13th International Research/Expert Conference
"Trends in the Development of Machinery and Associated Technology"
TMT 2009, Hammamet, Tunisia, 16-21 October 2009

WAITING LINE SIMULATION USING WinQSB SOFTWARE

Ranko Božičković, Ph.D. & Zdravko Božičković, M.Sc.
University of East Sarajevo, The Faculty of Transport and Traffic Engineering, Doboj
Bosnia and Herzegovina

Siniša Božičković, B.Sc. (Econ)
Raiffeisen bank, Doboj
Bosnia and Herzegovina

ABSTRACT

This paper focuses on possibility of Win QSB software application in dealing with waiting lines in banks. The application of queueing theory for solving waiting lines in the field of business has recently gained considerable attention. Modeling and simulation, carried out more often at a systematic high level of abstraction in a strategic decision-making phase, give design teams a unique opportunity to create a list of "potential layouts", that in a given or future time, provide the managerial team with the most optimal or near-optimal solutions to a material and human resource utilization. High-level simulation allows the design team to: a) supervise, analyze, and resume the critical paths of processing flows and information in each layout design, (b) refigure business system layout in order to reduce time losses and costs in the system c) eliminate inevitable bottlenecks present in almost every business, manufacturing and service system.

Key words: queueing theory, waiting line, simulation, model.

1. INTRODUCTION

Queueing theory is concerned with the study of processes in which service demands, on the one hand, and the possibilities of such demand fulfillment, on the other hand, are examined. The aim of queueing theory is to elaborate mathematical methods which can be used for assessment of waiting line system effectiveness. **The importance of queueing theory lies in the possibility of mathematical modeling of the real-life service processes and quantifying given process parameteres such as the service rate, the average time customer spends in the queue, the queue length, the number of servers, the service time, etc.** Such models in queueing theory allow service process analysis and offer a solution to improvement of given parameters essential for the service system and its optimization. The application of these models in practice is very useful for the society*: work processes are improved further influencing higher economic effects of the system, performance reliability and service management are increased, system performance in different situations is forecasted etc.*

A strong need for a more rational economic behavior and more efficient solutions has arisen from the current restrictive financing conditions, enterprises in transition and bank competitiveness. By increasing the number of servers, teller windows and tellers the real needs of the customers would be met. It is, therefore, essential to design an optimal waiting line model using scientific methods. Such model performs the defined tasks thus reducing total costs.

2. TYPES OF WAITING LINE SYSTEMS

Queueing theory deals with functional relationships between the factors determining service efficiency such as probability of service demands or service needs, cost of services on the one, and demand flow, service time and service strategy characteristics on the other hand. [1]

The subject of queueing theory can be described as follows: independent entities (input flow) from a population of customers arrive at a place (a service system) in a given time period (usually random) in order to obtain service, but with a possibility of waiting in the queue if the service is not available, and then leave the system (output flow).
With all the variety of entities, service demands and demand fulfillment methods, several elements and characteristics can still be identified as being common for all types or some groups of waiting line systems.

Waiting line models
Waiting line systems differ according to customer arrival rates, probability distribution of service time and „service discipline." Depending on the number of servers, waiting line systems can be: ***a***) ***single server (single channel) and b) multiserver (multichannel).***
Depending on customer behavior when a customer arrives in a busy period, waiting line systems can be divided into: a) queue management systems, b) balking systems and c) combined systems.
Customer service can be: ***a) single-phase*** (completed all at once) and ***b) multiphase*** (completed in a series of steps).
In case of a multiphase system, we are reffering to *a multiserver system* which consists of larger number of subsystems in which the input flow in one subsystem becomes the output flow of the customers in the next. Number of servers can differ from phase to phase. Phases, on the other hand, can be the same for all customers or can be randomized according to customer demands.
Classification of waiting line systems can also be made on a basis of **queue (service) discipline,** i.e. the method by which customers are selected from the queue, which will be discussed later.
A designed model should provide answers to the following important questions:
What is the probability a customer will be served without waiting in the queue? If a customer waits, what is the time spent in the queue? What is the expected service time? What is the expected total time a customer spends in the system? What is the expected average number of customers in the system? What is the queue length? What is the service capacity utilization?

3. AN ILLUSTRATIVE EXAMPLE
Two tellers (s=2) work at a teller window (with 2 servers) in two shifts on weekdays from 06.00 am to 08.00 pm (total of 14 hours). There could be 5 customers waiting at the teller's window (m=5). It has been observed that the average customer arrival rate per hour is 22 (λ=22), whereas the mean service time is 2.5 minutes. If there are five customers waiting at the teller's window, then the sixth customer cannot join the queue. This customer is being balked and needs to come back later or wait outside the bank.
Some costs were initially assumed: the busy server cost 200 BAM per hour, the idle server cost 200 BAM per hour, the customer being served cost 50 BAM per hour, the customer waiting cost 50 BAM per hour. We assume that the costumer waiting cost and the custumer being served cost are the same – 50 BAM per hour, the busy server cost and the idle server cost are also the same – 200 BAM per hour.
Based on input data, the analysis of the waiting line system will be performed in order to determine its basic characteristics and identify teller window utilization (i.e. the equipment utilization and worker utilization).

The following basic notation and results for WinQSB software are used:

n=1	-	Number of customers being served and queued	η=46,2	-	Overall system utilization
λ=22	-	Customer arrival rate (expected arrival per time unit),	μ=24	-	Service rate
s=2	-	Number of servers	λ_n	-	New customer arrival rate
μ_n=1	-	Service rate with *n* customers in the system	L=1,162	-	Average number of customers in the system
L_q=0,238	-	Expected queue length	W=0,052 9	-	Average time customer spends in the system
W=0,053l	-	Average time customer spends in the system	w_q=0,010 8	-	Average time customer spends in the queue

w_b=0,037 - Average time customer spends in the queue for a busy system

P_0=36,6 - The probability that all servers are idle

P_w=29,1 % - The probability an arriving custumer waits

Input data for a chosen software are shown in Figure 1., whereas Figure 2. shows the results.

Data Description	ENTRY
Number of servers	2
Service rate (per server per hour)	24
Customer arrival rate (per hour)	22
Queue capacity (maximum waiting space)	5
Customer population	
Busy server cost per hour	200
Idle server cost per hour	200
Customer waiting cost per hour	50
Customer being served cost per hour	50
Cost of customer being balked	
Unit queue capacity cost	

Figure 1. Input data

a) If the queue capacity *m* is increased on a 1-10 scale, observed parameters over 5 customers are approximately constant, i.e. increased waiting space does not affect waiting time or overall system utilization. Our queue capacity *m* is 5.

04-20-2009	Performance Measure	Result
1	System: M/M/2/7	From Simulation
2	Customer arrival rate (lambda) per hour =	22,0000
3	Service rate per server (mu) per hour =	24,0000
4	Overall system effective arrival rate per hour =	21,9778
5	Overall system effective service rate per hour =	21,9778
6	Overall system utilization =	46,2124 %
7	Average number of customers in the system (L) =	1,1622
8	Average number of customers in the queue (Lq) =	0,2380
9	Average number of customers in the queue for a busy system (Lb) =	0,8178
10	Average time customer spends in the system (W) =	0,0529 hours
11	Average time customer spends in the queue (Wq) =	0,0108 hours
12	Average time customer spends in the queue for a busy system (Wb) =	0,0372 hours
13	The probability that all servers are idle (Po) =	36,6789 %
14	The probability an arriving customer waits (Pw) or system is busy (Pb) =	29,1037 %
15	Average number of customers being balked per hour =	0,0740
16	Total cost of busy server per hour =	$184,8500
17	Total cost of idle server per hour =	$215,1500
18	Total cost of customer waiting per hour =	$11,8998
19	Total cost of customer being served per hour =	$46,2124
20	Total cost of customer being balked per hour =	$0
21	Total queue space cost per hour =	$0
22	Total system cost per hour =	$458,1122
23	Simulation time in hour =	1000,0000
24	Starting data collection time in hour =	0
25	Number of observations collected =	21978
26	Maximum number of customers in the queue =	5
27	Total simulation CPU time in second =	1,4350

Figure 2. Output data

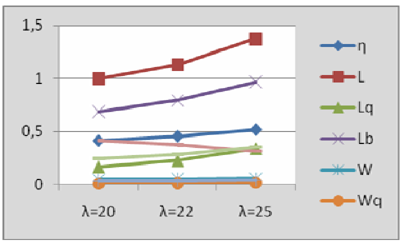
The screen with the results of this simulation is shown in Figure 3.

Sensitivity Analysis of Queue capacity for Sample M/M/1 Problem

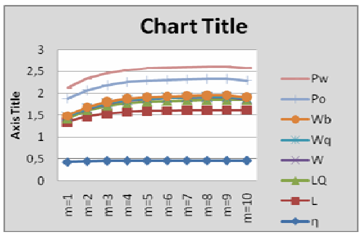
04-20-2009 Value	Effective Arrival Rate	System Utilization	L	Lq	Lb	W	Wq	Wb	P0	Pw	Average Balked	Busy Server Cost	Idle Server Cost	Waiting Customer	Served Customer	Balke Custom
1	20,3251	0,4234	0,9230	0,0761	0,3143	0,0454	0,0037	0,0155	0,3954	0,2422	1,6749	169,3760	230,6240	3,8066	42,3440	
2	21,2582	0,4429	1,0268	0,1410	0,5265	0,0483	0,0066	0,0248	0,3820	0,2678	0,7418	177,1519	222,8481	7,0499	44,2880	
3	21,6652	0,4514	1,0872	0,1845	0,6615	0,0502	0,0085	0,0305	0,3762	0,2789	0,3348	180,5433	219,4567	9,2254	45,1358	
4	21,8476	0,4552	1,1213	0,2109	0,7429	0,0513	0,0097	0,0340	0,3736	0,2839	0,1524	182,0634	217,9366	10,5468	45,5159	
5	21,9304	0,4569	1,1399	0,2261	0,7900	0,0520	0,0103	0,0360	0,3724	0,2862	0,0696	182,7531	217,2469	11,3046	45,6883	
6	21,9681	0,4577	1,1498	0,2345	0,8163	0,0523	0,0107	0,0372	0,3719	0,2872	0,0319	183,0678	216,9322	11,7228	45,7670	
7	21,9854	0,4580	1,1550	0,2389	0,8305	0,0525	0,0109	0,0378	0,3716	0,2877	0,0146	183,2117	216,7883	11,9472	45,8029	
8	21,9933	0,4582	1,1577	0,2413	0,8381	0,0526	0,0110	0,0381	0,3715	0,2879	0,0067	183,2776	216,7224	12,0652	45,8194	
9	21,9969	0,4583	1,1591	0,2425	0,8421	0,0527	0,0110	0,0383	0,3715	0,2880	0,0031	183,3078	216,6922	12,1262	45,8270	
10	21,9986	0,4583	1,1598	0,2431	0,8441	0,0527	0,0111	0,0384	0,3714	0,2881	0,0014	183,3216	216,6784	12,1573	45,8304	

Figure 3. Results simulation

b) If customer arrival rate (lambda) is increased from 20 to 25, with other values unchanged, it can be noticed that the other observed parameters are impaired, Figure 4a. Our λ=22.



a)



b)

Figure 4. Variation in arrival rate and service rate

c) If the service rate μ is increased from 20 to 30 customers per hour, other values being constant, all other characteristics are significantly improved, i.e. the values drop, Figure 5. In our case μ=24 customers per hour.

d) Service cost analysis shown in Figure 5b: number of servers is varied from 1 to 3, and queue capacity is varied from 4 to 6.

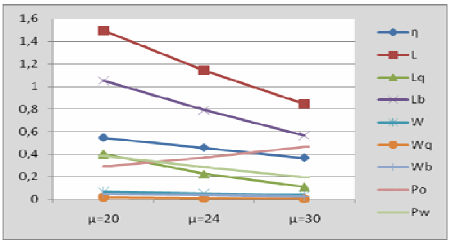


Figure 5. Service rate

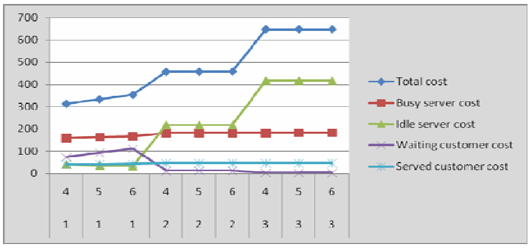


Figure 6. Variation in number of servers and capacity m

The analysis shows that the idle server cost increases with the increased number of servers, whereas the busy server cost shows a slight tendency to increase. On the other hand, an increase in the number of servers results in a lower customer waiting cost and a slight drop of service cost.

4. CONCLUSION

The acquired parameter values for the waiting line system at the bank teller window show that the mean arrival rate (i.e. customer arrival rate) is not high, the capacity utilization is consequently lower. The average percentage of an idle teller window (no customers being served or waiting in the queue) of 37.24 % and the average percentage of a busy teller window utilization of 45.68 % are unsatisfactory.

Optimization of this waiting line model in terms of efficiency improvement and better capacity utilization, could be performed in two directions:

1. Customer arrival rate at the teller's window per unit time could be increased by increasing the quality and range of services (this is one of the possible ways to attract new customers);
2. Statistically, the customer arrival rate varies, so it is advisable to reorganize and restructure a plan of hiring a teller. Instead of two tellers working in two shifts (total of 4 bank tellers), with two servers busy from 08.00 am to 08.00 pm, the total number of tellers could be reduced to three, one teller would work a split shift during the bank´s peak operating periods when the second server would be activated too.

We can conclude that the cost-effectiveness and the efficiency of the waiting line system for the described model is unsatisfactory, so there is still plenty of room for improvement and rationalization.

5. REFERENCES

[1] Nikolić I., Božičković R.: *Metode optimizacije;* Saobraćajni fakultet, Doboj, 2007
[2] Nikolić I., Borović S.: *Metode i softer za operaciona istraživanja*, CD ROM, FON, Beograd, 2001.
[3] Vukadinović S. : *Masovno opsluživanje*, Naučna knjiga, Beograd, 1988.