Republic of Iraq
Ministry of Higher Education
and Scientific Research
University of Babylon
College of Education for Pure
Sciences
Department of Mathematics

On Queueing Network and its Applications

A Research

Submitted to

The Faculty of the Department of Mathematics

College of Education for Pure Science

University of Babylon

In Partial Fulfillment

of the Requirements for the Degree of

High Diploma Education / Mathematics

by

Mohsen Mansour Abdullah Hammoud

Supervised by

Dr. Ali Hussein Mahmood Al-Obaidi

September 2022

بسم الله الرحمن الرحيم

(فتعالى الله الملك الحق ولا تعجل بالقرآن من قبل أن يقضى إليك وحيه وقل رب زدني علما)

صدق الله العظيم

(سورة طه 114)

Supervisor Certification

I certify that this research which is entitled "**On Queueing Network and its Applications**" was prepared by the student " **Mohsen Mansour Abdullah Hammoud**" under my supervision at the University of Babylon, College of Education for Pure Sciences as a partial requirement for the degree of High Diploma Education /Mathematics.

Signature:

Name: **Dr. Ali Hussein Mahmood Al-Obaidi**

Scientific grade: **Lect**.

Date: **/ /2022**.

According to the available recommendation, I forward this research for debate by the examining committee.

Signature:

Name: **Dr. Azal Jafar Musa**

Scientific grade: **Asst. Prof.**

Date: **/ /2022.**

Certification of Scientific Expert

I certify that I have read this research entitled "**On Queueing Network and its Applications**" and I found it is qualified for debate.

Signature:

Name:

Title:

Date: **/ / 2022.**

Certification of Linguistic Expert

I certify that I have read this research entitled "**On Queueing Network and its Applications**" and correct its grammatical mistakes; therefore, it has qualified for debate.

Signature:

Name:

Title:

Date: **/ / 2022.**

Examining Committee Certification

We certify that we have read the research entitled "**On Queueing Network and its Applications**", as a committee examined the student "**Mohsen Mansour Abdullah Hammoud**" in its contents and that in our opinion, it meets the standard of a research for the degree of High Diploma Education/Mathematics.

Chairman

Signature:

Name:

Title:

Date: **/ / 2022**

Member

Signature:

Name:

Title:

Date: **/ / 2022**

Member

Signature:

Name:

Title:

Date: **/ / 2022**

Supervisor

Signature:

Name:

Title:

Date: **/ / 2022**

Approved by the Dean of the College

Signature:

Name: **Dr. Bahaa Hussein Rabee Salih**

Scientific grade: **Prof.**

Address: **Dean of the College of Education for Pure Sciences**

Date: **/ / 2022**

Table of Contents

Dedication

I dedicate my thesis to

my family and many friends.

Special gratitude to

my loving parents, whose words of encouragement

and push for tenacity ring in my ears.

My sisters and brothers have never left

my side and are incredibly special.

Acknowledgments

I would like to acknowledge everyone who played a role in my academic accomplishments. First of all, my parents supported me with love and understanding. Without you, I could never have reached this current level of success.

Secondly, my advisor and committee members, each of whom has provided patient advice and guidance throughout the research process. Thank you all for your unwavering support.

Abstract

This research discusses the structure of the queueing network, and some related methods to solve it are shown. The different kinds of this queue are considered based on their form. The essential processes like the birth-death process and M/G/1 are regarded to form the behavior of queueing network. Some theorems are given with supporting examples to find the stationary distribution for the global balance equations.

Introduction

Queueing network systems have been widely employed to characterize and investigate source sharing systems, for instance, production, communication, and computer systems. They have been established to be a robust and flexible system performance evaluation and prediction method. This model is a group of facilities centers demonstrating the system sources that deliver the facility to customers by showing the users. According to the queueing discipline, customers compete for the source service and possibly wait to be served in the queue at the service centers.

This system's analysis evaluates performance procedures, such as source utilization, throughput, and customer response time. The dynamic behavior of this system can be explained by a set of random variables that define a stochastic process. Under some constraints on this system, it is possible to define an associated underlying stochastic Markov process and compute the performance indices by its solution. Product-form queueing networks have a brief closed-form appearance of the stationary state distribution that grants us to explain practical algorithms to calculate average performance measures with polynomial time complexity in the number of model components.

Queueing systems were initially used to evaluate telephonic systems' obstruction and then study computer and communication systems [1-5].

In this work, we introduce queueing network models and their properties. Queueing networks extend the basic queueing systems that are stochastic models first introduced to represent the entire system by a single service center.

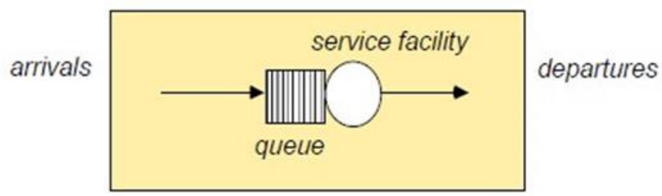
Chapter One

Elementary Queueing Systems

Chapter One

Elementary Queueing Systems

Queueing system is a system that customers arrive at the service center and ask for resource service. They possibly wait to be served in queueing discipline and leave the service center after getting their services.



Under exponential and independence assumptions, one can define an associated stochastic continuous-time Markov process to represent system behavior. The performance of system indices is derived from the solution of the Markov process. This section is included necessary concepts and the essential details to overhand our work in different areas. We assume that any random variable (in short, r.v.) is a real-valued function defined over the sample space Ω.

1.1. Markov Processes

We begin this section by defining some essential concepts related to Markov processes. **The stochastic process** is a set of random variables

$$\{X(t): t \in T\}$$

defined over the same probability space indexed by the parameter t, called **time**, and each $X(t)$ random variable takes values in the set Γ called **state space** of the process. Both T (time) Γ and (space) can be either discrete or continuous. A process is called a **continuous-time process** if the time parameter t is continuous, and it is called a **discrete-time process** if the time parameter t is discrete.

We also define the **joint probability distribution function** of the random variables $X(t_i)$ as

$$pr\{X(t_1) \le x_1, X(t_2) \le x_2, \dots, X(t_n) \le x_n\}$$

for any set of times, tie $t_i \in T, x_i \in \Gamma, 1 \le i \le n, n \ge 1$.

The discrete-time process $\{X_n: n = 1,2, \dots\}$ is called **discrete-time Markov process** (**Markov chain**) if the state at time $n+1$ only depends on the state probability at time n and is independent of the previous history

$$pr\,\{X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} = pr\{X_{n+1} = j | X_n = i_n\} = p_{ij},$$

$\forall n > 0, \forall j, i_0, i_1, \dots, i_n \in \Gamma$ where $P = [p_{ij}]$ is called **transition probability matrix**. We define the (i,j)th entry of the mth power of the transition matrix P by

$$p_{ij}^{n} = pr(X_{m+n} = j | X_n = i)$$

where P^0 is the identity matrix.

The continue-time process $\{X(t); t \in T\}$ is called the **continue-time Markov process** if

$$pr\{X(t) = j|X(t_0) = i_0; \dots; X(t_1) = i_1; \dots; X(t_n) = i_n\} = pr\{X(t) = j|X(t_n) = i_n\} = p_{ij},$$

$\forall t_0, t_1, \dots, t_n, t: t_0 < t_1 < \dots < t_n < t, \forall n > 0, \forall j, i_0, i_1, \dots, i_n, \in \Gamma.$

A state $j \in \Gamma$ is called **accessible** from a state $i \in \Gamma$if there is a number $m \in \mathbb{N}_0$ with $pr(X_m = j|X_0 = i) > 0$ and denoted it by $i \to j$. For any two states $i, j \in \Gamma$, if $i \to j$ and $j \to i$ hold, then i and j are said to **communicate** and indicate it by $i \leftrightarrow j$. The communication relation between states is an equivalence relation. Therefore, we can divide the space Γ of a Markov chain into a partition of countably many equivalence classes concerning the communication of states. Any such equivalence class shall be called a **communication class**. A communication class $C \subset \Gamma$ that does not allow access to states outside itself holds is called **closed**. If a Markov chain has only one communication class, i.e., if all states communicate, it is **irreducible**. Otherwise, it is called a **reducible**.

Define T_j as **the stopping time of the first visit to the state** $j \in \Gamma$, and denote by $T_j = \min\{n \in \mathbb{N}: X_n = j\}$. The conditional distribution of T_j given $X_0 = i$, can be established by for all $i, j \in \Gamma$.

$$F_k(i,j) = pr(T_j = k|X_0 = i) = \begin{cases} p_{ij}, & k = 1 \\ \sum_{h \neq j} p_{ih}\ F_{K-1}(h,j), & k \geq 2 \end{cases} \quad (1.2)$$

For all $i, j \in \Gamma$, we define f_{ij} as **the probability of ever visiting state *j* after beginning in state *i*** by

$$f_{ij} = pr\left(T_j < \infty \middle| X_0 = i\right) = \sum_{k=1}^{\infty} F_k(i,j).$$

Therefore, if we are summing up overall $k \in \mathbb{N}$ of the equation (1.2) leads to

$$f_{ij} = p_{ij} + \sum_{h \neq j} p_{ih} f_{hj}.$$

Now, we define the random variable N_j as **the total number of visits** to the state $j \in \Gamma$. The conditional probability of N_j given $X_0 = j$ is given by

$$pr\left(N_j = m \middle| X_0 = j\right) = f_{jj}^{m-1}\left(1 - f_{jj}\right) \quad (1.3)$$

and the distribution of N_j is **geometrically** distributed for $i \neq j$ as below expression

$$pr(N_j = m | X_0 = i) = \begin{cases} 1 - f_{ij}, & m = 0 \\ f_{ij}\, f_{jj}^{m-1}(1 - f_{jj}), & m \geq 1 \end{cases} .$$

For all $j \in \Gamma$, we sum up overall m in the above equation (1.3), we will get

$$pr(N_j < \infty | X_0 = j) = \begin{cases} 1, & f_{jj} < 1 \\ 0, & f_{jj} = 1 \end{cases} .$$

A state $j \in \Gamma$ is called **recurrent** if $f_{jj} = 1$ and **transient** otherwise. Define the **potential matrix** $R = (r_{ij})_{i,j \in E}$ of the Markov chain by its entries

$$r_{ij} = E(N_j | X_0 = i)$$

for all $i, j \in \Gamma$ such that the entry r_{ij} is the condition expected of N_j given $X_0 = i$. We can be computed r_{ij} by

$$r_{ij} = \sum_{n=0}^{\infty} p_{ij}^{n}$$

for all $i, j \in \Gamma$. From the above equations, we can conclude that for all $i, j \in \Gamma$,

$$r_{ij} = (1 - f_{jj})^{-1} \text{and } r_{ij} = f_{ij} r_{ij}.$$

In specific, r_{jj} is finite if j is **transient** and infinite if j is **recurrent**. Recurrence and transience of states are class properties concerning the relation ↔. Additionally, a recurrent communication class is permanently closed. If the state $j \in$

Γ is transient, then $\lim_{n=\infty} p_{ij}^n = 0,$ regardless of the initial $i \in \Gamma$. A recurrent state $j \in \Gamma$ with $r_{jj} < \infty$, will be called **positive recurrent**. Otherwise, j is called **null recurrent**.

We define **the period of the state** j as the greatest common divisor (in short, gcd) of the step n that makes the transition probability of state j to itself is not zero. In other words, this period is given by

$$d(j) = \text{gcd}\left\{n \geq 1: p_{jj}^{(n)} > 0\right\}.$$

The state j is **periodic** if p $d(j) > 1,$ it has periodic $d(j)$. If $d(j) = 1,$ then state j is called **aperiodic**.

The Markov chain $\{X_n: n = 1,2,\ldots\}$ is called **homogeneous** if the one-step conditional probability is independent on time n where this probability is given as

$$p_{ij} = pr\{X_1 = j|X_0 = i\} = pr\{X_{n+1} = j|X_n = i\}, \quad \forall n > 0, \forall i, j \in \Gamma.$$

If the stability conditions hold, we can compute the **stationary probability vector**

$$\pi \ = [\pi_0,\ \pi_1,\ \pi_2, \ldots], \pi_j \ = pr\{X = j\}, \forall j \in \Gamma$$

for **ergodic Markov chain** (irreducible and with positively recurrent aperiodic states) as below system of global balance equations

$$\pi = \pi P \text{ with } \sum_j \pi_j = 1.$$

A continuous-time Markov process $\{X(t): t \in T\}$ is called **homogeneous** if the one-step conditional probability only depends on the interval width where this probability is given as

$$p_{ij} = pr\{X(s) = j | X(0) = i\} = pr\{X(t+s) = j | X(t) = i\} \ \forall t > 0, \forall i, j \in \Gamma$$

We also define the matrix $Q = [q_{ij}]$ of **transition rates** (**infinitesimal generator**) where if the stability conditions hold, we compute the stationary probability vector $\pi = [\pi_0, \pi_1, \pi_2, \ldots]$ for **ergodic Markov process** as below system of global balance equations

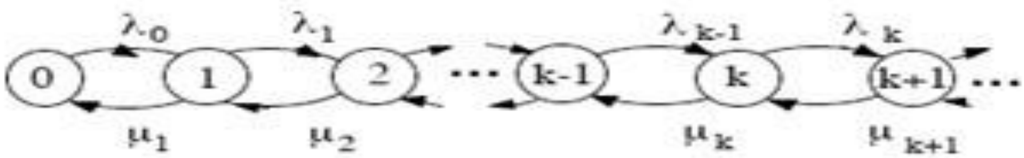
$$\pi Q = 0 \text{ with } \sum_j \pi_j = 1.$$

The **residence time** of the process in each state is distributed according to geometric for discrete-time Markov processes and exponential distribution for continuous-time Markov processes.

1.2. Birth-Death Markov Processes

Let state-space denote by $\Gamma = N$. The continuous-time Markov chain with discrete state space is called the birth-death process if the only non-zero state transitions are from state i to states $i-1, i, i+1, \forall i \in \Gamma$. Let P be probability transition matrix and let Q be the matrix of transition rates as tri-diagonal. Let λ_i be birth transition rate $i \geq 0$ and let μ_i be death transition rate $i \geq 1$ such that

$$q_{ij} = \begin{cases} \lambda_i, & j = i + 1. \\ \mu_i, & j = i - 1. \\ -(\mu_i + \lambda_i), & j = i, i > 0. \\ \lambda_0, & i = 0. \end{cases}$$



$$\pi_i = \pi_0 \prod_{j=0}^{i-1} \frac{\Lambda_j}{\mu_{j+1}}, \pi_0 = [\sum_{i=0}^{\infty} \prod_{j=0}^{i-1} \frac{\Lambda_j}{\mu_{j+1}}]^{-1}$$

The sufficient condition for the stationary distribution of this process is

$$\exists k_0 : \forall k > k_0, \Lambda_k < \mu_k$$

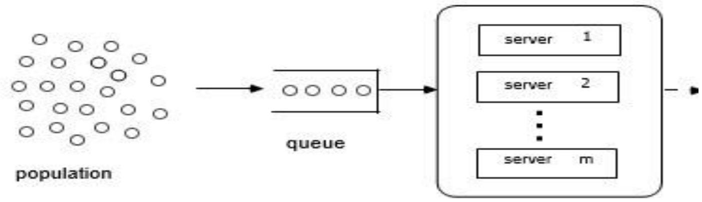
The birth and death rates are constant such that the birth transition rate is given as $\Lambda_i = \Lambda$, $i \geq 0$ and the birth transition rate is given as $\mu_i = \mu$, $i \geq 1$. Let $p = \Lambda / \mu$, then $p < 1$

$$\mu_0 = [\sum_k p^k]^{-1} = 1 - p, \qquad \pi_k = \pi_0 \, (\Lambda/\mu)^k, \qquad \pi_k = (1 - p)p^k, \qquad k \geq 0$$

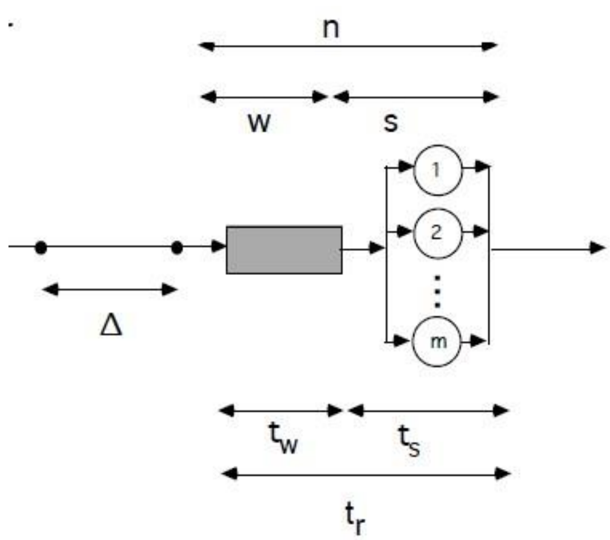
Therefore, π_k has the geometric distribution.

1.3. Queueing Systems

The queueing system is described by the arrival process, the service process, the number of servers and their service rate, the queueing discipline process, the system or queue capacity, and the population constraint.



In a single service center, we denote the following by Δ is an interarrival time, w is a number of customers in the queue, t_w is queue waiting time, s is number of customers in service, t_s is service time, n is number of customers in the system and t_r is response time.



Kendall’s notation is a way to symbolize the queue by A/B/X/Y/Z such that A is an interarrival time distribution (Δ), B is a service time distribution (ts), X is the number of servers (m), Y is system capacity (in the queue and in service), and Z is queueing discipline. We assume that $Y = \infty$, $Z = FCFS$ (default), A and B are “D” deterministic (constant), “M” exponential (Markov), “Ek” Erlang-k, or “G” general. Examples of queueing systems are D/D/1, M/M/1, M/M/m (m>0), M/G/1, G/G/1.

System analysis depends on the two kinds of systems. First of all, transient system, if for a time interval, we have given the initial conditions. Second of all, stationary system if we have a system in steady-state conditions (stable systems). Analysis of the associated stochastic process represents system behavior using Markov stochastic process or birth and death processes as an example. The average number of customers in the system is $N = E[n]$ and mean response time $R = E[t_r]$. Some fundamental relations in queueing systems

$$N = w + s$$

$$t_r = t_w + t_s$$

$$n = E[w] + E[s]$$

$$R = E[t_w] + E[t_s]$$

Theorem 1.1. (Little’s Formula)

the average number of customers in the system is equal to the throughput times the average response time

$$N = XR, E[w] = XE[t_w]$$

where X is throughput time.

The scheduling algorithms classify by the following roles

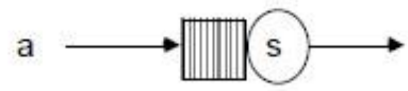
FCFS	first come first served
LCFS	last come first served
LCFSPr	idem with pre-emption (Immediate service)
Random	
Round Robin	each customer is served for a fixed quantum δ
PS *	Processor Sharing for $\delta \rightarrow 0$ all the customers are served at the same time for service rate μ and n customers, each receives service with rate μ / n
IS *	Infinite Serves no queue (delay queue)
SPTF	Shortest Processing Time First
SRPTF	Shortest Remaining Processing Time First

We can consider the following conditions

1. with/without priority.
2. Abstract priority/dependent on service time.
3. with/without pre-emption.

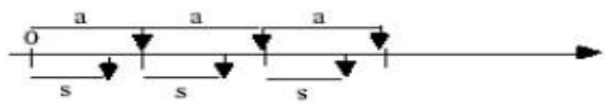
3.1.1. D/D/1 System

In this system, the deterministic arrivals are constant interarrival time (a), and the deterministic service for each customer has the same service demand (s). Let $n(t)$ be the number of customers in the system at time t.

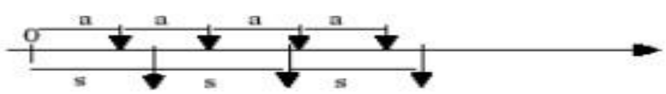


The transient analysis has the following conditions:

1. If $s < a$, then $n(t) = 0$ during $ia + s < t < (i+1)a, i \geq 0$.
2. If $s < a$, then $n(t) = 1$ during $ia < t < ia + s, i \geq 0$.



3. If $s = a$, then $n(t) = 1$ during $t \geq 0$.
4. If $s > a$, then $n(t) = [t/a] - [t/s]$ during $t \geq 0$.



The stationary analysis has happened when $s \leq a$, and the arrival rate less than or equal the service rate $(1/a \leq 1/s)$, then the system reaches the steady-state (stability condition). When $n \in \{0,1\}$, we have

$$pr\{n = 0\} = (a - s)/a$$

$$pr\{n = 1\} = s/a$$

$$w = 0 \quad t_w = 0 \quad t_r = s \quad \text{(deterministic r.v.)}$$

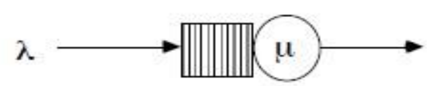
$$X = 1/a \quad \text{throughput}$$

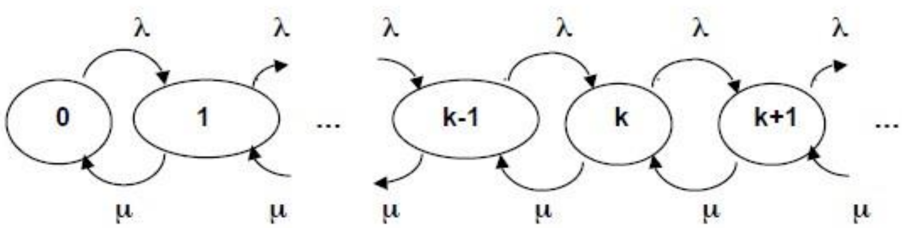
$$U = s/a \quad \text{utilization}$$

3.1.2. M/M/1 System

In this queue, we have the arrival Poisson process with rate λ and the exponential service time (exponential interarrival time) with rate μ. So, E[t_s] = 1/μ.

This system has a single server, and it includes n states. The associated stochastic process has the birth-death continuous-time Markov chain with constant rates λ and μ





The stationary analysis of this system appears when ʎ < μ (stability condition). This leads us to that the traffic intensity is $\rho = \lambda/\mu$ and stationary state probability

$$\pi_k = pr\{n = k\} \quad k \in \mathbb{N}$$

$$\pi_k = \rho^k(1-\rho) \quad k \geq 0$$

The essential characters of the queue are

$$n = \frac{\rho}{1-\rho}$$

R = 1 / (μ - ʎ) (Little's formula)

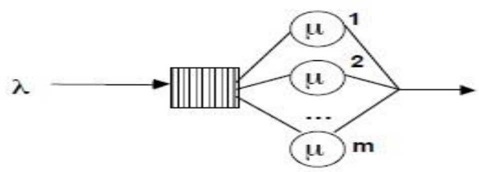
X=ʎ and $U = \rho$

E[w]= $\rho^2/(1-\rho)$

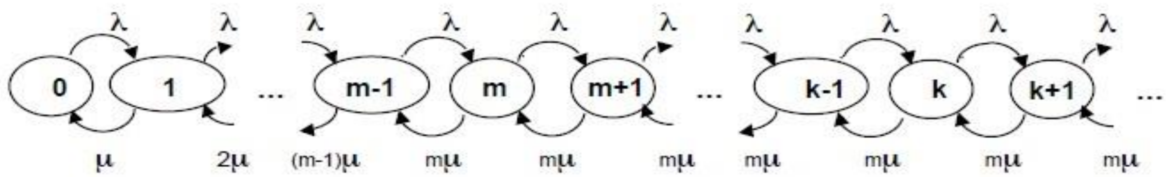
$$E[t_w]=(1/\mu)\rho/(1-\rho) \qquad \text{(Little's formula)}$$

3.1.3. M/M/*m* System

In this case, the arrival Poisson process is with rate λ and the exponential service time (exponential interarrival time) is with rate μ. Therefore, $E[t_s]= 1/\mu$.



The system has n state, and the associated stochastic process is birth-death continuous-time Markov chain with rates $\lambda_k = \lambda$ and $\mu_k = \min\{k, m\}\mu$.



The stationary analysis of the system is given by $\lambda < m\,\mu$ (stability condition). The traffic intensity is $\rho = \lambda/m\mu$, and the stationary state probability is

$$\pi_k = \pi_0 (m\rho)^k/k! \qquad 1 \le k \le m$$

$$\pi_k = \pi_0 m^m \rho^k/m! \qquad k > m$$

$$\pi_0 = \sum_{k=0}^{m-1} [(m\rho)^k/k! + (m\rho)^m/m!\,[1/(1-\rho)]]^{-1}$$

The essential characters of the queue are

$$N = m\rho + \pi_m\,\rho / (1 - \rho)^2$$

$$R = \pi_m / (m\,\mu\,(1 - \rho))^2) + 1/\mu \qquad \text{(Little's formula)}$$

$$X = \lambda,\ U = \rho$$

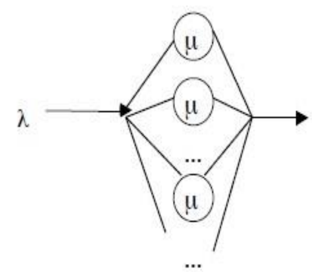
$$pr\{n(t) = m\} = \sum_{k\ge m} \pi_k = \pi_0\,(m\,\rho)^m\ /m!\ (1-\rho)$$

$$E[w] = \pi_m \rho / (1 - r \rho)^2$$

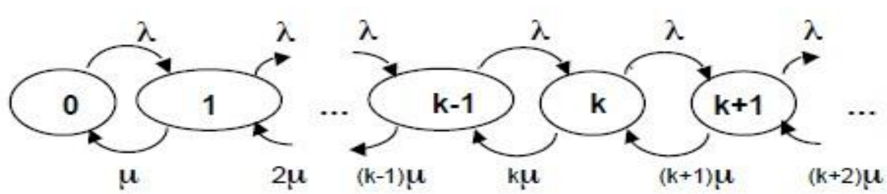
$$E[t_w] = \pi_m / ((1 - \rho)^2 \mu) \qquad \text{(Little's formula)}$$

3.1.4. M/M/∞ System

In this queue, the arrival process is Poisson with rate λ and the service time (interarrival time) is exponential with rate μ. Hence, $E[t_s] = 1/\mu$.



The system is with infinite identical servers, and the system has n states. The associated stochastic process of this system is birth-death continuous-time Markov chain with rates $\Lambda_k = \lambda$ and $\mu_k = k\,\mu$.



The stationary analysis of this model is always stable, and the traffic intensity $\rho = \lambda/\mu$. The stationary state probability is the Poisson distribution with the below formula

$$\pi_k = e^{-\rho}\rho^k/k! \qquad k \geq 0$$

The important roles of the queue are

$$n = \rho,\ R = 1/\mu,\ X = \Lambda,\ U = \rho$$

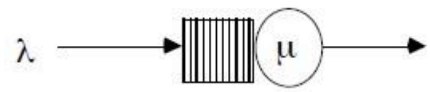
$$E[w] = E[t_w] = 0$$

where this queue has delay.

3.1.5. M/G/1 System

This system with a single server has the arrival Poisson process with rate ʎ and general service time (exponential interarrival time) with rate μ. So, we have

$$E[t_s] = 1/\mu \text{ and } C_B = (Var [t_s])^{1/2} / E[t_s]$$



The state defined as *n* (number of customers in the system) does not lead to a Markov process.

The state description *n* for system M/M/1 gives a (birth-death) continuous-time Markov chain because of the exponential distribution (memoryless property). We can use a different (more detailed) state definition to define a Markov process (e.g., the number of customers and the amount of service provided to the customer currently in service). The associated Markov process is not birth-death, and the analysis depends on an embedded Markov process and Z-transform technique.

Theorem 1.2. (Khinchine Pollaczeck)

For any queueing discipline independent of service time without pre-emption, we have

$$E[w] = \rho + \frac{\rho^2(1 + c_b^2)}{2(1-\rho)}.$$

The stability condition is $\lambda < \mu$, then

$$N = E[w] + I\, E[t_s]$$

$$R = N/X$$

Theorem 1.3. (Poisson Arrivals See Time Average, in short PASTA)

The state distribution and moments seen by a customer at arrival time are the same as those observed by a customer at arbitrary times in steady-state conditions.

Chapter Two

Queueing Networks

Chapter Two

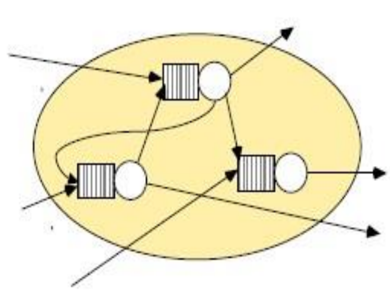
Queueing Networks

2.1. Preliminaries

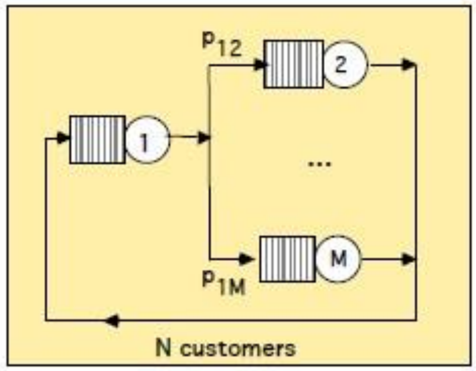
Queueing Network (in short, **QN**) is a system model set of service centers representing the system resources (that provide service) to a collection of customers (that represent the users). That means a queueing network describes the system as a set of interacting resources. For example, stochastic models of resource sharing systems computer, communication, traffic, manufacturing systems, and customers compete for the resource service. QN is a powerful and versatile tool for system performance evaluation and stochastic models based on queueing theory that represents the system as a set of interacting resources.

QN is a collection of service centers that provide service to a set of customers. There are several kinds of QN, such as

1. **open** external arrivals and departures,



2. **closed** constant number of customers (finite population),



3. **mixed** if it is open for some types of customers, closed for other types.

Customers arrive to a service center (**node**) (possibly external arrival for open QN), ask for resource service, and possibly wait to be served (queueing discipline) at completion time exit the node and immediately move to another node or leave the QN. In closed QN, customers are always in a queue or service.

Informally, a QN is defined by the set of service centers $\Omega = \{1, \ldots, M\}$, the set of customers, and the network topology. Each service center is defined by the number of servers (independent and identical servers), the service rate (either constant or dependent on the station state), and the queueing discipline. Customers are described by their total number (closed QN), the arrival process to each service center (open QN), the service demand to each service center (such as expressed in units of service), and the service rate of each server (units of service / units of time). The service time is given as the service demand over the service rate.

The **network topology** models the customer behavior among the interconnected service centers where we assume that a probabilistic model represents a non-deterministic behavior. Let p_{ij} be the probability that a customer completing its service in station i immediately moves to station $j, 1 \leq i, j \leq M$. We assume that p_{i0} is the probability that a customer completing its service in station i immediately exits the network from the station i for open QN. The **routing probability** matrix is denoted by $P = [p_{ij}]$,$1 \leq i\,, j \leq M$ where $0 \leq\ p_{ij} \leq 1$, $\sum_{i=1}^{M} p_{ij} = 1$ for each station i.

A QN is **well-formed** if it has a well-defined long-term customer behavior, where for a closed QN, every station is reachable from any other with a non-zero probability. At the same time, for an open QN, add a virtual station 0 that represents the external behavior, generates external arrivals, and absorbs all departing customers, so obtaining a closed QN.

2.2. Types of Customers

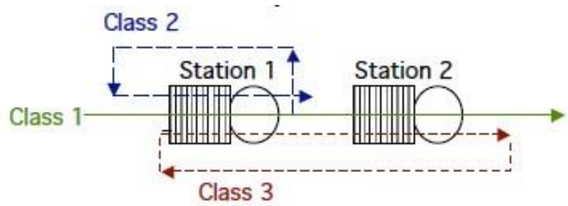
In simple QN, we often assume that all the customers are statistically identical. Modeling real systems can require to identify different types of customers depending on service time and routing probabilities. Several kinds of customers have many concepts, for example, **class** and **chain**. A **chain** forms a permanent categorization of customers, and a customer belongs to the same chain during its whole activity in the network. A **class** is a temporary classification of customers, and a customer can switch from a class to another during its activity in the network (usually with a probabilistic behavior).

The customer service time in each station and the routing probabilities usually depend on its class. There are other kinds of customers, multiple-class with single-chain QN and multiple-class with multiple-chain QN. Let R be a set of classes of the QN, R be a number of classes, C be a set of chains, and C be the number of chains. Classes can be partitioned into chains, such that there cannot be a customer switch from classes belonging to different chains. Let $\mathbf{P}^{(c)}$ be the **routing probability matrix** of customers for each **chain** $c \in C$. Let $p_{ir,js}^{(c)}$ be the probability that a customer completing its service in station i class r immediately moves to station j, class s, $1 \leq i, j \leq M, r, s$ in R, classes of chain c. Let $p_{ir,0}^{(c)}$ be the probability that a customer is completing its service in station i class r and immediately exits the network. Let $K^{(c)}$ be the population of a closed chain $c \in C$. Let $p_{0,ir}^{(d)}$ be the probability of an external arrival to station i class r for an open chain $d \in C$.

A QN is said to be open if all its chains are open, closed if they are all closed, and mixed otherwise.

Example 2.1. (multiple-class and multiple-chain QN)

Let M=2 (stations), R=3 (classes), C=2 (chains), and R ={1,2,3} such that chain 1 is open and formed by classes 1 and 2 and chain 2 is closed and formed by class 3.



Let R_i be a set of classes served by station i and define the set

$$E_c = \{(i,r): r \in R_i \text{ set}, 1 \le i \le M, \text{class } r \in chain\ c\}.$$

Let $R_i^{(c)}$ be a set of classes served by station i and belonging to chain c. Therefore, we have the following

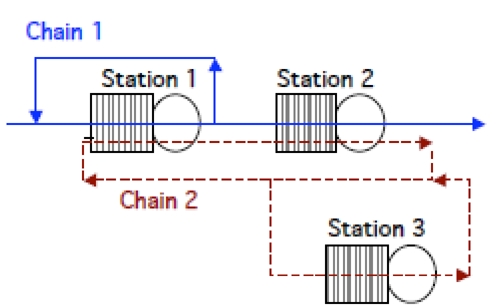
$R_1 = \{1,2,3\}$, $R_2 = \{1,3\}$

$E_1 = \{(1,1), (1,2),(2,1)\}$, $E2 = \{(1,3), (2,3)\}$

$R_1^{(1)} = \{1,2\}$, $R_2^{(1)} = \{1\}$, $R_1^{(2)} = R_2^{(2)} = \{ 3 \}$

Example 2.2. (single-class and multiple-chain QN)

There is only one in each chain, so R=C. Let M=3 (stations), R=2 (classes), C=2 (chains), and $R = \{1,2\}$ such that no class switching, chain 1 is open and formed by one class, and chain 2 is closed and formed by one class.



2.3. QN Performance Indices

The local performance indices related to single resource *i* (a service center) is given as

$\mathbf{n_i}$ number of customers **in the station i**

$\mathbf{n_{ir}}$ number of customers **in station i and class r**

$\mathbf{n_i}^{(c)}$ number of customers **in station i and chain**

Ui	utilization
Xi	throughput
Ni	mean queue lengt h
Ri	mean response time

Let $\mathbf{t_i}$ be customer passage time through the resource and let the distribution of $\mathbf{n_i}$ be $\boldsymbol{\pi_i}(\mathbf{n_i})$ at arbitrary times.

The **global performance indices** related to the overall network is given as

U	**utilization**
X	**throughput**
N	**mean population (for open networks)**
R	**mean response time**

The **Network model parameters** (single class, single chain) is denoted by

M number of nodes

K number of customers (closed network)

P=[p_{ij}] routing matrix

λ total arrival rate

μ_i service rate of node i

p_{0i} arrival probability at node i

$\mathbf{e_i}$ **visit ratio** at node i, and the solution of **traffic equations**

$$\boldsymbol{e_i} = \Lambda \boldsymbol{p_{0i}} + \sum_j \boldsymbol{e_j} \boldsymbol{p_{ji}}$$

$\mathbf{S} = (S_1,\ldots,S_M)$ system state

S_i node i state which includes n_i, $1 \leq i \leq M$

For example, if we have M nodes, *R* classes, and *C* chains, single-class multi-chain (*R*=*C*).

$\mathbf{n} = (\mathbf{n}_1,\ldots, \mathbf{n}_M)$ network state

$\mathbf{n}_i = (n_{i1},\ldots, n_{iR})$ station i state, $1 \leq i \leq M$

2.4. Markovian QN

Markovian network is the network behavior that a homogeneous continuous-time Markov process M can represent. Let **S** be a network state, **E** be a set of all feasible states of the QN, **E** be a discrete state space of the process, and **Q** be an infinitesimal generator. If *P* (network routing matrix) is irreducible, then there exists stationary state distribution $\boldsymbol{\pi} = \{\pi(\mathbf{S}), \mathbf{S} \in \mathbf{E}\}$ as a solution of the **global balance equations**

$$\pi Q = 0, \Sigma_{\mathbf{S} \in E} \pi(\mathbf{S}) = 1$$

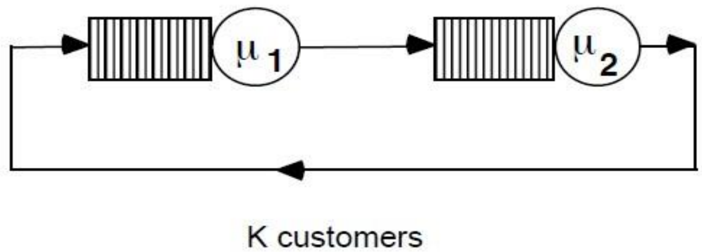
Solution algorithm for the evaluation of average performance indices and joint queue length distribution at arbitrary times (π) in Markovian QN considers the following steps:

1. Definition of **system state** and state-space E.
2. Definition of **transition rate matrix** Q.
3. **Solution** of global balance equations to **derive π**.
4. **Computation** from π of the average **performance indices**.

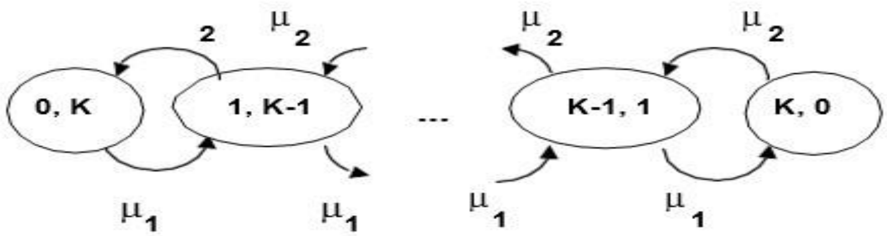
This method becomes unfeasible as |E| grows, i.e., proportionally to the dimension of the model (number of customers, nodes, and chains). For example, a single class closed QN with M nodes and K customers, then $|E| = \binom{M + K - 1}{K}$. Different methods deal with this case, like **exact product-form solutions** under special constraints and **approximate solution** methods.

Example 2.3. (two-node cyclic QN)

Assume that QN is a closed network with FCFS service discipline and exponential service time. Moreover, it has an independent service time. Let S = (S_1,S_2) be a system state with $S_i = n_i$.



The behave of QN govern by the birth-death Markov process with a closed-form solution



Since $\rho = (\mu_1/\mu_2)$, then

$$\boldsymbol{\pi}\,(n_1, n_2)=(1/G)\,\rho^{n2}\,,\ \ 0\leq n_1\leq K\ ,\ n_2=K-n_1$$

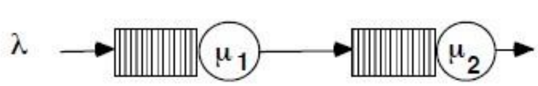
$$G\ =\ \Sigma_{0\leq k\leq k}\ \rho^{\,k}\ =\ (1-\ \rho^{k+1}\ )\ /\ (1-\ \rho\ )$$

$$\pi(\mathrm{K}-\mathrm{k},\mathrm{k}) = \pi(\mathrm{K},0)\rho^{\mathrm{k}} \qquad 0\leq k\leq K$$

$$\pi(K,0) = (1-\rho)/(1-\rho^{K+1})$$

Example 2.4. (Tandem QN)

Let QN be an open network with the arrival Poisson process and exponential service times. Furthermore, it has an independence assumption with FCFS discipline, and its state system is $S = (S_1, S_2)$ with $S_i = n_i$.



Let **E** = { (n_1, n_2) | $n_i\geq 0$} be a state space with $\boldsymbol{\pi}(n_1, n_2)$ stationary state probability. The behave of QN is not subjected to the birth-death Markov process. so it has a complex structure of global balance equations. However, node 1 can be analyzed independently as M/M/1 system with parameters λ and μ_1. Therefore, we have

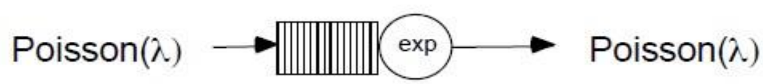
$$\pi_1(k)= \rho_1^k (1- \rho_1) \qquad k\geq 0 \text{ if } \rho_1 = \lambda / \mu_1<1.$$

The arrival process of node 2 is subjected to the following theorem

Theorem 2.5. (Burke's Theorem)

The departure process of a stable M/M/1 is a Poisson process with the same parameters as the arrival process.

Burke theorem also holds for M/M/m and M/G/∞



So, node 2 has a Poisson arrival process (λ), and isolated node 2 is an M/M/1 system with parameters λ and μ_2. Hence, we get

$$\pi_2(k) = \rho_2^k (1-\rho_2) \qquad k\geq 0 \text{ if } \rho_2 = \lambda / \mu_2<1$$

Moreover, for the independence assumption

$$\pi(n_1, n_2)= \pi_1(n_1)\, \pi_2(n_2)= \rho_1^{n1}\, \rho_2^{n2} (1-\rho_1) (1-\rho_2)$$

That satisfies the process global balance equation $\boldsymbol{\pi Q=0}$.

Some extensions as an immediate application of Burke theorem's together with the property of **composition and decomposition** of **Poisson** processes leads to a closed form solution of the state probability for a class of QN with exponential service time distribution, FCFS discipline, exponential interarrival time (Poisson arrivals, parameter γ_i), and independence assumption.

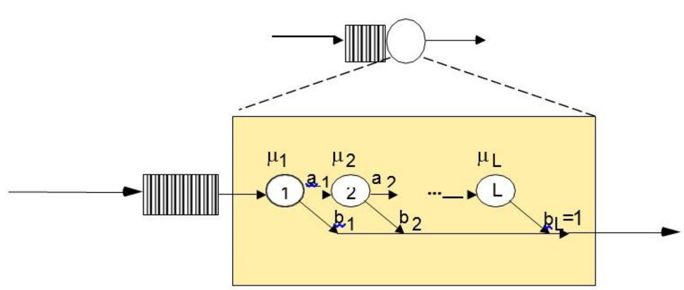
The acyclic probabilistic routing topology (triangular routing matrix P)

$$\pi(n_1, n_2, \dots, n_M) = \prod_{i=f}^{M} pr_i\{n_i\}$$

where

$$\begin{aligned} &pr_i(k) = \rho_i^k(1-\rho_i) \qquad && k \geq 0 \quad if \quad \rho_i = \lambda_i/\mu_i < 1 \\ &\text{and } \lambda_i = \gamma_i + \Sigma_j \lambda_j p_{ji} && 0 \leq i \leq M \end{aligned} \quad \text{(traffic equations)}$$

Coxian distributions are formed by a network of exponential stages such as L exponential stages. Stage ℓ has a service rate $\mu\,\ell$ and probabilities $a\ell, b\ell$ such that $a\ell + b\ell = 1$.



Coxian distributions have rational Laplace transform can be used to represent general distribution with rational Laplace transform, and it is helpful to approximate any general distribution with known bounds. PH-distributions (phase-type) have similar representations and properties.

2.5. Product-form Queueing Networks

The **product-form solution of** the stationary state probability $\boldsymbol{\pi}$ can be computed (under certain constraints) as the product of a set of functions, each dependent only on the state of a station. Other average performance indices can be derived by state probability $\boldsymbol{\pi}$, for example, **Jackson theorem** for open exponential-FCFS networks, **Gordon-Newell theorem** for closed exponential-FCFS networks, and **BCMP theorem** for open, closed, or mixed QN with various types of nodes. Moreover, the computationally efficient exact solution algorithms are considered, such as convolution algorithm and mean value analysis. The solution is obtained **as if** the QN is formed by independent M/M/1 (or M/M/m) nodes. The types of nodes are given

1) FCFS and exponential chain independent service time

2) PS

3) IS } and Coxian service time

4) LCFCPr

For types 2-4, the service rate may also depend on the customer chain. Let $\mu_i^{(c)}$ denote the service rate for node i and chain c. Let $\mu_i^{(c)} = \mu_i$ for each chain c, for type-1 nodes. Consider single-class multiple-chain QN such that open, closed, or mixed QN with M nodes of types 1-4 and Poisson arrivals with parameter $\lambda(n)$ dependent on the overall QN population n. Assume that we have R classes, C chains, and $K^{(c)}$ be the population for each closed chain $c \in C$. Let the external arrival probabilities denote by $p_{0,i}^{(c)}$ for each open chain $c \in C$ and the routing probability matrices denote by $\mathbf{P}^{(c)}$ for each chain $c \in C$. The **traffic equation system** derives the **visit ratio** of (relative) throughputs $e_i^{(c)}$

$$e_i^{(c)} = p_{0,i}^{(c)} + \Sigma_j e_j^{(c)} p_{ji} \qquad 1 \le i \le M \quad , \quad 1 \le c \le C$$

Theorem 2.6. (BCMP stand by Baskett, Chandy, Muntz, Palacios 1975**)**

For open, closed, or mixed QN with M nodes of types 1-4 and the assumptions above, let $\rho_i^{(c)} = e_i^{(c)} / \mu_i^{(c)}$ for each node i and each chain c. If the system is stable, i.e., if $\rho_i^{(c)} < 1 \; \forall i, \forall c$, then the steady-state probability can be computed as the product-form:

$$\pi(\mathbf{S}) = \frac{1}{G} \mathrm{d(n)} \prod_{i=1}^{\mathrm{M}} \mathrm{g_i(n_i)}$$

where G is a normalizing constant. The function d(n)=1 for closed network, and for open and mixed network, this function depends on the arrival functions as follows:

$$d(\mathbf{n}) = \prod_{k=0}^{n-1} \lambda(k), d(\mathbf{n}) = \prod_{c=1}^{C} \prod_{k=0}^{n^{(c)}-1} \lambda_c(k)$$

for arrival rates dependent on the number of customers in the network n, or in the network and chain c, functions $g_i(n_i)$ depend on node type as follows:

The functions $g_i(n_i)$ depend only on node i, parameters $e_i^{(c)}$, $\mu_i^{(c)}$ and state $n_i^{(c)}$

$$g_i(n_i) = n_i! \prod_{c=1}^{C} \frac{\left(\rho_i^{(c)}\right)^{n_i^{(c)}}}{n_i^{(c)}!}$$

for nodes of type 1, 2, and 4 (FCFS, PS, and LCFSPr)

$$g_i(n_i) = \prod_{c=1}^{C} \frac{\left(p_i^{(c)}\right)^{n_i^{(c)}}}{n_i^{(c)}!}$$

for nodes of type 3 (IS)

Where $n_i^{(c)}$ number of customers in node i and chain c

n_i number of customers in node i

The proof is based on the detailed definition of the network state and by substitution of the product-form expression into the global balance equations of the associated continuous-time Markov process.

Chapter Three

Conclusions and Future Works

Chapter Three

Conclusions and Future Works

3.1. Conclusions

In this research, we propose queueing network systems and their properties. The queueing network model extends the elementary systems as stochastic models were first established to characterize the entire system by a single service center.

3.2. Future Works

In the upcoming, we want to discover the queueing network with stochastic models but a non-Markovian version. Moreover, we hope to find some essential properties with some applications in real-life situations.

References

[1] Kleinrock, L. Queueing Systems, vol. 1 (Theory). John Wiley and Sons, 1975.

[2] Gelenbe, E., and Mitrani, I. Analysis and Synthesis of Computer Systems. Academic Press, New York, 1980.

[3] Lavenberg, S. S. Computer Performance Modeling Handbook. Academic Press, New York, 1983.

[4] Trivedi, K. S. Probability and statistics with reliability, queuing and computer science applications, second ed. Wiley-Interscience, 2002.

[5] Kant, K. Introduction to Computer System Performance Evaluation. McGrawHill, 1992.

الخلاصة

ناقشنا في هذا البحث البنى لطوابير الشبكات وبعض الطرق الخاصة بحلها. عدة أنواع مختلفة من هذه الطوابير درست اعتمادا على صياغتها. اعتبرت عمليات تصادفية مهمه المثال عملية الولادة والوفاة و M\G\1 لتشكل السلوك الخاص بطوابير الشبكات. بعض النظريات و الأمثلة المهمه لايجاد التوزيع الثابت لمعادلة التوازن الكلية.

جمهورية العراق

وزارة التعليم العالي والبحث العلمي

جامعة بابل

كلية التربية للعلوم الصرفة

قسم الرياضيات

حول طابور الشبكة وتطبيقاته

بحث مقدم

إلى مجلس قسم الرياضيات، كلية التربية للعلوم الصرفة، جامعة بابل

كجزء من متطلبات نيل درجة الدبلوم العالي في التربية / الرياضيات

مقدمة من قبل الطالب

محسن منصور عبدالله حمود

بإشراف الدكتور

علي حسين محمود العبيدي

2022 أيلول