Optimal Allocation of Cashiers and Pharmacists in Large Hospitals: A Point-wise, Fluid-based Dynamic Queueing Network Approach

Chung-Cheng Lu †
Department of Transportation and Logistics Management
National Chiao Tung University, Hsinchu, Taiwan
Tel: (+886) 2-23494963, Email: jasoncclu@gmail.com

Hui-Ju Chen
Department of Industrial Engineering and Management
National Taipei University of Technology (Taipei Tech), Taipei, Taiwan
Tel: (+886) 2- 2771-2171 ext. 2340, Email: vanillascody@gmail.com

Kuo-Ching Ying
Department of Industrial Engineering and Management
National Taipei University of Technology (Taipei Tech), Taipei, Taiwan
Tel: (+886) 2-2771-2171 ext. 2308, Email: kcying@ntut.edu.tw

Tsung-Chi Yu
Graduate Institute of Information Management
National Taipei University of Technology (Taipei Tech), Taipei, Taiwan
Tel: (+886) 2- 2771-2171 ext. 2340, Email: decavi400@hotmail.com

Abstract. This paper presents a bi-objective optimization model for determining the optimal time-varying numbers of cashiers and pharmacists in a large hospital. The two objectives of this model are to minimize the waiting cost incurred by patients and the operating costs incurred by the hospital. A point-wise fluid-based approximation approach is adopted to construct a dynamic queueing network that takes into account time-varying (or non-stationary) arrivals of patients and describes time-varying queue lengths. The dynamic queueing network is then encapsulated in the optimization model that determines optimal time-varying numbers of cashiers and pharmacists. A test problem instance is designed based on a large hospital in the city of Taipei, and the MINOS solver of GAMS is applied to solve the problem instance. Numerical results show that the optimization model can provide an optimal allocation of manpower that significantly reduces both waiting and operating costs.

Keywords: Queueing theory, Mathematical programming, Healthcare management, Manpower allocation

1. INTRODUCTION

In Taiwan, after the National Health Insurance system was implemented in 1995, hospitals sought to reduce operational costs and increase their revenues. Servicing a larger number of patients is one way to increase such revenues. Moreover, increasing satisfaction among patients becomes an important consideration for hospital managers. Most hospitals adopt various indices to evaluate their service quality, one of them being patient waiting time, which greatly affects patient satisfaction. Apart from outpatients' waiting times for diagnoses from doctors, the waiting times or queue lengths at cashier and pharmacy counters have long been a major concern for hospital managers and their patients, especially during rush hour when extended wait times generate numerous complaints. Although increasing the number of cashiers and pharmacists is a straightforward way of reducing patients' wait times, this approach inevitably raises hospitals' operating costs and may be inefficient during off-peak hours. Instead, allocating appropriate numbers of cashiers and pharmacists at different time periods during a given day to strike a balance between patient wait times and hospitals' operating costs is essential for hospital managers to maximize their revenues, while maintaining the level of service desired. Most hospital managers allocate manpower at cashier counters and dispensaries during peak and off-peak hours according to their past experiences. However, this manual approach may not be adequate for determining the optimal number of cashers and pharmacists in response to time-varying demands.

Very few previous studies in this area have developed optimization models for manpower allocation in hospitals, although some have addressed the optimization problems in allocating beds to inpatients from different divisions, and scheduling appointments with doctors, nurses, and pharmacists. In a study conducted by Creemers et al. (2012), an optimization model for assigning a set of predefined outpatient time slots to doctors from different divisions was presented in order to minimize the total expected weighted waiting time of patients. Xu and Liu (2011) developed an optimization model to solve the bed arrangement problem in one hospital, where a patient's priority is determined by the severity of his/her illness in order to reduce the amount of time beds spend unoccupied (in other words, to increase the utilization of beds). Some past studies have aimed to reduce the waiting times of patients, for instance, for cashiers (Chao, 2010), pharmacies (Afolabi and Erhun, 2003; Ndukwe et al., 2011), consulting rooms (Jerbi and Kamoun, 2009; Palvannan and Teow, 2010), and beds (Joustra et al., 2009; Hu et al., 2011). To the authors' knowledge, none of the previous research has developed optimization models for optimal manpower allocation for cashiers and pharmacists. The lack of a systematic approach for allocating optimal time-varying numbers of cashiers and pharmacists motivates us to develop the optimization model we present here.

This paper presents a bi-objective optimization model to determine optimal manpower allocation at the cashier counter, dispensary, and pharmacy at different time periods of a given day in a large hospital. The main decision variables are the time-varying numbers of cashiers and pharmacists and the two objectives are to minimize the total patient wait time and the total operating cost of the hospital. To effectively take into account time-varying or non-stationary arrival rates of patients and describe queue lengths in such a dynamic system, the point-wise fluid-based approximation method is adopted to construct a dynamic queueing network. Then, the bi-objective optimization model is formulated based on the dynamic queueing network.

We apply the weighted sum method (Zadeh, 1963; Hwang and Masud, 1979) to address the bi-objective optimization model, where the two objectives are transformed into a single objective by assigning weight parameters. In this transformation, the single objective is to minimize the weighted waiting costs of patients and operational costs of the hospital. The optimal solution obtained by the transformed method is a Pareto optimal solution of the original bi-objective optimization problem when a set of weight parameters is given. Readers can refer to Miettinen (1999) and Steuer (1989) for information on multi-objective programming. For further discussion about the choice of weights, please see Marler and Arora (2004).

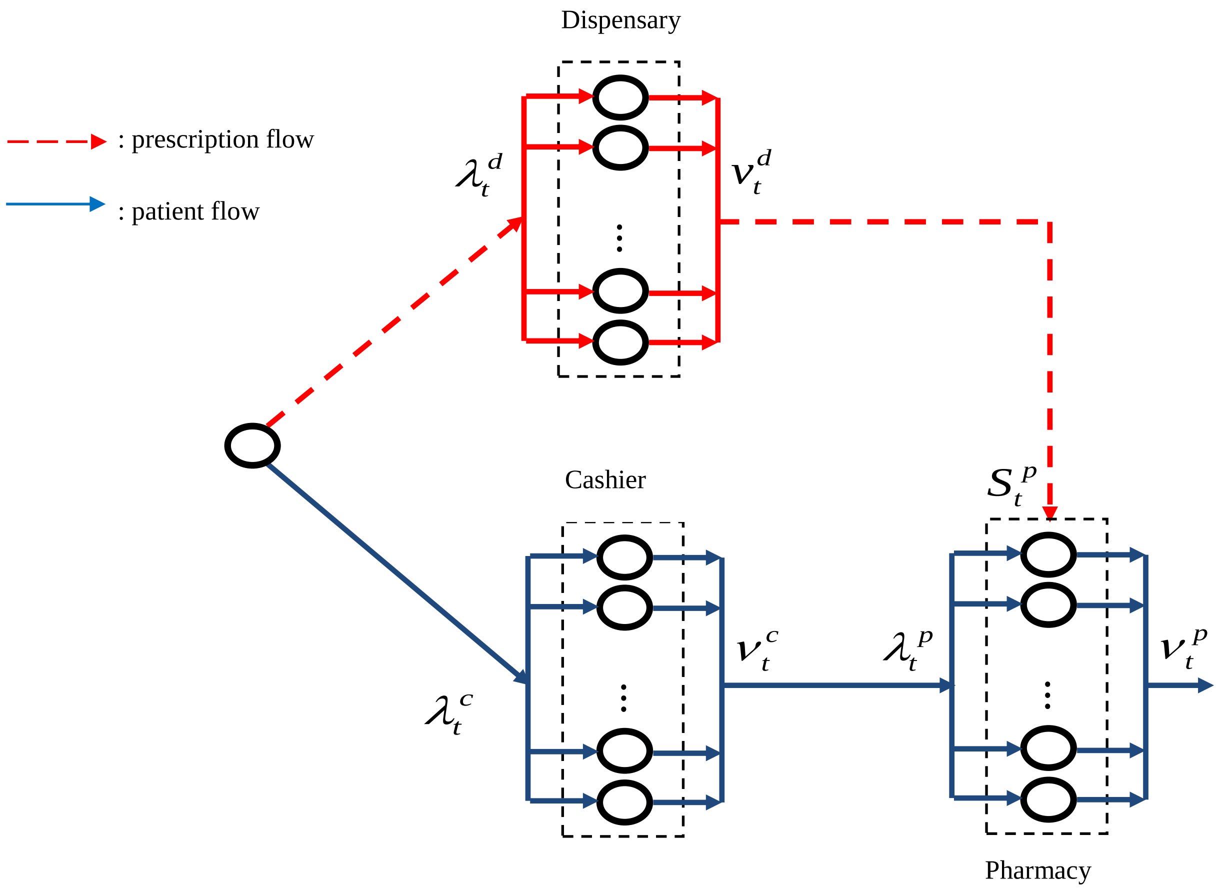

Figure 1: The dynamic queueing network constructed for the manpower allocation problem in this study

2. PROBLEM DESCRIPTION

Typically, in a large hospital in Taiwan, doctors will generate and transfer prescriptions to the dispensary after examining patients, where pharmacists fill prescriptions that will be delivered to the pharmacy for patients to pick up. In the meantime, patients leave consulting rooms and wait for cashiers to pay their bill before picking up their prescriptions according to their assigned serial number at the pharmacy counter.

This system process can be described as a dynamic queueing network, as shown in Figure 1. The queueing system consists of three waiting points (Cashier, Pharmacy, and Dispensary) and two flow subsystems (patient flow and prescription flow). The upper half of this system represented by dotted-line arcs is the prescription-flow subsystem, while the lower half of this system represented by solid-line arcs is the patient-flow subsystem. In the patient-flow subsystem, the arc coming from the left to the Cashier denotes the arrival of patients at the Cashier at time t (λ_t^c). The arc going from the Cashier denotes the serviced amount at Cashier at time t (v_t^c), which equals the arrival amount of patients at the Pharmacy at time t (λ_t^p). Patients will enter the queue at the Pharmacy after they have been serviced (i.e., pay their bill) at the Cashier. The arc leaving the Pharmacy denotes the serviced amount at the Pharmacy at time t (v_t^p) (i.e., the patients who have received their filled prescriptions and leave the hospital). In the prescription-flow subsystem, the arc coming from the left to the Dispensary denotes the arrival of prescriptions at the Dispensary at time t (λ_t^d). The arc going from the Dispensary denotes the serviced amount at the Dispensary at time *t* (ν_t^d). Since the maximum service capacity at the Pharmacy at time *t* is determined by the serviced amount at the Dispensary at time *t*, this arc also denotes the maximum service capacity of the Pharmacy at time *t* (S_t^p) (i.e., the filled prescriptions which will be delivered to the Pharmacy). We can make the following assumptions as we develop the optimization model:

- The time-varying arrival rates at each service facility, system service rates, employees' salaries per time unit, and patients' waiting costs per time unit are given.
- The time spent by patients on walking from consulting rooms to the Cashier and from the Cashier to the Pharmacy are ignored for simplicity and without loss of generality.
- The time spent on delivering prescriptions from the Dispensary to the Pharmacy is ignored for simplicity and without loss of generality.
- Prescriptions considered in this study involve common or chronic diseases drugs. Drugs for rare diseases drugs are not included.

3. POINT-WISE FLUID-BASED APPROXIMATION APPROACH AND OPTIMIZATION MODELS

A point-wise fluid-based approximation approach is adopted to construct the dynamic queueing network, which is encapsulated in the proposed optimization model for determining optimal time-varying numbers of cashiers and pharmacists in a large hospital. The notations used to present the optimization model are as follows.

Parameters:

λ_t^c	The patient arrival rate at the Cashier at time *t*
λ_t^d	The prescription arrival rate at the Dispensary at time *t*
λ_t^p	The patient arrival rate at the Pharmacy at time *t*
ρ_t^c	The capacity utilization ratio at the Cashier at time *t*
ρ_t^d	The capacity utilization ratio at the Dispensary at time *t*
C_{wait}	The waiting cost of patients(dollars/time)
W_{wait}	Weight for the waiting cost of patients
$T_{service}^c$	The service rate at the Cashier for each time
$T_{service}^d$	The service rate per unit time at the Dispensary
$T_{service}^p$	The service rate per unit time at the Pharmacy
$C_{service}^c$	The unit-time service cost of a counter at the Cashier (dollars/time)
$C_{service}^d$	The unit-time service cost of a pharmacist at the Dispensary (dollars/time)
$C_{service}^p$	The unit-time service cost of a counter at the Pharmacy (dollars/time)
$W_{service}$	Weight for the service cost of the hospital
I	The length of each time unit (minutes)
T	The number of time units
y_{max}^c	The upper bound of the number of counters at the Cashier
y_{max}^d	The upper bound of the number of pharmacists at the Dispensary
y_{max}^p	The upper bound of the number of the counters at the Pharmacy

Endogenous Variables：

x_t^c	The queue length of patients at the Cashier at time *t*
x_t^d	The queue length of prescriptions at the Dispensary at time *t*
x_t^p	The queue length of patients at the Pharmacy at time *t*
ν_t^c	The serviced amount at the Cashier at time *t*
ν_t^d	The serviced amount at the Dispensary at

	time t
v_t^p	The serviced amount at the Pharmacy at time t
S_t^c	The service capacity at Cashier at time t
S_t^d	The service capacity at Dispensary at time t

Decision Variables：

y_t^c	The number of Cashier counters open at time t (manpower allocation)
y_t^d	The number of pharmacists available in Dispensary at time t (manpower allocation)
y_t^p	The number of Pharmacy counters open at time t (manpower allocation)

The multi-objective optimization model for the M/M/1 manpower allocation problem:

$$\text{Minimize} \quad I \times C_{wait}\left(\sum_{t=1}^{T} x_t^c + \sum_{t=1}^{T} x_t^p\right),\ I \times (C_{service}^c \sum_{t=1}^{T} y_t^c + C_{service}^d \sum_{t=1}^{T} y_t^d + C_{service}^p \sum_{t=1}^{T} y_t^p) \quad (1)$$

$$\text{Subject to} \quad x_{t+1}^c = x_t^c + \lambda_t^c - v_t^c\ ,\ \forall t, \quad (2)$$

$$x_{t+1}^d = x_t^d + \lambda_t^d - v_t^d\ ,\ \forall t, \quad (3)$$

$$x_{t+1}^p = x_t^p + \lambda_t^p - v_t^p\ ,\ \forall t, \quad (4)$$

$$v_t^c - S_t^c \times \rho_t^c = 0,\ \forall t, \quad (5)$$

$$v_t^d - S_t^d \times \rho_t^d = 0,\ \forall t, \quad (6)$$

$$v_t^p = v_t^d = T_{service}^p \times y_t^p \quad \forall t, \quad (7)$$

$$S_t^c = T_{service}^c \times y_t^c,\ \forall t, \quad (8)$$

$$S_t^d = T_{service}^d \times y_t^d,\ \forall t, \quad (9)$$

$$\rho_t^c = x_t^c/(x_t^c + 1), \forall t, \quad (10)$$

$$\rho_t^d = x_t^d/(x_t^d + 1), \forall t, \quad (11)$$

$$1 \le y_t^c \le y_{max}^c,\ \forall t = 1, 2, \ldots, T, \quad (12)$$

$$x_t^c \ge 0,\ \forall t = 1, 2, \ldots, T, \quad (13)$$

$$1 \le y_t^d \le y_{max}^d,\ \forall t = 1, 2, \ldots, T, \quad (14)$$

$$x_t^d \ge 0,\ \forall t = 1, 2, \ldots, T. \quad (15)$$

$$1 \le y_t^p \le y_{max}^p,\ \forall t = 1, 2, \ldots, T, \quad (16)$$

$$x_t^p \ge 0,\ \forall t = 1, 2, \ldots, T, \quad (17)$$

The objective function (1) is to minimize the waiting cost for patients at the Cashier and Pharmacy (the first term) and the operational cost of the hospital at the Cashier, Dispensary, and Pharmacy (the second term). Note that the queue length at the Dispensary is not taken into account since patients do not wait for prescriptions at the Dispensary. Constraints (2)-(11) are established by the point-wise fluid-based approximation approach. Specifically, constraints (2)-(4) are the flow conservation equations at the Cashier, Dispensary, and Pharmacy, respectively. Constraints (5)-(7) determine the serviced amounts at the Cashier, Dispensary, and Pharmacy, respectively. Constraints (8) and (9) determine the maximum service capacities at the Cashier and Dispensary, respectively. Constraints (10) and (11) are the capacity utilization ratios at the Cashier and Dispensary, respectively, for the M/M/1 queueing system. Constraints (12), (14) and (16) present the minimum and maximum manpower requirements at the Cashier, Dispensary, and Pharmacy, respectively, set by the hospital, where the lower bounds all correspond to 1 and the upper bounds correspond to y_{max}^c, y_{max}^d and y_{max}^p, respectively. Constraints (13), (15) and (17) require nonnegative queue lengths.

By applying the weighted-sum method to solve the bi-objective optimization problem, we assign two weights (W_{wait}) and ($W_{service}$) to the waiting cost and operational cost in the objective function, respectively. The objective function (1) can be transformed into a single-objective with weight parameters as follows:

$$I \times [W_{wait} C_{wait}\left(\sum_{t=1}^{T} x_t^c + \sum_{t=1}^{T} x_t^p\right) + W_{service}(C_{service}^c \sum_{t=1}^{T} y_t^c + C_{service}^d \sum_{t=1}^{T} y_t^d + C_{service}^p \sum_{t=1}^{T} y_t^p)], \quad (18)$$

where the transformed objective (18) is the total weighted cost of this system.

In addition, we can establish the optimization model for M/G/1 queueing system. In this model, instead of (10) and (11), capacity utilization ratio function constraints with a general service time distribution (M/G/1 queue) are (19) and (20) as:

$$\text{(Cashier)}\ \rho_t^c = \frac{x_t^c+1-\sqrt{\left(x_t^c\right)^2+2\times(C_s)^2\times x_t^c+1}}{1-(C_s)^2}, \forall t, \quad (19)$$

$$\text{(Dispensary)}\rho_t^d = \frac{x_t^d+1-\sqrt{\left(x_t^d\right)^2+2\times(C_s)^2\times x_t^d+1}}{1-(C_s)^2}, \forall t. \quad (20)$$

The objective function and the other constraints are the same as those for the M/M/1 queueing system.

4. NUMERICAL EXPERIMENTS

4.1 Data preparation

To examine the proposed optimization model, we generate a problem instance based on real data provided by a large hospital (M) in Taipei. The patient arrival data were collected by the hospital between December 2013 and January 2014, including 43 weekdays (i.e., every Monday to Friday). Arrival data were collected every 10 minutes from 7AM to 11PM on each of those days.

A time interval (*I*) corresponds to 60 minutes. The service rate $T_{service}^c$ is 60 (persons/time). The upper bound of the number of Cashier counters (y_{max}^c) is 10. Based on an employee's monthly salary, the service cost at the Cashier counter can be computed as 4 dollars/minute per person. Then $C_{service}^c$ is 240 dollars/time unit per person. The service rate $T_{service}^d$ is 0.67. The upper bound of the number of pharmacists y_{max}^d is 17. The service cost at the Dispensary is 7 dollars/minute per person. This implies that $C_{service}^d$ is 420 dollars/time per person. Thus, the service rate $T_{service}^d$ is 120 (persons/time). The upper bound of the number of Pharmacy counters y_{max}^p is 8. The service cost at the Pharmacy is 7 dollars/minute per person, or $C_{service}^p$ is 420 dollars/time per person. Consider that the average salary per month is 45,888 dollars in the year 2014 provided by Directorate-General of Budget, Accounting and Statistics, Executive Yuan R.O.C. (Taiwan); the patients' waiting cost (C_{wait}) is 5 dollars/minute per person. Here, we assign the same weight for the patients' waiting cost and the hospital's operational cost. In other words, W_{wait} and $W_{service}$ are both set to be 0.5.

The MINOS solver of GAMS was applied to solve the optimization model in Section 3. The numerical experiments were implemented on a computer with Core i7-2600 3.4GHz CPU, 4.00GB Ram, equipped with Microsoft Windows 7. With the above parameters, the CPU time for GAMS to solve the problem instance was 0.109 seconds.

4.2 Numerical Results

We compare the optimal solution of the model with hospital M's current manpower allocation. The optimal numbers of cashiers do not significantly differ from those of hospital M's current allocation. At the Dispensary, the optimal numbers of pharmacists are less than those of the hospital's current allocation in the morning and afternoon; while the optimal number is more than that of the hospital's current allocation in the evening. In addition, the optimal number of Pharmacy counters is less than those of the hospital's current allocation in the morning and afternoon; while in the evening, they are almost the same.

The optimal manpower obtained by the model can reduce the queue length at the Cashier from 11.68 (persons) to 6.05 (persons). The numerical results show that at the Pharmacy, the optimal solution can reduce the queue length from 60.55 (persons) to 0 (persons), especially in the evening.

Furthermore, the service cost corresponding to the Cashier decreases by 2,380.87 dollars with the optimal solution, compared to the hospital's current allocation. The waiting cost decreases by 26,982.93 dollars. At the Pharmacy, the service cost decreases by 10,824 dollars while the waiting cost decreases by 290,631.9 dollars. The total service cost decreases from 105,813.49 dollars to 90,022.92 dollars if the optimal solution is used as opposed to the current allocation. That is, the hospital can save 15,790.57 dollars, or 15% of its current operating cost. The total waiting cost can be reduced from 34,675.53 dollars to 29,060.7 dollars (or 91.6%) using the optimal solution. Therefore, the optimal solution obtained by the model can significantly reduce the operating cost of the hospital as well as increase service quality.

5. CONCLUSION AND FUTURE RESEARCH

This paper adopts the point-wise fluid-based approximation approach to construct the dynamic queueing network, which is then encapsulated in the proposed bi-objective optimization model for determining the optimal number of cashiers and pharmacists in a large hospital. The numerical results of the test instance generated based on data for a large hospital in Taipei show that the proposed optimization model can provide the optimal allocation of manpower which significantly reduces waiting and operating costs. In sum, the hospital can save about 15% of its operating cost and decrease 91.6% of the waiting cost per day.

In future studies, the proposed dynamic queueing network modeling approach can be applied to determine the optimal manpower allocation in several other areas, such as check-in counters in airports and restaurants as well as service counters in banks.

REFERENCES

Afolabi, M.O. and Erhun, W.O. (2003) Patients' Response to Waiting Time in An Out-patient. *Tropical Journal of Pharmaceutical Research*, **2**, 207-214.

Asmussen, S.R. (2003) Random Walks. *Applied Probability and Queues*, in the series of *Stochastic Modeling and Applied Probability*, **51**, 220-243.

Chao, Y. (2010) Outpatient Queue Business Simulation Based on Acceptable Waiting Time. *International Conference On Computer Design And Appliations*, **1**, 120-123.

Chen, X., Zhou, X., and List, G.F. (2011) Using Time-varying Tolls to Optimize Truck Arrivals at Ports. *Transportation Research Part E*: *Logistics and Transportation Review*, **47**, 965-982.

Creemers, S., Beliën, J., and Lambrecht, M. (2012) The Optimal Allocation of Server Time Slots over Different Classes of Patients. *European Journal of Operational Research*, **219**, 508-521.

Grassmann, W.K. and Tavakoli, J. (2009) Transient Solutions for Multi-server Queues with Finite Buffers. *Queueing Systems*, **62**, 35-49.

Green, L. and Kolesar, P. (1991) The Pointwise Stationary Approximation for Queues with Nonstationary Arrivals. *Management Science*, **37**, 84-97.

Hu, R. and Chang, C. (2011) Using Mathematical Modeling to Rationally Arrange Hospital Beds. *International Conference on Electric Information and Control Engineering (ICEICE)*, Wuhan, China, 927-928.

Hu, R., Hu, J., and Wu, Y. (2012) A New Queueing Algorithm of Ophthalmic Hospital Beds Arrangement. *Advanced Materials Research*, **433-440**, 1971-1974.

Hwang, C.L. and Masud, A. (1979) *Multiple Objective Decision Making, Methods and Applications: A State-of-the-art Survey*. Springer-Verlag.

Jerbi, B. and Kamoun, H. (2009) Multiobjective Study to Implement Outpatient Sppointment System at Hedi Chaker Hospital. *Simulation Modelling Practice and Theory*, **19**, 1363-1370.

Jomon, A.P. and Li, L. (2012) Models for Improving Patient Throughput and Waiting at Hospital Emergency Departments. *The Journal of Emergency Medicine*, **43**, 1119-1126.

Joustra, P., Sluis, E., and Dijk, N.M. (2010) To Pool or Not to Pool in Hospitals: A Theoretical and Practical Comparison for A Radiotherapy Outpatient Department. *Annals of Operations Research*, **178**, 77-89.

Marler, R.T. and Arora, J.S. (2004) Survey of Multi-objective Optimization Methods for Engineering.*Structural and Multidisciplinary Optimization*, **26**, 369-395.

Miettinen, K. (1999) Nonlinear Multiobjective Optimization. International Series in *Operations Research & Management Science*, **12**, Kluwer Academic Publishers.

Ndukwe, H.C., Omale, S., and Opanuga, O.O. (2011) Reducing Queues in A Nigerian Hospital Pharmacyl. *African Journal of Pharmacy and Pharmacology*, **5**, 1020-1026.

Palvannan, R.K. and Teow, K.L. (2010) Queueing for Healthcare. *Journal of Medical Systems*, **36**, 541-547.

Steuer, R.E. (1989) *Multiple Criteria Optimization: Theory, Computation, and Application*. Krieger, Malabar.

Wang, W.P., Tipper, D., and Banerjee, S. (1996) A Simple Approximation for Modeling Nonstationary Queues. In *Networking the Next Generation Proceedings of the Fifteenth Annual Joint*

Conference of the IEEE Computer Societies (*INFOCOM*), **1**, San Francisco, CA, 255-262.

Xing, F. (2011) Research on Reasonable Arrangement of Ophthalmology Sickbeds in Hospital Based on Queueing Theory. *International Conference on Mechatronic Science, Electric Engineering and Computer* (*MEC*), Jilin, China, 19-22.

Xu, N. and Liu, J. (2011) Sickbed Arrangements Optimization Based on Patients' Prioritization. *International Conference on Management and Service Science* (*MASS*), Wuhan, China, IEEE, 1-3.

Zadeh L. A. (1963) Optimality and Non-scalar-valued Performance Criteria. *IEEE Transactions on Automatic Control*, **8**, 59-60.

Zhang, B. and Zwart, B. (2012) Fluid Models for Many-Server Markovian Queues in A Changing Environment. *Operations Research Letters*, **40**, 573-577.