

# Using Markov Chain to Analyze Production Lines Systems with Layout Constraints

Sadegh Abedi<sup>1,\*</sup>, Morteza Mousakhani<sup>2</sup>, Naser Hamidi<sup>3</sup>

<sup>1</sup> *Young research club Islamic Azad University, Qazvin Branch, Iran*

<sup>2</sup> *Department of management and accounting, Islamic Azad university of Qazvin, Qazvin, Iran*

<sup>3</sup> *Department of management and accounting, Islamic Azad university of Qazvin, Qazvin, Iran*

Received 20 Apr., 2009; Revised 15 May., 2009; Accepted 12 Jun. 2009

---

## ABSTRACT

There are some problems with estimating the time required for the manufacturing process of products, especially when there is a variable serving time, like control stage. These problems will cause overestimation of process time. Layout constraints, reworking constraints and inflexible product schedule in multi product lines need a precise planning to reduce volume in particular situation of line stock. In this article, a hybrid model has been presented by analyzing real queue systems with layout constraints as well as by using concepts and principles of Markov chain in queue theory. This model can serve as benchmark to assess queue systems with probable parameters of service. Here, the proposed model will be described drawing on the findings of a case study. Thus, production lines of a home application manufacturer will be analyzed.

*Keywords:* Queuing Theory, Markov Chain, layout, Line Balance.

---

## 1. INTRODUCTION

Queuing theory is one of the oldest and best-developed analysis techniques that are daily used in waiting line [5]. The main purpose of manufacturers and service providers is to satisfy customer [18]. This satisfaction is manifested by customer-desired characteristics. Gaining goods or services as soon as possible is a characteristic desired by the customer [19]. To optimize decisions and to reduce waiting time for customers, manufacturers and service providers have to use queuing theory. This will help them to specify the essential resource level, that should be allocated and to gain customers satisfaction as much as possible [9]. Resource allocation and customer satisfaction are very important for companies especially in a highly competitive environment. Therefore, investigating and describing queuing systems' performance in different environments is an essential issue. Queuing systems layout [16], servers' allocation [26], customer allocation to optimize the use of servers [4] and serving type [7] are all but some of factors that should be considered in real condition analysis. A layout that is not optimized can make long queues and consequently increase the waiting time.

Markov chain connection, with time and its correlation with Exponential distribution on the other has caused these stochastic models to be applied in real circumstances [10]. That is why the exponential distribution is usually a good fitness for real life problems linking process movements in time to Markov chains [13].

Next, different systems of physical layout of the performance assessment by line systems of Markov chains in the description of serving systems is investigated. Then, the situation of proposed line system regarding the complexity and constraints of the model in the present situation will be analyzed by using Markov chains.

Some new studies on the fields in which the Markov chain is used include: a single-server multi-station alternating queue where the preparation times and the service times are auto- and cross-correlated [15], the algorithmic development of the full busy period for the model under consideration [2], analysis of MAP/G/1 G-queues with possible preemptive resumption service discipline and multiple vacations wherein the arrival process of negative customers is Markovian arrival process [27]. Some studies on manufacturing systems

---

\*Corresponding author e-mail: sadegh.abedi@yahoo.com

field include Markov decision models for the optimal maintenance of a production unit with an upstream buffer [1], A semi-Markov decision algorithm

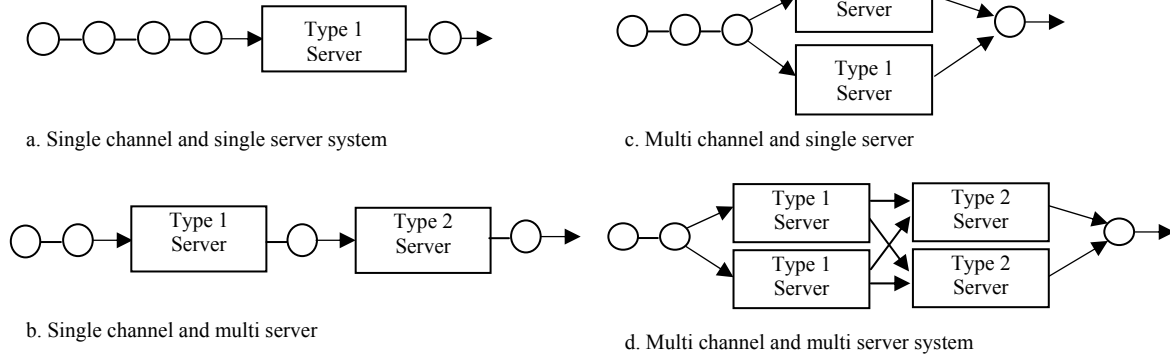


Fig.1. Queuing system types

For the maintenance of a production system with buffer capacity and continuous, repair times in 2008 year [23]. We will first look at queuing systems based on layout, performance appraisal and Markov chains in service systems, and then given the complexity and limitations of the purposed model, it is discussed for a system with two servers and a system with servers using Markov chain model.

## 2. THEORETICAL CONCEPTS

Queuing theory began in 1909 by studies of a Danish engineer, A.K.Erlang [4]. Studying and doing experiments on increasing and decreasing demand for telephone systems, he discussed factors and relations in this system. Eight years later, he published the studies details about telephone systems automation and the results of relations. This became the base of queuing theories. At the end of World War II, he developed the usage of queuing models in public sections and business [6]. Queuing theory is one of the oldest and best-developed techniques that can be used to analyze waiting lines [14].

### 2.1 Types of queuing systems based on layouts

In this classification, queuing systems are analyzed based on the type of Terms Channel and servers [9]. At simplest system, there is one terms channel and one server. This system is called Single channel and Single server system. At the most complicated system, there are several terms channels and several servers. This system is called multi- channel and multi- server queue. Fig. 1 shows types of queuing systems categorized by physical layout.

### 2.2. Utilization factor

In a single server system G/G/1 with an arrival rate  $\lambda$  and mean service rate  $E(B)$ , the work arrived in time unit is  $\lambda E(B)$ . If customers arrival rate is bigger than

service.  
rate,  $\lambda E(B) > 1$ , then the system's capacity would not be enough for all demands and the queue will go longer through infinity until  $\lambda E(B)$  equals to 1. So the equation  $\lambda E(B) < 1$  is steady state of most of the queuing systems. Unlike, D/D/1 system, non-stochastic state Systems and non-group systems do not need this condition. The usage of the equation is as follows [20]:  
 $\rho = \lambda E(B)$

If  $\rho < 1$  then  $\rho$  is Occupation rate or Utilization factor. This ratio is equal to the arrival rate into system to maximize the capacity of system to do the work.

### 2.3. The factors for appraising Queuing systems performance

To evaluate the performance of queuing systems, these factors are used [11]:

- The distribution of arrival time of people into system and the waiting time of customers in system. The waiting time for each customer in system will be the waiting time to get service plus time of serving.

$$T_s^{(n)} = T_l^{(n)} + S_n$$

- The distribution of the number of customers that are in system (involving people who are being served).

- The distribution of service time which involves service time for customers who are already serving plus serving time of other people who are in waiting line. Finally, we can evaluate the system's performance by using these parameters as well as mean waiting time and mean sojourn time now given the G/G/C model, if the random variable  $L(t)$  stands for the number of customers existing in system in time  $t$  and  $S_n$  stands for waiting time of  $n$  customers in system, it can be shown that the random variable  $L(t)$  has a limiting distribution by assuming utilization factor  $\rho < 1$ , for  $n \rightarrow \infty$  and,  $t \rightarrow \infty$  [8].

Assuming that L and S random variables have limiting distributions which are called L(t) and Sn , then the P<sub>k</sub>, probability of existence of k people in system for long time will be shown in equation(1):

$$p_k = P(L = K) = \lim_{t \rightarrow \infty} P(L(t) = K) \quad (1)$$

F(x) is the probability that waiting time for a customer is less than x time units:

$$F_s(X) = P(S \leq x) = \lim_{n \rightarrow \infty} P(S_n \leq x)$$

Also as can be seen below [23], mean number of customers in system for time period of [0,t] and also mean service rate in long time are E(L) and E(S) . These will be shown in equation (2, 3):

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_{x=0}^t L(X) dx = E(L) \quad (2)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n S_k = E(S) \quad (3)$$

These two parameters have basic applications in analyzing queuing systems [20]. Also according to Little's law, there is a very important relationship between E(L) (mean number of customers in system), E(S) (mean waiting time) and λ (mean arrival rate), that is shown in equation (4) below [24]:

$$E(L) = \lambda E(S) \quad (4)$$

2.4- Markov chains with continuous time

To analyze queuing models, we need to formulate them as Markov chains. Therefore, we will first define stochastic processes and basic concepts of Markov chains. A stochastic process [21] is a set of random variables X(t), if for each t ∈ T , there is a random variable, then X(t) can be considered as stochastic process state in time[10]. Stochastic process with continuous time t ≥ 0 takes integer and non-negative values. Process [N(t), t ≥ 0] is a Markov chain with continuous time, if we have the equation as below (5) [6].

$$p\{X(t+s)=j|X(s)=i, X(u)=x(u), 0 \leq u \leq s\} \quad (5)$$

$$p\{X(t+s)=X(s)=i\}$$

For all s, t ≥ 0 and non negative i, j, X(u), 0 < u < s. It follows that given the present state; X(t) and past state; X(u), a Markov chain with continuous time is a stochastic process with Markov characteristics whose future state conditional probability, X(t+s), depends on present state which have no relation to past state of process. If in a Markov chain, the below term is independent of s, then the Markov chain is called homogeneous. In other word if we have the equation (6) below [22].

$$P_{ij}(t) = p\{X(t+s)=j|X(s)=i\} = P\{X(t)=j|X(0)=i\} \quad (6)$$

Then the Markov chain is a homogeneous Markov. Note that P<sub>ij</sub>(t) is the probability of system's transition

from i state to j state in t time. The relation between Markov chain and time on one hand and exponential distribution on the other hand were made in real life problems[24]. This is because the Exponential distribution is usually a good estimation for real conditions and connects process changes in time to Markov chains.

3. RESEARCH PROBLRM DELINEATION

There are some problems with estimating time on the serviced manufacturing process of products, especially when there is variable serving time like control stage. These problems will cause overestimation of process time. For instance, a home application manufacturer can be used as a case. In production processes like that, products will be put on conveyers in vacuum and gas-charge station after assembling. Now after quality inspection, products will be delivered to cooling test station. Here if the workstation is not idle, then the products should be stocked. The flow process chart of production line is shown in figure 2.

The layout of cooling test machines has been designed in a way that the products (but those products that are served in the first row) cannot leave the system after test operation. These products will wait for the first product to leave the system and then will leave one by one. This will increase the product test mean time and therefore increases operation's mean waiting time. Queue model of this system involves a line stock that is overloaded (because of space limitation).

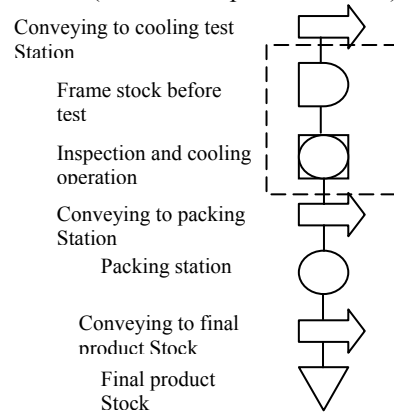


Fig. 2. The flow process chart of production line

After a time in this stock, products will enter the cooling test station. It should be mentioned that in this station, there are two test machines in each row, which work in parallel channels. In such a system if both machines are idle, then the product will be referred to first server in row. If the first machine is busy, the product will be referred to second server for cooling test operation. In this system while serving the first product is not finished, the second one cannot leave the system. Also in this system, since the second server is busy and the first server is idle, the next product in

waiting line cannot use the first server (because of physical limitations). In figure below, a schematic plan of layout of cooling test station is presented.

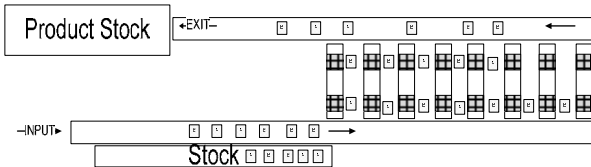


Fig. 3. A schematic plan of cooling test station

For this system, it is obvious that because of variable service rate and physical limitations, we cannot use available queuing models to analyze it. So we assume that product's arrival is based on Poisson process with  $\lambda$  parameter and two servers in each row as series (in special conditions) in which serving time is an exponential random variable with  $\mu_i$  parameter. Considering fact that the performance of each row is independence of others in practice, so firstly we will model the exits queuing system based on two servers in a row with  $\mu_1$  and  $\mu_2$  parameters and will develop M/M/C model using Markov chain concepts. At last, we will find out the ultimate model by developing the proposed model with i rows.

### 3.1 Modeling the exits queuing system based on one row with two servers

To model the proposed queuing system as shown in figure 4, we will define parameters, variables and assumptions as below:

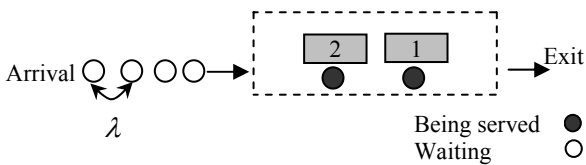


Fig. 4. Plan of a two server with one row system

- Products' arrival is based on Poisson process with  $\lambda$  parameter.
- Serving time for first product is considered,  $t_1$  and serving time for second product is considered,  $t_2$ .
- Just a one-row queue is possible.
- Two servers are serving as series in special conditions.
- If both servers were idle, the product would be referred to first server and if the first server were busy then the product would be referred to the second server.
- While serving to the first product is not finished, the second product cannot leave the system.

Period	$\Delta T$				$T$	
1 <sup>st</sup> server's output	1	2	3	4	n-1	n
Customer's output rate for	$t_1$	$2t_1$	$3t_1$	$4t_1$	$(n-1)t_1$	$(n)t_1$
1 <sup>st</sup> server and 2 <sup>nd</sup> server	$t_2$	$t_1+t_2$	$2t_1+t_2$	$3t_1+t_2$	$(n-2)t_1+t_2$	$(n-1)t_1+t_2$
2 <sup>nd</sup> server's output	1	2	3	4	n-1	n

Crosshatched area stands for second server idle time

Fig. 5. Serving system for first situation

- In this system if the second server is busy and the first server is idle, then the product in queue cannot be referred to first server. To model this queuing system, we will divide it in two classes based on serving time:

- o The conditions in which, serving time of server 1 is equal or more than serving time of server 2. It means product 1 is served in a time that is equal to or more than second product's time.
- o The conditions in which serving time of server 2 is more than that of server 1. It means product 2 is served in a time that is more than first product's time.

### 3.2 Modeling queuing system based on first assumption

In this condition, serving time of server 1,  $t_1$  is equal or more than that of server 2,  $t_2$ . It means  $t_1 \geq t_2$  and so  $\mu_1 \leq \mu_2$ . In fact, it is a variable serving system and serving rate of servers is not equal. If there is no limitation for departure from this system, then in  $\Delta T$  period, the first server will serve  $[\mu_1]$  products and second server will serve  $[\mu_2]$  products. It means that in  $\Delta T$  we expect  $[\mu_1]+[\mu_2]$  products to leave system. Nevertheless, in practice because of limitation of space to departure, the second server's product cannot leave the system until the first product leaves it. As it can be seen in figure 4, in  $\Delta T$  period, the first server can at most serve  $n$  products whereas  $nt_1 \leq \Delta T$ . So  $n$  is the same as  $\mu_1$  or first serving rate. The maximum products that second server can serve in  $\Delta T$  period is  $n$ , whereas  $(n-1)t_1+t_2 \leq \Delta T$ . Therefore,  $n$  is  $\mu_2$  or real second serving rate. Considering these two equations, it can be said; if  $X'(\Delta T)$  would be maximum products that can be served in  $\Delta T$  period and leave the system, then we have:  $X'(\Delta T) = 2\mu_1$

In figure 4, although the second server can serve more products, but because of limitation of space, the next product cannot arrive into system until the first product's serving is finished and should be idle until then. As equation below shows, the total idle time of second server in  $\Delta T$  period is:  $I(\Delta T) = (t_1 - t_2)\mu_1$

It can be concluded that if first server would spend  $t_1$  time to serve each product, then the second server spends the same time. That is why their real serving time is not equal but considering product waiting, serving time of previous product is the time it spends

on serving system. Therefore, if first product is been served with  $t_1$  time, the second product's serving time is also  $t_1$ . Here  $t_1$  for first product is his total serving time whereas for second product,  $t_1$  is  $t_2$ , time for serving, plus  $\beta$  which is waiting for first product's leaving; it is shown in equation(7):

$$t_1 = t_2 + \beta \tag{7}$$

Therefore as equation (8) shows, serving rate for both of them is  $\mu_1$ .

$$\mu_1 : \frac{\Delta T}{t_1} \tag{8}$$

$$\mu_2 : \frac{\Delta T}{t_2 + \beta} = \frac{\Delta T}{t_1} = \mu_1$$

This makes the explained model as an M/M/C=2 model in which serving rate of both servers is  $\mu_1$ . Therefore, we have equation (9) as below:

$$\lambda_n = \lambda \quad : n = 0,1,2,\dots$$

$$\mu_n = \begin{cases} \mu_1 & n = 1 \text{ There is no waiting line} \\ 2\mu_1 & n = 2, 3, \dots \end{cases} \tag{9}$$

To explain this model, we will define possible conditions as table 1. Being so in this system, always  $t_1 > t_2$ , so this (0, 1) condition\* will never happen.

Table1  
Possible situations for system in first condition

State	State description
0	System is empty
(1,0)	There is a product in system and this product is being served by server 1
(0,1)	There is a product in system and this product is being served by server 2*
(1,1)	There are two products in system and they are being served by server 1 and 2
$n \geq 2$	There are $n$ products in system

Based on model assumptions, limitations and transition matrix, the transition rate diagram will be as follows:

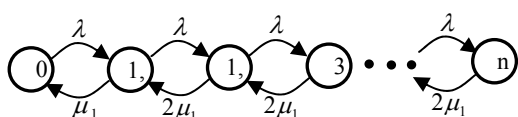


Fig. 6. The transition rate diagram for first situation

Moreover, by using this diagram and equality of arrival rate and departure rate for each state, the equilibrium equation set will be as follows in equation (10):

state	equilibrium equation set	
0	$\lambda P_0 = \mu_1 P_{10}$	
(1,0)	$(\lambda + \mu_1)P_{10} = \lambda P_0 + (2\mu_1)P_{11}$	(10)
(1,1)	$(\lambda + 2\mu_1)P_{11} = \lambda P_{10} + (2\mu_1)P_{31}$	
...	...	
...	...	
...	...	
$n > 2$	$(\lambda + 2\mu_1)P_n = \lambda P_{n-1} + (2\mu_1)P_{n+1}$	

As it can be seen in transition rate diagram (figure 5) for each state of  $n$ , except zero state, there are two

transition possible states:  $(n+1)$  and  $(n-1)$ . Therefore, it is easy to use one of variables to solve these equations. Here  $P_0$  is a good base to solve equations. We also have equation (11) as below:

$$C_n = \begin{cases} \left(\frac{\lambda}{\mu_1}\right) & n = 1 \text{ There is no waiting line} \\ \left(\frac{\lambda}{2\mu_1}\right)^{n-1} \frac{\lambda}{\mu_1} & n = 2, 3, \dots \end{cases} \tag{11}$$

Considering this definition, for  $n > 1$ , we will have:

$$P_n = C_n \cdot P_0$$

To calculate  $P_0$ , by assuming  $\sum_{n=0}^{\infty} P_n = 1$  we have

equation (12) as below:

$$P_0 = \left(1 + \frac{\lambda}{\mu_1} + \frac{\lambda^2}{\mu_1(2\mu_1 - \lambda)}\right)^{-1} \tag{12}$$

Therefore, to calculate the probability of existence of  $n$  people in system for long time (equation 13):

$$P_n = \begin{cases} \left(\frac{\lambda}{\mu_1}\right)^n P_0 & : n = 0,1,2 \\ \left(\frac{\lambda}{2\mu_1}\right)^{n-1} \frac{\lambda}{\mu_1} P_0 & : 2 \leq n \end{cases} \tag{13}$$

To calculate integrative criteria like averages number of products in queue ( $L_q$ ) and expected number of products in system ( $L$ ), considering little's relations and steady state condition of system that is:

$\rho = \frac{\lambda}{2\mu_1} < 1$  Then we have equation (14, 15) as below:

$$L_q = \frac{P_0}{2} \left(\frac{\lambda}{\mu_1}\right)^2 \frac{\rho}{(1-\rho)^2} \tag{14}$$

$$L = L_q + \frac{\lambda}{\mu_1} = \frac{P_0}{2} \left(\frac{\lambda}{\mu_1}\right)^2 \frac{\rho}{(1-\rho)^2} + \frac{\lambda}{\mu_1} \tag{15}$$

Also to calculate time criteria like, expected waiting time in queue ( $W_q$ ) and expected waiting time in system ( $W$ ), considering Little's relations and definitions, we have equation (16,17) as below:

$$W_q = \frac{L_q}{\lambda} = \frac{\pi_0}{2} \left(\frac{\lambda}{\mu_1^2}\right) \frac{\rho}{(1-\rho)^2} \tag{16}$$

$$W = W_q + \frac{1}{\mu_1} = \frac{\pi_0}{2} \left(\frac{\lambda}{\mu_1^2}\right) \frac{\rho}{(1-\rho)^2} + \frac{1}{\mu_1} \tag{17}$$

### 3.3 Queuing system modeling based on second assumption

Here, unlike the previous assumption, it is assumed that the serving time of server 2  $t_2$  is more than serving time of server 1,  $t_1$ . That is,  $t_1 \leq t_2$  and  $\mu_1 \geq \mu_2$ . If there is no departure limitation in the fore mentioned system, then in  $\Delta T$  period, the first server will serve  $[\mu_1]$  products and second server will serve  $[\mu_2]$  products. It means in  $\Delta T$  we expect  $[\mu_1] + [\mu_2]$  products

to leave system, but because of space Limitation for departure, if the first product would leave the system, the product waiting in line cannot be replaced. It is so because the second server is busy and so the whole system is busy. As it can be seen in fig. 6, in  $\Delta T$  period, the first server can serve at most  $n$  products whereas  $(n-1)t_2 + t_1 \leq \Delta T$ . Therefore  $n$  is  $\mu_1$  or real serving rate of first server. Also in  $\Delta T$  period, the second server can serve at most  $n$  products whereas  $nt_1 \leq \Delta T$ . Here  $n$  is the same as  $\mu_2$  serving rate of second server. Considering these two relations, it can be concluded that if maximum products, which can be served and leave the system, would be  $X'(\Delta T)$  in  $\Delta T$  period, then we have:  $X'(\Delta T) = 2\mu_2$

In this figure, although the first server can serve more products, because of limitation of space, the next product cannot arrive into system until the second product's serving is finished and therefore server 2 should be idle until then. As below equation, the total idle time of server 1 in  $\Delta T$  period is:

$$I(\Delta T) = (t_2 - t_1)\mu_2$$

From the viewpoint of product waiting at line, it can be concluded that if second product has been served with  $t_2$  time, the first product's serving time is also  $t_2$ , is totally several to the second product total serving time whereas for first product,  $t_2$  is  $t_1$ , time for serving, plus  $\alpha$  which is the time of the second product departure.

Therefore  $t_1 + \alpha$  is serving time of product 1 and  $t_2$  is serving time for product's departure 2. So serving rate for both of them will be  $\mu_2$  equation (18):

$$\mu_1 : \frac{\Delta T}{t_1 + \alpha} = \frac{\Delta T}{t_2} = \mu_2 \quad (18)$$

Period	0	1	2	3	4	n-1	n	T
1 <sup>st</sup> server's output	1	2	3	4	n-1	n		
Customer's output rate for	$t_1$	$t_1+t_2$	$2t_2+t_1$	$3t_2+t_1$	$(n-2)t_2+t_1$	$(n-1)t_2+t_1$	$(n)t_2+t_1$	
1 <sup>st</sup> server and 2 <sup>nd</sup> server	$t_2$	$2t_2$	$3t_2$	$4t_2$	$(n-1)t_2$	$(n)t_2$		
2 <sup>nd</sup> server's output	1	2	3	4	n-1	n		

Crosshatched area stands for first server idle time

Fig. 7. Serving system for second situation

At last, the model with assumption 2 will be as equation (20):

$$\lambda_n = \lambda \quad : n = 0,1,2,\dots$$

$$\mu_n = \begin{cases} \mu_1 & n = 1 \\ 2\mu_2 & n = 2,3,\dots \end{cases} \quad \text{There is no waiting line} \quad (20)$$

To explain model we will define possible conditions as table 2. Since in this system always  $t_1 < t_2$ , so this (1,0) condition\* will never happen

Table2  
Possible situations for system in second condition

State	State description
0	System is empty
(1,0)	There is a product in system and this product is being served by server 1*
(0,1)	There is a product in system and this product is being served by server 2
(1,1)	There are two products in system and they are being served by server 1 and 2
n > 2	There are n products in system

Based on model assumptions, limitations and transition matrix in arrival and processes departure, the transition rate diagram will be as follows:

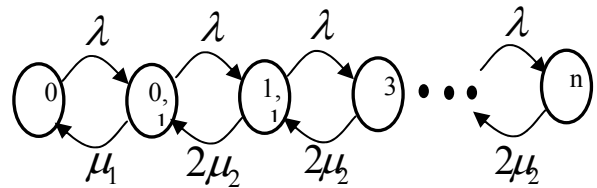


Fig . 8. The transition rate diagram for second situation

Moreover, drawing on this diagram and equality of arrival rate and departure rate for each state, the equilibrium equation set will be as follows in equation (21):

state	equilibrium equation set
0	$\lambda P_0 = \mu_1 P_{10}$
(1,0)	$(\lambda + \mu_1)P_{10} = \lambda P_0 + (2\mu_2)P_{11}$
(1,1)	$(\lambda + 2\mu_2)P_{11} = \lambda P_{10} + (2\mu_2)P_3$
...	...
...	...
...	...
n > 2	$(\lambda + 2\mu_2)P_n = \lambda P_{n-1} + (2\mu_2)P_{n+1}$

(21)

By solving this equations set and considering,

$$\sum_{n=0}^{\infty} P_n = 1$$

We have equation (22) as follows:

$$C_n = \begin{cases} \frac{\lambda}{\mu_1} & n = 1 \quad \text{no waiting line} \\ \left(\frac{\lambda}{2\mu_2}\right)^{n-1} \frac{\lambda}{\mu_1} & n = 2, 3, \dots \end{cases} \quad (22)$$

Considering this definition (equation 23) for  $n = 1,2,\dots$

$$P_n = C_n \cdot P_0 \quad (23)$$

To calculate  $P_0$ , we have equation (24) as follows:

$$P_0 = \left(1 + \frac{\lambda}{\mu_1} + \frac{\lambda^2}{\mu_1(2\mu_2 - \lambda)}\right)^{-1} \quad (24)$$

Therefore, to calculate the probability of existence of  $n$  people in system for long time, we have equation (25):

$$P_n = \begin{cases} \left(\frac{\lambda}{\mu_1}\right)^n P_0 & : n = 0,1 \\ \left(\frac{\lambda}{2\mu_2}\right)^{n-1} \frac{\lambda}{\mu_1} P_0 & : 2 \leq n \end{cases} \quad (25)$$

Considering Little's relations and steady state condition of system ,to calculate integrative criteria like expected number of products in queue ( $L_q$ ) and expected number of products in system ( $L$ ), we have equation (26,27) as follows:

$$L_q = \frac{\pi_0}{2} \left(\frac{\lambda^2}{\mu_1\mu_2}\right) \frac{\rho}{(1-\rho)^2} \quad (26)$$

$$L = L_q + \frac{\lambda}{\mu_2} = \frac{P_0}{2} \left(\frac{\lambda^2}{\mu_1\mu_2}\right) \frac{\rho}{(1-\rho)^2} + \frac{\lambda}{\mu_2} \quad (27)$$

Also considering Little's relations and definitions ,to calculate time criteria like, expected waiting time in queue ( $W_q$ ) and expected waiting time in system ( $W$ ), we have equation(28,29) as follows:

$$W_q = \frac{L_q}{\lambda} = \frac{P_0}{2} \left(\frac{\lambda}{\mu_1\mu_2}\right) \frac{\rho}{(1-\rho)^2} \quad (28)$$

$$W = W_q + \frac{1}{\mu_2} = \frac{P_0}{2} \left(\frac{\lambda}{\mu_1\mu_2}\right) \frac{\rho}{(1-\rho)^2} + \frac{1}{\mu_2} \quad (29)$$

3.4- The proposed hybrid model by assuming two servers

In this article, we first modeled the queuing system with two servers in one row with  $\mu_1$  and  $\mu_2$  parameters by developing M/M/C model and Markov chain concepts. However, in practice, these conditions do not occur one by one, but almost a combination of both states is possible. To have a hybrid model we need two new variables. We define  $\theta_1$  as probability of state one occurring and  $\theta_2$  as probability of state two occurring in  $\Delta T$  period. To calculate variables in hybrid conditions,  $Z_{(c=2)}$  we assume that the first condition will happen and then the second condition will happen and considering probability values  $\theta_1$  and  $\theta_2$  we will use weighted average in equation (30) as follows:

$$Z_{(c=2)} = \frac{\theta_1 X + \theta_2 Y}{\theta_1 + \theta_2} \quad (30)$$

X: decision variables if state one occurs  
Y: decision variables if state two occurs

3.5 Developing proposed hybrid model by assuming C servers:

The results of two server's conditions can be used to analyze the performance of whole system. As it can be

seen in figure 8, unlike M/M/C mode, when the product arrives at point A, it does not have the possibility of choosing one of C servers. Here he should choose one of serving rows i. therefore after arriving into one of rows; the product has nothing to do with other branches and just can choose one of two servers in his row. Therefore C server model will become a special kind of hybrid serving system with two servers.

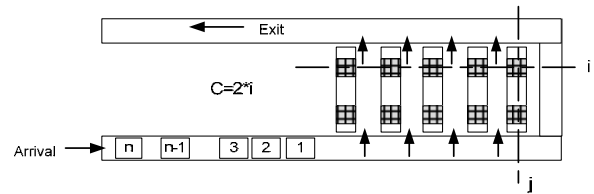


Fig. 9. Schematic plan of queuing system

If the number of products in the main queue (before decision point A) would be n and number of people in each branch (from point A to serving position) would be  $n_i$ , then , to calculate probabilities we need number of people in system, N, so we have:

Table 3  
Number of people in system

Number of production in system (N)	System state
n+n <sub>1</sub>	The production choose branch number 1
n+n <sub>2</sub>	The production choose branch number 2
n+n <sub>i-1</sub>	The production choose branch number i-1
n+n <sub>i</sub>	The production choose branch number i

To calculate parameters like  $\hat{L}$  and  $\hat{W}$  for queuing system, because of the similarity between serving conditions in all rows, we can modify product arrival rate ( $\lambda$ ) and introduce  $\hat{\lambda}$  as below and generalize the results to a C server system as equation (31).

$$\hat{\lambda} = \frac{\lambda}{i} \quad (31)$$

4. CASE STUDY

To study the relation between stock time in queue and number of servers (cooling test), data of a home factory was gathered as table 4. It should be mentioned that there are some factors related to time study and sampling that affect formation of waiting line. Therefore, by determining these factors and the design of experiments, the sample size and number of sampling were specified.

Table 4  
Gathered data of production line

Description	parameter	State 1	State 2	unit
		value		
Mean service rate of server 1	$\mu_1$	17	18	production per hour
Mean service rate of server 2	$\mu_2$	19	18	production per hour
Mean interracial time	$\lambda$	33		production per hour
Probability of events	$\theta_1, \theta_2$	45	55	percent
Number of cooling test (servers)	C	10		Numbers
Number of branches	i	5		Row
Maximum capacity of each branch	$n_i$	1		Numbers
Station area	S	200		square meters

Considering this information and by using hybrid model, system analysis results in a two-server row for long time below:

Table 5  
Information of production line

Description	parameter	State 1	State 2	unit
		value		
utilization factor	$\rho$	87	86	percent
Probability of existence of no production in system	$P_0$	2.23	2.38	percent
Expected number of production in queue	$L_q$	15	7	production
Expected number of production in system	L	17	9	production
Expected waiting time in queue	$W_q$	20	10	Minutes
Expected waiting time in system	$W$	23	13	Minutes

At last by using table 5 and multiplying probability factors  $\theta_1$  and  $\theta_2$  and modified arrival rate of  $\hat{\lambda}$  the results of total analysis for a six servers system are as follow.

Table 6  
Total analysis results

Description	parameter	value	unit
Modified arrival rate	$\hat{\lambda}$	7	production per hour
Total expected number of production in queue	$\hat{L}_q$	3	production
Total expected number of production in system	$\hat{L}$	5	production
Total expected waiting time in queue	$\hat{W}_q$	240	seconds
Total expected waiting time in system	$\hat{W}$	440	seconds

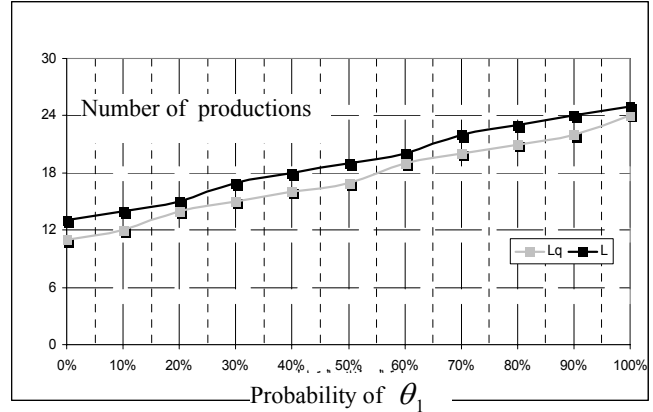


Fig. 10. Sensitivity analysis of number of productions in system based on  $\theta_1$

### 5. CONCLUSION

The close connection of Markov chain with time and its correlation with Exponential distribution has caused these stochastic models to be applied in real circumstances. The Exponential distribution is usually suitable for real life problems linking process movements in time to Markov chains. Here by analyzing a real queuing system with layout limitations in specific conditions and applying Markov chain concepts, a queuing model was developed. To model this queuing system, we divided it in two classes based on serving time. The conditions in which serving time of server 1 is equal or more than serving time of server 2 and the conditions in which serving time of server 2 is more than serving time of server 1. However, in practice, these conditions do not occur one at a time, but almost a combination of both states is possible. To have a hybrid model we need two new variables. We define  $\theta_1$  as a probability of occurring of state one and  $\theta_2$  as a probability of occurring of state two in  $\Delta T$  period. The results of last part for the two servers can be used to analyze whole system performance. This model can be a base for appraisal of queuing systems with probability parameters. By explaining a case study, we tried to describe the purposed model. To do this, production lines of a home application manufacturer will be analyzed. By analyzing the sensitivity on the occurrence probability of  $\theta_1, \theta_2$  on the average number of people in system and queue. As diagram 1 shows, the increase of the first  $\theta_1$  results in the increase of the second one in long time.

### REFERENCES

[1] E. G. Pavitsos, Markov decision models for the optimal maintenance of a production unit with an upstream buffer. Computers & Operations Research, 36(6), 1993-2006, 2009.  
 [2] B. Krishna Kumar, R. Rukmani, V. Thangaraj, On multi server



- feedback retrieval queue with finite buffer. *Applied Mathematical Modeling*, 33(4), 2062-2083, 2009.
- [3] B. D. Bunday, *An introduction to queuing theory*. Arnold, London, 1996.
- [4] C. ReVelle, K. Hogan, The maximum availability location problem. *Transportation Science*, 192-200, 1989.
- [5] R. B. Cooper, *Introduction to Queuing Theory*. 2<sup>nd</sup> ed., Elsevier North Holland, NY, 104-116, 1980.
- [6] D. Gross, C. M. Harris, *Fundamentals of queuing theory*. 2<sup>nd</sup> ed., John Wiley & Sons, NY, 1985.
- [7] F. S. Hillier, G. J. Lieberman, *Introduction to Operations Research*. 7<sup>th</sup> ed., McGraw-Hill, Singapore, 2001.
- [8] G. F. Newell, *Applications of queuing theory*. Chapman and Hall, London, 1971.
- [9] M. B. Gupta, R. B. Khanna, *Quantitative Techniques for Decision Making*. 2<sup>nd</sup> ed., Prentice Hall of India, New Delhi, 2006.
- [10] F. S. Hillier, G. J. Lieberman, *Introduction to Operation Research*. 3<sup>rd</sup> ed., Holden-Day Inc, Oakland, 1987.
- [11] I. vo & J. Resing, *Queuing Theory*. Eindhoven University of Technology, Netherlands, 2001.
- [12] J. D. Little, A proof of the queuing formula. *Operations Research*, 383-387, 1961.
- [13] S. Lipschutz, *Theory and Problems of Probability*. Mc Grow-Hill, 1968.
- [14] M. Abramowitz, I. A. Stegun, *Handbook of mathematical functions*. Dover, 1965.
- [15] M. Vlasidou, I. J. B. F. Adan, O. J. Boxma, A two-station queue with dependent preparation and service times. *European Journal of Operational Research*, 195(1), pp 104-116, 2009.
- [16] O. Berman, R. Mandowsky, Location-allocation on congested network. *European Journal of Operational Research*, 26, 238-250, 1986.
- [17] P. Bertail, S. Cl  men  on, Approximate regenerative-block bootstrap for Markov chains. *Computational Statistics & Data Analysis*, 52(5), 2739-2756, 2008.
- [18] A. Roper, T. Mason, F. Rossini, J. Banks, *Forecasting and Management of Technology*. John Wiley & Sons, NY, 1991.
- [19] B. Render, M. Stair, E. Hanna, *Quantitative Analysis Management*. 9<sup>th</sup> ed., Prentice Hall of India, New Delhi, 410-420, 2007.
- [20] S. M. Ross, *A First Course In Probability*. Macmillan Publishing, Inc, 1986.
- [21] S. M. Ross, *Introduction to Probability Models*. 2<sup>nd</sup> ed., Academic press, NY, 1980.
- [22] S. M. Ross, *Introduction to probability models*. 6<sup>th</sup> ed., Academic Press, London, 1997.
- [23] T. D. Dimitrakos, E. G. Kyriakidis, A semi-Markov decision algorithm for the maintenance of a production system with buffer capacity and continuous repair times. *International Journal of Production Economics*, 111(2), 752-762, 2008.
- [24] H. A. Taha, *Operation Research. An Introduction*, 2<sup>nd</sup> ed., Macmillan Publishing, Inc, 1976.
- [25] W. K. Ching, K. M. Ng, S. E. Fung, Higher-order multivariate Markov chains and their applications. *Linear Algebra and its Applications*, 428(2-3), 492-507, 2008.
- [26] Z. Drezner, *Facility Location: A Survey of Applications and Methods*. Springer, NY, 1995.
- [27] L. Zaiming, W. Jinbiao, MAP/G/1 G-queues with preemptive resume and multiple vacations. *Applied Mathematical Modeling*, 33(3), 1739-1748, 2009.

