Design and Implementation of an M/M/1 Queuing Model Algorithm and its Applicability in Remote Medical Monitoring

Ifeoma Oji+ Osuagwu O.E. *

+Petroleum Training Institute, Effurum, Delta State

*Department of Information Mgt Technology, Federal University of Technology, Owerri

Abstract

Remote Medical Monitoring is a component of telemedicine capable of monitoring the vital signs of patients in a remote location and sending the results directly to a monitoring station. Vital signs are collected by sensors attached to the human body and sent automatically to the server in the hospital. This paper focuses on the design and implementation of an M/M/1 queuing model capable of queuing the readings of the vital signs in the server according to how they arrive on a First In First Out (FIFO) basis and sending them in turn to the medical personnel when the need arises. The queuing model follows a Poisson distribution with parameter (β)t and a probability function called the negative exponential distribution. The obtained output is based on a simulation using the Queuing Model Simulator (QMS), simulation software which computes the mean, variance and the total cost of running the queue.

Keywords: M/M/1 queuing model, FIFO, QMS, simulator, mean, variance, total cost

---

1.0 Introduction

With the advancement of wireless technologies, wireless sensor networks can greatly expand our ability to monitor and track conditions of patients in the healthcare area [8]. A medical monitor or physiological monitor or display, is an electronic medical device that measures a patient's vital signs and displays the data so obtained, which may or may not be transmitted on a monitoring network. Physiological data are displayed continuously on a CRT or LCD screen as data channels along the time axis. They may be accompanied by numerical readouts of computed parameters on the original data, such as maximum, minimum and average values, pulse and respiratory frequencies, and so on [5].

In critical care units of hospitals, bedside units allow continuous monitoring of a patient, with medical staff being continuously informed of the changes in the general condition of a patient [1]. Some monitors can even warn of pending fatal cardiac conditions before visible signs are noticeable to clinical staff, such as arterial fibrillation or premature ventricular contraction (PVC). Old analog patient monitors were based on oscilloscopes, and had one channel only, usually reserved for electrocardiographic monitoring (ECG). So, medical monitors tended to be highly specialized [8]. One monitor would track a patient's blood pressure, while another would measure pulse oximetry, another ECG [2]. Later analog models had a second or third channel displayed in the same screen, usually to monitor respiration movements and blood pressure. These machines were widely used and saved many lives, but they had several restrictions, including sensitivity to electrical interference, base level fluctuations, and absence of numeric readouts and alarms. In addition, although wireless monitoring telemetry was in principle possible (the technology was developed by NASA in the late 1950s for manned spaceflight. It was expensive and cumbersome.

Typically a queuing model represents the system's physical configuration by specifying the number and arrangement of the servers, which provide service to the customers, and the stochastic nature of the demands, by specifying the variability in the arrival process and in the service process [9].

To achieve remote medical monitoring, patients' data on vital signs are collected via sensors attached to the patient's body and sent automatically to the server in the hospital. It is pertinent to note that several patient's information arrive to the server and therefore it is very important to let these patient data enter a queue from where they can be sent to the different doctors assigned to do that. The queue used here is the First In First Out (FIFO) queue.

2. Components of a basic Queuing System

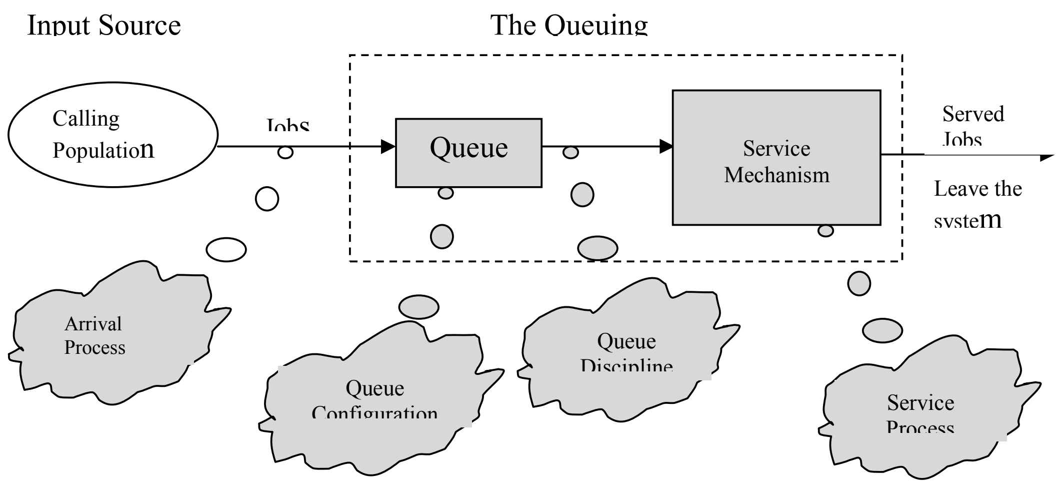


Fig. 2.1: Components of a queuing system

2.2 The Queuing Model for distribution of Patient data

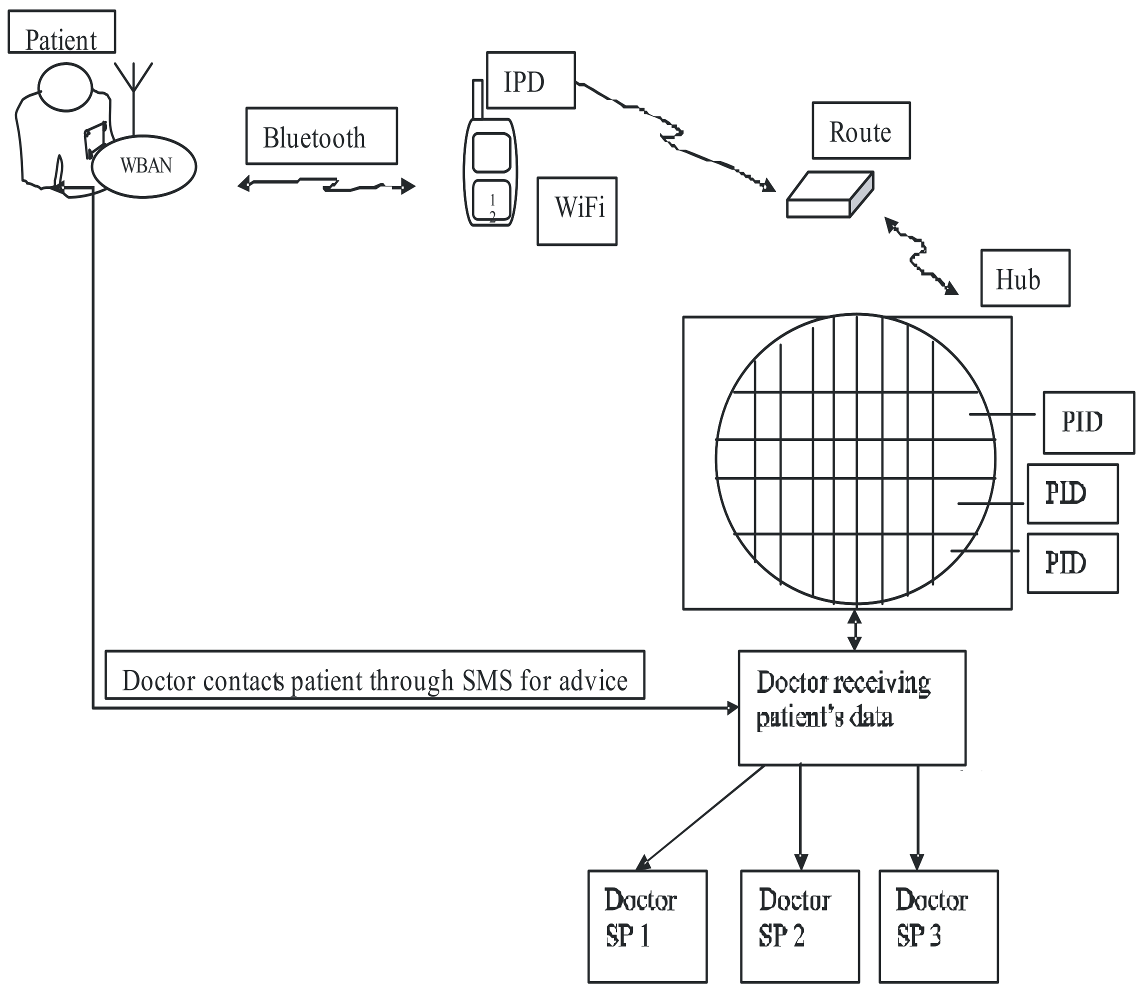


Fig. 2.2: The Queuing Model Architecture

PID 1 Patient ID 1: This can be from Pid 1 to Pid n.
SP 1 Specialist 1: This can also be form Sp 1 to Sp n.

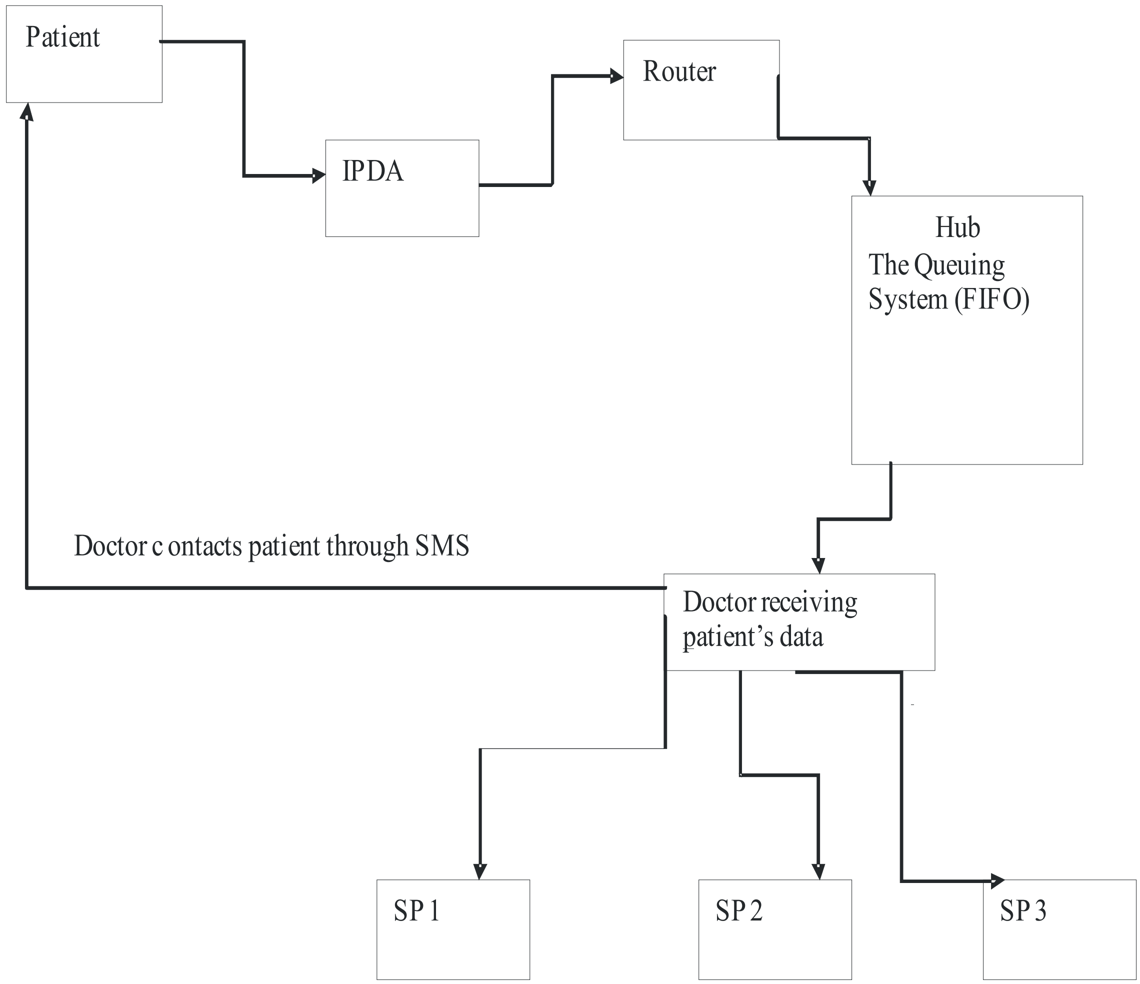


Fig. 2.3: Block diagram of the queuing model

The M/M/1 queue consists of a server which provides service for the packets of data from the patients who arrive at the system and depart. It is a single-server queuing system with exponential interarrival times, exponential service times and first-in-first-out queue discipline [4]. If a packet of data from a patient arrives when the server is busy, it joins the queue (the waiting line). There are two types of events: arrival events (A) and departure events (D). The following quantities are used in representing the model:

AT = arrival time

DT = departure time

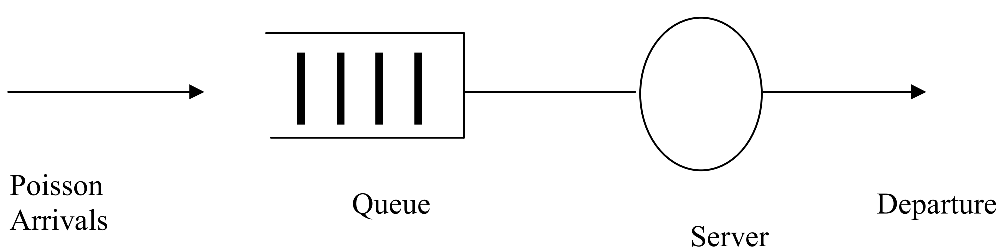


Fig 2.4: M/M/1 Queue

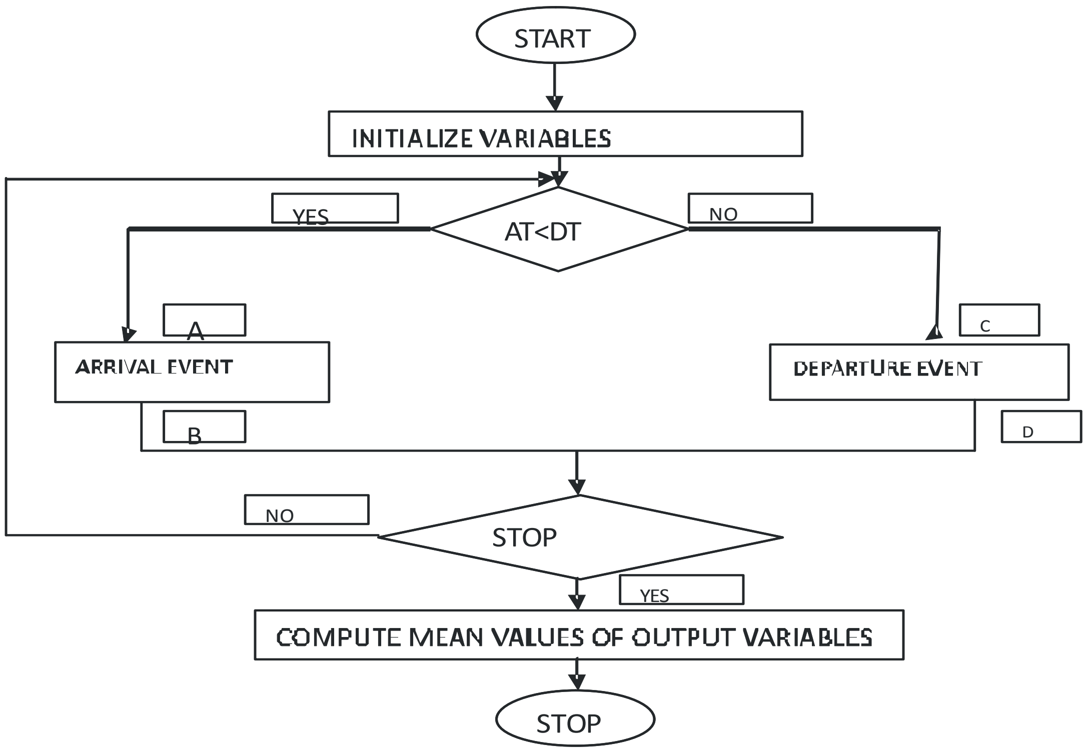


Fig 2.5: Flowchart of the queue simulation

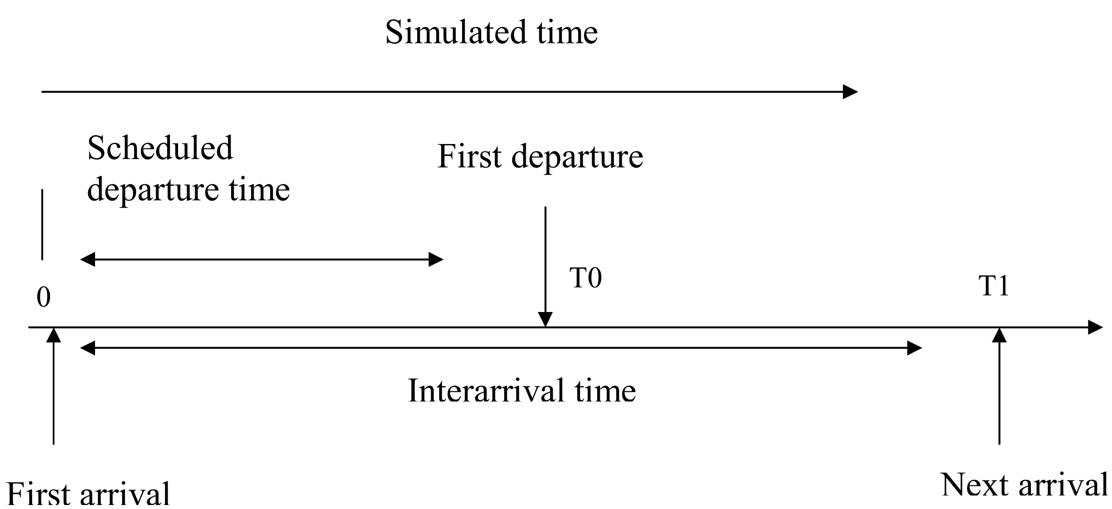


Fig. 2.6: The first few events in simulation

In the queuing model, vital signs are collected by the sensor on the patients' body, sent through the Bluetooth (this data is simulated in the IPDA) to the IPDA which transmits this data by WiFi to the router which further transmits the data wirelessly to the hub which is in the server.

The hub acts as a data repository where these data are stored and sent to the doctor when there is an abnormal situation.

2.1.2 **The Queuing Model**

Queuing models can be represented using Kendall's notation.

A/B/S/K/N/D [3].

where

A is the interarrival time distribution

B is the service time distribution

S is the number of servers

K is the system capacity

N is the calling population

D is the service discipline assumed

- **The Arrival Rate**

The data arrive as packets of data from different patients wearing the sensors into the hub.

Let C_i be the interarrival time between the arrivals of the (i – 1)th and the ith patients, the mean(or expected) inter-arrival time is denoted by E(C) and is called β; = 1/E(C) the arrival frequency.

- **Service Mechanism**

This is specified by the number of servers (denoted by S) each server having its own queue or a common queue and the probability distribution of the patient's service time [7].

Let S_i be the service time of the ith patient, the mean service time of a customer is denoted by E(S) = $\mu = \frac{1}{E(S)}$ the service rate of a server.

- **Queue Discipline**

Discipline of a queuing system means the rule that a server uses to choose the next patient from the queue (if any) when the server completes the service of the current patient [6].

The queue discipline for this system is Single Server- (FIFO) First In First Out i.e. patient's data are worked on according to when they came to the queue.

- **Measures of Performance for the**

Queuing System

Let

Di be the delay in queue of the ith patient

Wi be the waiting time in the system of the ith patient

F(t) be the number of patients in queue at time t

G(t) be the number of patients in the system at time t = F(t) + No of patients served at t.

Then the measures,

D = $\lim_{n\to\infty} \frac{\sum_{1=1}^{i=n} Di}{n}$ and

W = $\lim_{n\to\infty} \frac{\sum_{1=1}^{i=n} Wi}{n}$

are called the steady state average delay and the steady state average waiting time in the system. Also the measures,

F = $\lim_{n\to\infty} \frac{1}{T} \int_0^T F(t).dt$ and

G = $\lim_{n\to\infty} \frac{1}{T} \int_0^T G(t).dt$

are called the steady state time average number in queue and the steady state time average number in the system.

- **Single Channel Queue**

[M/M/1] : {FCFS or FIFO} Queue System

- **Arrival Time Distribution**

This model assumes that the number of arrivals occurring within a given interval of time t, follows a poisson distribution with parameter $(\beta)t$. This parameter $(\beta)t$ is the average number of arrivals in time t which is also the variance of the distribution. If n denotes the number of arrivals within a time interval t, then the probability function p(n) is given by

$$P(n) = \frac{(\beta)^n}{n!} e^{-x} \qquad n = 0,1,2\ldots. \qquad (1)$$

The arrival process is called poisson input

The probability of no(zero) arrival in the interval [0,t] is,

Pr (zero arrival in [0,t]) = $e^{-\beta t}$ = p(0)

Also

P(zero arrival in [0,t]) = P(next arrival occurs after t)

= P(time between two successive arrivals exceeds t)

Therefore the probability density function of the inter- arrival times is given by,

$e^{-\beta t}$ for t > 0

This is called the negative exponential distribution with parameter β or simply exponential distribution. The mean inter-arrival time and standard deviation of this distribution are both $1/(\beta)$ where, $(\beta)$is the arrival time.

3.0Analysis of the Queuing System

The state of the queuing system can be completely described by the number of units in the system. Thus the state of the process can assume values 0,1,2… (0 means none in the queue and the service is idle).

Let the steady state probabilities be denoted by P_n, n = 0,1,2,3,… where n refers to the number in the system. P_n is the probability that there are n units in the system. By considering a very small interval of time h, the transition diagram for this system can be seen as:

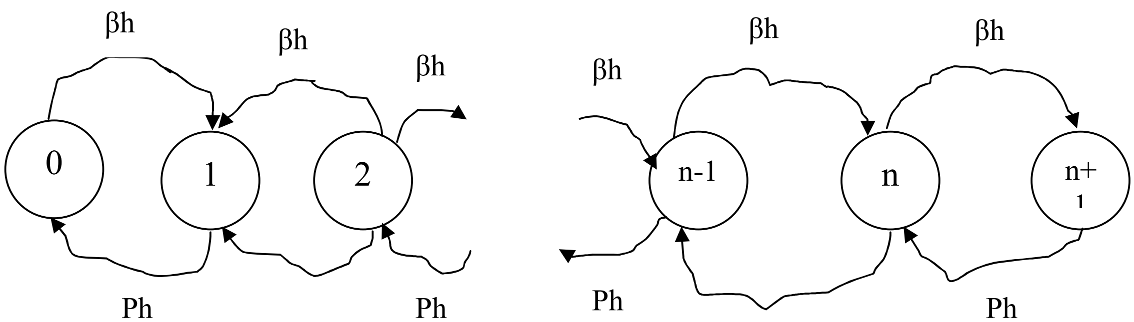


Fig. 4.18: The Transition Diagram

If h is sufficiently small, no more than one arrival can occur and no more than one service completion can occur in that time. Also the probability of observing a service completion and an arrival time in h is $\mu(\beta).h^2$ which is very small (approximately zero) and is neglected. Thus the following four events are possible:

1. There are n units and 1 arrival occurs in h
2. There are n units and 1 service is completed in h
3. There are n-1 units and 1 arrival occurs in h
4. There are n+1 units and 1 service is completed in h

For n > 1, (because of steady state and condition)

Pr (being in state n and leaving it) = Pr(being in other states and entering state n) = Pr(being in state n-1 or n+1 and entering state n).

Thus

$$P_n (\beta) * h + Pn * \mu h = Pn - 1(\beta) * h + Pn + 1 * \mu h \qquad (2)$$

This is the steady state balance equation.
For n = 0, only events 1 and 4 are possible,

$$P_o(\beta) * h = P_1 * \mu h$$

Therefore

$$P_1 = \frac{\beta}{\mu} P_0 \qquad P_n = \frac{\beta}{\mu} P_{n-1} \qquad P_n = \left(\frac{\beta}{\mu}\right)^n P_0$$

(3)

This can be determined by using the fact that the sum of the steady state probabilities must be 1. Therefore,

$$P_0+P_1+P_2+\ldots+P_n+P_{n+1}+\ldots= 1$$

$$P_0 + P_0\left[\frac{\beta}{\mu}\right] + P_0\left[\frac{\beta}{\mu}\right]^2 + \ldots + P_0\left[\frac{\beta}{\mu}\right]^n + P_0\left[\frac{\beta}{\mu}\right]^{n+1} + \ldots = 1$$

$P_0[1+P + P^2 + \ldots + P^n + P^{n+1} + \ldots] = 1 \quad P = \frac{\beta}{\mu}$

This is the sum of a geometric series. Therefore,

$$P_0\left[\frac{1-P^{n+1}}{1-P}\right] = 1 \text{ as } n \to \infty$$

Since P< 1,

$$P_0 = (1 - P) = \left[1 - \frac{\beta}{\mu}\right]$$

The term $P = \frac{\beta}{\mu}$ is equal to the probability that the service is busy, referred to as Pr (busy period).

4.0 Performance Measures

The average number of units in the system G can be found from

G = sum of [n*Pn] for n = 1 to ∞

$$G = \frac{\beta}{\mu-\beta} = \frac{P}{1-P} \text{ where } P = \frac{\beta}{\mu}$$

The average number in the queue is

$F = (G - (1 - P_0)$

Sum of [(n-1)*Pn] for n = 1 to ∞

$$F = \frac{\beta^2}{\mu(\mu-\beta)} = \frac{P^2}{(1-P)}$$

The average waiting time in the system (time in the system) can be obtained from

$$W = \frac{G}{\beta} = \frac{1}{\mu-\beta} \text{ and}$$

$$D = W - \frac{1}{\mu} = \frac{\beta}{\mu.(\mu-\beta)}.$$

The traffic intensity P (sometimes called occupancy) is defined as the average arrival rate (lambda) divided by the average service rate (mu). P is the probability that the server is busy.

$$P = \frac{\beta}{\mu}$$

The mean number of customers in the system (N) can be found using the following equation:

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

You can see from the above equation that as p approaches 1 number of customers would become very large. This can be easily justified intuitively. p will approach 1 when the average arrival rate starts approaching the average service rate. In this situation, the server would always be busy hence leading to a queue build up (large N).

Lastly we obtain the total waiting time (including the service time):

$$T = \frac{1}{\mu-\beta}$$

In a queuing system with the inter arrival time of 25 seconds and the service time of 10 seconds, the parameters are calculated thus:

Table 1

Ci	E(C) = 1/Ci β	Si	E(S) = 1/Si µ	P = β/ µ	N = P/1-P	T = 1/(µ - β)
25	0.04	10	0.1	0.4	0.666667	16.66667
50	0.02	20	0.05	0.4	0.666667	33.33333
75	0.013333	30	0.033333	0.4	0.666667	50
100	0.01	40	0.025	0.4	0.666667	66.66667
125	0.008	50	0.02	0.4	0.666667	83.33333
150	0.006667	60	0.016667	0.4	0.666667	100
175	0.005714	70	0.014286	0.4	0.666667	116.6667
200	0.005	80	0.0125	0.4	0.666667	133.3333
225	0.004444	90	0.011111	0.4	0.666667	150
250	0.004	100	0.01	0.4	0.666667	166.6667

This is illustrated in the line chart and column chart below:

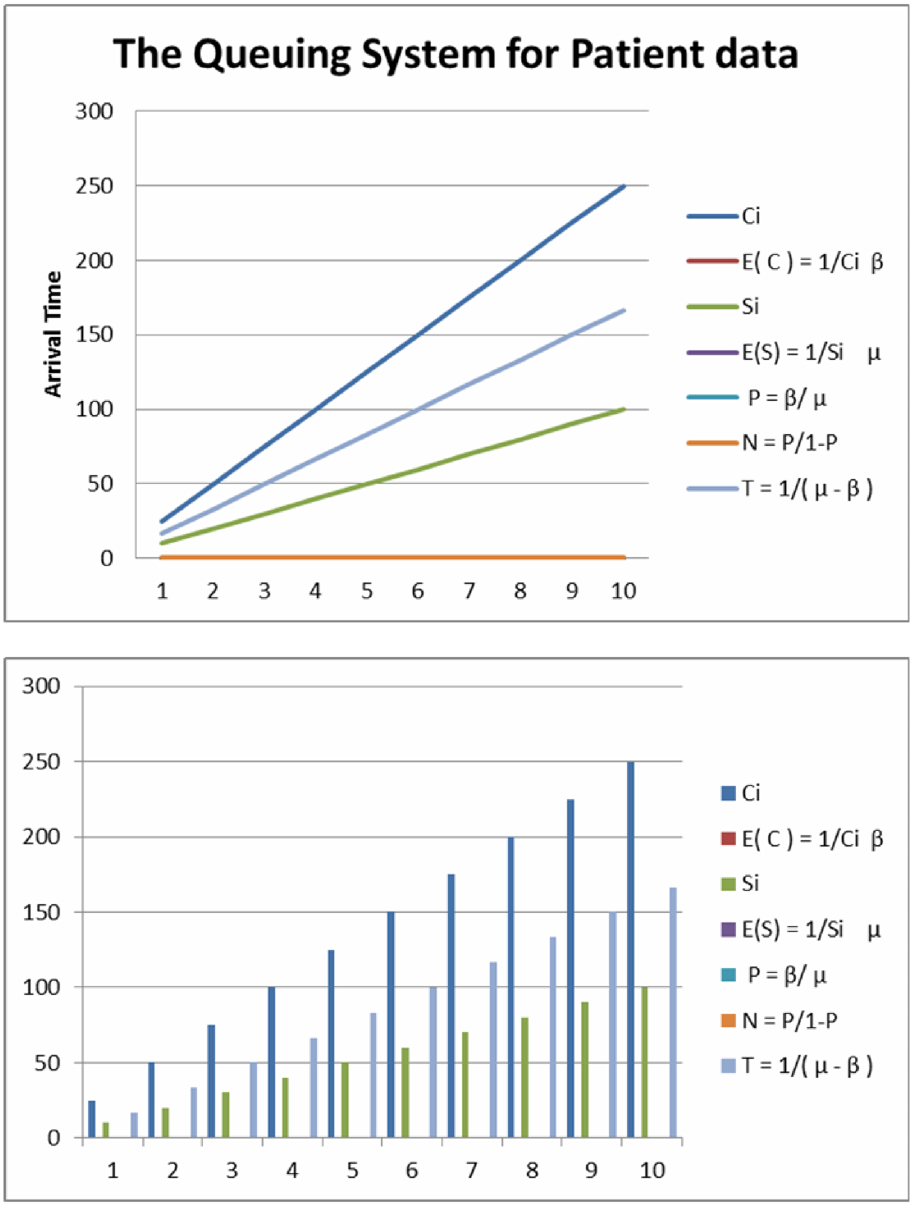
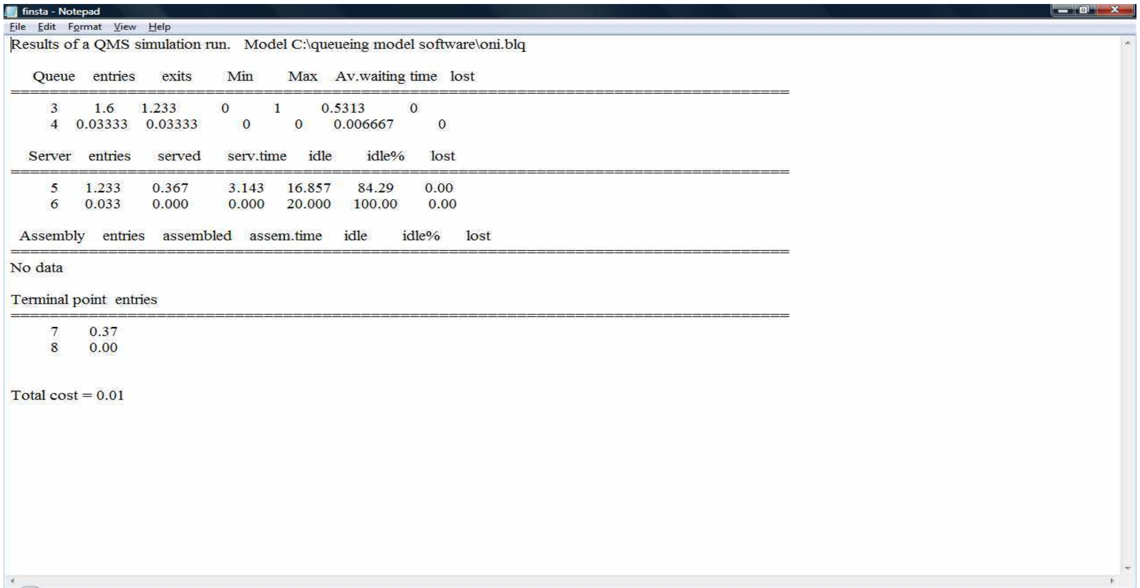


Fig. 5.0

5.0 Results

The output was obtained from the simulation done using the QMS simulator.



finsta - Notepad

File Edit Format View Help

Results of a QMS simulation run. Model C:\queueing model software\oni.blq

Queue	entries	exits	Min	Max	Av.waiting time	lost
3	1.6	1.233	0	1	0.5313	0
4	0.03333	0.03333	0	0	0.006667	0

Server	entries	served	serv.time	idle	idle%	lost
5	1.233	0.367	3.143	16.857	84.29	0.00
6	0.033	0.000	0.000	20.000	100.00	0.00

Assembly	entries	assembled	assem.time	idle	idle%	lost

No data

Terminal point	entries
7	0.37
8	0.00

Total cost = 0.01

Fig. 5.1: Results of a queuing model simulation

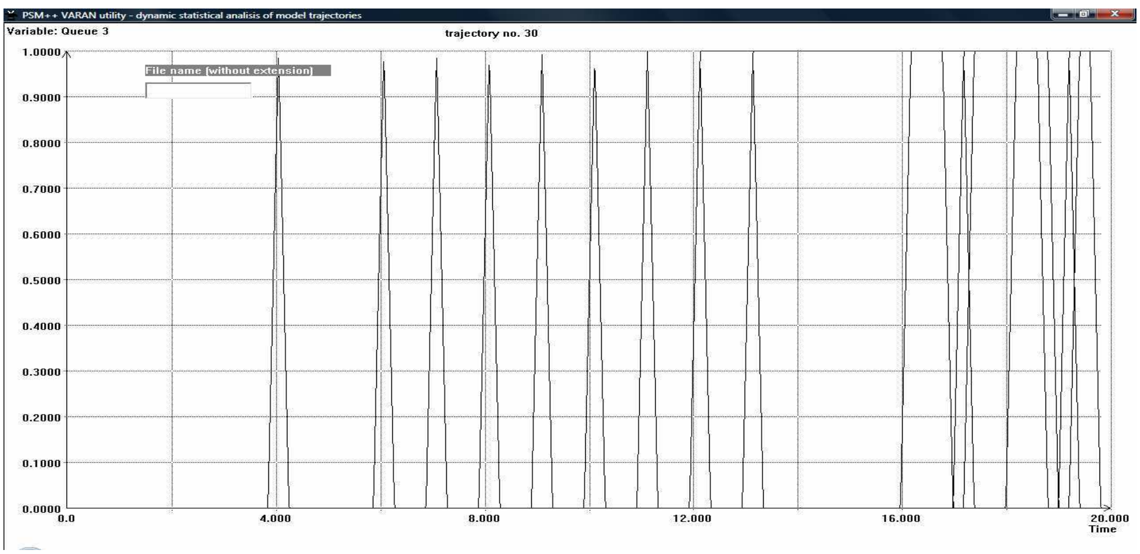


Fig. 5.2: Variance for the queue

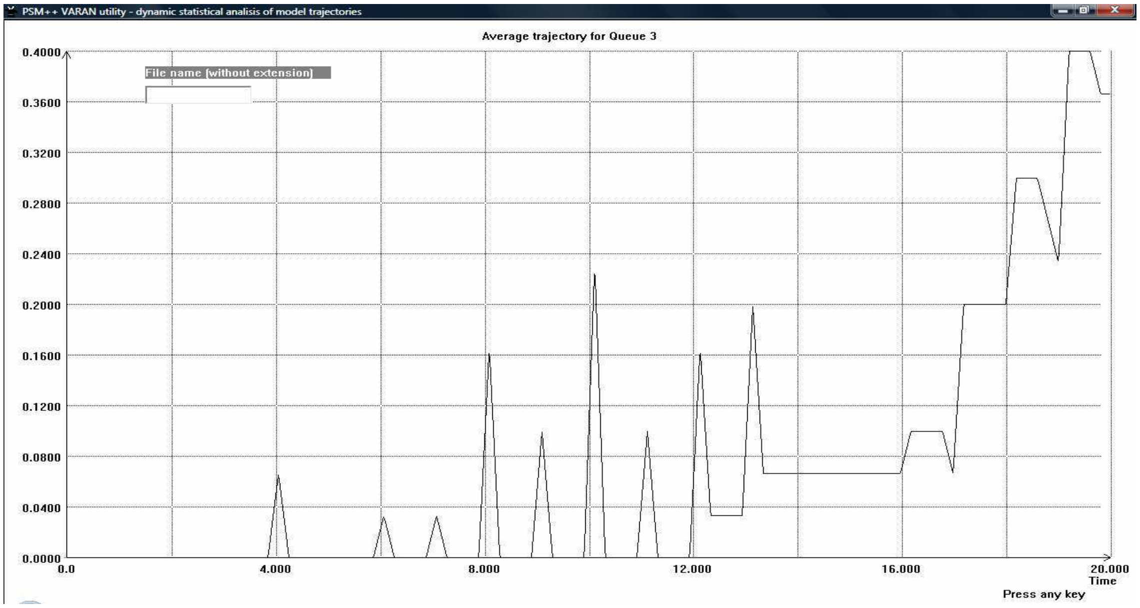


Fig. 5.3: Average for the queue

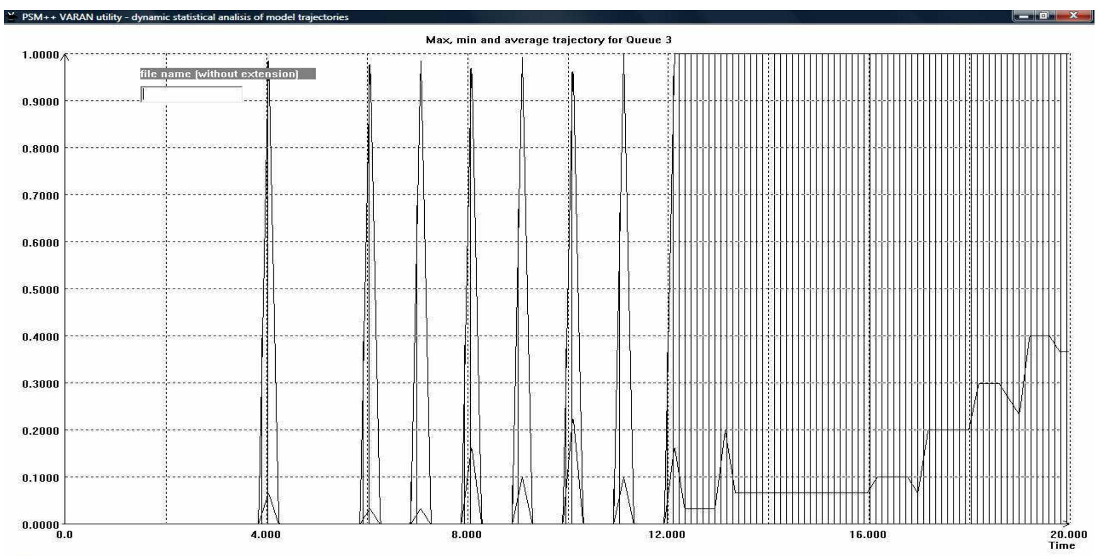


Fig. 5.4: Max, Min for the queue

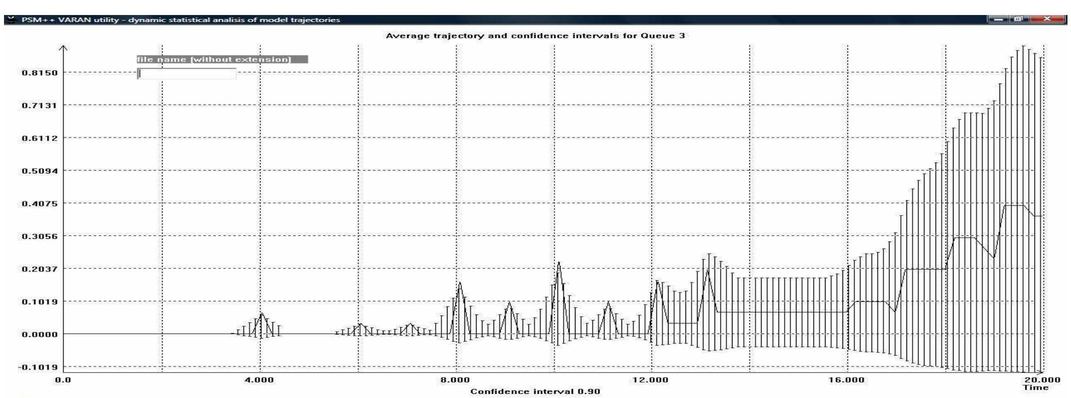


Fig. 5.5: Confidence Interval for the queue

6.0 Conclusion

A queuing model has been designed and simulated. The result of the simulation has been shown. It has been established that the packets of data from a patient's body arrives to the hospital's server and enter the queue from where they are kept according to the order in which they arrive and any data with abnormal readings are sent to the doctor for immediate medical intervention.

References

[1] Gao, T. (2005) "Vital Signs Monitoring and Patient Tracking Over a Wireless Network," IEEE- EMBS 27th Annual Int. Conference of the Eng. in Medicine and Biology.

[2] Jurik, A.D.; Weaver, A.C. (2008). Remote health care monitoring devices Computer 41(4).

[3] Kendall, D. (1953). "Stochastic Processes Occurring in the Theory of Queues and their Analysis By the Method of the Imbedded Markov Chain". Annals of Mathematical Statistics **24** (3).

[4] Lee, A., Miller D. (1966). "A Problem of Standards of Service ". Applied Queueing Theory. New York: MacMillan.

[5] Obrenovic, Z., Starcevic, D., Jovanov, E., & Radivojevic, V. (2002). An Agent Based Framework Medical Devices. Autonomous Agents & Multi-Agent Systems, Bologna, Italy. McGraw Hill Company.

[6] Sen, R.(2010). Operations Research: Algorithms and Applications. Prentice-Hall.

[7] Tijms, H.(2003), Algorithmic Analysis of Queues, A First Course in Stochastic Models, Wiley, Chichester,.

[8] Varshney U. (2008), "Improving Wireless Health Monitoring Using Incentive-Based Router Cooperation," Computer, 41(3).

[9] Zhou, Y., Gans, N.(1999). A Single-Server Queue with Markov Modulated Service Times". Financial Institutions Center, Wharton, UPenn. Retrieved from http://fic.wharton.upenn.edu/fic/papers/99/p9940.html. Retrieved 2011-01-11.