

# Using Queueing Theory-M/M/1 in Hospital Tetovo

Lazim Kamberi

*Department of Mathematics Faculty of Natural Science, University of Tetovo, Tetovo*  
*E-mail :lazim.kamberi@unite.edu.mk*

**Abstract:** Modern information technologies require innovations that are based on modeling, analyzing, designing and finally implementing new systems. The whole developing process assumes a well-organized team work of experts including engineers, computer scientists, mathematicians, physicist just to mention some of them. Modern infocommunication networks are one of the most complex systems where the reliability and efficiency of the components play a very important role. The aim of the paper is to present a basic method, approaches in a Markovian level for the analysis of not too complicated systems by using M/M/1 Queueing systems. Ours experience and advice are that if it is possible solve the same problem in different ways and compare the results. In this paper the concrete problem is expected in the laboratory blood tests at the hospital in Tetovo (Republic of Macedonia). Basis for achieving results was done an analysis on basis of a questionnaire and is studying the phenomenon for three months. From the data provided is concluded that on average 120 patient counter attainment of services on that day according to the model  $M / M / 1$  represents Poisson's membership ians with parameter  $\lambda$  (arrival/time). But at the time of service pose an exponential ians  $\mu$  (patient/time) of two parameters.

## 1. Introduction. Queue and Queue Theory

Basic mathematical model of the so-called "queue" is built as follows:

- It is a range of consumers expecting for service
- On arrival, each customer must counter while virtual absence of services to release, given priority to prior arrivals
- It is assumed that in the middle of inter-arrival time each customer present independent random variables with similar distribution, while service time are also independent random variables with other common distribution

The main interest lies in construction so-called random process which responds to the number of consumers in the queue. As a process is taken to include in the same time they who where served and they who are waiting for service. In the case where the waiting-times and service-times are random variables with the same distribution, the  $(X_t)_{t>0}$  process is called Markov chain continuous in time.

The basic mathematical model for- queues runs as follows: there is a succession of customers wanting service; on arrival each customer must wait until a server is free, giving priority to earlier arrivals; it is assumed that the times between arrivals are independent random variables of the same distribution, and the times taken to serve customers are also independent random variables, of some other distribution.

- The arrival process represent random process with special distribution called inter-arrival distribution of consumers  $A(t)$ ,  $A(t) = P(\text{inter - arrival time} < t)$
- The other variable is called the service time random variable with distribution  $B(x)$ ,  $B(x) = P(\text{service time} < x)$

The structure and the discipline of service inform us for the number of consumers in the queue, the capacity of the queueing system which represent the maximal number of consumer in the queue including they who are served in the meantime.

The service discipline include rules for serving consumers, they are:

- FIFO (first in, first out) – who comes first, leave first
- LIFO (last come, first out) – who comes last, leave first
- RS (random service) – consumer are served randomly
- Priority

For simplicity consider first a single-server system Let %, called traffic intensity, be defined as

$$\ell = \frac{\text{mean service time}}{\text{mean inter-arrival time}} = \frac{\lambda}{\mu}$$

The case when  $\ell > 1$ , the system is overloaded since the requests arrive faster than as the reserved. It shows that more servers are needed.

Let  $\chi(A)$  be the characteristic function of the event  $A$ , then:

$$\chi(A) = \begin{cases} 1 & A \text{ happen} \\ 0 & A \text{ not happen} \end{cases}$$

Theorem 1: Let  $(X(t))_{t>0}$  be an ergodic Markov chain and  $A \subseteq R_{X(t)}$  is a subset of its state space. Then with probability 1

$$\lim_{T \rightarrow \infty} \frac{1}{T} \left( \int_0^T \chi(X(t) \in A) dt \right) = \sum_{i \in A} P_i = \frac{m(A)}{m(A) + m(\bar{A})}$$

where  $m(A)$  and  $m(\bar{A})$  denote the mean sojourn time of the chain in  $A$  and  $\bar{A}$  during a cycle, respectively. The ergodic (stationary, steady-state) distribution of  $X(t)$  is given with  $P_i$ .

If we consider now the customers for a tagged customer the waiting and response times are more important than the measures defined above.

Let us define by  $W_j, T_j$  the waiting, response time of the  $j$ -th customer, respectively. Clearly the waiting time is the time a customer spends in the queue waiting for service, and response time is the time a customer spends in the system, that is

$$T_j = W_j + S_j$$

where  $S_j$  denotes its service time.

Other characteristic of the system is the queue length, and the number of customers in the system. Let the random variables  $Q(t), N(t)$  denote the number of customers in the queue, in the system at time  $t$ , respectively. Clearly, in an  $m$ -server system we have

$$Q(t) = \max\{0, N(t) - m\}$$

Kendall described a queuing system as follows:

$$A/B/m/K/n/D$$

where

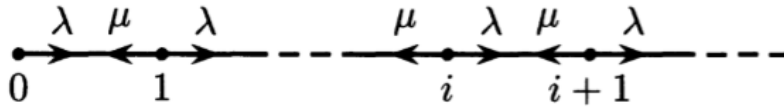
- $A$  - distribution function of the inter-arrival times,
- $B$  - distribution function of the service times,
- $m$  - number of servers,
- $K$  - capacity of the system, the maximum number of customers in the system including the one being serviced,
- $n$  - population size, number of sources of customers,
- $D$  - service discipline

## 2. The M/M/1 Queue

This is the simplest queue of all. The code means: *memory-less inter-arrival times / memory-less service times / one server*.

An M/M/1 queuing system is the simplest non-trivial queue where the requests arrive according to a Poisson process with rate  $\lambda$ , that is the inter-arrival times are independent, exponentially distributed random variables with parameter  $\mu$

Then the number of customers in the queue  $(X(t))_{t>0}$  evolves as a Markov chain with the following diagram:



Let  $N(t)$  denote the number of customers in the system at time  $t$  and we shall say that the system is at state  $k$  if  $N(t) = k$ .

In the next step we represent the transition probabilities during the time  $h$

$$P_{k,k+1} = (\lambda h + o(h))(1 - (\mu h + o(h))) + \sum_{k=2}^{\infty} (\lambda h + o(h))^k (\mu h + o(h))^{k-1} \quad k = 0, 1, 2, \dots$$

By using the independence assumption the first term is the probability that during  $h$  one customer has arrived and no service has been finished. The summation term is the probability that during  $h$  at least two customers has arrived and at the same time at least one has been serviced. It is not difficult to verify the second term is  $o(h)$  due to the property of the Poisson process. Thus

$$P_{k,k+1}(h) = \lambda h + o(h)$$

In the same way we calculate:

$$P_{k,k-1}(h) = \mu h + o(h)$$

for non-neighboring states we have:

$$P_{i,j}(h) = o(h), \quad |k - j| \geq 2$$

In summary, the introduced random process  $N(t)$  is a birth-death process with rates

$$\lambda_k = \lambda, \quad k = 0, 1, 2, \dots \text{ and } \mu_k = \mu, \quad k = 1, 2, 3, \dots$$

- Mean number of customers in the system

$$\begin{aligned} \bar{N} &= \sum_{k=0}^{\infty} k P_k = (1 - \ell) \ell \sum_{k=1}^{\infty} k \ell^{k-1} = \\ &= (1 - \ell) \ell \sum_{k=1}^{\infty} \frac{d \ell^k}{d \ell} = (1 - \ell) \ell \frac{d}{d \ell} \left( \frac{1}{1 - \ell} \right) = \frac{\ell}{1 - \ell} \end{aligned}$$

- Variance

$$\begin{aligned} \text{Var}(N) &= \sum_{k=0}^{\infty} (k - \bar{N})^2 P_k = \sum_{k=0}^{\infty} \left( k - \frac{\ell}{1 - \ell} \right)^2 P_k = \\ &= \sum_{k=0}^{\infty} k^2 P_k + \left( \frac{\ell}{1 - \ell} \right)^2 - \sum_{k=0}^{\infty} 2k \frac{\ell}{1 - \ell} P_k = \sum_{k=0}^{\infty} k(k-1) P_k + \frac{\ell^2}{(1 - \ell)^2} + \frac{\ell}{1 - \ell} - 2 \left( \frac{\ell}{1 - \ell} \right)^2 = \\ &= (1 - \ell) \ell^2 \frac{d^2}{d \ell^2} \sum_{k=0}^{\infty} \ell^k + \frac{\ell}{1 - \ell} - \left( \frac{\ell}{1 - \ell} \right)^2 = \frac{2 \ell^2}{(1 - \ell)^2} + \frac{\ell}{1 - \ell} - \left( \frac{\ell}{1 - \ell} \right)^2 = \frac{\ell}{1 - \ell} \end{aligned}$$

- Mean number of waiting customers, mean queue length

$$\bar{Q} = \sum_{k=1}^{\infty} (k - 1) P_k = \sum_{k=1}^{\infty} k P_k - \sum_{k=1}^{\infty} P_k = \bar{N} - (1 - P_0) = \bar{N} - \ell = \frac{\ell^2}{1 - \ell}$$

$$\text{Var}(Q) = \sum_{k=1}^{\infty} (k - 1)^2 P_k - \bar{Q}^2 = \frac{\ell^2 (1 + \ell - \ell^2)}{(1 - \ell)^2}$$

- Distribution of the response time of a customer

Before investigating the response we show that in any queuing system where the arrivals are Poisson distributed  $P_k(t) = \prod_k t$  where  $P_k(t)$  denotes the probability that at time  $t$  the system is a in state  $k$ , and

$\prod_k t$  denotes the probability that an arriving customer finds the system in state  $k$  at time  $t$ . Let  $A(t, t + \Delta t)$  denote the event that an arrival occurs in the interval  $(t, t + \Delta t)$ . Then:

$$\begin{aligned} \prod_k t &= \lim_{\Delta t \rightarrow 0} P(N(t) = k | A(t, t + \Delta t)) = \lim_{\Delta t \rightarrow 0} \frac{P(N(t) = k, A(t, t + \Delta t))}{P(A(t, t + \Delta t))} = \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(A(t, t + \Delta t) | N(t) = k)P(N(t) = k)}{P(A(t, t + \Delta t))} \end{aligned}$$

However, in the case of a Poisson process event  $A(t, t + \Delta t)$  does not depend on the number of customers in the system at time  $t$  and even the time  $t$  is irrelevant thus we obtain:

$$P(A(t, t + \Delta t) | N(t) = k) = P(A(t, t + \Delta t))$$

In the end we have:

$$\prod_k t = P_k = P(N(t) = k)$$

From Erlang distribution with parameters  $(n, \mu)$  we have:

$$f_n(x) = \frac{\mu(\mu x)^{n-1}}{(n-1)!}, \quad x \geq 0$$

Applying this theorem into the total probability, the density of response time of distribution is:

$$f_T(x) = \sum_{n=0}^{\infty} (1-\ell)\ell^n \frac{(\mu x)^n}{n!} \mu e^{-\mu x} = \mu(1-\ell)e^{-\mu(1-\ell)x}$$

And the distribution function

$$F_T(x) = 1 - e^{-\mu(1-\ell)x}$$

From those data mentioned above we conclude that the response time is just one continuous random variable with exponential distribution with parameter  $\mu(1-\ell)$ , then:

$$\bar{T} = \frac{1}{\mu(1-\ell)}$$

$$Var(T) = \left( \frac{1}{\mu(1-\ell)} \right)^2$$

- Distribution of the waiting time

Let  $f_W(x) = (1-\ell)\ell\mu e^{-\mu(1-\ell)x}$  denote the density function of the waiting time. According to the definitions mentioned above, is easy to calculate:

$$\bar{W} = \int_0^x x f_W(x) dx = \bar{N} \frac{1}{\mu}$$

$$Var(T) = Var(W) + \frac{1}{\mu^2} \Rightarrow Var(W) = \frac{\rho^2}{(\mu(1-\rho))^2}$$

The formula which connect those components mentioned above is  $T = W + S$ , and the formulas which represent the components of "Little Formula" or so-called "Little Theorem", they are:

$$\lambda \bar{T} = \bar{N} \quad \lambda \bar{W} = \bar{Q}$$

Let us consider a modification of an M=M=1 system in which customers are discouraged when more and more requests are present at their arrivals. Let us denote by  $b_k$  the probability that a customer joins to the system provided there are  $k$  customers in the system at the moment of his arrival. It is easy to see, that the number of customers in the system is a birth-death process with birth rates  $\lambda_k = \lambda b_k$ ,  $k = 0, 1, 2, \dots$

There are various candidates for  $b_k$  but we have to find such probabilities which result not too complicated formulas for the main performance measures:

$$b_k = \frac{1}{k+1}, \quad k = 0,1,2,\dots$$

$$\text{And: } P_k = \frac{\rho^k}{k!} P_0, \quad k = 0,1,2,\dots$$

Next we calculate the average and the variance of some fundamental characteristics modified on construction of M / M / 1 system:

- $\bar{N} = \rho, \quad \text{Var}(N) = \rho$
- $\bar{Q} = \rho + e^{-\rho} - 1, \quad \text{Var}(Q) = \rho - e^{-\rho}(e^{-\rho} + 2\rho - 1)$
- $\bar{T} = \frac{\rho}{\mu(1 - e^{-\rho})}, \quad \text{Var}(T) = \frac{\rho(2 - (\rho + 2)e^{-\rho})}{\mu^2(1 - e^{-\rho})}$
- $\bar{W} = \frac{1}{\mu} \left( \frac{\rho + e^{-\rho} - 1}{1 - e^{-\rho}} \right), \quad \text{Var}(W) = \frac{1}{1 - e^{-\rho}} (\rho + e^{-\rho} - 1)$
- $\bar{\lambda} = \mu(1 - e^{-\rho}), \quad \bar{\lambda T} = \bar{N}, \quad \bar{\lambda W} = \bar{Q}$

### 3. Some own results

Research and collection of data logging was done in the so-called "Home of Health" in Tetova, into the department of laboratory exactly in the first honor. For achieving basis results was done one basic analysis of one questionnaire and we takes a look a phenomenon studied about three months. The model that will be created does not represent other thing but one M / M / 1 queue model, because in laboratory is found only one server. From the data provided it is concluded that on average 120 patients in one day arriving in the server and according to the model M / M / 1 that represents one random variable with Poisson distribution with parameter  $\lambda$  (*arrival / hour*). Service times represent one random variable with exponential distribution with parameter  $\mu$  (*patient / hour*).

In one period of 6 hours the laboratory accepts patients, courses there arrives 100 patients for day. Those data allow us to evaluate the parameters of Poisson and Exponential distributions as:

$$\lambda = \frac{100}{6} \approx 17, \quad \mu = 20 \text{ and } \ell = \frac{17}{20} = 0,85. \quad \lambda = \frac{100}{6} \approx 17 \text{ tell us that 17 patients arrives for hour, } \mu = 20, 20 \text{ patients are served for one hour.}$$

Next Average number of customers in the system

$$\bar{N} = \frac{\ell}{1 - \ell} = \frac{0,85}{1 - 0,85} = 5,67$$

Mean number of waiting customers, mean queue length

$$\bar{Q} = \frac{\ell^2}{1 - \ell} = \bar{N} - \ell = 5,67 - 0,85 = 4,82$$

The response time of a customer (exponential distribution with parameter  $\mu(1 - \ell)$ )

$$\bar{T} = \frac{1}{\mu(1 - \ell)} = \frac{1}{20 \cdot (1 - 0,85)} = \frac{1}{3}$$

Distribution of the waiting time

$$\bar{W} = \bar{N} \frac{1}{\mu} = 5,67 \cdot \frac{1}{20} = 0,2835$$

If we want to calculate the probability that the number of patients waiting for service is bigger than 3, we have:

$$P(n > 3) = 1 - P(n \leq 3) = 1 - P_0 - P_1 - P_2 - P_3 =$$

$$= 1 - 1 + \ell - (1 - \ell)(\ell + \ell^2 + \ell^3) = \ell^4 = 0,85^4 = 0,522$$

If the arrival rate grows for 10%, then  $17 \cdot \frac{10}{100} + 17 = 19$  patients will arrive in one hour and the parameter  $\lambda = 19$ . Other calculations of the components which are mentioned above for one M/M/1 queue system are the same.

## References

- [1]. Basic Queueing Theory M/M/\* Queues Dr.Yih Huang of George Mason University.
- [2]. The Research and Application of M/M/1/N Queueing Model with Variable Input Rates, Variable Service Rates and Impatient Customers Quanru Pan World Academy of Science, Engineering and Technology International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering Vol:5, No:3, 2011
- [3]. Basic Queueing Theory Dr.JánosSztrik University of Debrecen, Faculty of Informatics