8-2012

Improving patient flow in emergency department through dynamic priority queue

Kar Way TAN
Singapore Management University, KWTAN@smu.edu.sg

Chao WANG
Singapore Management University, wangchao.2007@smu.edu.sg

Hoong Chuin LAU
Singapore Management University, hclau@smu.edu.sg

DOI: https://doi.org/10.1109/CoASE.2012.6386409

8-2012

Improving Patient Flow in Emergency Department through Dynamic Priority Queue

Kar Way TAN
Singapore Management University, KWTAN@smu.edu.sg

Chao WANG

Hoong Chuin LAU
Singapore Management University, hclau@smu.edu.sg

Improving patient flow in emergency department through dynamic priority queue

Kar Way Tan, Chao Wang and Hoong Chuin Lau
School of Information Systems
Singapore Management University
80 Stamford Road, Singapore 178902
{kwtan,wangchao.2007,hclau}@smu.edu.sg

***Abstract*— Most queuing problems are based on FIFO, LIFO, or static priority queues; very few address dynamic priority queues. In this paper, we present a case in a hospital's emergency department (ED) where the queuing process can be modeled as a time-varying M/M/s queue with re-entrant patients. In order to improve patient flow in the department, we propose the use of a dynamic priority queue to dispatch patients to consultation with doctors. We test our proposed model using simulation and our experimental results show that a dynamic priority queue is effective in reducing the length of stay (LOS) of patients and hence improving patient flow. Furthermore, we show that a hybrid scheme is effective in preventing starvation.**

I. INTRODUCTION

The competition in healthcare industry becomes increasingly intensive. In order to serve the emergency medical needs of the people while maintaining the quality of care, a public hospital needs to provide better service and seek ways to improve patient flow in the emergency department (ED) by improving the processes and queue management in the department.

There are several ways to manage queues. Two simplest models are the first-in-first-out (FIFO) and the last-in-first-out (LIFO) models. One of the key issues with the traditional structure of FIFO and LIFO is that they fail to recognize that the individuals in the queue might not have the same priority. In the context of a hospital, patients with more severe illnesses are generally treated at a higher priority than those with less severe illnesses. This is an example of a priority queue. A static priority queue model calculates the priority of an entity based on some attributes of the entity (e.g., severity of illness) when it enters the queue and remains the same with respect to the other entities in the queue throughout its life-time in the queue system.

In this paper, we introduce the concept and a case for a dynamic-priority queue model where the priorities of the entities within the queue system are recalculated when one or more resources serving the entities become available. As an application, we consider the case of a hospital ED where the entities are the patients and resources are the doctors. The queue in our context is modeled as a generalized time-varying M/M/s queue with customer-centric service time, and re-entrant customers. In this queuing system, patients arrive stochastically (arrival rate varies over different hours of the day), and a doctor is needed to evaluate a patient and order investigation tests or on-site treatment for the patient according to that patient's situation. The investigation tests include lab test such as blood test, x-ray and point-of-care tests such as urine test. The patient will leave the queue (temporarily) to consult with the doctor and if necessary, take the investigation tests or treatment. After taking the tests or treatment, that patient may re-enter the queue and wait for the same doctor to review the results before either being discharged or admitted to the hospital. We use the term length of stay (LOS) to refer to the duration a patient spends at the ED after triage and before discharge or admission. The hospital in our case-study aims to serve its patients in ED within a targeted LOS, e.g., within 60 minutes for a simple case without investigation or observation. We propose to recalculate the priorities of all patients in the queue each time whenever a doctor becomes available to see a patient. In our experiments, the priority of a patient is dependent one or more of the following factors: (a) his estimated consultation time with the doctor and/or (b) the remaining time to meet the targeted LOS. We propose such a queuing model that intelligently dispatches the patients to doctors so as to improve the patient flow and reduce the average LOS of all patients in the ED.

Even though this work is based on an on-site requirement and data gathered from a local hospital, we believe our approach is generic and applicable to ED operations in general.

II. MOTIVATING CASE-STUDY

Our motivating case-study is a process in the ED of a local hospital. There are four different severity levels of the patients' conditions, namely P1, P2, P3 and P4. The conditions of P1 patients are highly critical and require immediate medical attention. P2 patients are attended before P3 and P4 patients. P4 patients are least ill and they are seldom seen at the ED. As a requirement given by the hospital, we based our study on improving patient flow for P3 patients only.

In our process, each patient is modeled as a job and to be served within the service level agreement. The hospital targets to serve the patients within a stipulated time (targeted LOS of a patient). For a patient who requires no additional investigation tests (e.g., blood test, x-ray, point-of-care tests) or treatment, the targeted LOS is set to 60 minutes, which is a value based on a field-study at the hospital. Patients who require test(s) or treatment may take longer than the targeted LOS. A graphical representation of the business process for a typical P3 patient is provided as follows and in Figure 1.

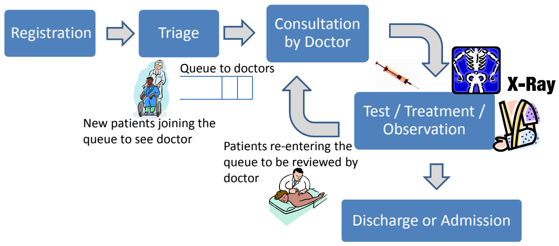


Fig. 1. Business Process in the Emergency Department

When a patient arrives, he/she first registers and then followed by a triage process where severity condition of the patient is determined by a nurse. The patient then consults a doctor. In some cases, the patient is required to take further investigative tests, treatment (e.g., medication), or observation. Some examples of these tests are blood test, x-ray or point-of-care-tests (e.g., urine test, electrocardiogram, eye test and hearing test). When the results are ready or the patient has received treatment, he/she is to be reviewed by the same doctor again. The patient will re-enter the queue to wait for the doctor. Such patients are called the re-entrants. After the review session, the patient may either be discharged or be admitted to the hospital.

In this paper, we consider only the queuing model for doctor consultation and review. The reason that we only focus on part of the process is based on our observation during our field study that this part of the process appears to be the bottleneck of the ED. Our queues are formed by two types of patients, firstly the new patients who have just arrived and completed triage, and secondly, the re-entrants.

Currently the patients are dispatched to doctors' consultation rooms in a FIFO manner. The doctors will select new patients based purely on their arrival times. For re-entrants, the doctors decide to call the patient in a somewhat unplanned fashion when their test(s)/treatment(s) are completed. Such an ad-hoc policy may not be optimal in minimizing the LOS for all patients. We propose a method of dispatching the patients based on dynamic priority queue. We implement three strategies in which the priorities of all patients in the queue are calculated each time a doctor becomes available. The strategies are shortest-consultation-time-first (SCON), shortest-remaining-time-first (SREM) and finally, a mixed strategy based on combination of SCON and SREM. The remaining-time of a patient is given by the difference between the targeted LOS and the total amount of time that the patient has spent in the sub-process of consultation/review by doctors. If the patient has to take tests or treatment, we allow the targeted LOS to extend beyond the original targeted LOS. This is required because the additional time spent may be substantial and it is reasonable for the patient to spend a time longer than original targeted LOS. When the patient re-enters the queue, his total time-spent does not start from zero, but it is accumulated and hence this patient can have a higher priority than another patient who is waiting to consult the doctor for the first time. In addition, we also observe in our field-study that the review consultation time could be shorter than the first consultation time. As such, the priorities of all the patients within the same queue are dynamically set. In more complex scenarios for future extension, the severity of the patient's illness may also change the priority of the patient. A patient who arrives with a fever may develop a seizure at the later stage and hence it is critical to allow the patient to be evaluated by the doctor sooner. We argue therefore that a dynamic priority queue is important in the context of an ED and intelligent dispatch policies need to be developed to allow the patients with the greatest (life-threatening) medical need to be attended, as well as to minimize LOS by considering each patient's time spent in the queue.

III. RELATED WORK

In this section, we provide a sketch of the literature in improving patient flow or patient waiting time, from the computational perspectives, and position our work against the literature.

Mayhew and Smith (2008)[1] used queuing theory to analyse a 4-hour completion time target in ED in UK. This is similar to our target LOS. We, however, have a more challenging target of 1 hour and hence we find a need for intelligence in improving the process such as dynamic priority to dispatch the right patient to the doctors.

The works in [2], [3], [4], [5] and [6] showed that the ED department can be modeled using discrete-event simulation. Simulation is used to perform what-if analysis to improve processes or have more effective staff planning in ED. In [6], authors allowed patient's priority to be dynamically updated based on the waiting time and the patient's underlying clinical priority. Our work differs from these works as we consider the situations that some patients may see the same physician multiple times (re-enter the queue) due to investigation tests and treatment. Re-entrance adds to the complexity because the arrival of the patients at consultation is no longer purely based on an arrival rate which a basic queue model assumes.

King et al. (2006)[7] introduced ways of reducing patient waiting time in ED department, such as triage systems to categorize patients into different group and treat them differently. Our work do not categorize the patients further into sub-category, instead we dynamically compute their priorities based on some criteria (such as their estimated consultation time with the doctor).

Chakravarthy et al. (1992)[8] used a threshold to dynamically switching between two types of customers to enter the service. Our work implements different strategies to determine which patient is desirable to enter the next service (i.e., consultation with the doctor).

IV. THE DYNAMIC-PRIORITY QUEUING MODEL WITH RE-ENTRANT ENTITIES

The overview of our dynamic priority queuing model is shown in figure 2. We assume that patients arrive according to a Poisson process with time-varying arrival rates. Each doctor is a server, and all servers are identical with an exponential service time distribution whose rate is dependent on the type of consultation. The service rate of an investigation test/treatment is also exponentially distributed and dependent on the type of test/treatment.

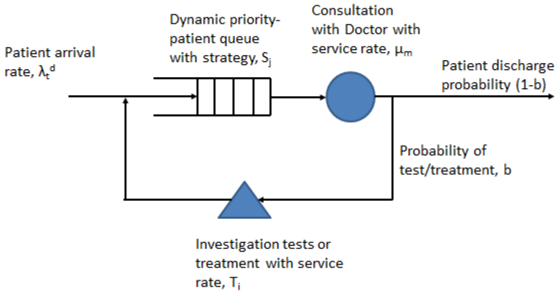


Fig. 2. The dynamic-priority queuing model

A. Model Parameters

We define the following parameters in our model.

- λ_t^d - Time-varying arrival rate of patients where t is the hours of the day in 24-hours format and d represent the day of week (i.e., Monday to Sunday).
- μ_m - Service rate of the doctors where m is the type of consultation, i.e., either first consultation or review consultation after test or treatment.
- T_i - Service rate of investigation test or treatment where i is the type of investigation test or treatment.
- S_j - Patient prioritization strategy where j is the strategy type as explained in section IV-B.
- b - Probability of re-entrance. Patients who require test(s) and/or treatment are required to be reviewed by a doctor, hence re-entering the queue.

B. Strategies in calculating priorities of patients

We propose three strategies to calculate priorities of patients in the queue, namely shortest-consultation-time-first (SCON), shortest-remaining-time-first (SREM) and mixed strategy incorporating both consultation-time and remaining-time factors (MIXED).

The formal definitions of the variables in our model are as follows: Each patient k who has registered has a targeted LOS d_k and an elapsed time e_k that he has spent in the department. At the end of triage (i.e. beginning of the queuing model), d_k is set to 60 (for simplicity of discussion) and the elapsed time is set to zero. If a patient requires an investigation test and/or treatment, he incurs additional time t_k based on the type of test or treatment based on treatment service rate T_i. The *remaining time* for the patient, r_k, is then defined as $d_k + t_k - e_k$. Note that we added t_k so as to allow the patient to spend more time in the ED to take the test(s) or treatment. In the consultation queue, the patient is either be a new patient or re-entrant. Each patient has an estimated *consultation time* c_k, computed based on doctors service rate μ_m against the type of consultation.

1) Shortest-Consultation-Time-First (SCON): In the SCON strategy S_1, we rank (give priority to) patients according to their estimated consultation times with the doctor. The intuition for this strategy is based on our observation that some re-entrants have very short estimated consultation time, e.g., doctor reviewed the blood test result which is clear of any suspicion with the patient takes only about 2 minutes. We suggest to clear such patients earlier and allowing them to exit the ED. Let c_k be the estimated consultation time of patient k, we use an exponential function (with ρ_1 as the constant parameter to set the gradient of the exponential function) to determine the priority of a patient p_k:

$$p_k^{S_1} \leftarrow e^{\frac{\rho_1}{c_k}} \tag{1}$$

2) Shortest-Remaining-Time-First (SREM): In the SREM strategy S_2, we rank patients according to their remaining times. This is based on the intuition that we want to maintain the reputation and confidence from the patients that the hospital could serve them within the targeted LOS. We like to assign the patient a priority, which tends to a large number when the remaining time tends to zero, and tends to 1, when the remaining time is sufficiently large. Furthermore, since the remaining time may even become negative (i.e. when the patient is yet to be served after the targeted LOS has elapsed), his/her priority should be set to an even larger value. For this purpose, we propose an exponentially decreasing function (with ρ_2 as a constant parameter to set the gradient of the function) when the remaining time is a positive number. When the remaining time tends to zero, we set it to a large constant to avoid division by zero. And when the remaining time becomes more negative, it increases linearly. As such, we propose a 3-segment function as shown in the following equation. Let c be a small value (e.g., 0.1) to ensure priority is very high (only few cases fall into such category). We let $f_c = e^{\frac{\rho_2}{c}}$ when $r_k = [-c, c]$. When $r_k < 0$, we use a negative linear function with a constant slope $m > 0$.

$$p_k^{S_2} \leftarrow \begin{cases} e^{\frac{\rho_2}{r_k}} & \text{if } r_k > c \\ f_c & \text{if } r_k \in [-c, c] \\ f_c - m * r_k & \text{if } r_k < -c \end{cases}$$

3) Mixed strategy (MIXED): We consider a MIXED strategy S_3 to take into considerations multiple factors in determining the priorities of the patients. We also observed (see experiments below) that a pure SCON strategy may potentially face the problem of starvation (patients who have long estimated consultation times repeatedly get preempted by patients who have shorter estimated consultation times). Hence, we propose a mixed strategy to prevent the starvation problem. Note that although we consider only two factors in this paper, namely, consultation-time and remaining-time, we believe that the model can be extended to more factors. We use a weighted scheme so that weights can be assigned to the various factors that make up to the MIXED strategy. The weights are normalized such that the sum of weights equals to 1. The weights are constant parameters that are calibrated via a local search algorithm. Hence, in general, suppose if we have n factors, each having a weight of a_n contributing to the MIXED strategy, then we have the priority of patient k:

$$p_k^{S_3} \leftarrow \sum_n a_n . p_k^{S_n} \tag{2}$$

where $\sum_n a_n = 1$ and $a_n \epsilon [0, 1]$

To find the weights a_n, we propose use of a heuristic multi-dimensional binary search algorithm. We begin the search with a_n set to $\frac{1}{n}$ which gives the centroid of the multi-dimensional search space. In each iteration of the search, we test mid-points of the 2^n neighboring points. We run simulations for current point and those of the neighboring points. We then take the average LOS over all hours and obtain the weights that provide the best average LOS. We continue the search until the algorithm converges to a point such that difference (δ) between the current average LOS and minimum neighboring average LOS is less than a targeted delta δ_g.

V. EXPERIMENTAL EVALUATION

A. *Experimental Setup*

We set up our experiments using 6 months data from the hospital in which we built our motivating case-study. Based the data, we derive the parameters λ_t^d, μ_m, T_i and b. We wrote simulation software using Java™ and ran the simulation over 10 iterations for static FIFO queue and each of the three proposed dynamic priority queue strategies. The result is an average over the iterations. We use two sets of time-varying arrival rates, one for running simulation for Tuesday to Saturday and another set for Sunday and Monday. The latter are the two days of the week when the ED experience higher volume of patients, hence the arrival rates are different. In each set, exponential distributions with mean $\frac{60}{\lambda_t^d}$ are used. In our time-varying arrival rates, the low peak period is between 1am to 8am daily. The peak period is between 9am to midnight. Midnight to 1am and 8am to 9am periods are moderate.

Similarly, we also derived two sets of consultation time represented by exponential distributions for the service by the doctors. One set of consultation time is used for new patient (first consultation) and another for the re-entrants (review consultation). The first consultation time consists of only a single exponential distribution with mean $\frac{60}{\mu_1}$ where μ_1 is the doctor's service rate. The review consultation time consists of a set of 4 exponential distributions with corresponding probability of occurrence. The service time for investigation test or treatment is a set of 4 exponential distributions with corresponding probability of patients' going through the test or treatment. Finally, we set the probability of re-entrance b to 40 percent based on historical data.

For all simulations (except the sensitivity test on number of doctors), we set the number of doctors at consultation to 3 and we allow one day of simulation to pass before we start collecting the results. We then run the simulation over additional 2 days and collect the results for the 2 days.

B. *Experimental Results*

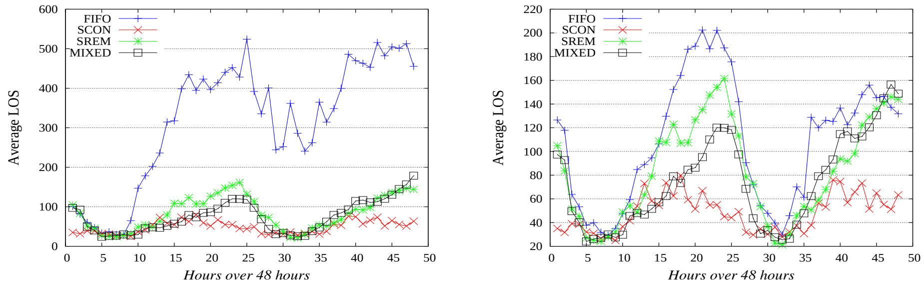
We have our first set of results as shown in Figure 3, showing comparisons between the proposed strategies and FIFO. The number of doctors is 3 and the service rate of the doctors (μ_1) is set to 6 new consultations per hour (i.e., doctors' service time is an exponential distribution with mean $(60/6) = 10$ minutes). This is the current settings in the ED of the hospital we studied. The average service rate information was provided by the hospital.

From Figure 3, we observed the proposed strategies clearly out-performed the FIFO queue, both in terms of lower average LOS of all patients as well as having consistent and almost stable average LOS over peak and non-peak hours (after 9am to midnight). The pattern of SREM and MIXED strategies are similar to FIFO, average LOS increases during the peak hours and decreases during non-peak hours. On days which the ED has higher demand (Sundays and Mondays), the 3 strategies performed far better than FIFO; On days that do not have such high demand, the performance of SREM and MIXED are closer to FIFO. We observe that SCON copes very well with various types of demands in the ED and yield consistent relatively low average LOS with smaller variance compared to other strategies.

Next, we performed what-if analysis by setting the service rate of the doctors (μ_1) to only 5 new consultations per hour (i.e., doctors' service time is an exponential distribution with mean $(60/5) = 12$ minutes). This is to test what happens if the doctors take more time to serve a patient. We show in Figure 4 the results from this analysis. On Sundays and Mondays as shown in Figure 4(a), the queue becomes unstable and the average LOS became increasingly high using FIFO, SREM or MIXED strategies. Unstable queue appears when the service is slower than arrival (which is consistent with standard FIFO queuing theory). In contrast, although the average LOS for SCON is also high at peak hours (reaching 400 minutes at some points), it is interesting to see that it remains fairly stable under such condition. We draw conclusion here that SCON strategy scales better under high load conditions. SREM and MIXED strategies still perform better than FIFO.

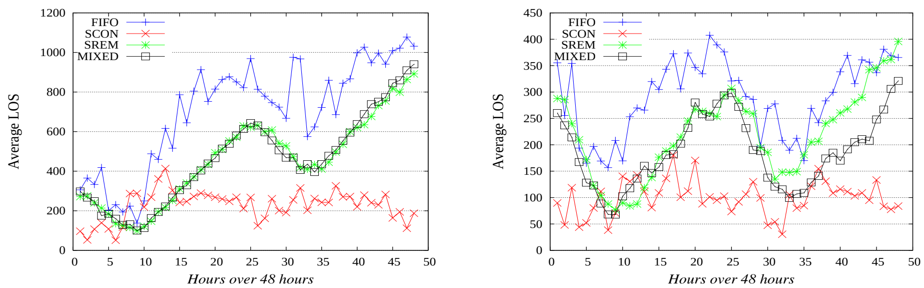
Another what-if analysis we have conducted is when the service rate of the doctors (μ_1) is set to 7.5 new consultations per hour (i.e., doctors' service time is an exponential distribution with mean $(60/7.5) = 8$ minutes). We test the situation when the hospital increases the efficiency of the doctors to serve each patient faster. We show in Figure 5 results of our investigation. An interesting observation is in Figure 5(b) (results of simulation for Tuesday to Saturday). We see that when the ED is not heavily loaded (i.e., during off-peak hours where there are sufficient resources such as doctors), the various strategies including FIFO, perform similarly. The proposed strategies have less advantage over FIFO in low peak situations.

Since a hospital must assure quality of care, increasing service rate of doctors may not always be feasible (doctors need to thoroughly examine a patient). We examine the



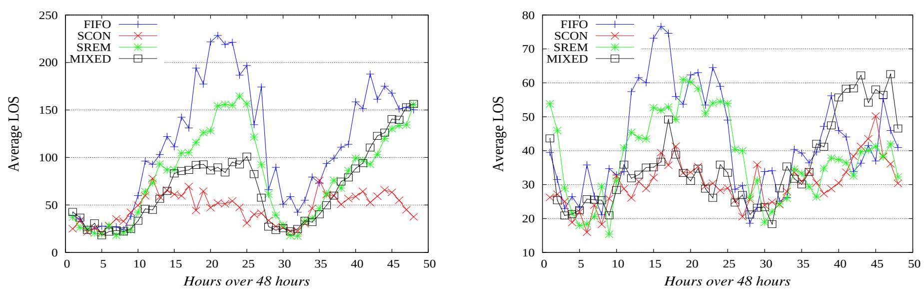
(a) Sunday and Monday (b) Tuesday to Saturday

Fig. 3. Comparison of proposed strategies against FIFO for 3 Doctors with $\mu_1 = 6$



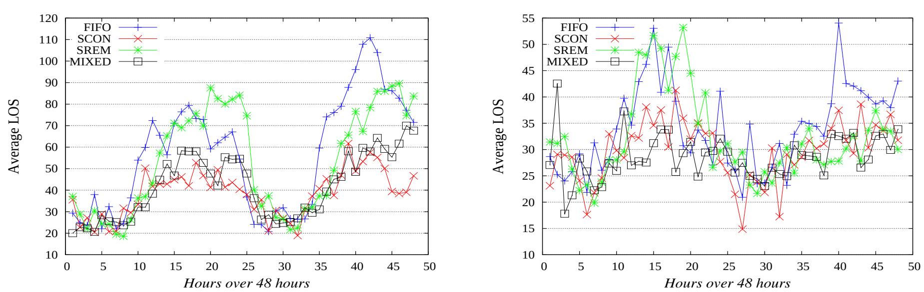
(a) Sunday and Monday (b) Tuesday to Saturday

Fig. 4. Comparison of proposed strategies against FIFO for 3 Doctors with $\mu_1 = 5$



(a) Sunday and Monday (b) Tuesday to Saturday

Fig. 5. Comparison of proposed strategies against FIFO for 3 Doctors with $\mu_1 = 7.5$



(a) Sunday and Monday (b) Tuesday to Saturday

Fig. 6. Comparison of proposed strategies against FIFO for 4 Doctors with $\mu_1 = 6$

situation that the hospital increases the number of doctors to 4 and the result is in Figure 6. Figure 6(b) further verifies our observation that if there is no contention for resources (doctors) on non-peak days like Tuesdays to Saturdays, any strategy works rather similarly. Our intelligent dispatch strategies become useful with high volume of patients which compete for expensive resources such as doctors.

Across all the various sensitivity tests on doctor's service rate and number of doctors, we found that the SCON strategy yields a very consistent and optimistic result, regardless of whether it is the peak hours. From the management perspective, this strategy not only has the best performance but also has the ability to provide the hospital a confidence to ensure the public that a certain service level can be achieved at its ED. E.g. hospital could convey to the public that x% of patients can be served between 20 to 80 minutes on Tuesday to Saturday based on doctors' service rate of 6 patients per hour.

	Normal Treatment						Long Treatment					
	3 Doctors, μ=7.5			4 Doctors, μ=6			3 Doctors, μ=7.5			4 Doctors, μ=6		
	SCON	SREM	MIXED	SCON	SREM	MIXED	SCON	SREM	MIXED	SCON	SREM	MIXED
120 < w < 150	1	0	5	0	0	0	6	84	106	3	14	0
150 < w < 180	0	0	0	0	0	0	3	63	10	0	0	0
180 < w < 210	3	0	0	2	0	0	4	35	55	1	1	0
> 210	8	0	0	1	0	0	27	66	16	5	0	0

Fig. 7. Comparison of starvation phenomenon among the 3 strategies

Despite having great performance, we suspect that SCON has a potential limitation of having starved patients (patients who repeatedly get preempted due to high estimated consultation time). We present the results in Figure 7 to show this phenomenon. The symbol w in the table indicates the length of time (in minutes) that patients have waited in the queue for a doctor at the end of our simulation runs. The number in each column shows the total number of such patients over the 10 iterations. We run the experiments over two types of treatment duration distributions, *normal treatment* and *long treatment*. The duration distributions for normal treatment comprises of 90% of re-entrants who has less than or equal to the estimated consultation time of a new patient. The duration distributions of long treatment only comprises of 70% of such re-entrants. From our table, we found that even under non-peak conditions ("3 doctors $\mu_1 = 7.5$ normal treatment", "4 doctors $\mu_1 = 6$ normal treatment" and "4 doctors $\mu_1 = 6$ long treatment"), the SCON has cases of starvation, while SREM and MIXED are cleared of starvation. When the ED is slightly crowded ("3 doctors $\mu_1 = 7.5$ long treatment"), the MIXED strategy still out-performs SCON in terms of starvation cases. SREM did not perform well because the strategy does not perform well on its own in the first place. In this starvation study, we observe a trade-off between the performance and the risk of having starved patients.

VI. CONCLUSIONS AND FUTURE WORKS

In this paper, we proposed several strategies for dynamic priority queuing of patients at an ED. We found that the proposed strategies result in shorter average LOS compared to using FIFO method and they are able to scale well over peak days and peak hours. Introduction of the SREM strategy also helps the hospital to meet the desired service level. With the advancement of IT that enables patient real-time movements to be tracked and data to be mined quickly, we believe that our proposed concept of dynamic priority queue is implementable. Although SCON has the best performance, the challenge with implementing SCON is the accuracy of estimating each patient's consultation time. Further studies such as mining of patients' data and motion-study are required to be able to estimate the consultation time. The SREM strategy, on the other hand, is readily implementable as information (arrival time and treatment time) for calculating the priorities of patients is easily available. The performance of SREM, however, is not as promising as SCON. In this aspect, we observed a trade-off between performance and implementation readiness.

In summary, the SCON strategy yields the best performance with lower variance over the average LOS for all the scenarios, but the drawback is that the strategy requires additional efforts to implement and faces the possibility of starvation. The SREM strategy is readily implementable but its performance is less promising. We see the potential of the MIXED strategy because it addresses the problem of starvation (faced by SCON) and gives better results compared to SREM. More importantly, the MIXED strategy reduces the reliance on the accuracy of estimating the consultation time of the patients; it also has the flexibility of incorporating other important contributing factors (e.g., patient's illness type - patients with a fish bone in the throat may be more uncomfortable than another patient with cold and fever) to setting patient's priority.

As part of future work, it would be interesting to explore hybrid strategies such as using SCON strategy for few hours in the day and the MIXED strategy for the next few hours to reap benefits of both strategies. We conclude that despite the various pros and cons of the proposed strategies, a dynamic priority queue has many advantages over the standard FIFO queue to improve the patient flow in an emergency department.

REFERENCES

[1] L. Mayhew and D. Smith, Using queuing theory to analyse the Government's 4-H completion time target in Accident and Emergency departments, *Health Care Management Science*, Volume 11, No. 1, 2008, pp. 11-21.

[2] Alexander Komashie and Ali Mousavi, Modeling Emergency Departments Using Discrete Event Simulation Techniques. *Proceedings of the 2005 Winter Simulation Conference*, 2005.

[3] Pajouh, FoadMahdavi and Kamath, Manjunath, Applications of Queueing Models in Hospitals. *Midwest Association for Information Systems (MWAIS)*, 2010, Paper 23, http://aisel.aisnet.org/mwais2010/23.

[4] Samaha, S.; Armel, W.S.; Starks, D.W. The use of simulation to reduce the length of stay in an emergency department. *Proceedings of the 2003 Winter Simulation Conference*, vol.2, no. 7-10 Dec. 2003, pp. 1907- 1911.

[5] Murat M. Gunal and Michael Pidd. Understanding accident and emergency department performance using simulation. *Proceedings of the 38th conference on Winter simulation*, 2006, pp. 446-452.

[6] Andrew M. Hay, Edwin C. Valentin, and Rienk A. Bijlsma. Modeling emergency care in hospitals: a paradox - the patient should not drive the process. *Proceedings of the 38th conference on Winter simulation conference*, 2006, pp. 439-445.

[7] King, D. L., Ben-Tovim, D. I. and Bassham, J. Redesigning emergency department patient flows: Application of Lean Thinking to health care. *Emergency Medicine Australasia*, 18(4), August 2006, pp. 391-397.

[8] S. Chakravarthy. A finite capacity dynamic priority queuing model. 1992. *Computer Industry Engineering*. Vol 22, 4, October 1992, pp. 369-385.