ARCC 2009 - Leadership in Architectural Research, between academia and the profession, San Antonio, TX, 15-18 April 2009

Use of discrete event simulation in hospital capacity planning

Zhengwei Li, Yeonsook Heo, and Godfried Augenbroe

Georgia Institute of Technology, Atlanta, Georgia

ABSTRACT: In recent years, the healthcare industry is undergoing a rapid expansion in the United States. For healthcare facilities, resource planning at early design stage is a critical step before architectural design. The ‘resources’ here refer to both long term resources (pods, rooms, beds, configuration of one pod) in terms of capacity and configuration, and short term resources(staffs, equipments) in terms of capacity and allocation. To achieve performance targets defined by the clients, such as staff/equipment/bed utilization efficiency, average waiting time of all patients, turn away rate, an assessment and verification at the preliminary planning stage is necessary. There are at least two methods to solve this problem. The first is analytical in nature, relying on queuing theory, and falls under the industrial engineering field. The other is computational in nature, relying on process simulation, and specifically discrete event simulation. While queuing theory is easier to conduct, usually requiring less data, and providing more generic rules than simulation, simulation methods result in detailed information about patient flow modeling and deliver more accurate results. This paper is divided into three parts. The first part introduces queuing theory and discrete event simulation in terms of their principles, features and applications in healthcare planning. This is followed by a case study in the ED using discrete event simulation to plan pod configuration and number of pods for an emergency department. During this process, the simulation tool is introduced as an example instrument for advanced DES simulation. The paper ends with a discussion of outcomes. (1) DES is capable to differentiate between alternatives with small changes, and can be widely used to do capacity planning for healthcare facilities. (2) the chosen simulation tool supports the modelling and analysis steps well.

Conference theme: Innovations in architecture for health and related facilities
Keywords: healthcare, capacity planning, discrete event simulation

INTRODUCTION

In the United States more than $20 billion will be spent on construction per year by the end of the decade (Roger Ulrich, 2004). Increasing patient satisfaction has become an active research theme, one of the goals set by these researches is to reduce waiting time patients spend in hospitals (Jones, 1994; Anthony, 2001). On the other hand, maximizing resource utilization efficiency is a critical operational target of hospitals. Some resources like rooms/pods depend on long term static investments, whereas other resources are based on short-term and flexible investment decisions such as equipment and staff. These two goals typically conflict with each other, therefore, a detailed capacity planning process needs to be conducted before preliminary architectural design stage.

Traditionally, there are two methods to do resource capacity planning for healthcare units. The first method uses queuing theory. Methods in queuing theory can further be divided into either simple queuing theory (typically there is only one queue), or queuing network theory (composed of several queues). A simple queuing model has only one queue and one server, but it can be extended to multiple servers and priority queuing. As early as 1971, Gupta et.al (Gupta, 1971) used queuing theory to solve the problem of multi-server, single queue problem. Daniel G. Shimshak et.al (Shimshak, 1981) conducted a staffing analysis for a pharmacy unit using three models (normal queue without priority, priority queue with non-preemptive service, priority queue with preemptive service). Green(2002) concluded that as a general model M/M/s is widely applied to service industries for capacity planning purposes. This model assumes an unlimited queue with n servers. Typical assumptions are that arrivals occur at a Poisson process, and service time has an exponential distribution. On the other hand, queuing network theory is a more complex theory, in which servers are represented as nodes, all the nodes are connected in a network, the status of the network is represented by the number of customers in each node. It has been used to help analyze a regional hospital system in Philadelphia (Koizumi N, 2005). Compared to one queue problem, the solution process of queuing network theory requires much more effort.

Compared to queuing theory, discrete event simulation has the advantage that simulation can optimize resources or schedules by comparing different alternatives, while queuing theory can only get an optimized value (Lowery, 1998). Jun et.al (1999) divided this research area into three categories: patient scheduling and admissions, patient routing and flow schemes, and scheduling and availability of resources.
In the patient scheduling/arrival study, Smith and Warner (1971) found that with a uniform arrival rate, the average length of stay can be decreased by over 40%. Ming Guo et.al (2004) used the average/maximum patient waiting time, and staff utilization level as main performance characteristics, to compare several out-patient scheduling types. Similar work can be found in Klassen(1996), but with different simulation approaches. Blake et.al(1996) used DES to analyze the emergency room of a children's hospital, they found that using 'fast track' and increasing physician hours can decrease the patient waiting time.
Another active application of DES is resource planning. Giorgio et.al (1987) used DES to simulate a children ward in Italy, to give optimal number of beds and number/organization of nurses. Harris et.al(1985) use DES to plan the bed requirements for a operation theater in South Wales. Klafehn et.al (1989) find that moving a nurse from a regular emergence area to a triage position reduces the number of waiting patients.
However, the simulations mentioned above basically run behind the screen, they didn't deploy visual animation. Advances in software industry have made this possible in recent extensions. Some of them have been particularly applied to healthcare industry, such as MedModel (Deney,1997; Charles ,2000), ProModel(Charles,1992) and Arena(Guo,2004).
To summarize, both DES and queuing theory can be applied to solve resource planning problems. Queuing theory is more straightforward, requires less effort, but has limitations in handling complex healthcare environments. On the other hand, DES is capable to solve complex problems but requires more time to deploy, in particular to build the process model. However, with the advances in software industry, the time taken to build the model is tending to reduce. This paper explores the use of a more recent simulation package on healthcare planning processes, using the real life ED planning for a new hospital as a case study.

1. CASE STUDY

1.1. Problem Statement

The new hospital is in the design programming stage. We focus on the ED of the hospital, where both ambulance patients and walk-in patients arrive randomly. The total patient arrival rate is assumed to be known, based on the data from previous years (Fig. 1). The ratio of ambulance patients and walk-in patients is assumed to be 1:9. Also assumed to be known are the patient disease distribution (Table 1) and patient ESI (Emergency Severity Index) level distribution (Table 2). An ESI level is an index associated with a patient, to group him/her into five groups, from 1(most urgent) to 5(least urgent). The patient flow process is shown in Fig. 2. In all the queues of the system, patients with ESI level 2 have higher priority than patients with ESI level 3 and 4.
The first design question is: what is the best pod configuration? One possible choice is 'pod without sub-waiting room', where each patient occupies the bed until he/she is discharged. The competing choice is 'pod with sub-waiting room', where each patient releases the bed when he/she leaves the pod for other care processes, after comes back, he/she will occupy another bed in the same pod. In case that there is no bed temporarily available, he/she waits in the sub-waiting room.
The second question is related with the first one. After the best pod configuration is chosen, how many pods in the ED and how many beds in a pod are optimal? Because the hospital needs at least one pod for gynecology patients and one pod for psychiatric patients, these two pods are fixed in this study.

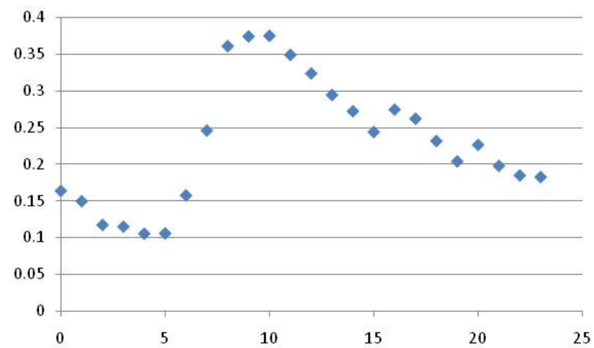

Figure 1: Total Patient Arrival Rate (unit: per/min)

Table 1: Patient Disease Type Distribution

Disease Type	***Walk-in***	***Ambulance***
Psychiatric	0.15	0.15
Gynecology	0.04	0.04
Trauma	0	0.005
Other	0.81	0.805

Table 2: Patient ESI Level Distribution

ESI Level	*Walk-in*	*Ambulance*
1	0	0.005
2	0	0.295
3	0.107	0.7
4	0.763	0
5	0.13	0

1.2. Simulation Approach

We conducted a discrete event simulation to solve these two problems using Anylogic 6.3.1, a java based simulation package from X-logic. Fig.3 illustrates the process flow charts of this model. The top shows the overall flow chart. Each patient enters the process from a source, then goes through care processes based on

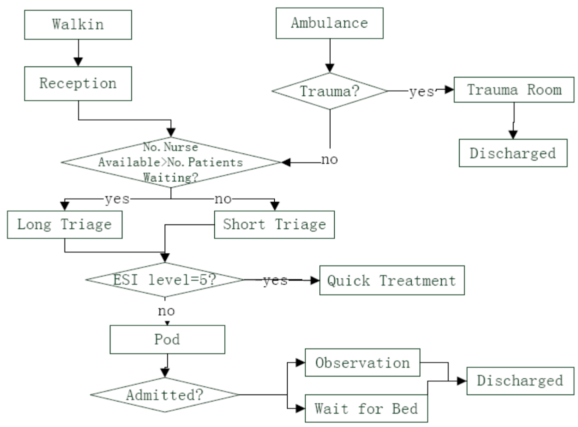

Figure 2: Patient Flow Process Chart (walkin and ambulance)

his/her disease type and ESI level. Sub-processes (reception, triage, examination…) are modeler created blocks, composed by software built-in components. For example, reception block is composed by two smaller blocks: waiting_checking and action, both of them are made by basic components(See Fig.3). Using this modularized modeling approach, debugging becomes much easier.

Another advantage of the software is the ability to change resource parameters real-time. One can, for example, change the number of nurse back and forth. In this way, it is easy to find where the bottleneck is. An example of animation screen is shown in Fig. 4: on the left is the animation screen, and on the right top are the radio buttons, by which the number of nurses can be changed during the runtime, and at the bottom shows the average utilization rate of doctors and nurses during the running simulation.

1.3. Simulation Results

As said above, simulations are set up to provide answers to two questions:

Figure4: Model Runtime Screenshot

(1) What is the best pod configuration?

There are three options as candidates:

- 10 examination room without sub-waiting spaces
- 8 examination rooms with 5 sub-waiting spaces
- 6 examination rooms with 5 sub-waiting spaces.

(2) For the best pod configuration, how many pods are most cost effective?

It is recognized that different pod configurations should be compared under the assumption that the total number of beds is the same or at least close. To answer question 1 and 2, a total number of 14 scenarios are modeled and simulated:

- 10/11/12/14/16 pods, with each pod having 8 examination rooms with sub-waiting space for 5 patients
- 12/13/14/16 pods, with each pod having 6 examination rooms with sub-waiting space for 5 patients
- 10/11/12/14/16 pods, with each pod having 10 examination rooms.

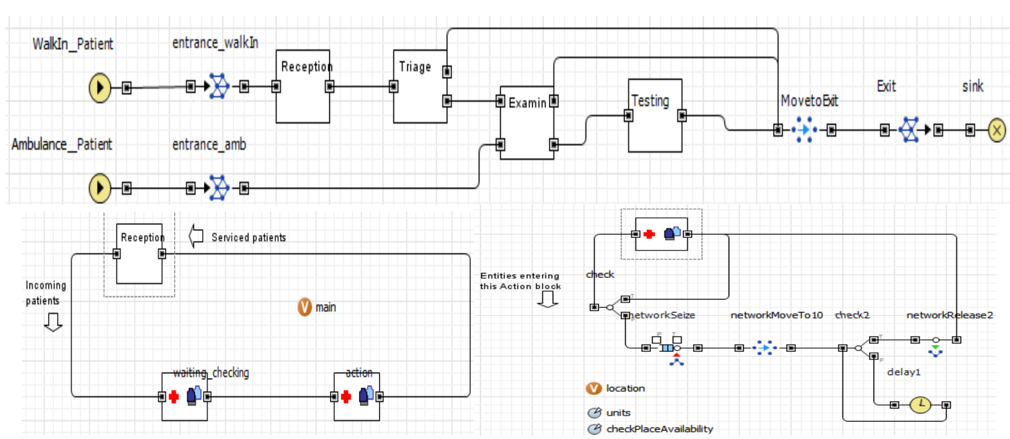

Figure 3: Process Flowcharts in Anylogic

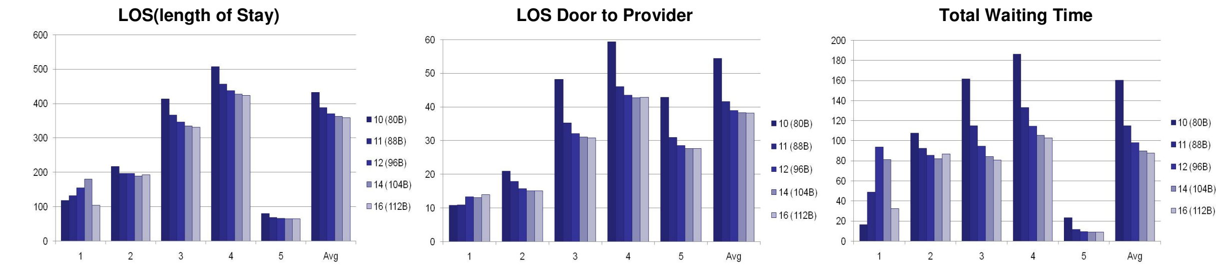

Figure5: LOS, LOS Door to Provider, total Waiting Time for 8+5(units: mins)

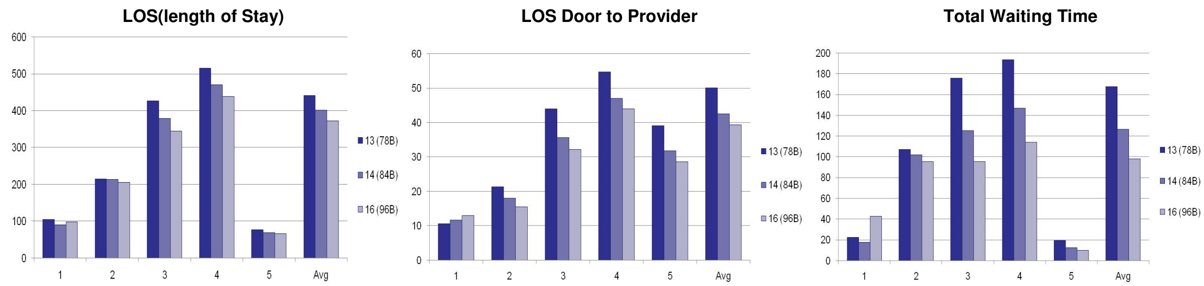

Figure6. LOS, LOS Door to Provider, Total Waiting Time for 6+5(units: mins)

To measure the through-put performance,six criteria are introduced. They are

- Average LOS (length of stay)
- Average LOS from Door to Provider
- Average total waiting time
- LOS for each ESI level patient
- LOS Door to Provider for each ESI level patient
- Total waiting time of each ESI level patient

1.3.1. First Configuration: 8 exam rooms with 5 sub-waiting spaces

The performance results are presented in Fig.5. Level 1(trauma) patients are treated differently compared with other patients, so they are not discussed in the following part. Level 5 patients don't go through any testing process, and the treatment process is very simple, so their LOS and waiting time are the lowest. Among level 2, 3 and 4 patients, level 2 patients have the lowest LOS, LOS door to provider, and waiting time, due to their higher priority in getting resources.

From Fig 5, it can be seen that increasing the number of pods from 10 to 11 has the biggest impact on performance, a further increase of pods has basically no effect on the performance. Therefore, for this configuration, the optimal number of pods Is 11, the corresponding bed number is 88.

1.3.2. Second Configuration: 6 exam rooms with 5 sub-waiting spaces

The performance results are shown in Fig.6. It can be seen that from 13 to 14, and from 14 to 16, there is a steady increase in the performance. It is estimated that further increase from 16 will have only minor impact on the performance. This is not studied here, because the maximum number of pods was initially limited to 16 based on other considerations in the planning phase

1.3.3. Third Configuration: 10 exam rooms without sub-waiting spaces

The results for 10 exam rooms without sub-waiting spaces are shown in Fig7. Increase from 10 pods to 11 pods makes the LOS door to provider and total waiting time apparently lower, however, the total LOS is almost not affected. Increase after 11 pods has no further effects.

1.3.4. Comparison Of Different Pod Configurations

- 8+5 v.s. 6+5

From Fig.8, it can be seen that the configuration 8+5 and 6+5 basically follow the same rule: bed number determines the capacity, although there is one exception that 8+5 with 80 beds has higher LOS door to provider than 6+5 with 78 beds.

Considering the overall performance, it is fairly safe to conclude that there is no significant difference between these two configurations

- 8+5 v.s. 10+0

For configuration with 10+0, only 10 and 11 pods cases were simulated. It can be seen that although 10+0 with 10pods has higher bed number than 8+5 with 12pods, it has poorer performance in both LOS and LOS door to provider. In terms of total waiting time, there is no significant difference between these two configurations. Based on this result, it is concluded that 10+0 configuration is not as good as the other two configurations.

To summarize, pod configuration with sub-waiting spaces have higher performance than those without sub-waiting spaces. There is no obvious performance difference between configurations with 8 beds per pod and 6 beds per pod.

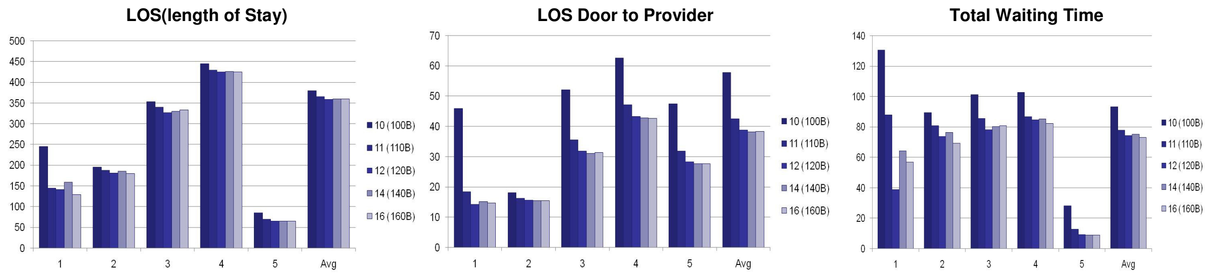

Figure7. LOS, LOS Door to Provider, total Waiting Time for 10+0(units: mins)

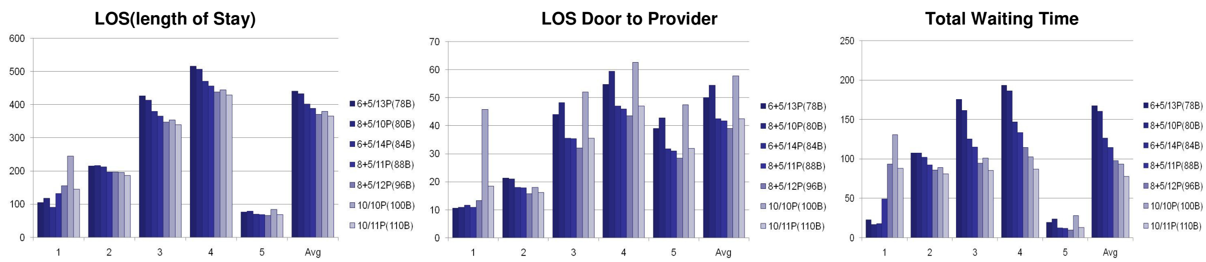

Figure8. Comparison of all three configurations (units: mins)

Based on the findings from the simulation, our answers to the question 1 and 2 is that, using either the 6+5 or the 8+5 case, the optimum number of pods is either 14(6+5) or 11(8+5).

CONCLUSIONS

In this study, an ED case study is used to demonstrate how DES can be used to support general capacity planning problem. Three performance measures: Length of Stay, Length of Stay from door to provider and total waiting time are used to compare different alternatives. It is found that configurations with sub-waiting spaces perform better in terms of LOS and LOS door to provider. The total waiting time is affected dominantly by the total number of beds. Based on the performance, either 14 pods with 6+5 configuration or 11 pods with 8+5 configuration is recommended. A further check to whether the 6+5 or 8+5 variants score better on other design outcome criteria requires more detailed information, which can't be known at the planning and design programming stage.
Further work is planned that will deal with those stages, inspecting cost benefits of different layouts, staffing, etc.

ACKNOWLEDGEMENT

The authors sincerely thank Prof. David Cowan in the Health System Institute at Georgia Tech for his great help and contribution. Also thank all colleagues and students that are involved in this project.

REFERENCES

Blake, J. T. 1996. *An Analysis of Emergency Room Wait Time Issues Via Comupter Simulation*. INFOR, 34(4), 12

Deney, S. H. 1997. *MedModel-Healthcare Simulation Software.* Paper presented at the1997 Winter Simulation Conference, Atlanta, GA, United States

Green, L. V. 2002-2003 Winter. *How many hospital beds?* Inquiry, 39(4), 12

Guo, M., and West, C.. 2004. *Outpatient Clinic Scheduling-A Simulation Approach.* Paper presented at the 2004 Winter Simulation Conference, Washington D.C., United States.

Gupta, J. Z., and Kramer, N. 1971. *Hospital manpower planning by use of queuing theory,* Health Services Research, 6(1), 7

Harrell, C. R. 2000. *Healthcare Simulation Modeling And Optimization Using MedModel.* Paper presented at the 2000 Winter Simulation Conference, Orlando, FL, United State.

Harrell, C. R. 1992. *ProModel Tutorial.* Paper presented at the 24th conference on Winter Simulation, Arlington, Virginia.

Harris. R.A. 1985. *Hospital bed requirements planning.* European Journal of Operation Research, 25(1),6

Jones, P. 1994. *Improving Service: Managing Response Time in Hospitality.* Production Management, 14(5), 7

Jun J. B., and Swisher J.R. 1999. *Application of discrete-event simulation in health care clinics: A survey*. Journal of the Operational Research Society*,* *50*(2), 15

Klafehn, K. A. 1989. *Evaluating the linkage between emergency medical services and the provision of scarce resources through simulation.* Paper presented at the 13th Annual Symposium on Computer Applications in Medical Care, Washington DC, Nov 5-8)

Klassen, K. J. 1996. *Scheduling outpatinet appointments in a dynamic Environment*. Journal of Operations Management,14(2), 19

Koizumi N, and Smith T. E. 2005. *Modeling patient flows using a queuing network with blocking*. Health Care Management Science,8(1),12

Lowery, J. C. 1998. *Getting Started In Simulation In Healthcare.* Paper presented at the 1998 Winter Simulation Conference, Washington, D.C., United States

Nosek, R. A., and Wilson, J. P. 2001. *Queuing Theory and Customer Satisfaction: A Review of Terminology, Trends, and Applications to Pharmacy Practice*. Hospital Pharmacy, 36(3)

Romanin-Jacur, G. 1987. *Optimal planning of a pediatric semi-intensive care unit via simulation*. European Journal of Operational Research, 29(2), 7

Shimshak, D. G., and Burden, H. D. 1981. *A priority queuing model of a hospital pharmacy unit*. European Journal of Operational Research*,* 7(4), 5

Smith, E. A. 1971. Simulation of a multiphasic screening procedure for hospital admissions. Simulation*,* 17(2), 8

Ulrich R., et al 2004. *The Role of the Physical Environment in the Hospital of the21st Century: A Once-in-a-Lifetime Opportunity*: The Robert Wood Johnson Foundation, Published in Centre of Health Design Website