
      IF (WB(2).LT.0.) WB(2)=0.
```

```
      WB(3)=YM(3)-3.*YM(1)*YM(2)+2.*YM(1)**3
      WB(4)=YM(4)-4.*YM(1)*YM(3)+6.*YM(1)**2*YM(2)-3.*YM(1)**4
      GT=SQRT(WB(2))
      GCV=GT/WB(1)
      GS=WB(3)/GT**3
      GK=WB(4)/GT**4
      WRITE(6,75)
   75 FORMAT(1X,'CENTRAL MOMENTS OF Y: ')
      WRITE(6,955) (WB(I),I=1,4)
      WRITE(6,76)
   76 FORMAT(1X,'MEAN, CV, SKEWNESS, AND KURTOSIS OF Y: ')
      WRITE(6,955) WB(1),GCV,GS,GK
      GO TO 539
  962 FORMAT (2X,'BL,BR ARE',2F10.5)
  901 FORMAT(8F11.7)
  920 FORMAT (I3,I1,F11.6,3F9.6,F7.4,F8.4,2F7.4,F8.5,I1)
  903 FORMAT(2X,'CT=',F10.8,3X,'C1=',F11.7,3X,'NI=',I4,3X,'CF=',F10.6)
  539 CONTINUE
  500 CONTINUE
      GO TO 615
  904 FORMAT(1X,'P-VALUES ARE',4F12.5)
  910 FORMAT(8F10.7)
  911 FORMAT(7F11.7)
  913 FORMAT(1X,'MEAN,CV,SB1,B2 ARE',4F15.7)
  955 FORMAT(4F20.8)
  950 FORMAT(' ')
  951 FORMAT(1X,'CV,A1,A2 ARE',3F8.2)
  960 FORMAT (1X,'CT,T,SS,M ARE',F15.9,2F15.7,I7)
  615 STOP
      END

*STATISTICS*   SOURCE STATEMENTS = 129, PROGRAM SIZE = 5864 BYTES, PROGRAM NAME = MAIN

*STATISTICS*   NO DIAGNOSTICS GENERATED.

**MAIN** END OF COMPILATION 1 ******
```

```
C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      SUBROUTINE FDSD CALCULATES M(T) IN (6.2) AND   C
C      (6.3) WITH STEP LENGTH SS, FROM BL TO BR.      C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      SUBROUTINE FDSD(SS,P1,P2,P3,P4,BR,BL)
      REAL*8 SS,X2
      COMMON D(25001),M,YY
      P5=1./P3
      DO 2 K=1,M
    2 D(K)=0.
      Y1=0.
      X2=BL
    5 DO 40 K=1,M-1
      X2=X2+SS
      IF (X2.GT.P1) GO TO 30
      Y2=P4-((P1-X2)/P2)**P5
      GO TO 31
   30 Y2=P4+((X2-P1)/P2)**P5
   31 D(K)=Y2-Y1
   40 Y1=Y2
      D(M)=1-D(M-1)
      RETURN
      END

*STATISTICS*   SOURCE STATEMENTS = 19, PROGRAM SIZE = 1292 BYTES, PROGRAM NAME = FDSD

*STATISTICS*   NO DIAGNOSTICS GENERATED.

**FDSD** END OF COMPILATION 2 ******
```

```
MOMENTS OF R:
        5.00000000          34.00000000         360.00000000        4500.00000000
CENTRAL MOMENTS OF R:
        5.00000000           9.00000000         100.00000000         525.00000000
LEMDA=
      100.00000000

CV,A3,A4 ARE    0.60    1.20    6.00
MOMENTS OF S:
        1.50000000           3.05999947           7.89479828          25.18289180
CENTRAL MOMENTS OF S:
        1.50000000           0.81000006           0.87479997           3.93659973
P ARE   1.39450  45.13611    4.63836   0.47117 BR/BL=    3.7451    0.0186  SS  0.00075
INTEGRAL KI:
        0.05160357           0.16335613           0.56675237           2.04813004
MOMENTS OF Y:
        1.75801754           6.44808102          51.63566590         649.21142600
CENTRAL MOMENTS OF Y:
        1.75801754           3.35745621          28.49490360         377.02148400
MEAN, CV, SKEWNESS, AND KURTOSIS OF Y:
        1.75801754           1.04227352           4.63181877          33.44609070
```

APPENDIX D

THE COMPUTER PROGRAM AND PRINTOUT FOR THE NUMERICAL EXAMPLE 9.1

```
      REAL*8 SS
      DIMENSICN Y(5,5),B(5),TH(7),XI(7),CX(12,12),DX(12,12),R(5),CV(5)
      COMMON D(100),THO(7),M,MM,M2,QC(5),IA(7),WA(5),WB(5),P3,P4
      DIMENSICN PV(10),DP(10),DG(10),VL(5,4),X(5,4),W(5,4),F(5),YB(5,5)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     XK IS A SMALL POSITIVE NUMBER TO CONTROL THE ACCURACY.   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      XK=1.0E-4
      K=0
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     N - NUMBER OF STATIONS.    C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      N=4
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     XIN - SYSTEM INPUT RATE.  C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      XIN=0.3
      WRITE (6,951) XIN
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     VECTOR B CONTAINS BLOCKING RATES.    C
C     NOTE: STATICN N+1 IS STATION 0,      C
C     THE CENTRAL STORAGE/MHS, AND         C
C     STATION N+2 IS "OUTPUT" STATION.     C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      B(N+1)=0
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     VECTOR XI CCNTAINS BRANCH STREAMS     C
C     FLOW OUT OF STATION 0, VECTOR THO     C
C     CONTAINS STATIONS INPUT RATE, AND     C
C     VECTOR TH CCNTAINS OUTPUT RATES.      C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      XI(N+2)=XIN
      TH(N+1)=XIN
      THO(N+1)=XIN
      THO(N+2)=XIN
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     CX AND DX ARE 7X11 TABLES CCNTAINING    C
C     S-D DISTRIBUTION'S THIRD AND FOURTH     C
C     MOMENTS PARAMETERS.                     C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      N3=6
      N4=7
      N3T=2*N3-1
      M3=N3-1
      READ(5,910) ((CX(I,J),J=1,N4),I=N3,N3T)
      READ(5,910) ((DX(I,J),J=1,N4),I=N3,N3T)
      DO 10 J=1,N4
      DO 10 I=1,M3
      IT=N3T-I+1
      CX(I,J)=CX(IT,J)
```

```
   10 DX(I,J)=1.-DX(IT,J)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      MATRIX Y CONTAINS THE FIRST FOUR MOMENTS   C
C      OF EACH STATION'S EFFECTIVE PROCESSING     C
C      TIME, VECTOR QC CONTAINS BUFFER CAPACITY,  C
C      VECTOR R CONTAINS BRANCH PROBABILITIES     C
C      RI'S, AND VECTOR F CONTAINS STATION 1-4'S  C
C      LEAVING PROBABILITIES.                     C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      DO 11 I=1,N+1
      IA(I)=0
   11 READ(5,901) (Y(I,J),J=1,4)
      READ(5,811) (QC(I),I=1,N+1)
      WRITE (6,918) (QC(I),I=1,N+1)
      READ(5,911) (R(I),I=1,N)
      WRITE (6,926) (R(I),I=1,N)
      READ(5,911) (F(I),I=1,N)
      WRITE (6,928) (F(I),I=1,N)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      INITIALIZATION: STEP 1 OF PROCEDURE II.  C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      DO 12 I=1,N
      XI(I)=R(I)*TH(N+1)
   12 THO(I)=XI(I)
      DO 14 I=1,N+1
   14 CV(I)=SQRT(Y(I,2))/Y(I,1)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      YB CONTAINS Y'S FIRST FOUR       C
C      MOMENTS ABOUT ZERO.              C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      DO 5 I=1,N+1
      YB(I,2)=Y(I,2)+Y(I,1)**2
      YB(I,3)=Y(I,3)+3.*Y(I,1)*Y(I,2)+Y(I,1)**3
    5 YB(I,4)=Y(I,4)+4.*Y(I,1)*Y(I,3)+6.*Y(I,1)**2*Y(I,2)+Y(I,1)**4
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      CALCULATE FIRST THREE MOMCENTS OF   C
C      X (TIME IN STATION), W (WAITING     C
C      TIME IN THE QUEUE), AND L (LENGTH   C
C      OF THE QUEUE) FOR STATION I.        C
C      SEE PROCEDURE I AND SECTION 8.2.2   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
   54 DO 15 I=1,N+1
      XN=XI(I)
      RO=XN*Y(I,1)
      RN=1-RO
      IF (RN.LE.0) GO TO 22
      W(I,1)=XN*YB(I,2)*0.5/RN
      X(I,1)=W(I,1)+Y(I,1)
      VL(I,1)=XN*X(I,1)
      W(I,2)=2*W(I,1)*W(I,1)+(XN/3)*Y(I,3)/RN
      X(I,2)=W(I,2)+2*W(I,1)*Y(I,1)+YB(I,2)
      VL(I,2)=XN*XN*X(I,2)+VL(I,1)
      W(I,3)=XN*(3*YB(I,2)*W(I,2)/2+YB(I,3)*W(I,1)+YB(I,4)/4)/RN
      X(I,3)=W(I,3)+3*W(I,2)*Y(I,1)+3*W(I,1)*YB(I,2)+YB(I,3)
      VL(I,3)=XN*XN*XN*X(I,3)+3*VL(I,2)-2*VL(I,1)
      GO TO 15
```

```
C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      IA(I)=1 MEANS QUEUE I IS UNSTABLE.   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
  22  IA(I)=1
  15  CONTINUE
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      DETERMINE THE SKEWNESS AND KURTOSIS   C
C      OF THE QUEUE LENGTH DISTRIBUTION.     C
C      SEE SECTION 3.2.                      C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       DO 100 I=1,N+1
       IF (IA(I).GE.1) GO TO 34
       DO 16 J=1,N-1
  16   WA(J)=VL(I,J)
       WB(1)=WA(1)
       WB(2)=WA(2)-WA(1)**2
       IF (WB(2).LE.0) WB(2)=0.2
       WB(3)=WA(3)-3.*WA(1)*WA(2)+2.*WA(1)**3
       B1=WB(3)/WB(2)**1.5
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      ADJUSTMENT FOR SPECIAL CASES.   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       IF (B1.LT.-2) B1=-1.9999
       IF (B1.GT.2) B1=1.9999
       IC=6
       IB=B1/0.4+6
       P3=CX(IC,IB)
       XB=(IB-6)*0.4
       JB=IB
       IF (B1.LT.XB) JB=IB-1
       IF (B1.GT.XB) JB=IB+1
       IF (P3.GT.0) GO TO 20
       IF (IB.GT.6) GO TO 18
       DO 17 KK=IB+1,6
       P3=CX(IC,KK)
       IF (P3.GT.0) GO TO 20
  17   CONTINUE
  18   DO 19 KK=IB-1,6,-1
       P3=CX(IC,KK)
       IF (P3.GT.0) GO TO 20
  19   CONTINUE
  20   PX=CX(IC,JB)
       IF (PX.LE.0) GO TO 21
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      SIMPLE INTERPOLATION IF NECESSARY TO   C
C      DETERMINE P3 AND P4 FOR QUEUE LENGTH   C
C      DISTRIBUTION.                          C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       P3=P3+(PX-P3)*ABS((IB-6)*0.4-B1)/0.4
  21   P4=DX(IC,IB)
  23   CONTINUE
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      DETERMINE P1 AND P2 FOR QUEUE LENGTH   C
C      DISTRIBUTION.                          C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
```

```
   26  ZP=DX(IC,JB)
   27  PA=P3+1
       PB=PA**2
       PC=2.*P3+1
       PD=(1.-P4)**PA-P4**PA
       P2=PB*(P4**PC+(1.-P4)**PC)-PC*PD**2
       PW=PC*PB/P2
       IF (PW.LT.0.) PW=0.
       P2=CV(I)*SQRT(PW)
       P1=WB(1)-PD*P2/PA
       BR=P1+P2*(1.-P4)**P3
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      CALCULATE THE BLOCKING RATE. (SEE SECTIO 8.4.)  C
C      HERE CDF IS GIVEN IN EXAMPLE 1 (SECTION 6.1.1). C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       B(I)=0.
       IG=QC(I)
       IF (IG.GE.BR) GO TO 100
       RVC=1/P3
       IF (IG.GT.P1) GO TO 30
       B(I)=P4-((P1-IG)/P2)**RVC
       GO TO 32
   30  B(I)=P4+((IG-P1)/P2)**RVC
   32  B(I)=1-B(I)
       GO TO 100
   34  TH(I)=1/Y(I,1)
       B(I)=(XI(I)-TH(I))/XI(I)
  100  CONTINUE
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      CALCULATE EACH STATION'S OUTPUT RATE    C
C      AND INPUT RATE (STEP 3 OF PROCEDURE II) C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       Z1=0
       DO 40 I=1,N
   40  Z1=Z1+R(I)*(B(I)+(1-B(I))*(1-F(I)))
       TH(N+1)=XIN*(1-B(N+1))/(1-Z1)
       DO 50 I=1,N
       IF (IA(I).GE.1) GO TO 50
       TH(I)=TH(N+1)*R(I)*(1-B(I))
   50  XI(I)=R(I)*TH(N+1)
       XI(N+2)=0
       DO 60 I=1,N
       IA(I)=0
   60  XI(N+2)=XI(N+2)+TH(I)
       IA(N+1)=0
       TH(N+2)=XI(N+2)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C      CALCULATE LAPDA-F AND LAMCA-B.      C
C      STEP 4 OF PRCCEDURE II.             C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
       FLD=0
       FB=0
       DO 65 I=1,N
       FLD=FLD+TH(I)*(1-F(I))
   65  FB=FB+XI(I)*B(I)
       XI(N+1)=FLD+FB+XIN*(1-B(N+1))
```

```
          B(N+1)=1-(TH(N+1)-FLD-FB)/XIN
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C         CHECK THE CONVERGENCE AND DETERMINE   C
C         IF MORE ITERATIONS ARE NECESSARY.     C
C         STEP 6 OF PROCEDURE II.               C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
          DO 70 I=1,N+2
          AX=ABS(TH(I)-THO(I))
          IF (AX.GT.XK) GO TO 80
   70     CONTINUE
          GO TO 90
   80     K=K+1
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C         THE MAXIMUM NUBER OF ITERATIONS IS 800.        C
C         IF IT IS NOT CONVERGE AFTER 800 ITERATIONS,    C
C         THE PROCEDURE WILL BE FORCED TO STOP.          C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
          IF (K.GT.800) GO TO 90
          DO 85 I=1,N+2
   85     THO(I)=TH(I)
          GO TO 54
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C         PRINT OUT EACH STATION'S INPUT RATE, OUTPUT    C
C         RATE, BLOCKING RATE, MOMENTS OF WAITING TIME   C
C         AND QUEUE LENGTH, AND ACCURACY MESSAGE.        C
C         OUTPUT STEP OF PROCEDURE II.                   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
   90     WRITE (6,916)
          WRITE(6,913) (TH(I),I=1,N+2)
          WRITE (6,917)
          WRITE(6,913) (THO(I),I=1,N+2)
          WRITE (6,912)
          WRITE(6,913) (B(I),I=1,N+1)
          WRITE (6,822)
          WRITE(6,913) (W(I,1),I=1,N+1)
          WRITE (6,823)
          WRITE(6,913) (VL(I,1),I=1,N+1)
          WRITE (6,914)
          WRITE(6,913) (XI(I),I=1,N+2)
          WRITE (6,916)
          WRITE(6,913) (TH(I),I=1,N+2)
          GO TO 615
   901    FORMAT(4F10.4)
   910    FORMAT(8F10.7)
   911    FORMAT(4F7.4)
   811    FORMAT(5F7.4)
   913    FORMAT(1X,6F15.7)
   955    FORMAT(4F20.8)
   950    FORMAT(' ')
   912    FORMAT(1X,'BLOKING RATE:')
   822    FORMAT(1X,'WAITING TIME IN THE GUEUE:')
   914    FORMAT(1X,'INPUT RATE:')
   823    FORMAT(1X,'GUEUE LENGTH:')
   916    FORMAT(1X,'CUTPUT TH:')
   917    FORMAT(1X,'CLTPUT THO:')
   918    FORMAT(1X,'BLFFER CAPACITY:',5F7.2)
```

```
      B(N+1)=1-(TH(N+1)-FLD-FB)/XIN
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     CHECK THE CONVERGENCE AND DETERMINE   C
C     IF MORE ITERATIONS ARE NECESSARY.     C
C     STEP 6 OF PROCEDURE II.               C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      DO 70 I=1,N+2
      AX=ABS(TH(I)-TH0(I))
      IF (AX.GT.XK) GO TO 80
   70 CONTINUE
      GO TO 90
   80 K=K+1
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     THE MAXIMUM NUBER OF ITERATIONS IS 800.       C
C     IF IT IS NOT CONVERGE AFTER 800 ITERATIONS,   C
C     THE PROCEDURE WILL BE FORCED TO STOP.         C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      IF (K.GT.800) GO TO 90
      DO 85 I=1,N+2
   85 TH0(I)=TH(I)
      GO TO 54
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     PRINT OUT EACH STATION'S INPUT RATE, OUTPUT    C
C     RATE, BLOCKING RATE, MOMENTS OF WAITING TIME   C
C     AND QUEUE LENGTH, AND ACCURACY MESSAGE.        C
C     OUTPUT STEP OF PROCEDURE II.                   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
   90 WRITE (6,916)
      WRITE(6,913) (TH(I),I=1,N+2)
      WRITE (6,917)
      WRITE(6,913) (TH0(I),I=1,N+2)
      WRITE (6,912)
      WRITE(6,913) (B(I),I=1,N+1)
      WRITE (6,822)
      WRITE(6,913) (W(I,1),I=1,N+1)
      WRITE (6,823)
      WRITE(6,913) (VL(I,1),I=1,N+1)
      WRITE (6,914)
      WRITE(6,913) (XI(I),I=1,N+2)
      WRITE (6,916)
      WRITE(6,913) (TH(I),I=1,N+2)
      GO TO 615
  901 FORMAT(4F10.4)
  910 FORMAT(8F10.7)
  911 FORMAT(4F7.4)
  811 FORMAT(5F7.4)
  913 FORMAT(1X,6F15.7)
  955 FORMAT(4F20.8)
  950 FORMAT(' ')
  912 FORMAT(1X,'BLOKING RATE:')
  822 FORMAT(1X,'WAITING TIME IN THE QUEUE:')
  914 FORMAT(1X,'INPUT RATE:')
  823 FORMAT(1X,'QUEUE LENGTH:')
  916 FORMAT(1X,'OUTPUT TH:')
  917 FORMAT(1X,'OUTPUT TH0:')
  918 FORMAT(1X,'BUFFER CAPACITY:',5F7.2)
  926 FORMAT(1X,'DISPER. PROB.:',4F6.2)
  928 FORMAT(1X,'LEAV. PROB.:',4F6.2)
  951 FORMAT(1X,'LAMDA IS',1F8.2)
  615 STOP
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     THE FOLLOWING DATA FORM TABLE CX AND TABLE DX.  C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      END
```

LAMDA IS 0.30
BUFFER CAPACITY: 3.00 3.00 3.00 3.00 10.00
DISPER. PROB.: 0.30 0.30 0.20 0.20
LEAV. PROB.: 0.35 0.35 0.25 0.25

OUTPUT TH:

0.2903225	0.2903225	0.1935483	0.1935483	0.9677417	0.9677416

OUTPUT THO:

0.2903225	0.2903225	0.1935483	0.1935483	0.9677417	0.9677416

BLOKING RATE:

0.0000000	0.0000000	0.0000000	0.0000000	0.0000000

WAITING TIME IN THE QUEUE:

0.2781816	0.2781816	0.1631998	0.1631998	0.0259800
0.2018971	0.2018971	0.0809163	0.0809163	0.0013499
0.9621100	0.9621100	0.4382698	0.4382698	0.0011339

QUEUE LENGTH:

0.3710847	0.3710847	0.2251353	0.2251353	0.2186902
0.5496266	0.5496266	0.2913405	0.2913405	0.2702382
1.0322027	1.0322027	0.4511033	0.4511033	0.3871995

INPUT RATE:

0.2903225	0.2903225	0.1935483	0.1935483	0.9677415	0.9677416
3.4444447	3.4444447	5.1666679	5.1666679	1.0333328	1.0333328

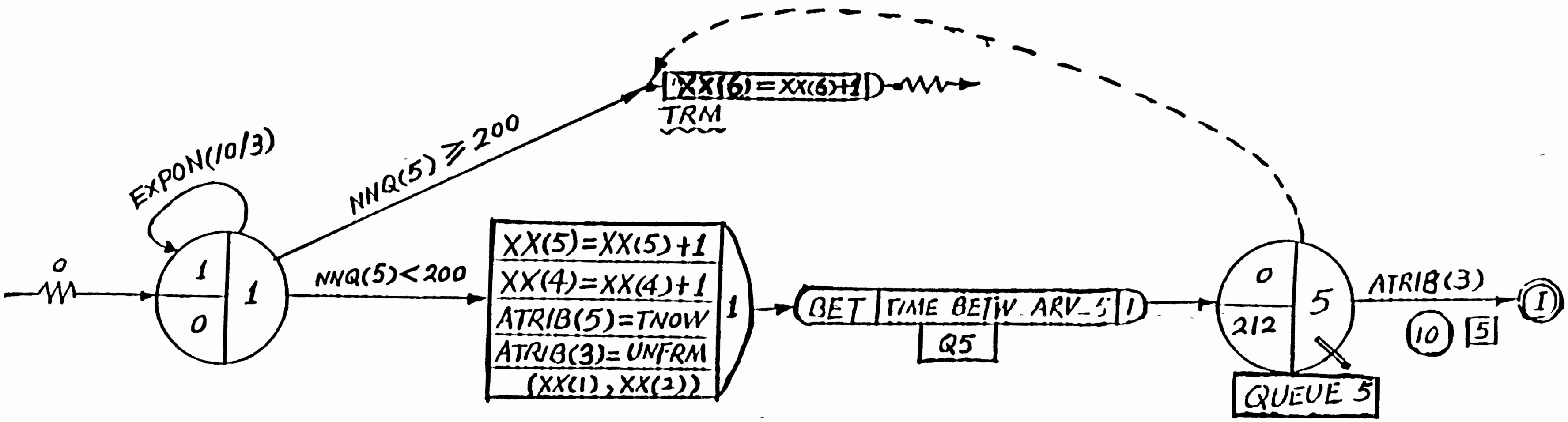
OUTPUT TH:

0.2903225	0.2903225	0.1935483	0.1935483	0.9677417	0.9677416
3.4444447	3.4444447	5.1666679	5.1666679	1.0333328	1.0333328

APPENDIX E

SIMULATION MODELS AND SAMPLE PRINTOUT

SLAM II NETWORK MODEL: AN FMS WITH FOUR STATIONS (Page 1; Total 3 Pages)

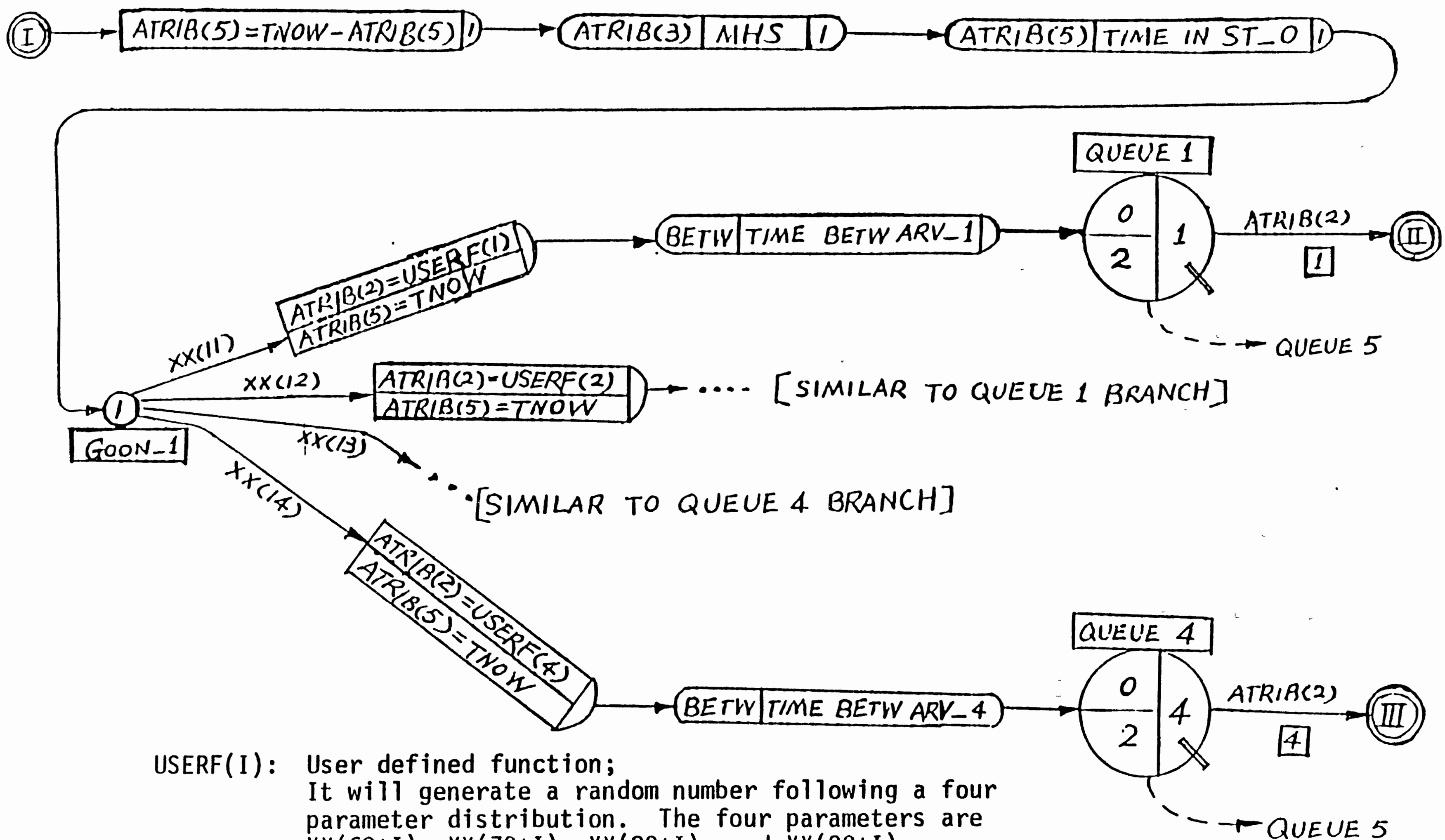


Queues 1 - 4: Stations 1 - 4.
Queue 5: Central Storage & MHS.

TNOW: Present Time.
NNQ(I): Number of Jobs in Queue I.

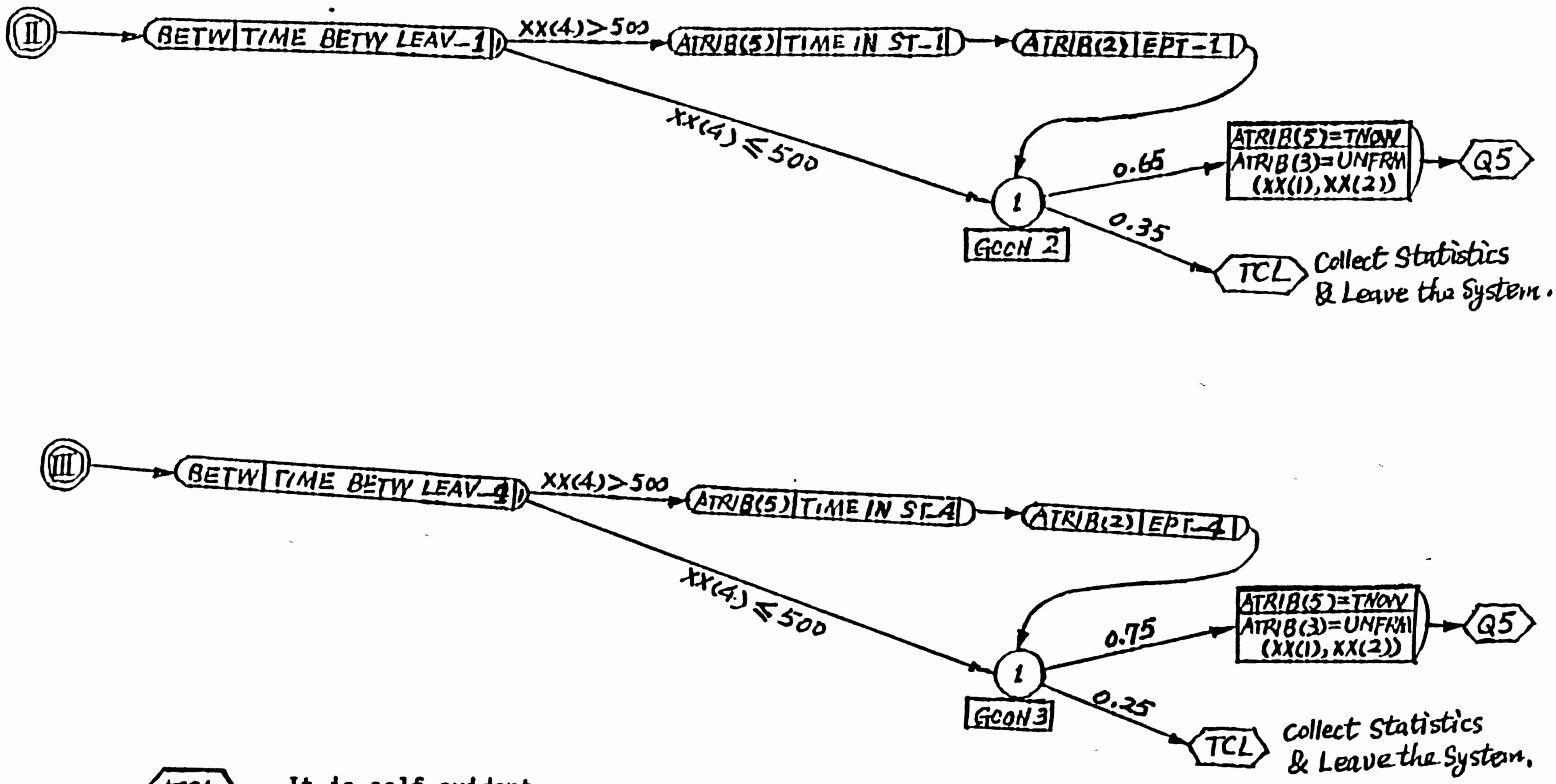
MHS Delay: UNFRM(XX(1),XX(2)), where XX(1)=0.1 (Lower Bound) and XX(2)=0.3 (Upper Bound), and UNFRM means uniform distributed.

XX(4): Number of Arrivals.
XX(5): Number of Jobs in the System.
XX(6): Number of Jobs Rejected by the System.
Vector ATRIB describes a job's attributes:
ATRIB(1) - Arrival Time. ATRIB(2) - Processing Time. ATRIB(3) - MHS Delay.
ATRIB(5) - Time in the Station.



USERF(I): User defined function;
It will generate a random number following a four parameter distribution. The four parameters are XX(60+I), XX(70+I), XX(80+I), and XX(90+I).

SLAM II NETWORK MODEL: AN FMS WITH FOUR STATIONS (Page 2; Total 3 Pages)



TCL: It is self-evident.
See the SLAM II printout.

SLAM II NETWORK MODEL: AN FMS WITH FOUR STATIONS (Page 3; Total 3 Pages)

U14345A.SIMX2.DATA
VPSPRINT 5.1.002 MONDAY OCTOBER 28,1991 15:15:58 U14345A MVS1
VPSPRINT SIMX2.DATA LOCAL FORMS(9001)

1

SLAM II VERSION 4.03

PRITSKER AND ASSOCIATES, INC.
P.O. BOX 2413
WEST LAFAYETTE, INDIANA 47906
(317)463-5557

1

```
GEN,ZHAO,Dissertation,10/25/91,1,,NO,,NO;
LIMIT,5,9,1200;
INTLC,XX(4)=0,XX(5)=0,XX(6)=0;
INTLC,XX(61)=1.68141,XX(71)=1.19402,XX(81)=0.22996,XX(91)=0.83996;
INTLC,XX(62)=1.68141,XX(72)=1.19402,XX(82)=0.22996,XX(92)=0.83996;
INTLC,XX(63)=1.68141,XX(73)=1.19402,XX(83)=0.22996,XX(93)=0.83996;
INTLC,XX(64)=1.68141,XX(74)=1.19402,XX(84)=0.22996,XX(94)=0.83996;
INTLC,XX(1)=0.1,XX(2)=0.3;
INTLC,XX(11)=0.3,XX(12)=0.3,XX(13)=0.2,XX(14)=0.2;
TIMST,XX(5),SYSTEM SIZE;
TIMST,XX(6),BLOCK FROM SYSTEM;
NETWORK;
          CREATE,EXPON(3.3333,7),,1,,1;
          ACT,,NNQ(5).GE.200,TRM;
          ACT,,NNQ(5).LT.200;
          ASSIGN,XX(5)=XX(5)+1,XX(4)=XX(4)+1,1;
Q5        ASSIGN,ATRIB(3)=UNFRM(XX(1),XX(2),3),ATRIB(5)=TNOW,1;
Q5A       COLCT,BETWEEN,TIME BETW ARV_5;
          QUEUE(5),,212,BALK(TRM);
          ACT(10)/5,ATRIB(3);
          ASSIGN,ATRIB(5)=TNOW-ATRIB(5),1;
          COLCT,ATRIB(3),MHS;
          COLCT,ATRIB(5),TIME IN ST_0;
          GOON,1;
          ACT,,XX(11),Q1;
          ACT,,XX(12),Q2;
          ACT,,XX(13),Q3;
          ACT,,XX(14),Q4;
Q1        ASSIGN,ATRIB(2)=USERF(1),ATRIB(5)=TNOW,1;
          COLCT,BETWEEN,TIME BETW ARV_1;
Q1A       QUEUE(1),,2,BALK(Q5);
          ACT/1,ATRIB(2);
          COLCT,BETWEEN,TIME BTW LEAV_1;
          GOON,1;
          ACT,,XX(4).LE.500,N1;
          ACT,,XX(4).GT.500;
          ASSIGN,ATRIB(5)=TNOW-ATRIB(5),1;
          COLCT,ATRIB(5),TIME IN ST_1;
          COLCT,ATRIB(2),EPT_1;
N1        GOON,1;
          ACT,,0.65,Q5;
          ACT,,0.35,TCL;
Q2        ASSIGN,ATRIB(2)=USERF(2),ATRIB(5)=TNOW,1;
          COLCT,BETWEEN,TIME BETW ARV_2;
Q2A       QUEUE(2),,2,BALK(Q5);
          ACT/2,ATRIB(2);
          COLCT,BETWEEN,TIME BTW LEAV_2;
          GOON,1;
          ACT,,XX(4).LE.500,N2;
          ACT,,XX(4).GT.500;
          ASSIGN,ATRIB(5)=TNOW-ATRIB(5),1;
          COLCT,ATRIB(5),TIME IN ST_2;
          COLCT,ATRIB(2),EPT_2;
N2        GOON,1;
```

```
        ACT,,0.65,O5;
        ACT,,0.35,TCL;
O3      ASSIGN,ATRIB(2)=USERF(3),ATRIB(5)=TNOW,1;
        COLCT,BETWEEN,TIME BETW ARV_3;
O3A     QUEUE(3),,2,BALK(O5);
        ACT/3,ATRIB(2);
        COLCT,BETWEEN,TIME BTW LEAV_3;
        GOON,1;
        ACT,,XX(4).LE.500,N3;
        ACT,,XX(4).GT.500;
        ASSIGN,ATRIB(5)=TNOW-ATRIB(5),1;
        COLCT,ATRIB(5),TIME IN ST_3;
        COLCT,ATRIB(2),EPT_3;
N3      GOON,1;
        ACT,,0.75,O5;
        ACT,,0.25,TCL;
O4      ASSIGN,ATRIB(2)=USERF(4),ATRIB(5)=TNOW,1;
        COLCT,BETWEEN,TIME BETW ARV_4;
O4A     QUEUE(4),,2,BALK(O5);
        ACT/4,ATRIB(2);
        COLCT,BETWEEN,TIME BTW LEAV_4;
        GOON,1;
        ACT,,XX(4).LE.500,N3;
        ACT,,XX(4).GT.500;
        ASSIGN,ATRIB(5)=TNOW-ATRIB(5),1;
        COLCT,ATRIB(5),TIME IN ST_4;
        COLCT,ATRIB(2),EPT_4;
N4      GOON,1;
        ACT,,0.75,O5;
        ACT,,0.25,TCL;
TCL     GOON,1;
        ACT,,XX(4).LT.500,TM;
        ACT,,XX(4).GE.500;
        COLCT,BETWEEN,TIME BETW APART;
        COLCT,INT(1),TIME IN SYSTEM;
        ACT;
TM      ASSIGN,XX(5)=XX(5)-1;
        TERM,10000;
TRM     ASSIGN,XX(6)=XX(6)+1;
        TERM;
        END;
FIN;
```

1

```
      PROGRAM MAIN
      COMMON/SCOM1/ATRIB(100),DD(100),DDL(100),DTNOW,II,MFA,
     1 MSTOP,NCLNR,NCRDR,NPRNT,NNRUN,NNSET,NTAPE,SS(100),SSL(100),
     2 TNEXT,TNOW,XX(100)
      DIMENSION NSET(30000)
      COMMON QSET(30000)
      EQUIVALENCE (NSET(1),QSET(1))
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     THIS HEAD IS REGUIRED BY THE SLAM II SYSTEM   C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      NNSET=30000
      NCRDR=5
      NPRNT=6
      NTAPE=1
      CALL SLAM
      STOP
      END
C
C
      FUNCTION USERF(I)
      COMMON/SCOM1/ATRIB(100),DD(100),DDL(100),DTNOW,II,MFA,MSTOP,
     1 NCLNR,NCRDR,NPRNT,NNRUN,NNSET,NTAPE,SS(100),SSL(100),TNEXT,
     2 TNOW,XX(100)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C     THIS FUNCTION GENERATES A RANDOM SAMPLING   C
C     FROM S-D DISTRIBUTION WITH 4 PARAMETERS     C
C     GIVEN AS XX(60+I), XX(70+I), XX(80+I),      C
C     AND XX(90+I).                               C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C
      RX=UNFRM(0.0,1.0,2)
      IF (RX.LT.XX(I+90)) GO TO 17
      RX=RX-XX(I+90)
      RX=XX(I+60)+XX(I+70)*RX**XX(I+80)
      GO TO 18
 17   RX=XX(I+90)-RX
      RX=XX(I+60)-XX(I+70)*RX**XX(I+80)
 18   IF (RX.GT.0.) GO TO 19
      RX=0.005
 19   USERF=RX
      RETURN
      END
```

S L A M   I I   S U M M A R Y   R E P O R T

SIMULATION PROJECT Dissrtation BY ZHAO

DATE 10/25/1991 RUN NUMBER 1 OF 1

CURRENT TIME 0.3513E+05
STATISTICAL ARRAYS CLEARED AT TIME 0.0000E+00

STATISTICS FOR VARIABLES BASED ON OBSERVATION

	MEAN VALUE	STANDARD DEVIATION	COEFF. OF VARIATION	MINIMUM VALUE	MAXIMUM VALUE	NUMBER OF OBSERVATIONS
TIME BETW ARV_5	0.1032E+01	0.1577E+01	0.1528E+01	0.0000E+00	0.4067E+02	34033
MHS	0 1995E+00	0.5761E-01	0.2888E+00	0.1000E+00	0.3000E+00	34034
TIME IN ST_0	0.1995E+00	0.5761E-01	0.2888E+00	0.9961E-01	0.3008E+00	34034
TIME BETW ARV_1	0.3477E+01	0.4546E+01	0.1307E+01	0.0000E+00	0.6057E+02	10099
TIME BTW LEAV_1	0.3524E+01	0.4507E+01	0.1279E+01	0.5342E+00	0.5898E+02	9966
TIME IN ST_1	0.1304E+01	0.9698E+00	0.7437E+00	0.5342E+00	0.1472E+02	9480
EPT_1	0.1019E+01	0.7034E+00	0.6900E+00	0.5343E+00	0.1472E+02	9480
TIME BETW ARV_2	0.3434E+01	0.4374E+01	0.1274E+01	0.0000E+00	0.4678E+02	10225
TIME BTW LEAV_2	0.3490E+01	0.4345E+01	0.1245E+01	0.5342E+00	0.4667E+02	10061
TIME IN ST_2	0.1303E+01	0.9910E+00	0.7605E+00	0.5342E+00	0.1594E+02	9623
EPT_2	0.1016E+01	0.7133E+00	0.7022E+00	0.5343E+00	0.1594E+02	9623
TIME BETW ARV_3	0.5124E+01	0.6433E+01	0.1256E+01	0.0000E+00	0 6955E+02	6856
TIME BTW LEAV_3	0.5149E+01	0.6409E+01	0.1245E+01	0.5347E+00	0.6910E+02	6823
TIME IN ST_3	0.1181E+01	0.8813E+00	0.7461E+00	0.5342E+00	0.1667E+02	6497
EPT_3	0.1007E+01	0.7090E+00	0.7044E+00	0.5343E+00	0.1667E+02	6497
TIME BETW ARV_4	0.5127E+01	0.6514E+01	0.1270E+01	0.3906E-02	0.9738E+02	6850
TIME BTW LEAV_4	0.5150E+01	0.6481E+01	0.1258E+01	0.5352E+00	0.7779E+02	6819
TIME IN ST_4	0.1219E+01	0.9455E+00	0.7759E+00	0.5342E+00	0.1696E+02	6515
EPT_4	0.1030E+01	0.7625E+00	0.7405E+00	0.5344E+00	0.1696E+02	6515
TIME BETW APART	0.3355E+01	0.3357E+01	0.1000E+01	0.7324E-03	0.5442E+02	10004
TIME IN SYSTEM	0.4699E+01	0.4243E+01	0.9030E+00	0.6406E+00	0.4032E+02	10005

STATISTICS FOR TIME-PERSISTENT VARIABLES

	MEAN VALUE	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	TIME INTERVAL	VALUE
SYSTEM SIZE	0.1403E+01	0.1292E+01	0.0000E+00	0 9000E+01	0.3513E+05	0.0000E+00
BLOCK FROM SYSTE	0.0000E+00	0.0000E+00	0.0000E+00	0.0000E+00	0.3513E+05	0.0000E+00

FILE STATISTICS

FILE NUMBER	LABEL/TYPE		AVERAGE LENGTH	STANDARD DEVIATION	MAXIMUM LENGTH	CURRENT LENGTH	AVERAGE WAITING TIME
1	Q1A	QUEUE	0 0817	0.3177	2	0	0 2879
2	Q2A	QUEUE	0 0818	0.3205	2	0	0.2856
3	Q3A	QUEUE	0 0336	0.2009	2	0	0 1731
4	Q4A	QUEUE	0.0366	0.2099	2	0	0 1883
5		QUEUE	0 0000	0 0000	0	0	0 0000
6		CALENDAR	2.1693	0.9877	8	0	0 3470

SERVICE ACTIVITY STATISTICS

ACTIVITY INDEX	START NODE OR ACTIVITY LABEL		SERVER CAPACITY	AVERAGE UTILIZATION	STANDARD DEVIATION	CURRENT UTILIZATION	AVERAGE BLOCKAKE	MAXIMUM IDLE TIME/SERVERS	MAXIMUM BUSY TIME/SERVERS	ENTITY COUNT
1	Q1A	QUEUE	1	0.2898	0.4537	0	0.0000	58 4180	20.7305	9967
2	Q2A	QUEUE	1	0.2905	0.4540	0	0.0000	46.0713	23 6006	10062
3	Q3A	QUEUE	1	0.1960	0.3970	0	0 0000	68 5537	18.7676	6824
4	Q4A	QUEUE	1	0.1998	0.3999	0	0.0000	77.0859	19.1084	6820
5		QUEUE	10	0.1933	0.4374	0	0.0000	10.0000	5.0000	34034

2

VITA

Long-Geng Zhao

Candidate for the Degree of

Doctor of Philosophy

Thesis: PERFORMANCE EVALUATION OF FLEXIBLE MANUFACTURING SYSTEMS WITH STATION BREAKDOWNS, MATERIAL HANDLING SYSTEM DELAY, AND GENERAL PROCESSING TIMES

Major Field: Business Administration

Biographical:

Personal Data: Born in Shanghai, China, March 1, 1950, the son of Yan Zhao and Yueming Zhang.

Education: Graduated from Shanghai Second Polytechnic University majoring in Computer Engineering in June, 1982; received Master of Science degree in Computer Science from Fudan University, China, in July, 1985; received Master of Business Administration from Oklahoma State University in December, 1991; enrolled in doctoral program in January, 1987, and completed requirements for the Doctor of Philosophy degree at Oklahoma State University in July, 1992.

Professional Experience: Lecturer, Computer Science Department, Fudan University, 1985-86; graduate research assistant, Management Department, Oklahoma State University, 1987-88; graduate teaching associate/instructor, Department of Management, Oklahoma State University, 1988-91; five papers published in major academic journals, 1987-92; presented three papers in the national meetings of Decision Sciences Institute (DSI, 1990/1991) and the Institute of Management Science (TIMS, 1991); member of DSI and TIMS since 1987.