

# CONTENTS

## PREFACE

i

## CHAPTER ONE

### DEFINITIONS AND SOME CONCEPTS FOR THE QUEUEING SYSTEMS

1.1	Introduction	1
1.2	Probabilistic Queueing Models	8
1.2.1	Model (I) Single Channel Poisson Arrivals with Exponential Service, Infinite Population: Model [(M/M/1): (FCFS/ $\infty/\infty$ )]	9
1.2.2	Model (II) Multi Channel Queueing Theory: Model [(M/M/C): (FCFS/ $\infty/\infty$ )]	10
1.3	Queueing Network Modeling	11
1.3.1	Capacity Of Queueing Networks with Batch Servers	13
1.4	Simulation of Queueing System	18
1.5	Simulation Techniques	23
1.5.1	Simulation of A Single Server Queue	25
1.6	Efficient Parallel Queueing System Simulation	30

## CHAPTER TWO

### OPTIMAL TREATMENT OF QUEUEING MODEL FOR HIGHWAY

۲,۱	Introduction	۳۳
۲,۲	Description of Queueing Model with Traffic Flow Theory	۳۵
۲,۳	Analysis of M/M/۱ Model	۳۹
۲,۴	Analysis of M/G/۱ Model	۴۲
۲,۵	Analysis of G/G/۱ Model	۵۱
۲,۶	Newtown - Raphson Method	۵۳
۲,۷	<b>Main Results and Numeric Applications</b>	۵۴
۲,۷,۱	Analyzing for Highway of the Model G/G/۱	۵۴
۲,۷,۲	Analyzing for Highway of the Model M/G/۱	۵۷
۲,۸	Conclusion	۶۰

## CHAPTER THREE

### OPTIMAL TREATMENT OF QUEUEING SYSTEM FOR SUPPLY CHAIN IN MULTIPRODUCT

۳.۱	Introduction	۶۱
۳,۲	Description of Inventory Queueing System	۶۲
۳,۲,۱	The Bulk Queue $M^{Cj}/M/1$	۶۳
۳.۲,۲	Notations	۶۷
۳,۲,۳	Definition	۶۹
۳,۲,۴	Assumptions	۷۱
۳,۳	Modeling of The Inventory Queueing System	۷۵
۳,۴	Model Performance Measures	۷۹
۳,۴,۱	Total Cost and workload	۸۳
۳,۵	Numerical Example and Discussion	۸۴
۳,۵,۱	Computing the Characteristics of Inventory Queueing System $M^{Cj}/M/1$	۸۵
۳,۶	Main Results	۸۷
۳,۷	Conclusion	۹۲

## CHAPTER FOUR

### REDUCING CONSUMED TIME IN QUEUEING SYSTEM FOR RUNWAY

ॡ, 1	Introduction	93
ॡ, 2	Characteristics of a Queueing System	94
ॡ, 2, 1	Arrival Characteristics	95
ॡ, 2, 2	Waiting Line Characteristics	97
ॡ, 2, 3	Service facility characteristics	98
ॡ, 3	Multiterminal Markovian Queueing Model with Poisson Arrivals and Exponential Service Times (m/m/s)	98
ॡ, 3, 1	Equations for the Multiterminal Queueing Model	99
ॡ, 3, 2	Cost Strategy Analysis	102
ॡ, 3, 3	Efficiency of M/M/S queueing model	104
ॡ, 4	Numerical Example and Discussion	105
ॡ, 4, 1	Computing the Characteristics of Queueing System	105
ॡ, 5	Main Result	106
ॡ, 6	Conclusions	110

## **CHAPTER FIVE**

### **ANALYSIS PERFORMANCE MEASURES FOR QUEUEING NETWORK IN MANUFACTURING SYSTEM**

۵,۱	Introduction	۱۱۲
۵,۲	Description of Queueing Network Model	۱۱۳
۵,۳	Basic Assumptions of Queueing Network	۱۱۵
۵,۴	Analyzing of Queueing Network Model	۱۱۷
۵,۵	Model Performance Measures	۱۱۹
۵,۶	Illustrative Example	۱۲۲
۵,۷	Main Results	۱۲۳
۵,۷,۱	Analyzing of Queueing Network Model Using Simulation Technique	۱۲۳
۵,۷,۲	Analyzing of Queueing Network Model Using Discrete Review Policy	۱۳۱

## CHAPTER SIX

### WAITING, RESPONSE AND COMPUTING TIMES OF MATRIX OPERATIONS USING MULTI THREADING USED IN PARALLEL QUEUEING

٦,١	Introduction	١٤١
٦,٢	Parallel Programming Model	١٤٤
٦,٢,١	Threads model	١٤٤
٦.٢,٢	Program Transformations	١٤٥
٦,٢,٣	Analysis of Parallel Computing Reduction	١٤٥
٦,٢,٤	A Thread State Diagram	١٤٧
٦,٣	The Queueing Model of MTA	١٤٨
٦,٣,١	Mean Thread Analysis MTA	١٥٣
٦,٣,٢	Efficiency of the MTA	١٥٦
٦,٣,٣	Efficiency of the MTA with Fixed Number of $L$ Servers	١٥٨
٦.٤	Deception of Queueing Parallel Model for Transpose Matrix	١٥٩
٦,٤,١	Traffic Intensity of Device	١٦٠
٦,٤,٢	Normalization constant	١٦٥

٦,٤,٣	Algorithm	١٦٦
٦,٤,٤	Data Dependency Analysis	١٦٧
٦,٤,٥	Types of Dependencies	١٦٧
٦,٤,٦	Loop Dependence Analysis	١٦٨
٦,٤,٧	Using Queueing Parallel Computing of Shuffle Transpose	١٦٩
٦,٥	Deception of Queueing Parallel Model for Inverse Matrix	١٧١
٦,٥,١	By Partitioning for Inverse Matrix	١٧١
٦,٥,٢	Gauss Elimination Algorithm	١٧٣
٦,٥,٣	Linear Solving Systems for Inverse Matrix	١٧٥
٦,٦	The Performance Measures of Queueing Parallel Computing	١٧٥
٦,٧	Numerical Results	١٧٨
٦,٨	Conclusions	١٨٥

<b>REFERENCES</b>	١٨٦
-------------------	-----

## **ARABIC SUMMARY**

## ACKNOWLEDGEMENTS

*First*, thanks a lot for my God. I wish to express my thanks to many people who have assisted me in the completion of my Ph.D degree. My advisors, my family, my friends and all provided endless support in my research. I would like to express my deepest gratitude to **Prof.Dr. G.S.Mokaddis** Professor Emeritus of Mathematical Statistics Faculty of Science, Ain Shams University for suggesting the topics of the thesis, sponsoring this work, his guidance and providing me the direction and motivation to complete my studies. I would like to thank **Prof.Dr.I.A.Imail** Professor of Computer science my major advisor for his valuable suggestions during the preparation of this thesis. I also thank him for his support and encouragement, Also I would like to thank **Dr. Samia Awwad** associate Professor of Mathematical Statistics for her patience, endless support and encouragement for her valuable support and advice. Without the incredible love and support backing from my family, the dream of receiving my Ph.D degree could never be attained. And extremely special thanks are due to my mother and my father for their pure love and deepening support, in conclusion, I hope to get all your dear confidence in my research and hope it will be satisfied and achieved the goal. I am deeply indebted to mathematical department, Faculty of Science, Ain Shams University for kind supervisions. *Finally* this work is dedicated to my father (**Kamal Metry**) and my mother (**Makarim Gergis**), without whose love and this support would not have been possible. And also thanks for my brothers and my sister to help me. Thanks for Chairman, Director of the production sector and Director of quality sector of my working in *Kader Factory* and also many thanks for the president of the Arab Organization for Industrialization encourage joining the higher studies.

Mariam K. Metry



# **CHAPTER ONE**

## **DEFINITIONS AND SOME CONCEPTS FOR THE QUEUEING SYSTEMS**

### **1.1. INTRODUCTION**

In general we do not like to wait. But reduction of the waiting time usually requires extra investments. To decide whether or not to invest, it is important to know the effect of the investment on the waiting time. Queues help facilities or business provide service in an orderly fashion. Forming a queue being a social phenomenon, it is beneficial to the society if it can be managed so that both the unit that waits and the one that serves get the most benefit. For instance, there was a time when in airline terminals passengers formed separate queues in front of check-in counters. But now we see invariably only one line feeding into several counters. This is because of the realization that a single line policy serves better for the

passengers as well as the airline management. Such a conclusion has come from analyzing the mode by which a queue is formed and the service is provided.

The analysis is based on building a mathematical model representing the process of arrival of customers who join the queue, the rules by which they are allowed into service, and the time it takes to serve the customers. The unit providing service is known as the server. Queueing theory is mainly seen as a branch of applied probability theory. Its applications are in different fields, e.g. communication networks, computer systems, machine plants and so forth. For this area exist a huge body of publications, a list of introductory or more advanced texts on queueing theory is found in the bibliography.

The subject of queueing theory can be described as follows: consider a service center and a population of customers, which at some times enter the service center in order to obtain service. It is often the case that the service center can only serve a limited number of customers. If a new customer arrives and the service is exhausted, he enters a waiting line and waits until the service facility becomes available. So we can identify three main elements of a service center: a population of customers, the service facility and the waiting line. Also within the scope

of queueing theory is the case where several service centers are arranged in a network and a single customer can walk through this network at a specific path, visiting several service centers. Since queueing theory is applied in different fields, also the terms client and task are often used instead customer. The service center is often named processor or machine. Queueing theory tries to answer questions like e.g. the mean waiting time in the queue, the mean system response time, mean utilization of the service facility, distribution of the number of customers in the queue, distribution of the number of customers in the system ([9]Bhat, 2008).

These questions are mainly investigated in a stochastic, where e.g. interarrival times of the customers or the service times are assumed to be random. A basic model of a service center is the customers arrive to the service center in a random fashion. The service facility can have one or several servers, each server capable of serving one customer at a time; the service times needed for every customer are also modeled as random variables. Queueing systems may not only differ in their distributions of the inter arrival and service times, but also in the number of servers, the size of the waiting line (infinite or

finite), the service discipline and so forth. Some common service disciplines are:

**FIFO:** (First Input, First Output): a customer that finds the service center busy goes to the end of the queue. **LIFO:** (Last Input, First Output): a customer that finds the service center busy proceeds immediately to the head of the queue. It will be served next, given that no further customers arrives. Service in random order (**SIRO**): the customers in the queue are served in random order.

**Priority Disciplines:** every customer has a (static or dynamic) priority; the server selects always the customers with the highest priority. This scheme can use preemption or not. The Kendall notation is used for a short characterization of queueing systems. A queueing system description looks as follows:

$$a/b/c/d/e/f,$$

where  $a$  denotes the distribution of the interarrival time,  $b$  denotes the distribution of the service times,  $c$  denotes the number of servers,  $d$  denotes the maximum size of the waiting line in the finite case and the optional  $e$  denotes the service discipline used (FIFO, LCFS) and  $f$  : Size of the source feeding the system with customers or size of calling source.

If S is omitted the service discipline is always FIFO. For A and B the following abbreviations are very common:

- $M$  (Markov): this denotes the exponential distribution with

$$A(t) = 1 - \exp(-\lambda t), \quad a(t) = \lambda \exp(-\lambda t),$$

where  $\lambda > 0$  is a parameter of interarrival time. The name  $M$  stems from the fact that the exponential distribution is the only continuous distribution with the Markov property, i.e. it is memoryless.

- $D$  (Deterministic): all values from a deterministic “distribution” are constant, i.e. have the same value.
- $E_k$  (Erlang-k): Erlangian distribution with  $k$  phases  $k \geq 1$ . For the Erlang-k distribution we have

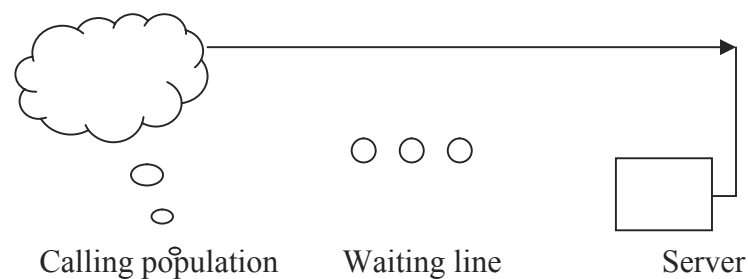
$$A(t) = 1 - \exp(-k\mu t) \sum_{j=1}^k \frac{(k\mu t)^j}{j!},$$

where  $\mu > 0$  is a parameter of service time. This distribution is popular for modeling telephone call arrivals at a central office.

- $G$  (General): general distribution, not further specified. In most cases at least the mean and the variance are known. The most simple queueing system, the  $M/M/1$  system (with FIFO) can then be described as follows: we have a single

server, an infinite waiting line, the customer interarrival times are independent and identically (iid) and exponentially distributed with some parameter  $\lambda > 0$  and the customer service times are also (iid) and exponentially distributed with some parameter  $\mu > 0$ . We are mainly interested in steady state solutions, i.e. where the system after a long running time tends to reach a stable state, e.g. where the distribution of customers in the system does not change ([16] Gross, D., Harris, 1998).

A queueing system is described by its calling population, the nature of the arrivals and services, the system capacity, shown in **Figure (1.1.1)** and the queueing discipline.



**Figure (1.1.1):** Queueing system

In this system the calling population is infinite; that is, if a unit leaves the calling population and joins the waiting line or enters service, also the system capacity is unlimited. (The

system includes the unit in service plus those waiting in line). Finally, callers are served in the order of their arrival (often called FIFO, for first input, first output) by a single server, or channel. Arrivals and services are described by the distributions of the time between arrivals and service, or the waiting line will grow without bound. Exceptional situation would be arrival rates that are greater than service rates for short periods of time. However, such a situation is  $\lambda < \mu$  to reach to steady state ( $t \rightarrow \infty$ ) more complex. As shown as in **Figure (١.١.٢)**. If the server is busy, then the arriving unit enters the queue. If the server is idle and the queue is empty, while the unit enters the server. It is impossible for the server to be idle and the queue to be not empty.

		Queue system	
		Not empty	Empty
Server status	Busy	Enter queue	Enter queue
	Idle	Impossible	Enter Service

**Figure (١.١.٢):** Potential unit actions upon arrivals.