2013-6

Towards Operations Excellence: Optimising Staff Scheduling For New Emergency Department

Amr Arisha
Technological University Dublin, amr.arisha@tudublin.ie

Waleed Abo Hamad
Technological University Dublin, waleed.abohamad@tudublin.ie

Towards operations excellence: Optimising staff scheduling for new emergency department

Waleed Abo-Hamad (wabohamad@dit.ie)
Amr Arisha
3S Group, College of Business
Dublin Institute of Technology
Dublin 2, Ireland

Abstract
This paper presents a case study of an Emergency Department of a public hospital in Dublin, and uses an integrated approach to determine optimal staffing levels to meet the challenges of its dynamic patient demand levels. A comprehensive stochastic model is developed to incorporate patients care pathways and the resources required along their treatment journeys. Analytical Hierarchical Process is utilised to enable decision makers to set their preferences for the facility's strategic objectives. Evolutionary algorithms are applied to optimise staff schedules. The resulted optimized schedules maintains continuity of care delivery for patients while ensuring a balanced equilibrium among available staff.

Keywords: Decision support systems, Artificial intelligence, Staff Scheduling Optimisation

Introduction
Overcrowding in hospital emergency departments (EDs) can be described as an international crisis that negatively affects patient safety, the quality of their care and their satisfaction (Hwang *et al.,* 2011). The problem was declared a 'National Emergency' in Ireland in 2006,with more than 500 patients waiting on trolleys every day for admission to Irish hospitals. according to the latest report of Health Service Executive (HSE), 40% of patients waiting between 10 and 24 hours and 18% waiting over 24 hours (HSE Performance Monitoring Report, 2010). Consequences of this situation on patients, staff and the healthcare sector across the State as a whole. Improved staff scheduling is commonly proposed as a solution that enables enable managers to increase capacity utilization, minimise costs and improve the tactical and operational efficiencies of services within such facilities (Rocha, Oliveira and Carravilla, 2012).

Staff scheduling can be defined as assigning staff with different skill sets to different shifts to guarantee operational cover while still satisfying as many soft constraints and personal preferences as possible (Brucker and Burke, 2011). Given the dynamic nature of typical ED environments, scheduling ED staff is a very challenging task, so applications supporting operational decision-making are widespread and have become increasingly crucial (Eldabi *et al*., 2006, Katsaliaki and Mustafee, 2010). Simulation models can be effective as tools that can take into account the uncertainty of patient arrival patterns and predict the maximum demand levels that the ED staff are likely to have to handle (Fletcher *et al*,. 2006*)*, and determine the staffing levels required to meet those demands and still keep patients' average waiting times below certain thresholds (Abo-Hamad *et al.,* 2011). Simulation models cannot determine the optimum values of decision variables in terms of predefined objective function(s), hence optimisation models are required to be integrated with simulations to provide the best possible solutions (Abo-Hamad and Arisha, 2011). ED service quality can be improved by utilizing simulation together with a genetic algorithm (GA) to adjust staff schedules appropriately and avoid hiring additional staff (Yeh and Lin, 2007). Staff members can be assigned to different duties dynamically according to a variety of constraints, such as working patterns, staff qualifications and preferences, as well as costs (Gutjahr and Rauner, 2007). However, very little of the literature has examined the nature of the trade-offs and inter-dependencies between performance measures when evaluating the impact of the use of such tools on staff schedules (Neely *et al.*, 2000).

This study integrates simulation modelling and meta-heuristic algorithms to find optimal staff schedules that improve system performance, while incorporating an analytical hierarchical process (AHP) to assess the trade-offs between different system performance measures.

Proposed integrated framework

Figure 1 shows an overview of the proposed framework.

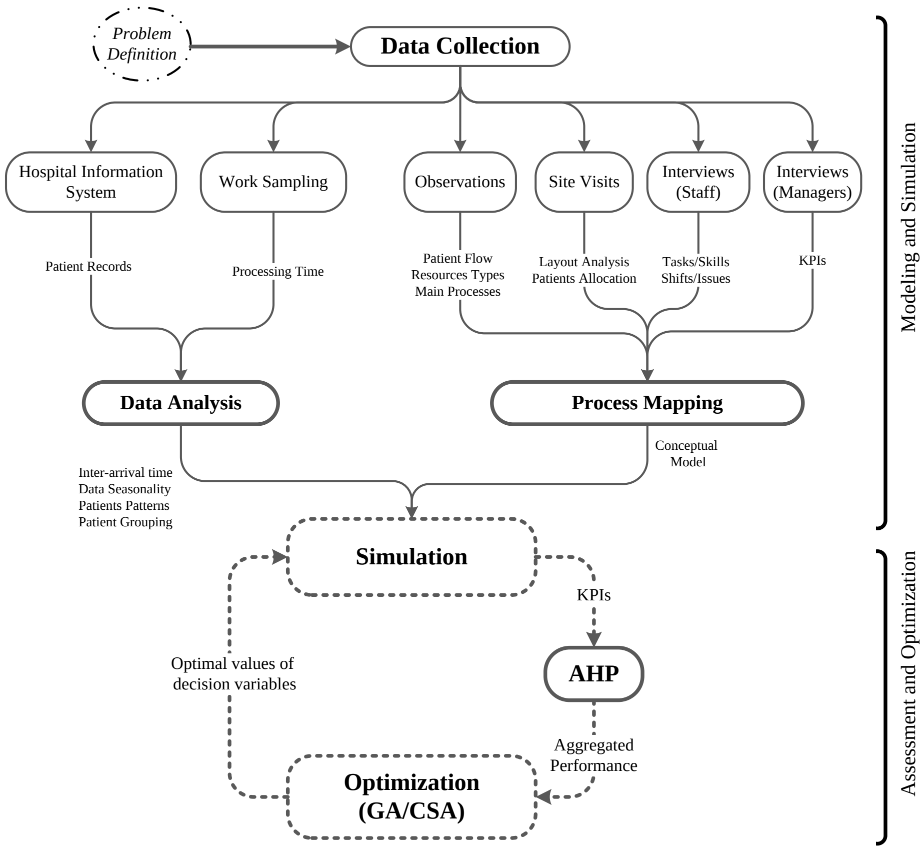

Figure 1 An optimisation-based decision support framework

Modelling and simulation

Conceptual model development begins with a data collection phase which incorporates information gathered through mapping the underlying processes. Patient records from the hospital information system (HIS) provide the required data about patients' arrival times, and care paths and other information required for the simulation model.

Assessment and optimisation

Once the simulation model is verified and validated, the decision makers can use it to investigate a number of alternative decisions to foresee their consequences. However, the number of key performance indicators (KPIs) used as criteria can affect the analysis and evaluation of the simulation results, and, if different objectives conflict with one another, an analytical hierarchical process (AHP) (Saaty, 1990) will be needed to analyse the trade-offs between them. AHP was also used to aggregate the marginal performance of the KPIs considering decision makers' weighted preferences about achieving different strategic objectives.

Analytical hierarchical process

Analytical Hierarchical Process (AHP) is based on paired comparisons and uses ratio scales to represent judgments about preferences illustrated in a comparison matrix. The decision maker expresses their preferences as ratios by weighting the main performance criteria in the form:

$$r_{ij} = \frac{\mathrm{w_i}}{\mathrm{w_j}}$$

where r_{ij} is the ratio between the weightings of criterion i (w_i) and criterion j (w_j). The elements on the diagonal of the comparison matrix are 1, and, where

$$r_{ij} = \frac{1}{r_{ji}}$$

The ratio scale of weights ranges from 1 (equally important) to 9 (where one of a pair of indicators is judged extremely more important than the other). The weights are then by normalising the elements of the eigenvector corresponding to the largest Eigen value of the comparison matrix, and the marginal performance of each alternative scenario is calculated using the results of the preference model aggregated as:

$$\mathrm{V}(x) = \sum_{i=1}^{N} \mathrm{w_i} v_i(x_\mathrm{i})$$

where w_i, $i \in (1,2,..., N)$ corresponds to the relative weight of the KPI_i and $v(x_i)$ is the desirability value of x_i for the corresponding KPI. The marginal performance then represents the desirability level of the performance given the current values of the decision variables. This is then considered as the objective function in the optimisation process, where their initial values are used in the optimisation model with the aim of generating new values towards improving this objective function.

The optimisation model

Due to its combinatorial nature, scheduling problem is a challenge for any local search algorithm. Therefore meta-heuristics optimisation techniques were used for such problems. Genetic algorithm (GA) (Goldberg and Holland 1988) and Clonal Selection Algorithm (CSA) (Abo-hamad *et al*., 2010). A hybrid between GA and CSA were used in medical application (Korayem *et al*., 2010). This approach shows better results in terms of the quality of solutions and conversion rate. Optimal staff scheduling is a new application for this hybrid approach (Figure 2).

Development and implementation

The hospital studied in this case is large public hospital (570-bed) with a 24hr Emergency Department (ED) which services over 55,000 patients annually.

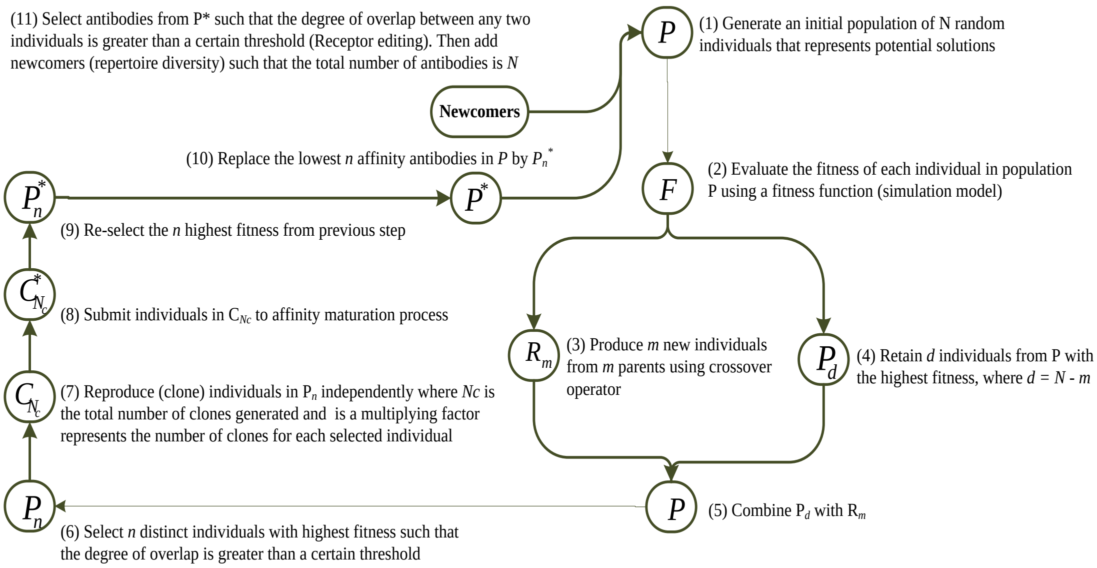

Figure 2 A description of a hybrid Genetic Algorithm and Clonal Selection Algorithm

The hospital is unable to comply with patient waiting time targets (which specify a 6 hour maximum). Average time from registration to discharge is 9.16 hrs with 2.58 hrs standard deviation, i.e. 3.16 hrs over the HSE metric (0-6 hrs). Besides, the average time from registration to acute admission is 21.3 hrs with a standard deviation of 17.2 hrs, which is 15.3 hrs above the national metric. ED figures show clear evidence of its overcrowding - on average 17 % of its patients leave without even being seen. The ED has three consultants (who provide cover between 9am-5pm (or 8am-8pm) with 24/7 on-call provision), two nursing managers, and three grades of physician - registrars/specialist registrars; Senior House Officers (SHOs), and interns - distributed as follows (when the roster allows): three registrars per day working 10hr shifts starting at 8am, 12pm, and 10pm, twelve SHOs working fixed shifts during the day and night to keep the ED running and two interns working 8am to 5pm day shifts Monday to Friday, so that the numbers of doctors on duty varies between 2 and 7 depending on the time of day or night. In addition, eleven nurses are scheduled during the day and nine at night.

Modelling and Simulation

When walk-in patients (self- or GP-referred) arrive and register at the ED, they usually stay in the waiting area until they are called to be triaged, which (depending on triage staff availability) generally means being assessed by a triage nurse. Patients then follow the care path as required. Hospital managers provided the research team with a total of 59,986 anonymous patient historical records from the ED information system over a 16-month period. Patients were then grouped according to their triage category. Table 1 summarises patient information for each of the five triage categories, as well as their arrival mode.

Table 1 Summary of the analysis of patients' records

Triage Category	% of Patients	Arrival Mode	
		Walk-in	Ambulance
IMM	01.1 %	05 %	95 %
VURG	16.5 %	40 %	60 %
URG	58.0 %	61 %	39 %
STD	23.9 %	81 %	19 %
NURG	00.5 %	72 %	28 %

IMM: *Immediate* ***VURG****: Very Urgent* ***URG****: Urgent* ***STD****: Standard* ***NURG:*** *Non-Urgent*

Analytical hierarchical process for the emergency department

The research team made repeated visits to the ED and interviewing its senior management team, and (working with the ED manager) identified four main performance measures: layout efficiency (LE), patient throughput (PT), productivity (PR), and resource utilization (RU). The layout efficiency measures the average daily distances travelled by doctors (Avg. Doctor Distance) and nurses (Avg. Nurse Distance), while patient throughput is measured via three indicators: average waiting time to first clinical contact (Avg. Doctor WT), average length of stay (LOS) times for discharged patients and for admitted patients (Avg LOS Dis. and Avg LOS Ad.). ED productivity is measured in terms of three indicators: the ratio of patients per doctor (Patient/Doctor Ratio), the ratio of patients per nurse (Patient/Nurse Ratio), and the percentage of patients who present at the Department who are actually treated by its staff (% Patients Treated), while Resource utilization is measured for two types of resources: ED staff (i.e., doctors and nurses) and ED assets - trolleys, ambulatory care units (ACUs), and resuscitation rooms (CPRs). A comparison matrix for each of these pairs of criteria was then constructed to obtain the weights of individual KPIs. A number of pair comparisons between KPI's for each main criterion was repeated until the last level was reached. Figure 4 shows the final weights for all the levels as a performance 'value tree'.

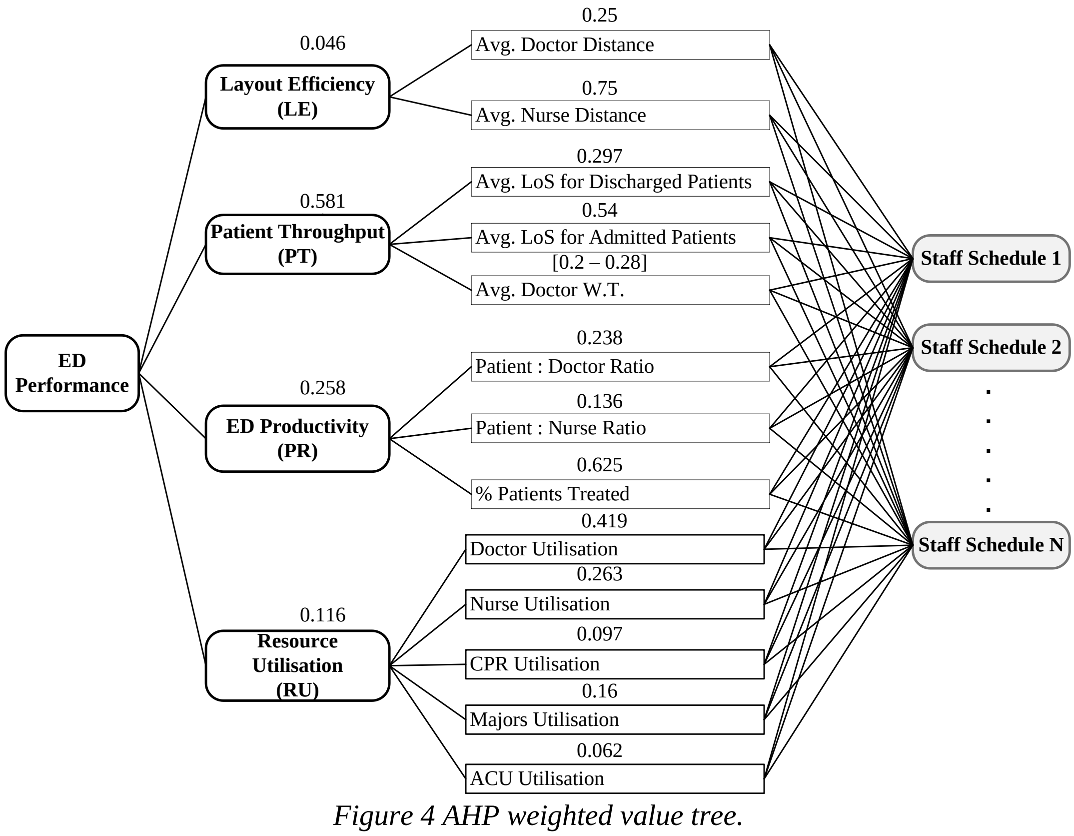

Figure 4 AHP weighted value tree.

The ED director then assigned acceptable ranges for each KPI. For example, the staff utilization had a range between 50% and 85%, the lower level to avoid resource under-utilization, and the upper to avoid staff burnout. Similarly, he specified range of between 0 and 6 for the LOS KPI's for both admitted and discharged patients to measure the achievement level in each scenario considering the 6-hour maximum LOS HSE target. After the acceptable ranges had been assigned, a *value function* was designated to describe the *desirability* level for each individual KPI.

Emergency department staff optimisation
The ED's scheduling problem is to arrange weekly schedules involving up to 50 SHOs to meet the variable demand of patient arrivals, assigning each doctor a shift pattern that aligns with their work contracts. Patient demand has a particular day-evening-night pattern, but most SHOs' working contracts mean they work either days/evenings or nights in any one week, but not both, and also the numbers of days worked are not usually the same as the numbers of nights. At the ED manager's suggestion, we based the model's examination of re-scheduling possibilities on the number of SHOs currently employed at the ED, and taking into account work-shifts that were feasible in their current contracts. The aim was to keep the ED running with its current staff of only twelve SHOs, and a roster is used to rotate the remaining SHOs' work-stretches. A work stretch is the set of consecutive shifts doctors work between having at least two days off, and a shift the period within a working day during which a doctor is assigned to ED duties. Table 2 shows range of feasible 10-hour work-shifts ED managers currently assign to doctors - 4 day shifts, 2 evening shifts and a night shift - with their start and end times.

Table 2 Feasible work-shifts in the emergency department.

Work-shift	Time	Shift Name
Day shifts	06 - 16:00	D1
	08 - 18:00	D2
	10 - 20:00	D3
	11 - 21:00	D4
Evening shifts	14 - 00:00	E1
	16 - 02:00	E2
Night Shift	22 - 08:00	N

Solution representation and encoding
To illustrate and address the contractual rules, constraints and assumptions noted earlier in our schedule optimisation procedure, we produced a binary representation of the roster for staff work-shifts (Table 3).

Table 3 Binary representation of feasible work shifts.

	Day shift				Evening shift		Night Shift
	06 - 16:00	08 - 18:00	10 - 20:00	11 - 21:00	14 - 00:00	16 - 2:00	22 - 8:00
D1	1	0	0	0	0	0	0
D2	0	1	0	0	0	0	0
D3	0	0	1	0	0	0	0
D4	0	0	0	1	0	0	0
E1	0	0	0	0	1	0	0
E2	0	0	0	0	0	1	0
N	0	0	0	0	0	0	1
OFF	0	0	0	0	0	0	0

A doctor's work stretch is represented in terms of these work shifts as a binary vector W_{7x1} over the whole week - e.g., D1, D1, E2, N, N, OFF, OFF, and Table 6 replaces each of these consecutive work shifts by its binary representation. Thus, the full doctor roster is a Vector R_{12x7}, for the 12 SHOs in the ED. After encoding the scheduling problem in this way, the hybrid GA/CSA is then used to find the optimal doctors' roster – that which allocates work stretches and work shifts for the ED's full complement of SHOs in a way that maximises the department's performance.

The optimisation process

Randomly fixed-length binary strings for N solutions were first generated to build up the initial population of solutions. According to (Haupt and Haupt, 1998), the number of initial solutions in our case is $N = 84$, as a multiple of the binary string length, which represents each solution. Then, the simulation model combined with the AHP is used to calculate the fitness of each one. Solutions from the current population are then selected for a crossover process according to their fitness. The better the solutions the more chances to be selected. Crossover greatly accelerates the search process early in evolution of a population, and guides it towards promising regions of the search space. To avoid losing best founded solutions, a few best solutions are copied into the new population, and the rest replaced by the offspring solutions resulting from the crossover process. The Clonal selection principle is then applied on the resultant new population. The concentration of high fitness solutions is increased by a process known as Cloning. Solutions with the highest fitness were selected to be cloned independently, and the reproduced solutions are then mutated with a rate inversely proportional to their fitness, allowing us to explore local areas around each specific solution by making small steps towards a solution with even higher fitness. The fitness of these mutated cloned solutions is then calculated, and the best fitted are then inserted into the new population in place of those with lowest fitness. Retaining multiple suitable solutions is desirable, as many can have high fitness levels: this is accomplished by first creating a pool of distinct solutions and then adding entirely new solutions to this pool in place of the least well-fitted, thus allowing the model to ‘escape’ from unsatisfactory local optima.

Analysis of Results

Steady state analysis of simulation results

Simulation variables – patients’ inter-arrival times, arrival modes, medical complaints presented, processing time, routing and triage category allocation, etc – were initialised based on the analyses of empirical data and of the ED layout and patient flows given in previous sections. Queues at each stage of patient care were set as empty and idle, and a two month ‘warm-up’ period used to mitigate against any bias introduced by the simulation model’s initial conditions. To validate the simulation model results, only three KPIs were used - average waiting times to see a doctor and average lengths of stay for discharged admitted patients, with their actual values calculated from the patient records provided. The simulation model was run for 10 independent replications to obtain independent and identically distributed of ED's KPIs, with each run being re-initialised by a different random number seed. Table 4 shows the average simulation outputs for the 3 KPIs for all 10 replications.

Table 4 ED simulation output

	Avg. Doctor WT (mins)	Avg. LOS (hrs)	
		Discharged Patients	Admitted Patients
μ_{av}	189.02	9.16	21.98

Avg.W.T: Ave Waiting Time to see Doctor **Avg. LOS:** Ave Length of Stay **μ_{av}:** Actual Value Mean

Staff scheduling optimisation analysis

The main rationale for employing the optimisation process was to try to generate an optimal (or near-optimal) schedule for the ED's SHO staff that could improve its performance and so minimise average patient LOS times. The final output of the optimisation procedure is the near-optimal SHO schedule for the ED, as shown in Table 5, which details the optimal weekly work stretches and the total numbers of physician hours the schedule provides each day of the week.

Table 5 Optimal weekly work stretches for SHO staff members.

Work stretch no.	**M**	**T**	**W**	**T**	**F**	**S**	**S**
1	OFF	D2	D2	D2	D2	OFF	
2	E1	E1	E1	E1	E1	OFF	
3	N	N	N	N	N	N	N
4	OFF						
6	D4	D4	OFF		D4	D4	D4
7	E2	E2	E2	E2	OFF		E2
8	E2	E2	E2	X	X	E2	E2
9	OFF						
10	D1	D1	D1	OFF		D1	D1
11	D2	D2	D2	D2	D2	OFF	
12	N	N	N	N	N	N	X
Daily staffing level (hrs)	90	100	90	70	70	50	50

Using the ED simulation model has resulted in obtaining working stretches for SHO doctors that match the demand (i.e., patient arrivals) which is the highest during weekdays, and at its lowest levels over weekends (Figure 5).

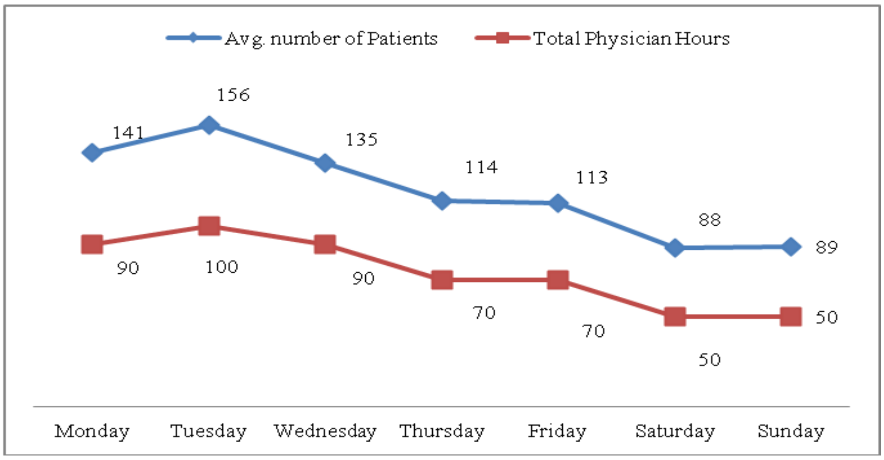

Figure 5 The ED optimal staffing levels matching the weekly patient arrival rate.

This modelled schedule would reduce average patient waiting times by 57%, with nearly 92% of treated patients converging on the HSE 6-hours target, as Table 6 shows.

Table 6 Simulation results of the optimal staff schedule vs. baseline scenario

KPI's	**Base Line**	**Optimal schedule**	**↓↑**
Avg. Doctor W.T. (mins)	177.43	75.68	-57%
Avg. LOS Discharged Patients (hrs)	8.95	7.13	-20%
% Patients Treated	83%	92%	11%

To statistically compare the optimal schedule's performance with the current (i.e., base line) ED performance, a confidence interval was constructed for the difference between μ_1 and μ_2 with an overall confidence level of $1 - \alpha$, where $\alpha = 5\%$. After two-tailed t-test computations, the results showed that there are significant differences between the optimal and the current schedules, with the former showing the potential for yielding significant improvements in the quality of care at the ED. The optimal work shifts account for the time-varying characteristics of the daily patient arrival patterns by allocating optimal staffing levels at the ED over the 24-hours period. For example, Figure 6 shows the overlapping between staff working shifts on Tuesday, which is one of the busiest days with high patient arrival rate at the ED.

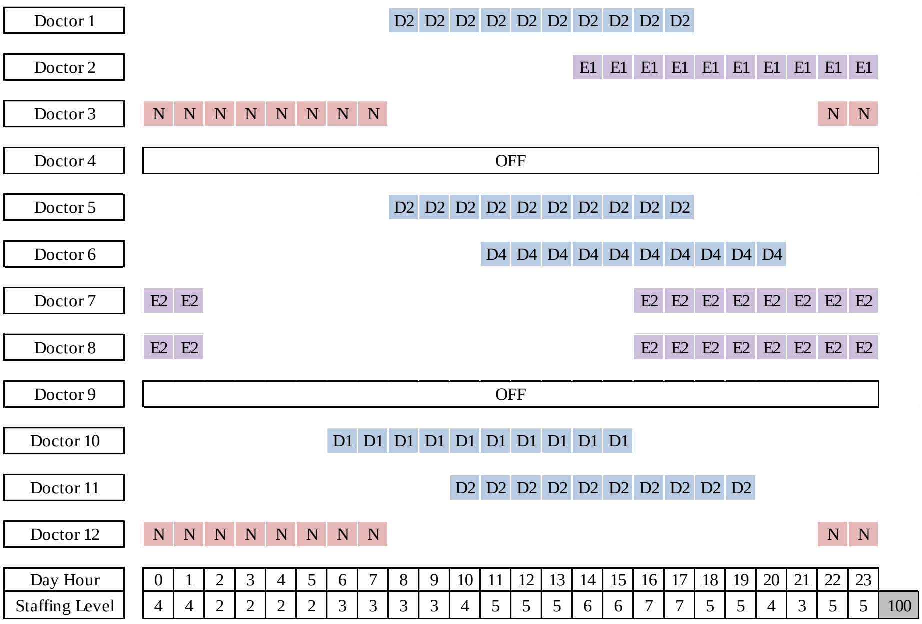

Figure 6 Overlapping staff work shifts to cope with daily demand fluctuation.

The currently-used schedule leads to staff shortages during peak times (between 14:00 and 18:00), so contributing to overcrowding. However, the optimal schedule effectively overcomes this problem by scheduling dynamically overlapping staff working shifts to meet such demand fluctuations, providing 6 to 7 doctors during this peak demand period. At the same time, the optimal schedule arranges the overlapping shifts to avoid under-utilising staff by reducing staff levels to adapt to the slowly lessening patient arrivals rates, which reach their lowest levels during the night time (for which the schedule provides just 2 or 3 doctors).

Conclusion

The challenges of managing healthcare facilities are increasing are significantly in line with the pressures from economic downturns. Timely access to care, prompt responses to patient needs, and the availability of adequate resources to deliver quality service are the key priorities of healthcare systems, particularly hospitals. To meet these challenges, this study applies simulation to the internal processes of a hospital Emergency Department to evaluate the effect of various physician schedules on key performance measures. The framework integrates AHP to incorporate decision makers' preferences in evaluating different possible schedules, and then applies a hybrid genetic algorithm/artificial immune system a near-optimal schedule for the work patterns of the department's physicians.

The proposed framework intelligently produces optimal staffing patterns that match the available human resources to the fluctuating patient demand for services. Optimal staffing levels can significantly contribute to the efficacy of healthcare managers' decision making process. The proposed model allows management to allocate its physician staff resources more accurately to the proper patients at the proper peak times - despite the variability in patient arrival rates, the optimised staff schedules reduced average patient waiting times by up to 57%, and yielded a significant increase in productivity (92% treated patients within the HSE time benchmarks). These results show that the quality of ED care can be improved by dynamically configuring physicians' schedules without the ED having to recruit additional doctors.

References

Abo-Hamad, W. and Arisha, A. (2011), "Simulation optimisation methods in supply chain applications: a review", *Irish Journal of Management*, Vol. 30, No. 2, pp. 95-124.

Abo-Hamad W., McInerney, J. and Arisha, A. (2011), "Virtual ED: utilisation of a discrete event simulation-based framework in identifying 'real-time' strategies to improve patient experience times in an emergency department", *Emergency Medicine Journal*, Vol. 28, pp. A3-A4.

Abo-hamad, W., Korayem, M. and Moustafa, K. (2010), "Biclustering of DNA microarray data using artificial immune system", *Proceedings of International Conference on Intelligent Systems Design and Applications (ISDA'10),* 2010 Cairo, Egypt. IEEE Computer Society.

Bond, K., Ospina, M. B., Blitz, S., Afilalo, M., Campbell, S. G., Bullard, M., Innes, G., Holroyd, B., Curry, G., Schull, M. and Rowe, B. H. (2007), "Frequency, determinants, and impact of overcrowding in emergency departments in Canada: a national survey of emergency department directors". *Healthcare Quality,* Vol. 10, No. 4, pp. 32-40.

Brucker, P., Qu, R. and Burke, E. (2011), "Personnel scheduling: models and complexity", *European Journal of Operational Research,* Vol. 210, pp. 467-473.

Eldabi, T., Paul, R. and Young, T. (2006), "Simulation modelling in healthcare: reviewing legacies and investigating futures", *Journal of the Operational Research Society,* Vol. 58, pp. 262-270.

Fletcher, A., Halsall, D., Huxham, S. and Worthington, D. (2006), "The DH accident and emergency department model: a national generic model used locally", *Journal of the Operational Research Society*, Vol. 58, pp. 1554-1562.

Goldberg, D. E. and Holland, J. H. (1988), "Genetic algorithms and machine learning", *Machine Learning*, Vol. 3, pp. 95-99.

Gutjahr, W. J. and Rauner, M. S. (2007), "An ACO algorithm for a dynamic regional nurse-scheduling problem in Austria", *Computers and Operations Research*, Vol. 34, pp. 642-666.

Haupt, R. L. and Haupt, S. E. (1998). *Practical genetic algorithms*. Wiley Interscience: New York.

HSE (2010), "Health service executive emergency departments - patient profiles, experiences and perceptions", http://www.hse.ie, accessed 3 September 2011.

Hwang, U., Mccarthy, M. L., Aronsky, D., Asplin, B., Crane, P. W., Craven, C. K., Epstein, S. K., Fee, C., Handel, D. A. and Pines, J. M. (2011), "Measures of crowding in the emergency department: a systematic review", *Academic Emergency Medicine,* Vol. 18, pp. 527-538.

Katsaliaki, K. and Mustafee, N. (2010), " Applications of simulation within the healthcare context", *Journal of the Operational Research Society,* Vol. 62, No. 8, pp. 1431-1451.

Korayem, M., Abo-hamad, W. and Moustafa, K. (2010), "A hybrid genetic algorithm and artificial immune system for informative gene selection", *International Journal of Computer Science and Network Security (IJCSNS),* Vol. 10, pp. 76-83.

Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M. and Kennerley, M. (2000), "Performance measurement system design: developing and testing a process-based approach", *International Journal of Operations & Production Management,* Vol. 20, pp. 1119-1145.

Rocha, M., Oliveira, J. F. and Carravilla, M. A. (2012), " Quantitative Approaches on Staff Scheduling and Rostering in Hospitality Management: An Overview", *American Journal of Operations Research,* Vol. 2, pp. 137-145.

Saaty, T. L. (1990), "How to make a decision: the analytic hierarchy process", *European Journal of Operational Research*, Vol. 48, pp. 9-26.

Yeh, J. Y. and Lin, W. S. (2007), "Using simulation technique and genetic algorithm to improve the quality care of a hospital emergency department", *Expert Systems with Applications,* Vol. 32, pp. 1073-1083.