Journal of Social Sciences (COES&RJ-JSS) ISSN (E): 2305-9249 ISSN (P): 2305-9494 Publisher: Centre of Excellence for Scientific & Research Journalism, COES&RJ LLC Online Publication Date: 1st April 2015 Online Issue: Volume 4, Number 2, April 2015 http://www.centreofexcellence.net/J/JSS/JSS Mainpage.htm

APPLICATIONS OF QUEUES IN HOSPITALS IN ISTANBUL

Assist.Prof. Dr. Imran Aslan

The University of Bingöl Business Administration Department, 12000, Bingöl E-Mail: imranaslan@gmail.com, iaslan@bingol.edu.tr

ABSTRACT

Queues are one of the worst problem in hospitals in almost all World. Patients have to wait long time even to get a small operation in hospitals in Turkey. Turkey healthcare system has recently improved, but it is still not at expected level. The quantity oriented healthcare services meaning caring more patients are not good at meeting quality parameters of healthcare services. The private sector partly subsidized by governments is more profit oriented and its share has been increasing. The quality of care can be increased by low waiting and better performance of doctors. In this study, the queuing theory applications are applied two big hospitals in Istanbul during 2013-2014 years by measuring waiting time and services time and systems during different times in six months. It is found that the performance of doctors are not at expected level when total utilization is considered. Waiting long time in queues is ignored by hospitals and these long queues create high stress over patients especially at public hospitals. Patient categorization based on appointments such as scheduled patients, checking patients, urgent patients, priority patients and new patients of day, serving different patients at district time of day, combining system of all hospitals in the same city for scheduling of surgery and other operations, decreasing waiting queues and workload on doctors can help to improve the current healthcare system.

Keywords

Healthcare Management, Queuing Theory, Hospitals

Citation

Aslan, Dr. Imran, (2015); Applications of queues in hospitals in Istanbul, Journal of Social Sciences (COES&RJ-JSS), Vol. 4, No. 2, pp:770-794. http://dx.doi.org/10.17632/rv3p3bvf8b.1

© Author(s) 2015. Licensed under the Creative Commons License - Attribution 4.0

INTODUCTION

Healthcare management can be improved by knowing how the dynamics of queues are working in real life. Serving patients quickly, reliably and efficiently is the main goal of healthcare management by providing less waiting and delays with improved staff satisfaction. In hospitals, there are long queues and waiting lists in Turkey. Thus, patients have to wait for months to get even small operations due to having enough staff and inefficient usage of resources in some places like Istanbul. The purpose of this study is to define queues and models that can be used to decrease waiting an important indicator of measuring system quality and queues. The models are partly supported with real data from different hospitals and examples to improve the practical side of theory. Simulations are carried out to realize the reality of queues and get results. Tandem and priority queues are two most significant types of queues that are used in real life. To prevent congestions and to protect the rights of priority patients such as old patients over 65 ages, and severe ill patients in Turkey, priority queues can be used. There is a main system named "Hastane Randevu Merkezi-Hospital Scheduling System" for Turkish hospitals to get appointments. Patients take the queue number from online web page of Health Ministry or via phone. But, there is no priority order in the system. Thus, it is believed that priority system can be added to the system. Besides, other types of queues models are widely used in hospital services. Simple M/M/1 queue can be used to model the arrivals of patients and service of a doctor for a single policlinic. The nature of queues is stochastic and thus the probabilities of each state are calculated and compared to measure performance values.

One of the main problems in hospitals is long queues due to low capacity and high rate of patients. Queuing theory gives important performance measures of servers and queues in order to decrease waiting and queues. By measuring expected waiting times, waiting and delays can be decreased in many ways such as by adding more capacity to servers or decreasing service times to treat more patients. If the service time is decreased, the quality may be affected negatively. To providing a better service, some useless processes can be eliminated. More trained staff can be used to increase performance measures. The main indicator of performance is the number of patients treated in Turkey, but the quality dimension is totally ignored and the increasing costs on government have also increased. Controlling crowds is not easy and systems balancing the servers and patients are very helpful to prevent misuses in queues. Some patients try to go into service soon by taking the place of other patients and this is not a fair situation. The complexity of service and the capacity increase the stress rate for doctors, staff and patients. Patients have to wait in long queues sometimes with pain. Urgency cases have the first priority in queues. Later pregnant women and old people have further priority. The structure and behaviors of queues are different in each department of hospitals. Emergency room has different priority system and capacity. This department is one of the main streams to the whole hospital. Patients can be sent from here to every department and service points such as testing blood section or radiology. Patients can go one department and they can be sent to another department if it is understood that they should be checked by another branch. The number of transfers is high among departments and this can be modeled by open queuing theory.

In healthcare, M / M / s / ∞ / FCFS model can be applicable since the doctors are mainly limited and there are different patients having taken queue from internet or via telephone (preordered) in Turkey. Moreover, patients having checked before by doctor can go and visit the doctor after doing some tests such as blood tests. Then, they show these

results the doctor and they take a queue order from operator or by waiting in front doctor room. These kinds of patients are more unplanned and they cannot be well controlled by hospital. However, some estimation can be drawn to estimate their expected arrivals per unit time of day. The checking can be carried out or not according to test results. The second types of patients mainly visit the same doctor or another one in 10 days after getting the treatment. Some tests take long time and thus patients take a new number after 10 days from internet. Thus, there are two arrivals: planned ones and checked patients' arrivals. Having two arrivals at the same time creates problems for checked patients since they do not know exactly when they will be treated by doctor if the queue number is not given. Priority is decided by doctor, but it is considered that both arrivals are independent. Moreover, "no-shows" and unplanned appointments not having any appoint on that day can be added to the model. The unused capacity of policlinics due to No-Shows is filled with new patients coming on that day in Turkey.

THEORETICAL BACKGROUND

Waiting costs can be decreased by increasing number of servers(s), but the service cost increases in this way. Total cost shows a different structure as shown below. It decreases until a point and after that point, it increases again. Finding minimal total cost is the optimal solution in that model (Wang, 2009). Expected cost per unit time has two dimensions; one is the costumer time and the second is service time. Waiting has a cost and service providers have also a cost of giving service. The system success can be measured by minimizing costs of waiting and servers. Space costs, loss of business of the healthcare organization and incurred costs of health situation due to not getting treatment are other critical costs. Balking going another queue and reneging leaving queue are other consideration to be analyzed in queues for performance measurement (MDE, 2013).

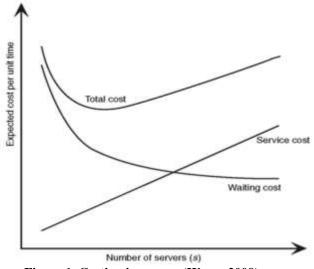


Figure 1. Costing in queues (Wang, 2009).

Healthcare systems operation such as patient scheduling, resource scheduling, queue length, Limited Queue Discipline (LQD), blocking, healthcare systems design and analysis are crucial parts of study in healthcare management. The effective usage of resources, high quality of service and less queues are the main aims that needed to evaluated for the successes of policlinics and hospitals (Lakshmi & Sivakumar, 2013).

There are four types service disciplines: FCFS (First-Come-First-Serve), LCFS (Last-Come-First-Serve), RSS (Random-Selection-for-Service) and PRI (Priority). A/S/c/m/N/SD is the Kendall's notation for classifying queues. A for the arrival process, S for service, c for the number of servers, m for the capacity, N for customer population, and SD for service discipline represent (Hillston, 2012). Arrivals occur according to random process and modeled as Poisson with λ rate:

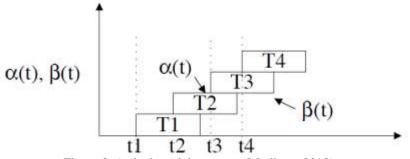


Figure 2. Arrivals and departures (Modiano, 2013)

- $\alpha(t)$ = number of arrivals by time t, $\beta(t)$ = number of departures by time t
- t_i = arrival time of ith customer, T_i = amount of time ith customer spends in the system • N(t) = number of customers in system at time t = $\alpha(t) - \beta(t)$
- N(t) = number of customers in system at time $t = \alpha(t) \mu$

$$\lambda_t = \alpha(t)/t, \ \lambda = Limit_{t \to \infty} \lambda_t = arrival rate$$

$$N = \frac{\sum_{i=1}^{\alpha(i)} T_i}{t} = (\frac{\alpha(t)}{t}) \frac{\sum_{i=1}^{\alpha(i)} T_i}{\alpha(t)} = \lambda T$$
 (Modiano, 2013)

The independent replications running of the simulation with different initial random seeds is used for systems with short transient period to optimize some system performance measures in long run. After a time period, the steady state is reached in queues and after that, performance measures of queues can be found. (Mousavi, 2011) Knowing interactions between the queues like operative and post-operative units is useful to look at behaviors of queues to prevent bed blocking and to improve patients recovery procedure. M/M/c/K parameters are analyzed to investigate interactions. Blocking probabilities are beneficial to estimate congestions and further capacity increases through simulation (Osorio & Bierlaire, 2009). No-show rates are not cancelling appointments or cancelling too late not allowing for a new appointment request and causing to more unused appointment slots. Wastage of physician time can be decreased by sending preappointment reminders, using financial penalties, and providing services like transportation vouchers and free or low-cost childcare. No-shows because of personal or work-related problems seeking treatment elsewhere instead of waiting can result in appointment backlogs in further times. The expected patient backlog or queue length, the probability of getting a same-day appointment, and a feasible range of panel sizes can help to get better utilization of physicians. An M/D/1/K Queue with State-Dependent No-Shows and patient panel size N reasonably handled by a practice forming a Poisson Process with rate λN is modelled and it is found that excessive backlogs search service elsewhere. Capacity requirements and patient panel sizes for any outpatient facility can be found in that model and they result in savings (Green& Savin, 2008). There are s busy

servers independent from each other and n costumers (can be up to infinity). Arrival comes according to Poisson distribution and service time is exponential distribution. The framework of the birth-and-death queuing model:

$$\begin{split} \lambda_n &= \lambda, \quad n = 0, 1, 2, \dots, \\ \mu_n &= n\mu, \quad n = 1, 2, \dots, s - 1, \\ &= s\mu, \quad n = s, s + 1, \dots \\ \\ A &= \begin{bmatrix} -\lambda & \lambda \\ \mu & -(\lambda + \mu) & \lambda \\ \vdots & \vdots \\ s + 1 \\ \vdots \end{bmatrix} \begin{bmatrix} -\lambda & \lambda \\ \mu & -(\lambda + \mu) & \lambda \\ \vdots & \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ s\mu & -(\lambda + s\mu) & \lambda \\ \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ s\mu & -(\lambda + s\mu) & \lambda \\ \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ \vdots \\ s\mu & -(\lambda + s\mu) & \lambda \\ s\mu & -(\lambda + s\mu$$

(Al-Nowibet, 2013)

Tandem Queues

Tandem queues named as queues in series have a wide application in production of parts in series, in registration of university students and healthcare. These tandem queues can have zero queue capacity or fine capacity. (Chandrasekar et al., 2011) Delays in queues can be decreased by re-sequencing buffer size up to a degree. The high service rate ratio in heterogeneous re-sequencing systems means more delay and fewer outputs. The balanced two servers perform better than unbalanced servers. The slow server has a better utilization than faster server due to blocking causing from that server. (Liu et al., 2013) These queues are classified as a Markovian queue (M/M/s) queues. Q₁(t) and Q₂(t) are number of customers at time t in M/M/2. {Q₁(t),Q₂(t)} with n₁, n₂ is the markow vector where n₁, n₂ = 0, 1, 2, . . . for M/M/1 queue. P_{n1,n2}= P(Q₁ = n₁,Q₂ = n₂), n₁, n₂ = 0, 1, 2, . . . , where p_i= λ/μ_i <1(Bhat, 2000) In tandem queues, the costumer can join another queue with p_{i, j} ≥ 0, ∀(i, j), or leave the system with $1-\sum_{i=1}^{N} p(i, j) \leq 1$ at node Q_i as shown below. (Alfa, 2010) Arrival theory is developed to investigate the behaviors of arrivals to the queues in finite and infinite queues for open and closed networks.(Boucherie & Dijk, 1997) The two-server queue operation is more efficient than the two single server

operations (Bhat, 2000; Al-Nowibet, 2013). In this case, unused capacity of a server can be used by other costumers of server. By using the equations above, all performance parameters of system can drawn.

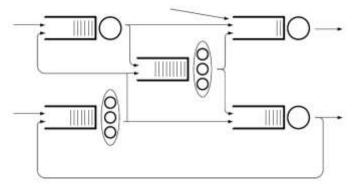


Figure 3. Example of a network of Queues. (Alfa, 2010)

For on-line ticketing, electronic banking, on-line shopping, location-based services etc., the response time is unimportant performance measure with mean and its variability. This is modeled as end-to-end response time as the sojourn time of a customer in an open queuing network. The end user sends the request to the server-side script and the server sends the request to different providers to give the requested service. (Gijsen et al., 2006) Maximum and minimum of the total virtual waiting time of a customer in open queues in M/G/1 and GI/M/1 service systems can be found. The method of strong approximations for several different queuing processes can be used to find maximum and minimum parameters in queues (Minkevi£ius, 2009). Blocking and starvations are two main considerations in tandem queues. If there is no costumer in the queue at the time service finished, then the server with zero queues will be idle. If the queue that the costumer will go is full, then the server will wait for opening the place. (Kim & Ayhan, 2010) In the blocking case, patients may not get service and leave the system if there is no place in beds. In this case, patients may search for another hospital or less costly services. Unnecessary stays at hospital can be decreased to open place for new comers. This kind of system can be named as tandem two-station system with no buffer. (Koizumi, 2002) With no buffers, flexible servers can be analyzed to maximize the total outputs and policies can be developed based on different rates of servers (heterogeneous mean service times). Holding cost, switching costs and loss of costumers are also the main centre of observation in tandem queues. Policy iteration algorithms are developed to maximize throughputs by applying different policies (Yarmand & Down, 2012). Open models have wide applications in healthcare management. There are many clinics in each government hospitals. After a patient is checked, he or she can be sent to another clinic or labs. There are 36 clinics in sample hospital- IEAH. Patients are sent to blood lab, urine laboratory or radiology lab mainly for tests. They can be sent also other special labs to get information about health situation. They can do some tests on the same day and others can be done in further dates. Open networks are beneficial to see the movements in the hospital to make better plans and decisions. If it is known that the blood centre is the most visited lab then, some new servers can be employed to decrease the queues. They may leave the system after checked if there is no need to get tests or may go another clinic to get a different kind of observation.

Bulk arrivals ($M^{[X]}/M/S/K/FCFS$) are expected to be in emergency department, lab and other unscheduled cases. In bulk arrivals, the queues' size can increase dramatically and in this rush times, the extra servers can be needed to decrease the queues. $M/M^{H}/1$ and $M^{H}/M/1$ batch queuing systems probabilities are found by a Lattice Path Combinatorics approach. Batch sizes with different sizes are also found. (Krinik & Mohanty, 2010) Multiple Poisson bulk arrival queuing system with several classes of customer enters the system independently of each other. Lambda is demand of j items or costumers in that system where $N_i(t)$, $t\geq 0$. (Armero & Conesa, 2004)

A doctor can be interrupted by a phone call or emergency situation during service, which is an outage. This causes to decreasing the performance measures and dissatisfaction in many ways. It is expected that a doctor checks a patient in a suitable time without interruptions. However, the scarcity of doctors in reality prevents that; there can be many outages in hospitals. In the sample hospital in internal medicine department, it was observed that during checking, the doctor was called and he talked about another case with another person. I have waited for him to finish his call around 2-3 minutes at HH hospital. It was noticed that he checked my data again in his computer and I had to talk little more to repeat again some things. Moreover, this caused dissatisfaction on me. This case is introduced in that section based on previous observations. Another case was met while taking the heart film. Nurses went to toilet for 7 minutes. Patients had to wait until she came.

Outages may have significant effect on over loaded hospitals. They are mainly ignored in models. However, even a doctor may not have high outages, the total of all doctor may be significant on total service times and waiting. Reducing the effect of outages can decrease flow times by taking some administration outsourcings in hospitals. Overtime can be used to decrease peaks in hospitals due to outages. Variation in outages can be decreased by managers and found that it is beneficial to increase efficiency in hospitals. (Creemers& Lambrecht, 2007) Outages on patient flow cause to large problems such as congestion and instability in healthcare. Unplanned absences of medical staff and interruptions are major problems in healthcare due to high workloads on doctors. Vacation models are used in queuing theory when servers are unavailable. Non-preemptive outages occur not during job but at the beginning of service. (Hopp&Spearman, 2000). Vacations are categorized as customer that remains in service, customer that leaves behind a non-empty queue, customer that leaves behind an empty queue and no customer. Based on these types, a model is developed in an infinite capacity buffer with geometrically distributed inter-arrival times. Effective service time of a customer and the partial probability generating functions are defined in that model. There are three operation modes: continue after interruption mode, repeat after interruption mode and repeat after interruption with re-sampling mode. (Fiems & Bruneel, 2013)

MMAP[K]/PH[K]/1: A multi-type queue with adaptive arrivals and general impatience is analyzed by Benny Van Houdt(2012). K different types costumers with Markovian inter-arrival times are served by one person. Different possibilities with customer impatience form and workloads are considered in this type of queue. (Houdt, 2012) This model is not analyzed in that study. However, it can be used in further studies.

CASE STUDIES

Case studies are prepared by analyzing sample hospitals' systems(HH and IEARH). These hospitals are two biggest hospitals in Istanbul. These hospitals systems and queues were investigated around 1.5 years. Time for queues were taken at different time periods. The system in hospitals was new and they had some problems such as not reaching main system all time. In these cases, patients had to wait few hours for system returning. These kinds of extreme situations are not included in this study.

Case Study 1:

The Ear, Nose and Throat Clinic (KBB) of IEARH hospital has 5 doctors and they work the from 8:00 AM until 17:00 PM. There is one hour break (12:00-13:00) and they accept patients as planned ones (ordered before: pre-ordered) and unplanned patients (take order in the day for rechecking). 29 patients are pre-ordered for each clinic and 10 patients are cared per hour. The total arrival rate for each clinic is expected to 8 patients per hour. Even, there is a restriction on the arrival of pre-ordered patients; the unplanned patients can be served without any restriction during the day. Thus, the patient's arrival is considered as infinite. Each server works in the same manner and they are considered as identical. Pre-ordered patients can take their queues' number by internet as shown below from Health Minister Web Page developed to prevent long queues. Before that system, queues are given early and patients had to be at hospital early in the morning to take a queue number without any exact time. This scheduling program has brought great benefits to patients. The main problem in this system is unplanned patients. The average service rate is determined from patients types. It is expected that unplanned patients take more time of a doctor since test results are controlled and then they are prescript. One of the problems is that the pre-ordered patients can be served at the time given by the system. Pre-ordered patients have to wait until the end if they could not catch their queue on the time doctor announced. It was investigated that that situation makes great stress over both patients and doctors. Patients want to get service as soon as possible while doctors are under pressure of supplying service on time. Unexpected patients having checked and new ones come to ask questions and this takes also the time of doctors. Urgent patients come and enter the room of doctor without a queue number.

STANBUL Region Name Hospital Name monori S Hustaliklari (Dahiliyo)		08:00	100-00	CONTRACTOR	ACCRETES 1		-		
			Street and						10
			THE OWNER WATER	and a constraint	Particular and a second				1.1
- Hestaldheet (Dabilises)		09.00		1000100	10-001	10,210	100000000	COLUMN 1	5
f tomania (continue) of	100	10:00	STATISTICS.	percent and					-
farkounaz. farkounaz.		11:00	-	(CENTER)	ACCOUNTS.	acciercia	(Internet)	100000	Ķ
and a second	atevu Ara	15.00	anteita	- BETTECON	15:20	18:00	16:40	15:50	1
	and the second se				attents of	1.11	-		
	(and a		ann an ta	_	al forme of	ai la	-		ļ
06.06.2013				(Dabilityer)					
04.06.20				Hastahk	are N	Name of			
				Hastakk					
(0.06.20) #			le Hantabou	iam i					
					The second se				
	Kan Kan Contraction 100. Attraction 400.000.000 Contraction 000.000.000 Contraction 000.000.000 Contraction 000.000.000 Contraction 000.000.000 Contraction 000.000.000 Contraction 000.000.000	Ranutexu Ara 10.06.2013 Attract (Lastandar) 06.06.2013 06.06.2013 06.06.2013 06.06.2013	Instructures Bestificity Annual 15:00 Occurrence 10:00:2013 10:00 Arterna Corrence 10:00 Arterna Corrence Arterna Oc. 00:0013 Corrence Oc. 00:0013	Bankleve Ara 16:00 16:00 Contraction 10:00:2013 10:00 10:00 Contraction Contraction 10:00 10:00 Contraction Contraction Contraction 10:00 10:00 Contraction Contraction Contraction Contraction Contraction Contraction Contraction Cont	Carbonics Reinflow 16:00 10:00 10:00 Carbonics Carbonics 10:00 10:00 10:00 Carbonics Carbonics Carbonics 10:00 10:00 Carbonics Carbonics Carbonics 10:00 10:00 10:00 Carbonics Carbonics Carbonics Name 06:00:00:00 10:00 10:00 10:00 10:00 10:00 10:00 10:00 10:00 10:00 10:00 10:00 10:00 10:00	Arkennez Bankleve An	Tarketment T	Tarketmer Tarketmer	Tarketmer Tarketmer

Figure 4. Hospital Appointment System

Based on queuing theory, performance measures are calculated by using Excel Macros. Model 4- M/M/s Queue: Multiple servers, Infinite population, Poisson arrival, FCFS, Exponential service time, unlimited waiting room. The parameters of systems are shown in table below for IEARH hospital.

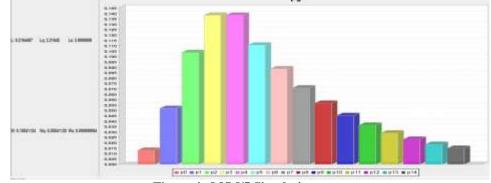
Unit of time	Hour
Arrival rate (λ)(5*8=40)	40 customers per hour
Service rate (mean)	10 customers per hour
Number of identical servers (s)	5 servers
Results	
Mean time between arrivals	0.025 hour
Mean time per service	0.1 hour
Traffic intensity	0.8
Performance measures	
Average utilization rate of server	80.0%
Average number of customers waiting in line (L _q)	2.216 customers
Average number of customers in system (L)	6.216 customers
Average time waiting in line (W _q)	0.055 hour
Average time in system (W)	0.155 hour
Probability of no customers in system (P ₀)	0.012
Probability that all servers are busy	55.4%
Probability that at least one server is idle	44.6%
Distribution customers in system	
n (customers)	P(n in system)
1	0.051948
Distribution of time in queue	
t (time in queue)	P(wait > t)
0.33333	0.019767

Table 1. Model 4- M/M/s Queue

 P_0 = 0.01 is the probability of being empty in the system. With 55.4% probability, all servers are busy and 44.6% is the probability that at least one server is idle. Average utilization rate of server is 80.0%, which is well but can be improved. Waiting probability (P(wait > t)) of 0.33 hour is 0.019767. 3.32 minutes are the average time of waiting in queue. 9.32 minutes are the average waiting time in system. These measures can be generalized to the whole day and the utilization can be improved. The existing system does not work well due to unplanned patients as mentioned above. Some patients urgent and unplanned want to enter the queue earlier and this creates problem to catch the schedule done by system. Overloading the doctors creates stress and not treating patients well. A fixed time can be given these unplanned patients to prevent extra congestions such

as from 9:00-10:00 and 14:00-15:00. Moreover, some doctors can care urgent and unplanned patients to allow the schedule work on time. Patients are also busy people and they mainly work. Too much waiting and not catching the queue is the loss of time and money. One patient had to travel 1-2 hours to come the hospital in Istanbul with her car and she missed her schedule and she had to wait until all planned patients cared.

The same model can be simulated by java based simulatron. The number of patients in queue in time can be seen and the probabilities of each sate can be determined. Based on probabilities and waiting times, different scenarios can be run by changing service time, servers or arrivals. Arrivals are random and cannot be changed by management.P3 and P4 are the highest probabilities as shown in figure below. Ls with 3.9 and Ws with 0.09 are shown in that simulation. P(n in system) decreases as n increases.





To improve the system above, some scenarios are carried out. The service time is increased to 25 patients per hour in the second model, which means more patients can be cured over time. Thus, the crucial queues parameters decrease with high P(n in system) order probabilities of patients(P₀-P₄). In this model, the quality dimension of service should be considered. In third model, mean and servers are not changed but, the λ is 30 in that time. That means less patients will arrive to the system and the queues will be decreased in that model as well. There is a beneficial improvement in performance parameters of queue. If we want to see the how the queue size can be increased, the λ should be increased. In that model, the λ is 75 and servers parameters are the same. However, it is decided to increase service time from 10 to 16 in that model as expected in real life. Lq is 12.7246 and the p(n) are high for n>4.

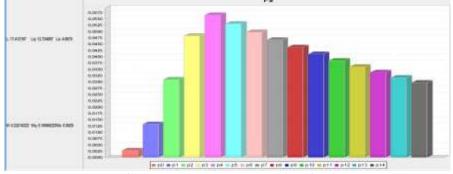


Figure 5. Increased queues with high λ

Case Study 2: M/M/1 Queue

This queue type is one of the most used one in healthcare to simulate the queues for a single doctor or staff. There are many independent single desk operators and doctors in hospitals.

Arrivals - Poisson process with rate λ

Service times - exponentially distributed with rate μ

Table	2.	M /	M/1	Queue
-------	----	------------	-----	-------

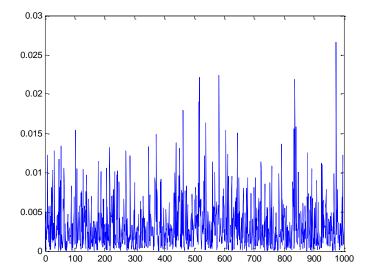
_		
P	arameters	
	Units of time	Day
		Day
	Arrival rate (λ)	250 customers
		200
	Service rate (µ)	300 customers
	Results	
r		
	Mean time between arrivals	0.004 day
		0.00222.1
	Mean time per service	0.00333 day
	Traffic intensity	0.83333
P	erformance Measures	
	Utilization rate of server	83.3%
	Average number of customers waiting in line (Lq)	4.166 customers
	Average number of customers in system (L)	5 customers
	Average time waiting in line (Wq)	0.01666 Day
	Average time in system (W)	0.02 Day
	Probability of no customers in system (P0)	16.7%
	Distribution of number of customers in system	
	n (customers)	P(n in system)

20	0.004347
Distribution of time in queue	
t (time in queue)	P(wait > t)
0.2	0.000038

With 16.7%, the system is empty. The average number of costumers is 5 that means there should be 5 or more chairs for patients to sit in that situation. More than 4 patients wait in the queue to take the service. Decreasing these both parameters is beneficial for hospital management to have fewer patients in the hospital and smooth the movements of patients in aisles. Especially, it was seen that many old patients had to stand up in HH and IEAH hospitals due to not having enough chairs. 83.3% of server utilization is high and some more patients can be treated by that server. Server utilization main performance parameter can be increased in that model by increasing λ . Decreasing λ results in less waiting and patients in the system.

The same model is simulated by Matlab by simulating 1000 patients to analyze arrivals and service times. The simulation results are shown below.

Average service time = 0.0032; Average wait time = 0.0157; Average total time = 0.0189





The service times of 1000 patients are shown above. If more patients are treated per period time, there will be fewer queues. There are some peak times of wait as shown below. During these peak times, some extra servers may be needed.

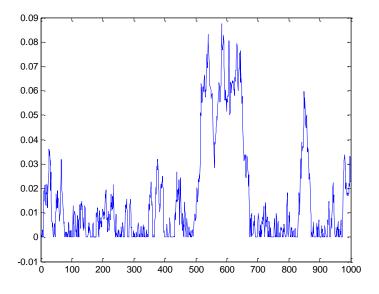
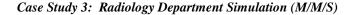


Figure 7. Waiting Time



Data were collected from Radiology department for one day. It was found that patients wait 67.3 minutes in queue. However, patients do not stay in the queue and they go to other tests or doctors in the hospital during this time. Thus, this waiting time does not reflect real waiting in the queue at radiology. The service time is 2.25 minutes. This time reflects the real service time of films with 1.53 standard deviation. Hands films have the highest service time and later, legs take the most time of service. Lungs take the least time of service at the department as shown below.

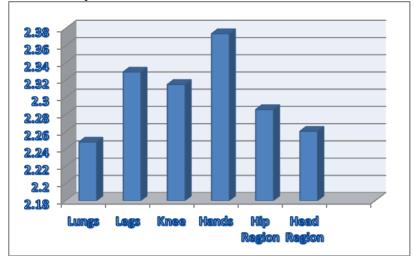


Figure 8. Service Time(In Min.)

M/M/1 Simulation at Policlinic of Radiology Room

There is one radiology machine and 9 hours simulation carried out for a day. The parameters taken from daily collected data are shown below and discrete simulation is used to simulate that case. The arrivals change during the day. There are more arrivals early in the morning and afternoon while there are few arrivals in some hours. Especially, it is noticed that there are few patients after 15:00 PM. More patients are sent during policlinic time by doctors.

Lamda:0.5 per minute(0.5/60 per sec),Mu=0.4 per minute(0.4/60 per sec), Simulating 9 hours(9*60*60 sec).

There are 278 arrivals with none drops and 99.3% server utilization. Average 39.280 patients wait in the queue. 4781.442 seconds are spent in the queue. With K= 30, average 21.521 patients wait in the system with 42 drops from the queue. The server utilization is 99.4%.3102.059 seconds are spent in queue. With s=2, K=50, the servers are parallel and have different queues in that model. The first server has 54.3% utilization and 0.402 patients in the queue while server 2 has 58.5% utilization and 0.625 patients in queue. 117.812 seconds are spent the first queue and 139.681 seconds are spent in the second queue.

Case Study 4: M/G/1 Queue

Costumer arrives in a poisson process with parameter λ and there is a server. Service times (S_n) are identically independently distributed. Q(t) is the number of costumers in the system at time t. X_n is the number of costumers arriving during S_n. Finding performance measures are showed here in details. M/G/1 queuing system with multiclass customer arrivals and fixed feedback is carried out. The means queue size of each class and sojourn time of a customer in each class are determined in this work. Costumers take more than one time services in this model. After a customer gets a service k^{th} (k < k_i), he joins the tail of queue to take another service. A costumer leaves the system after k_ith service. The newly arriving customer or the feedback customer is not considered differently, they are served according to FCFS policy. A function is developed for service times of costumers (Qi-zhi, 2008). The model with Single server, Infinite Population, Poisson Arrival, FCFS, General Service Time, and Unlimited Queue Length is defined per minute. The probability that there is no costumer is 16% and utilization rate is 84 % as other model. The only change is the service time stochastic with 0.3 minute standard deviation. This model can be more preferable for healthcare management. Average number of patients waiting in line is less than other two models. However, other findings are close to each other.

The arrivals and service time are exponentially distributed and it is decided to simulate the queue with 10.000 patients. Inter-arrival times mean is 250 and Service times mean is 300 with 1 server. 95% Confidence Interval for parameters of queue is expected to be found. The Excel Macro Solver is used to simulate that case as shown below. Negative probabilities are not possible.

Table 3. Confidence Interval for estimates for M/G/1 queue

	Parameters	95% Confidence Interval	
	Estimate	Lower Bound	Upper Bound
L =	3.0668096	0.393016752	5.740602448

$L_q =$	2.33705432	0.010860142	4.663248503
W =	876.467818	579.3500514	1173.585585
$W_q =$	667.910034	362.4525117	973.3675559
$\mathbf{P}_0 =$	0.27024472	-0.085354565	0.62584401
$P_1 =$	0.23186301	0.062385397	0.401340622
$P_2 =$	0.0685182	0.0003184	0.136718009
$P_3 =$	0.02986302	0.000138772	0.059587264
$P_4 =$	0.06862514	0.000318897	0.136931382
$P_{5} =$	0.03573259	0.000166047	0.071299124
$P_{6} =$	0.06725904	0.000312548	0.134205524
$P_7 =$	0.09001687	0.000418303	0.17961543

M/C/1 Queue has constant service time and arrivals are passion distribution. This case may not be seen in healthcare commonly. There is not any section to assume that the service time is constant but some machines at labs. Constant service time is mainly possible with machines of blood testing. But, it is possible to compare the performance measures with other models., Constant service time decreases the variability in the model. The results are close to other models with small differences.

Case Study 4: M/M/S/K(S)/FCFS- M/M/1/K(S) Queues

Servers are randomly selected and there is a limited space for patients. The main property of this model is finite queue capacity. When a patient cannot find a place in the queue, he has to leave the system and this is a loss of a patient in the system. The model M/M/s/K with multiple servers, infinite population, poisson arrival, FCFS, exponential service time, limited buffer is solved below. Buffer size is equal to number of servers plus waiting room size. The model of the hospital is solved with M/M/s/K formulas. Queue is closed with K-system size different from other models. Some patients will be lost and lambda effective will be a parameter of performance to see how well the hospital works.

Parameters				
Unit of time	Day			
Arrival rate (lambda)	250			
Service rate (mu)	60			
Number of identical servers (s)	5			
Buffer (waiting room) size	150			
Results				
Direct outputs from inputs				
Mean time between arrivals	0.004			
Mean time per service	0.016666			
Traffic intensity	0.833333			
Performance Measures				

Average utilization rate of servers	83.3%
Average number of customers waiting in line (L _q)	3.1007
Average number of customers in system (L)	7.2674
Average time waiting in line (W _q)	0.0124
Average time in system (W)	0.0291
Probability of no customers in system (P ₀)	0.0099
Probability of rejecting a customer (balking rate)	0.00%
Effective arrival rate	250

As noticed above, when the buffer size is high, the rejecting or balking rate is low. Effective entering rate is 250 and average waiting is low. As the buffer size decreases, the balking rate increases whereas effective arrival rate decreases. M/M/s/s is a loss model where the system size is finite. M/M/5/5 results show that rejecting rate increases, but the waiting decreases. Each model defined with M/M/s/K queues has some advantages over each other and the model can be determined based on objectives of hospital. Waiting times can be increased with increased servers while the costs of treatments will increase. There is no best model and each model serves to some purposes. There are some other models with different distributions. However, models presented above are the most widely ones used to optimize queues.

Case Study 5: M/M/1- M/M/1(c)/k Tandem Queues

Patients come to the hospital with λ to a special department. First, they go to registration section to get barcodes (5-6 barcodes are given each time) at HH and IEAH hospitals that will be used during the treatments to attach documents. Later, they go to doctors' room or tests sections to get treatment. The patients have to wait in the first queue and after treatment, they have to wait at second queue. This case can be presented as M/M/1/k with limited k people in the queue. Patients will get service μ_i rate. There are one server and one doctor at each station and there can be many identical stations at the same hospital. This situation can be modeled in three ways: In the first case, service station is full, patients are not blocked but have limited queue capacity. In the last case, each station can have infinite capacity and all patients can enter the queue. Expected waiting times, λ_e , λ_b , expected number of people and percentage of working can be found in these models.

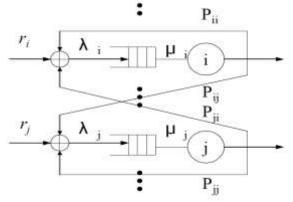


Figure 9. Open Network example(Yan& Malathi , 2004)

M/M/1/k Queuing Model with blocking

Blocking occur due to limited capacity of serial queues in the network. Blockingbefore-service, blocking-after-service, and repetitive-service with random or fixed destination are types of blocking in the literature. Bottlenecks in the system can cause more queues. GI/G/m queues are used to model arrival and service rates with their variability. Based on approximations, the expected waiting times and losses are calculated for the model. The throughput in steady State is driven from model developed. Different scenarios of service time, servers and arrivals are simulated to measure performances and find best model. (Shi, 1995) Extending k queues of M/M/1 system with μ_i ; $\rho_i = \lambda/\mu_i < 1$, i = 1, 2, ..., k. A simple case of 1 arrival per time (5 minutes) and 2 patients can be given barcodes by the counter and the doctor can cure 3 patients per time. This case is modeled by finite queue, K=2 and the queue will be blocked if there are already two patients in the queue. Moreover, this model can be used in some daily operations

 $\rho_i \!=\! \lambda\!/\mu_i \text{ ; } \lambda\!\!=\!\!1,\, \mu_1\!\!=\!\!2,\, \mu_2\!\!=\!\!3;\, \rho_1 \!=\!\!0.5;\, \rho_2\!\!=\!\!1/3$

M/M/1/2:

 $\lambda = 1, \mu_1 = 2$

In this model, M/M/1/2 is considered as independent queue and the probabilities are found. There are 3 states(0,1 and 2-number of patients in queue).

 $p(0)=(1-0.5)/(1-(0.5)^{3})=(1/2)/(7/8)=4/7$ p(1)=(4/7)*(1/2)=2/7

 $p(2)=(1/2)^{2}*(4/7)=1/7$

The doctor can give service in μ_2 meaning longer than first station. Probabilities, through-put rate are determined down.

$$\begin{split} \lambda = 1, \ \mu_2 = 3 \\ p(0) = (1 - 1/3)/(1 - (1/3)^3) = (2/3)/(26/27) = 9/13 \\ p(1) = 9/13 * 1/3 = 3/13 \\ p(2) = 1/13 \end{split}$$

Case Study 6: Simulation of Queues in Series

The simulator developed by SIM.JS – Discrete Event Simulation in JavaScript is used to simulate networks queues in series over long period time such as 4,318,500 arrivals per time. The model below shows the movements of patients in series. This model is not so common in healthcare but to get health reports in Turkey. However, there are some cases that patients have to see doctors in series in case of general checking.

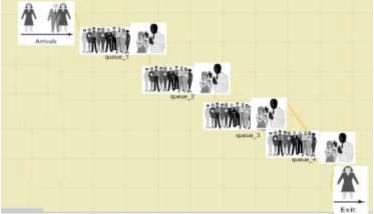


Figure 10. Queues in series

The values are approximated for example hospital of IEAH. The importance of bottlenecks can be seen more clearly in this model. If there is a slow doctor, the whole doctors performance decreases due to not having patients. There can be some large queues in front of bottlenecks.

Arrivals 300 per unit time (day)

Mu1= 400 per unit time;Mu2=350 patients per unit time;Mu3:350 patients per unit time; Mu4 400 patients per unit time and infinite queue size for each queue in the system Results:

Stay duration mean is 0.060 with 0.032 standard deviation. There are 4, 318,475 departures from system. None has dropped the queues and they get services at the end. First and last servers have less utilization while other servers have more server utilization due to averages per unit time. Time spent in system and length of queue increase for second and third queues. Simulating with M/M/1/50, there are not big differences in values however there are 271 patients drops at queue 2 and 231 patients drops at queues 3. Dropping the queues happen due to the limited capacity of queues (50 patients).In the first case, the queues are infinite. Decreased capacity of queues causes not entering the system. To decrease the queues, the number of identical servers can be increased.

In this open queue, there are two doctors from different branches and some patients visit each doctor. 50% of patients from doctor 1 needs to see the second doctor and the rest of patients leave the system. After seeing the second doctor, 30% of needs to see the first doctor again to show results. Two patients are served by each doctor per unit time and there are 2 arrivals per unit time.

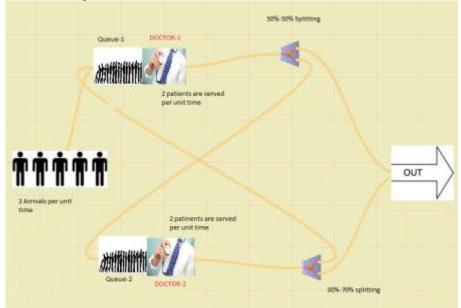


Figure 11. Two tandem open queues

There is a 29.7% service utilization of doctor whereas this is 59.2% for second doctor. The main reason for the second doctor is that there are not patients in queue 2 when it is empty. Hence, it has to be idle and wait for the patients. The second queue depends on the first doctor. If more patients can be sent per unit time, the utilization of

second doctor will increase. 0.702 unit time is spent in queue 1 and this is 0.212 per unit time for second queue. There are no drops due to infinite of capacity of queues.

In this model, four doctors are modeled and patients can visit other doctors based on structure of IEAH hospital. Some patients from each doctor can leave the hospital after getting treatment. Each queue has a limited capacity to see the effect of drops and the model was simulated 480.000 unit times. Other information (arrivals, splitting, exits) about the model is shown below.

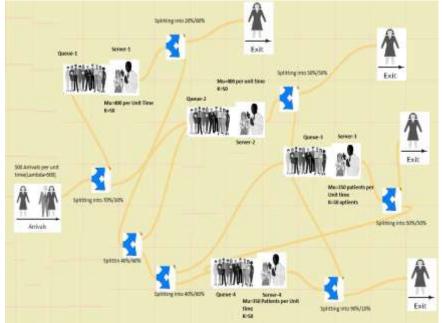


Figure 12. Four servers open queues

Doctor 3 serves fewer patients than others and thus 12,855,194 patients have visited that doctor and there are 2,774,581 drops from that queue. This queue has 100.0% utilization and it is a sign that another doctor can be added to that section to decrease drops. There are 1,655 drops from first queue with 87.5% server utilization. The capacity and arrivals of queue can be increased to raise the utilization. This queue takes 10, 079,117 arrivals. Queue 2 gets 7,545,242 arrivals. The second server takes least arrivals with 65.5% utilization. In the same manner, number of arrivals for the fourth server is not enough to increase utilization. 0.135 time unit of patients is spent in queue 3 with 0.021 standard deviation. This queue has 47.159 length of queue and this is the highest among them. 5, 043,892 patients leave the system from exit 3.

Case Study 7: Priority Queues

A priority system in the healthcare of arrangements has been made by the Health Ministry to protect and prevent further problems of some patients in Turkey. According to September 12, 2010 "People Rating" No. 5982 adopted by presenting to the "Law on the Amendment of Certain Provisions of the Constitution of the Republic of Turkey" in the provisions of paragraphs added to Article 10 of the Constitution, priorities of outpatient examinations are re-determined without causing any inequality among men and women. Priority order is given below:

1) Emergency cases (with an acute onset illness, accident, and injury, immediate medical attention should be done in such situations that are decided by the physician)

2) Severely disabled people with disabilities in the report,

3) Pregnant women,

4) Over 65 years old people,

5) Children under seven years of age,

6) Task force together with the disabled and war veterans and widows and orphans of the martyrs. (SB, 2013)

In the observations and talking with some patients, they stated that the priority requirements are not applied well in hospital and other healthcare institutions. Old and Severely disabled people have been seen in queues with long waiting in HH and IEAH hospitals. Based law above, a priority model is tried to be developed for hospitals by using Markov chains and Queuing Theory. In the sample hospital selected, there is not any polyclinic priority system. They have tried to give priority to some patients, resulted in more problems and misunderstandings. They way to overcome it is to define a model defined below and order is classified according to that order. Then patients should follow just its priority gueue instead of following all queues. Numerators can sort the priority of queues and a priority section can be seen on it.

Single-server, markov preemptive model with up to five priority classes and one service rate and preemption with lower number having higher priority are simulated according Turkey priority system as shown in Appendix section. The overall service rate is 0.01 as shown below. The utilization rate is 74.00% for that system. Mean number of customers in the queue is 2,106,153,846 patients to determine waiting room place. With 26%, the system will be empty. This rate can be increased to decrease idle time of doctor. Arrivals rate can be increased to raise utilization of doctor. Different types of priority classes and services are applied in that study to show priority that queues can improve services in hospitals. Over observations, there was a cancer patient waited in the line to see the doctor in policlinic. The same case was investigated for other priority patients in queues at sample hospitals. The mathematics and logic behind priority queues are explained in details in that study. The right of some patients should be protected by management of hospitals and it should be affected to the performance of hospital to increase the treatment quality of these patients.

DISCUSSION

The sample hospitals' queues system was analyzed from Hastane Randevu System of Health Ministry started in 2011. Doctor, hospital branches and times can be seen from this online system. Patient can select suitable time to go doctor from that system. Appointments of patients can be seen from system and they can cancel over time if they have other plans. Doctors' names and free times can be seen online in that system. Moreover, alternative hospitals, region of hospitals and other days are shown on the system to see alternatives. Furthermore, an appointment can be give by phone also as shown below in figure. Moreover, complaints can be sent by mail and surveys, news, announcements, rules about appointments and other information can be seen in web wage. Patients do not go doctor at the day they get appointment, which is one of the weaknesses of that system. This results in not taking other patients. Some hospitals give new queue

numbers to patients without online appointment. This also creates stress and some problems in waiting line. Patient without online appointment can go doctor earlier than patient with online fixed time appointment. Other patients can be angry due to not having checking on the time given by system. Unplanned arrivals and rechecking are the main trouble makers in this system. With effective planning, the trouble can be decreased. Some patients are taken to be rechecked in some time duration. Moreover, some fixed doctors can be assigned to for unplanned patients. Stochastic nature of queues explained in that chapter can be effective to optimize the system.

Ten minutes checking time is given a doctor for scheduled patients and two hours afternoon is given to rechecking previous patients. The system is actually effective on paper. The main problem is that the number of rechecking patients and unplanned patients. During the observation at hospital, unplanned patients created great stress over scheduled patients and the rechecked patients came not just afternoon from the same day. They came also in the morning and other times or day during the service. There were no sign in front of doctor room and desk operator to sort patients as rechecked, unplanned and unscheduled patients. At HH hospital, 16 minutes per patient is given to psychiatry patients. It was observed that patients were treated in 5-10 minutes. Psychiatry tests are carried out by psychologist at the same department. Patients were collected at the same room and all tests were carried at the same time to all patients. 16 minutes is given for checking at IEAH hospital. Between 10-11 hours, five patients are checked. Not used time here can be used for other types of patients. Moreover, 16 minutes is given for neurology department, Ear, Nose and throat and urology departments. Furthermore, 10 minutes checking time is given at HH hospital for cardiology. Both hospitals are at the same city, but the checking times at policlinics are different.

CONCLUSION

Queues are big problem in public hospitals. Patients have to wait long times under sometimes undesirable conditions. In front of policlinic, many patients wait in lines and there are stress and unnecessary waiting. This chapter shows mathematically how the queues can be decreased. It provides beneficial results to be used in hospitals to make performance measures based on less waiting and less patients on queues. The current number based performance management ignores the quality side. Some patients come from far regions of city and they cannot sometimes get treatment at HH Hospital in Istanbul. Thus they have to go back. The management of patients is not well designed at the hospital. Based on my try to get an endoscopy at HH Hospital, they gave me a 72 days later queue number. It was too long and then I took that film at a private hospital in two days. Normally, one can get that film in one day in private hospital. Based on some interviews, some hospitals outside of that region in Istanbul have shorter waiting time in public hospitals. The way to decrease these queues for limited films is to connect all hospitals scheduling system online and patients can select suitable hospital. Thus the waiting time can be decreased for endoscopy and similar films. There are misuses by giving the shorter queue to some patients. This system is to be changed and patients can get films at other hospitals in shorter time by some legal changes allowing to taking at other hospitals films. The average waiting time around Istanbul can be decreased in that manner by using not used capacity at other hospitals of endoscopy. The current online appointment system is a kind of new inventions for patients, but the unorganized rechecking and new patients without online appointment create great problems in queues.

They are to be separated to different lines. Many hospital staffs try to get prior treatment for their friends or relatives. This kinds of misuses are not to be allowed and these things are seen by other patients. It is clear that there should be some penalties to hospital management and staff for not providing enough services to patients.

There is no priority order in Hastane Randevu System. Thus, it is believed that priority system can be added to the system. A fixed time can be given these unplanned patients and re-checking patients to prevent extra congestions such as from 9:00-10:00 and 14:00-15:00. Especially, it is noticed that there are few patients after 15:00 PM at hospitals. Normal physicians can check patients until 17:00. Furthermore, 10 minutes checking time is given at HH hospital for cardiology department different from 16 minutes checking at IEAH Hospital. Both hospitals are at the same city, but checking times at policlinics are different.

It is found that open queues are the most applicable types of queues in healthcare. In this study, open queues are analyzed in details. There are not any queues based performance system in hospitals but average waiting. Fewer queues can be a success parameter for hospitals. Not outages calculations are done at hospitals and there are not statistics about it. It was noticed that some staffs went outside at radiology department at HH Hospital to smoke.

A new system measuring the time of waiting from taking barcode from operator until doctor start service time, service time at doctor room and sending another doctor with waiting time and service time by considering the full capacity of the other doctor and scheduling another day. Then, the automatic performance measures of queues and efficiency of doctors can be measured. In this system, patients coming for rechecking are to take a queue number or can be served at definite times of day.

REFERENCES

Alfa A.S. (2010) "Queueing Theory for Telecommunications- Discrete Time Modelling of a Single Node System", Springer Science+Business Media, ISBN 978-1-4419-7313-9, e-ISBN 978-1-4419-7314-6, DOI 10.1007/978-1-4419-7314-6)

Armero C.& Conesa D.(2004) "Statistical performance of a multiclass bulk production queueing system", European Journal of Operational Research 158, pp.649–661.

Al-Nowibet K. M/M/s Queueing Model,2013. .<u>http://faculty.ksu.edu.sa/9766/OR372/Lec12_MM1k_Queueing%20System2.pdf</u>, Access Date:25.03.2013.

Bhat U.N.(2000). An Introduction to Queueing Theory Modeling and Analysis in Applications. Birkhauser, Boston, 2000. ISBN: 978-0-8176-4724-7.

Boucherie R.J. & Dijk N.M. (1997) "On the arrival theorem for product form queueing networks with blocking", Performance Evaluation, 29, pp.155-176.

Creemers, S. and Lambrecht, Modeling a Healthcare System as a Queueing Network: The Case of a Belgian Hospital(Research Report 0710) Department of Decision Sciences & Information Management, Research Center for Operations Management, Katholieke Universiteit Leuven, 2007.

Chandrasekar B., Chandrasekhar P., Ramesh Kumar N., VSS Yadavalli(2011) "A Tandem Queue With Dependent Structure For Interarrival And Service Times: An

Estimation Study" Proceedings of the 41st International Conference on Computers & Industrial Engineering, p.145-150.

Gijsen B.M.M., Mei R.D., Engelberts P., Berg J.L., Wingerden K.M.C. (2006) "Sojourn time Approximations in Queueing Networks with Feedback", Performance Evaluation,63, pp.743–758.

Green L.V.& Savin S(2008) "Reducing Delays for Medical Appointments: A Queueing Approach", Operations Research, 56(6), pp. 1526–1538

Fiems D. & Bruneel H.(2013) "Discrete-time queueing systems with Markovian preemptive vacations", Mathematical and Computer Modeling, 57, pp.782–792.

Hillston J.Queuing Networks (Lecture Notes) School of Informatics The University of Edinburgh, Scotland, 2012.

Hopp, W. and Spearman, M. (2000) Factory Physics, Foundations of Manufacturing Management, New York: Irwin/McGraw-Hill.

Houdt B.V.(2012) "Analysis of the adaptive MMAP[K]/PH[K]/1 queue: A multi-type queue with adaptive arrivals and general impatience", European Journal of Operational Research, 220, pp.695–704.

Kim K.J. and Ayhan H.(2010) "Tandem Queues with Subexponential Service Times and Finite Buffers" Queueing Systems: Theory and Applications, Volume 66 (2), pp. 195-209.

Koizumi N., (2002) A queueing Network Model with Blocking: Analysis of Congested Patients Flows in Mental Health Systems, *Dissertations available from ProQuest*. Paper AAI3073021.

http://repository.upenn.edu/dissertations/AAI3073021.

Krinik A. & Mohanty S.G.(2010) "On batch queueing systems: A combinatorial approach" Journal of Statistical Planning and Inference, 140, pp.2271–2284.

Modiano E., Lectures 5 & 6 Introduction to Queueing Theory, Massachusetts Institute of Technology, 2013.

http://ece.ut.ac.ir/Classpages/S87/ECE658/Notes/Note%207.pdf(Access date: 22.03.2013))

Mousavi A., Introduction to systems modelling & simulation(Course notes), School of Engineering and design, MSc Programme, Brunel University, UK, 2011.

Minkevi£ius S.(2009) "On extreme values in open queueing Networks", Mathematical and Computer Modelling, 50, pp.1058-1066

MDE-Lab(2013). Queuing Models And Capacity Planning.

http://labs.fme.aegean.gr/decision/images/stories/docs/HealthCare_Queuing&Capacity Planning.pdf(Acess Date:01.04.2013.)

Osorio C. & Bierlaire M.(2009) "An analytic finite capacity queuing network model capturing the propagation of congestion and blocking" European Journal of Operational Research, 196, pp.996–1007.

Lakshmi C. & Sivakumar A.I.(2013) "Application of queueing theory in health care: A literature review", Operations Research for Health Care, http://dx.doi.org/10.1016/j.orhc.2013.03.002)

SB-Sağlık Bakanlığı.Poliklinik Hizmetlerinde Öncelik Sırası Hakkında Genelge. (2010/73). Ankara: Sağlık Bakanlığı, 2013. http://www.saglikaktuel.com/haber/poliklinik-hizmetlerinde-oncelik-sirasi-hakkinda-genelge-201073-12701.htm(Access Date:24.04.2013)

Yarmand M.H.and Down D.G.Maximizing Throughput in Zero-buer Tandem Lines with Dedicated and Fexible Servers., McMaster University, Department of Computing

and Software, Canada, 2012. <u>http://www.cas.mcmaster.ca/~downd/yardow.pdf(Access</u> Date: 12.04.2013.

Yan T. and Veeraraghavan M. Networks of Queues, 2004. http://www.ece.virginia.edu/mv/edu/715/lectures/QNet.pdf(Acess Date: 21.03.2013) Wang J.Y .Chap17 Queuing Theory. College of Management, NCTU., 2009.

APPENDIX:

Input Parameters:	
Mean arrival rate for class 1- Emergency cases	4
Mean arrival rate for class 2- Severely disabled people with	5
disabilities in the report	
Mean arrival rate for class 3- pregnant women	20
Mean arrival rate for class 4 - over 65 years old people	15
Mean arrival rate for class 5- children under seven years of age	30
Overall mean service time	0.01
Results of All System:	
Overall arrival rate	74
Mean interarrival time	0.013513514
Overall service rate	100
Server utilization	74.00%
Probability of an empty system (p0)	0.260000
Mean number of customers in the system (L)	2.846153846
Mean number of customers in the queue (Lq)	2.106153846
Mean wait time (W)	0.038461538
Mean wait time in the queue (Wq)	0.028461538
Priority Class 1- Emergency cases:	
Expected time in the system (W1)	0.010416667
Expected waiting time in the queue (Wq1)	0.000416667
Expected number in the system (L1)	0.041666667
Expected number in the queue (Lq1)	0.001666667
Priority Class 2- Severely disabled people with disabilities in the	
report:	
Expected time in the system (W2)	0.011446886
Expected waiting time in the queue (Wq2)	0.001446886
Expected number in the system (L2)	0.057234432
Expected number in the queue (Lq2)	0.007234432
Priority Class 3- pregnant women:	
Expected time in the system (W3)	0.01547748
Expected waiting time in the queue (Wq3)	0.00547748
Expected number in the system (L3)	0.309549605
Expected number in the queue (Lq3)	0.109549605
Priority Class 4- over 65 years old people:	
Expected time in the system (W4)	0.025150905

Expected waiting time in the queue (Wq4)	0.015150905
Expected number in the system (L4)	0.377263581
Expected number in the queue (Lq4)	0.227263581
Priority Class 5- children under seven years of age:	
Expected time in the system (W5)	0.068681319
Expected waiting time in the queue (Wq5)	0.058681319
Expected number in the system (L5)	2.06043956
Expected number in the queue (Lq5)	1.76043956

------ XXXXXXXXXXXX -------