Hindawi Publishing Corporation
Advances in Operations Research
Volume 2016, Article ID 3090758, 16 pages
http://dx.doi.org/10.1155/2016/3090758

Research Article

A Stochastic Location-Allocation Model for Specialized Services in a Multihospital System

Khadijeh Naboureh[1] and Ehram Safari[2]

[1]*Department of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran*
[2]*Research Faculty for ICT Policy Making and Strategic Management, Iran Telecommunication Research Center, Tehran 1594667711, Iran*

Correspondence should be addressed to Khadijeh Naboureh; kh.naboureh@gmail.com

Received 24 September 2015; Accepted 23 December 2015

Academic Editor: Konstantina Skouri

Copyright © 2016 K. Naboureh and E. Safari. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Rising costs, increasing demand, wasteful spending, and limited resources in the healthcare industry lead to an increasing pressure on hospital administrators to become as efficient as possible in all aspects of their operations including location-allocation. Some promising strategies for tackling these challenges are joining some hospitals to form multihospital systems (MHSs), specialization, and using the benefits of pooling resources. We develop a stochastic optimization model to determine the number, capacity, and location of hospitals in a MHS offering specialized services while they leverage benefits of pooling resources. The model minimizes the total cost borne by the MHS and its patients and incorporates patient service level, patient retention rates, and type of demand. Some computational analyses are carried out to gauge the benefits of optimally sharing resources for delivering specialized services across a subset of hospitals in the MHS against complete decentralization (CD) and full centralization (FC) policies.

1. Introduction

The number of multihospital systems (MHSs) has increased rapidly in recent years [1]. This development leads to a number of opportunities and threats for MHS managers. Nowadays, they are under increasing pressure to improve the quality of service, decrease costs, and become as efficient as possible in all aspects of their operations [2]. Rising costs, increasing demand, wasteful spending, and limited resources in the hospital industry reveal the necessity of paying more attention to this subject [1, 3, 4]. Note that both public and private expenditures on US healthcare are rising faster than any other sectors. Healthcare spending in the US was more than $2.8 trillion in 2012 [5]. This situation will be worse in the next years; according to the study conducted by the Bipartisan Policy Center, US healthcare spending will be approximately $5 trillion by 2021 [6]. In addition, 16.9% of GDP was spent on the healthcare in the US in 2012 [7] and it has been projected to be 27% of GDP by 2040 [8]. These rising expenditures can influence the economy and its various sectors [9]. The upsetting fact is that more than half of the US healthcare spending ($1.2 trillion of the $2.2 trillion in 2007) was wasteful. Consumers, government, and healthcare organizations are responsible for such wastes [10].

Another problem facing the US healthcare system is increasing demand. The Association of American Medical Colleges (AAMC) has projected that the annual demand for physician visits will increase by 53% from 2000 to 2020 [11]. The situation for the urgent demand will be even worse because there has recently been a noticeable decrease in the number of emergency departments. Demand for the specialized services is also increasing in the US. For example, based on OECD report, the numbers of Magnetic Resonance Imaging (MRI) tests per capita were highest in the US in 2013 [12]. An aging population also causes an increase in the demand of the health services where the 85 and older population is projected to increase from 5.4 million in 2008 to 19 million by 2050 in the US [13].

Another problem is limited resources. Inadequate resources are not limited to equipment and facilities. Furthermore, the healthcare systems are facing insufficient treatment staff. It has been projected that in 2020, deficiencies in the practicing nurses of the US will be at least 400,000 more than today and vacancy rates for the registered nurses (RN) will increase from 7% in 2005 to 29% by 2020 (nearly 1 million lack) [14].

In the light of the above, the hospital administrators have practically no other choice except to constantly try to become as efficient as possible in all aspects of their operations including location-allocation. Some promising strategies for tackling these challenges are joining some hospitals to form MHSs, specialization, and using the benefits of pooling resources in such systems. There are compelling reasons to apply these strategies. MHSs admit that they have a variety of opportunities to attain cost savings from system-wide economies of scale [15]. In addition, Christensen et al. indicated that the only industry where factories (hospitals) are not specialized and attempt to offer everything to everybody is the healthcare. They revealed that setting up specialized hospitals has significant opportunity to decrease cost for the hospitalizations by 15–20% [16]. This approach can lead to a decrease of the cost rather than providing these services at all hospitals. For example, Narayana Hrudayalaya's Bangalore Cardiac Care Unit (CCU) which provides specialized cardiac surgery services was set up for admitting significantly a large number of patients. Since surgeons receive fixed salary (instead of per operation) in the unit, hospital cost has dropped by increasing the number of operations and using the benefits of economies of scale [17]. As stated above, the third strategy for the efficient use of resources and reducing the costs is to utilize the benefits of pooling resources in the designed MHS. Sharing resources for specialized services (e.g., MRI, transplants, CT scans, and neonatal intensive care) due to the expensive cost of setting up new facilities is more crucial. Therefore, in this paper for proper utilization of the medical capacity in MHSs, we adopt all of the three mentioned strategies and intend to restrict offering of specialized services to a subset of hospitals instead of all hospitals of the network. In such MHSs, expensive equipment could be provided in one or more of the larger hospitals. An example of resource sharing for specialized services in MHSs is Northland Healthcare. This network involves 25 hospitals and has enabled opportunity of sharing some services such as MRI across 11 hospitals [18]. While applying shared resources policy has some advantages, it may result in longer travel distances for a patient to access the service. Therefore, the research described in this paper takes into account accessibility and service level as well as financial aspects of the problem.

In spite of the potential benefits of resource sharing in MHS, there has been limited research on this field. This paper considers locating and allocating of specialized services in a MHS. Specifically, we develop a mixed integer nonlinear stochastic optimization model to determine how many and which locations in a MHS should be set up for offering specialized service, the capacity levels of hospitals, and directions for which hospitals should handle demand of each district in the system. The model minimizes per period total cost borne by the MHS and its patients (i.e., by fixed cost of setting up locations, variable cost of maintaining the capacity across the sites, diversion cost of diverting patients demand to another location in the MHS, penalty cost of not being able to satisfy patients demand at a location, treatment cost related with met demand, transportation cost of traveling from a district to a location, and lodging costs related to the travel and hotel lodging of patient's family members who go with the patient to a given hospital) and incorporates patient service levels at each site and overall system, type of demand (urgent or nonurgent), and patient retention rates by willingness of patient to travel from a district to a location.

This paper is organized as follows. The following section discusses healthcare location-allocation and capacity planning literature relevant to our research. Section 3 develops a stochastic optimization model to determine the number, capacity, and location of hospitals in a MHS for offering specialized services. Section 4 describes experimental design. Section 5 details and discusses the results. Last, concluding remarks and directions for future research are provided in Section 6.

2. Literature Review

The research literature on facility location-allocation is vast. Daskin and Dean conducted an extensive review on facility location models in the healthcare field [19]. Location-allocation models can aid the planners to determine the number and location of facilities from which services can be provided as well as their capacities. Each of the studies based on its purpose has taken into account several factors such as social, geographical, financial, and political considerations [20–23]. Some studies have assumed geographical consideration as a most significant factor in the utilization of the health services. Most of these models have been concerned with the traveling cost and supposed that patients always patronize the nearest location to their residence [24], but most of the times, other related parameters such as types of services, quality, or adequacy influence the attraction of a hospital [25, 26]. Some researchers believe that the key characteristic in locating healthcare facilities is their multiobjective nature [27]. Using multiobjective modeling, one can take several factors into account in the modeling. For instance, Mitropoulos et al. developed a biobjective model and considered fair distribution of facilities among dispersed population and minimization of distance from patient's home to hospitals as the objectives of the model [24]. Other researchers considered the problem as a multicriteria decision making problem and tried to take into account several criteria by imposing suitable constraints [28].

Several studies concentrated on the advantages of specialization. They concluded that specialized care units can raise the number of patients, decrease costs, and enhance quality [29–34]. Phibbs et al. showed how mortality rates in neonatal intensive care units reduce as the number of the procedures increases [29]. Ketabi concluded that service level in CCUs is affected by human resources, population index, and regional concern [30]. Côté et al. developed an optimization model to locate traumatic brain injury (TBI) treatment units and provided available and accessible specialized services to qualified veterans under budget constraints. They examined the cost and expected service implications for a range of TBI treatment unit location options [31]. In another recent study, Syam and Côté showed that the degree of centralization of services, geographic density of the patient population, and retention rates are the decisive factors in the specialized services location-allocation modeling [32]. Later, they developed a comprehensive model to minimize the total cost under common resource constrains at each location and incorporated acuity level, retention rates, target service level, and overloading penalty cost in the model [33].

Despite the positive aspects of the model [33], they assumed that the demand over the time is fixed. Nevertheless, Harper and Shahani stated that the demand function is often dependent on time and distance and stochastic traveling times differ from patient to patient. Therefore, applying a deterministic approach may underestimate capacities of resources [35]. For this reason, Harper et al. developed a discrete-event location-allocation simulation model and considered demand as a function of time. The simulation model determined the facility that patients must refer to and mode of transportation [36]. Demand for some services may be seasonal, such as elderly care through the winter months or pediatric services over the school holidays. Ndiaye and Alfares focused on locating public services for nomadic population groups. He formulated a binary integer programming model to determine the optimal number and locations of primary health units for satisfying a seasonally varying demand [37]. Ndiaye and Alfares conducted a statistical technique to estimate county based demand for some healthcare services. They developed an optimization model to determine the location and number of new Community Health Centers (CHCs) in a geographical network. They concluded that optimizing the overall network can result in 20% improvement in considered measures [37].

Several studies concentrated on the benefits of MHSs [15, 34, 38, 39]. MHSs obtain significant opportunities to attain cost savings from system-wide economies of scale. Based on most of the researches, the most important action step of the MHSs to be sustainable and highly effective in the future is to specify the suitable balance between centralized and decentralized elements of the system [15]. According to [2], the integrative healthcare models taking into account location, size, and service mix across MHSs would be helpful. Mahar et al. developed an optimization model to evaluate how MHSs can use benefits of pooling resources in location-allocation models. They considered cost and service level in the model and concluded MHSs can better use resources by delivering service across fewer locations [34]. The research offers that models such as ours that share specialized hospital operations in a MHS may not only leverage economies of scale to decrease cost but also result in an increase in actual and perceived quality of service. Increasing in the number of patients for a specialized operation leads medical staffs to build their cumulative experience in administering the service and this can enhance the quality of service.

The original study that this work extends is based on an increasing need for improving the efficiency and effectiveness of delivering specialized services at a MHS [34]. We assume normally distributed stationary demand at each location. The specific improvements to make the proposed model more realistic in comparison with the model in [34] are as follows: we consider a cluster of patient demand areas (as the origin of demand) expected to be satisfied by a hospital is labeled as a "district." Therefore, the transportation costs are related to the cost of traveling from district to hospital. However, the hospitals could divert patient demand to another location in the network and incur diversion cost. Furthermore, we incorporate patient retention rates, treatment cost, and lodging cost in the model.

3. A Model

The purpose of the periodic model presented here is to develop a mixed integer nonlinear stochastic optimization model to determine how many and which existing hospitals in a MHS should be set up for specialized service, the capacity levels, and directions for which hospitals should provide specialized service for which districts in the system and which type of patients. This model includes two primary criteria: (1) the MHS's cost of providing specialized service and (2) the service level at each site and the overall system. It also minimizes the total cost borne by the MHS and its patients and incorporates patient service level at each site and overall system, type of demand, and retention rates.

The subscripts in the model are as follows:

h: hospital location index;

k: specialized services patient district index;

i: retention rate index.

We consider a general hospital network with H hospital locations. Subscript h identifies potential locations for offering specialized capacity. Probably, not all of the hospitals will be prepared to deliver specialized services due to high fixed operating costs or limited medical personnel resources. For subscript k, "district" is the term for separate geographical units such as counties, ZIP codes, or neighborhoods. With subscript i, "retention rate" is the term for the proportion of potential patients for specialized services that are willing to travel from a district to a hospital. The retention rate assists in modeling the accessibility to a hospital by considering how distance to a hospital may impact on the willingness of patient to travel to that hospital. For example, the number of patients looking for care services may be decreased as the distance to a hospital increases.

We presume normally distributed stationary demand for the specialized service at each district k. In the model, patient demand according to its type can be classified as urgent and nonurgent. We use the term "urgent" to refer to the emergency (nonflexible) demand where the procedure must be dealt with immediately and use the term "nonurgent" to refer to the nonemergency (flexible) demand that may be postponed or scheduled ahead. In the model, the demand in each district based on retention rate is allocated to the hospitals which cause transportation cost. This cost may differ according to the type of the demand. The cost of traveling the patients from district k to hospital h for nonurgent demand (Tn_{kh}) will tend to be slightly lower than for that of urgent demand (Ts_{kh}) since patients with urgent demand may choose more expensive modes of transportation.

As we noted above, hospitals face the demand of the districts. A hospital may send a patient to other hospital locations in the network and cause some cost (to the hospital and patient satisfaction and well-being). This cost may also differ according to the type of the demand. Let h, h' indicate subscripts ranging from 1 to H; the cost of diverting a patient with nonurgent demand from hospital h to hospital h' ($Dn_{hh'}$) will tend to be much lower than for that of urgent demand ($Ds_{hh'}$). When ($Ds_{hh'}$) is quite high, offering the service at more hospitals would be a reasonable choice.

Let y_h be a binary decision variable signifying if hospital h is set up for delivering the specialized service or not, and let c_h be the variable capacity level at the location h each period. Moreover, let u'_{khi} and e'_{khi} be the binary decision variables indicating whether or not district k's urgent and nonurgent demand is allocated to location h with retention rate i, respectively, and let $u_{hh'}$ and $e_{hh'}$ denote the fraction of urgent and nonurgent demand that hospital h' satisfies for hospital h (i.e., by sending patients from hospital h to hospital h' for the service), respectively. Actually, by centralizing scheduling activities in a MHS, a hospital in the system can allocate some of its demand for the specialized services. For example, by centralizing scheduling an appointment for a specialized procedure, the procedure can be routed to one of the hospitals in the network. Furthermore, when the closest hospital faces a lack of medical personnel and/or room/bed or because of the complexity of the cases, ambulance diversion can be applied to divert patients to another facility in the network. One study estimated that, in 2003, about 500,000 ambulances (an average of about one ambulance every minute) were diverted from their initial hospital destination [40].

The following model allocates the urgent and nonurgent demand and service capacity (staff, equipment, etc.) to hospitals where per period total cost (i.e., by fixed cost of setting up locations, variable cost of maintaining capacity, diversion cost of diverting patient demand to another location in the MHS, penalty cost of not being able to satisfy patient demand at a location, treatment cost associated with met demand, transportation cost of traveling from a district to a location, lodging costs related to the travel, and hotel lodging of patient's family members who go with the patient to a given hospital) is minimized while meeting target service level and ensuring the retention rate applied corresponds to the distance between the district and the hospital. To make the model more tangible, we first predefine the model's decision variables, system parameters, and quantities which can be described in terms of the decision variables.

3.1. Model Variables. The model variables are as follows:

c_h = variable specialized capacity level at location h each period;

e'_{khi} = 1 if district k's nonurgent demand is allocated to location h with retention rate i, 0 otherwise;

$e_{hh'}$ = the fraction of location h's nonurgent demand allocated to location h';

u'_{khi} = 1 if district k's urgent demand is allocated to location h with retention rate i, 0 otherwise;

$u_{hh'}$ = the fraction of location h's urgent demand allocated to location h';

y_h = 1 if location h is set up for offering the specialized service, 0 otherwise.

3.2. Model Parameters. The model parameters are as follows:

$\mu n'_k$ = expected nonurgent demand for specialized service at district k per period;

$\mu s'_k$ = expected urgent demand for specialized service at district k per period;

$\sigma n'_k$ = expected deviation of nonurgent demand for specialized service per period at district k;

$\sigma s'_k$ = expected deviation of urgent demand for specialized service per period at district k;

H = number of hospitals in the system;

K = number of districts in the system;

I = number of different retention categories in the system;

p_h = penalty cost associated with per unit of not satisfied demand at location h per period;

F_h = fixed cost of setting up location h for delivering specialized service per period;

f_h = variable cost per unit of variable capacity at location h per period;

Uc_h = upper bound of the capacity if positioned at location h;

T_h = treatment cost associated with per unit of satisfied demand at location h;

L_h = average lodging cost per day for patient family at location h;

g_h = average days of staying per patient at location h;

$d_{hh'}$ = average distance in miles between location h and location h';

d'_{kh} = average distance in miles between district k and location h;

Tn_{kh} = average transportation cost per mile between district k and location h for nonurgent demand;

Ts_{kh} = average transportation cost per mile between district k and location h for urgent demand;

$Dn_{hh'}$ = average cost per unit of diverting nonurgent demand from location h to location h';

$Ds_{hh'}$ = average cost per unit of diverting urgent demand from location h to location h';

Us_i = upper bound in miles of the ith retention rate for an urgent demand;

Un_i = upper bound in miles of the ith retention rate for a nonurgent demand;

Ls_i = lower bound in miles of the ith retention rate for an urgent demand;

Ln_i = lower bound in miles of the ith retention rate for a nonurgent demand;

α_m = minimum fraction of location h's demand that is met;

α_{system} = minimum fraction of overall demand of the system that is met;

Rs'_{khi} = proportion of urgent demand from district k retained at location h for retention rate i;

Rn'_{khi} = proportion of nonurgent demand from district k retained at location h for retention rate i.

3.3. Quantities Being a Function of the Decision Variables. The decision variables are as follows:

μs_h = expected urgent demand for specialized service at location h per period (after district's demand allocation);

μn_h = expected nonurgent demand for specialized service at location h per period (after district's demand allocation),

$$\mu s_h = \sum_{k=1}^{K}\sum_{i=1}^{I} \mu s'_k \cdot Rs'_{khi} \cdot u'_{khi},$$
$$\mu n_h = \sum_{k=1}^{K}\sum_{i=1}^{I} \mu n'_k \cdot Rn'_{khi} \cdot e'_{khi}; \quad (1)$$

μ_h = expected demand for specialized service at location h per period (after allocation between locations),

$$\mu_h = \sum_{h'=1}^{H} \left(\mu s_{h'} \cdot u_{h'h}\right) + \left(\mu n_{h'} \cdot e_{h'h}\right); \quad (2)$$

σs_h = standard deviation of urgent demand for specialized service per period at location h;

σn_h = standard deviation of nonurgent demand for specialized service per period at location h,

$$\sigma s_h = \sqrt{\left(\sum_{k=1}^{K}\sum_{i=1}^{I} \sigma s'_k \cdot Rs'_{khi} \cdot u'_{khi}\right)^2},$$
$$\sigma n_h = \sqrt{\left(\sum_{k=1}^{K}\sum_{i=1}^{I} \sigma n'_k \cdot Rn'_{khi} \cdot e'_{khi}\right)^2}; \quad (3)$$

σ_h = standard deviation of demand for specialized service per period at location h,

$$\sigma_h = \sqrt{\sum_{h'=1}^{H} \left(\mu s_{h'} \cdot u_{h'h}\right)^2 + \left(\mu n_{h'} \cdot e_{h'h}\right)^2}; \quad (4)$$

$R(z_h)$ = the unit normal right-tail linear loss functions,

$$R\left(z_h\right) = \int_{z(h)}^{\infty} \left(w - z_h\right) \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-w^2/2} \cdot dw$$
$$\varphi\left(z_h\right) = 1 - R\left(z_h\right) \quad (5)$$

or

$$R\left(z_h\right) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-z_h^2}{2}\right) - \frac{z_h}{2} + \frac{z_h}{2} \cdot \operatorname{erf}\left(\frac{z_h}{\sqrt{2}}\right), \quad (6)$$

where

$$z_h = \frac{\left(c_h - \mu_h\right)}{\sigma_h}. \quad (7)$$

In order to estimate error function $\operatorname{erf}(y)$, we use the following approximation provided by Winitzki [41]:

$$\operatorname{erf}\left(y\right) \approx \frac{y}{|y|} \cdot \sqrt{1 - \exp\left(-y^2 \frac{4/\pi + by^2}{1 + by^2}\right)};$$
$$b = \frac{8\left(\pi - 3\right)}{3\pi \cdot \left(4 - \pi\right)}. \quad (8)$$

3.4. Model Formulation. Consider the following:

$$\begin{aligned}
\text{(P)}\ \ \text{Mininmize}\quad & \sum_{h=1}^{H} F_h \cdot y_h + \sum_{h=1}^{H} f_h \cdot c_h + \sum_{h=1}^{H} p_h \cdot R\left(z_h\right) \cdot \sigma_h + \sum_{h=1}^{H} \left(\mu_h - R\left(z_h\right) \cdot \sigma_h\right) \cdot V_h \\
& + \sum_{h=1}^{H} \left(\mu_h - R\left(z_h\right) \cdot \sigma_h\right) \cdot g_h \cdot L_h + \sum_{h=1}^{H}\sum_{k=1}^{K}\sum_{i=1}^{I} \mu s'_k \cdot Rs'_{khi} \cdot u'_{khi} \cdot d'_{kh} \cdot Ts_{kh} \\
& + \sum_{h=1}^{H}\sum_{k=1}^{K}\sum_{i=1}^{I} \mu n'_k \cdot Rn'_{khi} \cdot e'_{khi} \cdot d'_{kh} \cdot Tn_{kh} + \sum_{h=1}^{H}\sum_{h'=1}^{H} \mu s_h \cdot u_{hh'} \cdot d_{hh'} \cdot Ds_{hh'} \\
& + \sum_{h=1}^{H}\sum_{h'=1}^{H} \mu n_h \cdot e_{hh'} \cdot d_{hh'} \cdot Dn_{hh'}
\end{aligned} \quad (9)$$

$$\text{s.t.} \quad \sum_{i=1}^{I}\sum_{k=1}^{K} u'_{khi} = 1, \quad h = 1,2,\ldots,H \tag{10}$$

$$\sum_{i=1}^{I}\sum_{k=1}^{K} e'_{khi} = 1, \quad h = 1,2,\ldots,H \tag{11}$$

$$u'_{khi} \le y_h, \quad k = 1,2,\ldots,K;\ h = 1,2,\ldots,H;\ i = 1,2,\ldots,I \tag{12}$$

$$e'_{khi} \le y_h, \quad k = 1,2,\ldots,K;\ h = 1,2,\ldots,H;\ i = 1,2,\ldots,I \tag{13}$$

$$\sum_{h'=1}^{H} u_{hh'} = y_h, \quad h = 1,2,\ldots,H \tag{14}$$

$$\sum_{h'=1}^{H} e_{hh'} = y_h, \quad h = 1,2,\ldots,H \tag{15}$$

$$u_{hh'} \le y_{h'}, \quad h,h' = 1,2,\ldots,H \tag{16}$$

$$e_{hh'} \le y_{h'}, \quad h,h' = 1,2,\ldots,H \tag{17}$$

$$c_h \ge \mu_h, \quad h = 1,2,\ldots,H \tag{18}$$

$$c_h \le y_h \cdot Uc_h, \quad h = 1,2,\ldots,H \tag{19}$$

$$\varphi\left(z_h\right) \ge y_h \cdot \alpha_h, \quad h = 1,2,\ldots,H \tag{20}$$

$$1 - \left(\frac{\sum_{h=1}^{H} R\left(z_h\right) \cdot \sigma_h}{\sum_{h=1}^{H} \mu_h}\right) \ge \alpha_{\text{system}} \tag{21}$$

$$\left(d'_{kh} - Us_i\right) \cdot e'_{khi} \le 0, \quad k = 1,2,\ldots,K;\ h = 1,2,\ldots,H;\ i = 1,2,\ldots,I \tag{22}$$

$$\left(d'_{kh} - Un_i\right) \cdot u'_{khi} \le 0, \quad k = 1,2,\ldots,K;\ h = 1,2,\ldots,H;\ i = 1,2,\ldots,I \tag{23}$$

$$\left(d'_{kh} - Ls_i\right) \cdot e'_{khi} \ge 0, \quad k = 1,2,\ldots,K;\ h = 1,2,\ldots,H;\ i = 1,2,\ldots,I \tag{24}$$

$$\left(d'_{kh} - Ls_i\right) \cdot e'_{khi} \ge 0, \quad k = 1,2,\ldots,K;\ h = 1,2,\ldots,H;\ i = 1,2,\ldots,I \tag{25}$$

$$0 \le e_{hh'} \le 1, \quad h,h' = 1,2,\ldots,H \tag{26}$$

$$0 \le u_{hh'} \le 1, \quad h,h' = 1,2,\ldots,H \tag{27}$$

$$y_h, e'_{khi}, u'_{khi} \in \{0.1\}, \quad k = 1,2,\ldots,K;\ h = 1,2,\ldots,H;\ i = 1,2,\ldots,I. \tag{28}$$

The objective function (9) minimizes the total expected cost for the MHS. The first two terms in the objective function signify the fixed cost of setting up locations for delivering the specialized services and variable cost associated with maintaining the capacity across the sites. The third term represents the penalty cost of not being able to satisfy demand at a location in the current period. The fourth and fifth terms represent the treatment cost associated with met demand and lodging costs of family members. The sixth and seventh terms show the transportation cost of traveling patients with urgent and nonurgent demand from a district to a location. Terms 8 and 9 represent diversion cost of rerouting patients with urgent and nonurgent demand to another location in the MHS. Since there are the loss function and fractional division of decision variables in terms 3, 4, and 5, the objective function is nonlinear. The nonlinear term $R(z_h) \cdot \sigma_h$ denotes the expected units of not satisfied demand at location h.

Constraints (10) and (11) ensure that the urgent and nonurgent demand in each district is assigned to exactly one location with a single retention rate. Constraints (12) and (13) put capacity at a site if any urgent or nonurgent demand from a district with the corresponding retention rate is allocated to that site. Constraints (14) and (15) guarantee that all of site h's demand is assigned to a location if any urgent or nonurgent demand is allocated to that site. Constraints (16) and (17) put capacity at a site if any urgent or nonurgent demand from a location is allocated to that site. Constraints (18) guarantee that variable capacity at any site h is enough to satisfy expected demand at the site, so $z_h \geq 0$. Constraints (19) restrict capacity at site h to either 0 if the site is not set up to offer the service or Uc_h. The Uc_h value may be bounded by staff or equipment capacity and size and space of the location and so forth. Constraints (20) and (21) guarantee minimum service levels at each location and overall network. Constraints (20) guarantee that site h's chance for handling its demand in each period is at least $(100\alpha_h)$% if site h is set up to offer the specialized service and constraint (21) guarantees that at least α_{system}% of the total demand is expected to be satisfied across all locations in each period. Constraints (22) to (25) identify the lower and upper bounds for each retention rate category. Constraints (26) and (27) guarantee that the fractional allocations of urgent and nonurgent demand lie between 0 and 1. Constraints (28) necessitate that the specialized capacity is either positioned at a site ($y_h = 1$) or not ($y_h = 0$); district k's urgent demand is either allocated to location h with retention rate i ($u'_{khi} = 1$) or not ($u'_{khi} = 0$); district k's nonurgent demand is either allocated to location h with retention rate i ($e'_{khi} = 1$) or not ($e'_{khi} = 0$).

Problem (P) is a mixed integer nonlinear optimization problem (MINLP) with nonlinear objective function subject to nonlinear constraints (i.e., (20) and (21)) and integer decision variables (y_h, e'_{khi}, u'_{khi}). It is necessary to note that in general, a 0/1 integer programming problem is NP-complete [42]. Therefore, the model is difficult to solve. There are many binary variables of the types e'_{khi} and u'_{khi} in the model. This feature makes a major obstacle to implementation. As even for medium-sized problems, solving models coded in the GAMS (http://www.gams.com/) needs many hours. In order to overcome these barriers, we precalculate the retention rates that potentially apply between districts and locations and replace e'_{khi} and u'_{khi} variables altogether with e'_{kh} and u'_{kh}, respectively. Now, by imposing retention rates in this manner we can eliminate constraints (22) to (25) and redesign the model. This artifice results in reduction in the significant number of binary variables and constraints. This in turn leads to reduction in the computing times.

Similar to [34], we let z_h differ across the hospitals for relaxing the balance assumption and allowing patient service levels differ by site. In the next section, some computational analyses are carried out to guarantee that the analytic total cost gives rational results and gauge the benefit of optimally sharing resources for offering the specialized services across a subset of hospitals in the MHS against two alternate (CD and FC) policies: (1) locating the service at all of the sites where each of them handles its own demand and (2) offering the service at one site to satisfy all of the network demands.

4. Experimental Design

When we want to locate the specialized services in a MHS, a range of different settings happen in practice. For some services, the main cost is the cost of purchasing the expensive specialized equipment, such as a positron emission tomography (PET) scan machine, whereas other services such as transplant procedures need a wide combination of costs related to equipment, specialized personnel and physicians, and availability of surgical beds and rooms on certain days of the week. Examples mentioned reveal that based on the service there are quite different values for the cost of satisfied or unmet demand. Geographic layout of the hospital system is another key factor in locating the specialized services. The candidate hospitals in the system can be located quite near to each other (e.g., in one large city) or spread across several cities or states. Type of demand and service levels are other important characteristics that may provide some guidelines to decision makers as they do location decisions. To investigate the trade-offs between different decisions and generalize the results to a range of MHS settings, the impact of service level at each site and overall system, diversion cost, treatment cost, and penalty cost and percentage of urgent demand on the allocation of the demand and placement of the capacity decision for three MHSs are examined in greater detail in the next section. Figures 1, 2, and 3 represent the geographic locations for MHSs in North Carolina, Michigan, and Georgia, respectively. Each of the networks has 100, 83, and 159 districts, respectively. In this paper, a district corresponds to a county. In other studies, it corresponds to a ZIP code or larger geographic areas such as counties, states, or regions [31, 33]. Each network has 5 hospital locations denoted by numbers $1, 2, \ldots, 5$. The potential hospital locations are considered in the centers of some of the districts. Demand originates from the center of a district. As it has been explained in previous section, we precalculate the retention rates that potentially apply between districts and locations. This artifice significantly results in reduction in the number of binary variables. Therefore, increasing in the number of districts does not make barriers to solving the model. The expected periodic total cost is obtained under the optimal position of service capacity (MP policy) for each combination of factor levels and compared against two alternate (CD and FC) policies: (1) locating the service at all of the sites where each of them handles its own demand and (2) offering the service at one site to satisfy all the network demands.

For each of the problem sets, the percentage of the total demand originating from each district (district demand) is set to be proportional to the population density in the county. We assume that all the unsatisfied demands make a diversion cost, and the standard deviation of the mixed demand (urgent plus nonurgent) at each district is set at 25% of the expected demand of the region. The fixed operating and variable costs are estimated to be \$900 per period and \$60 per unit, respectively. Distance between a district and a hospital is considered according to the distance between a county and the potential hospital location. The transportation cost of a patient from his\her district to a hospital in the network is assumed to be proportional to the Euclidean distance between the district and the hospital location and is set at \$0.20 per mile. The similar way is applied to calculate the distance between hospital locations. The average diversion cost for nonurgent demand is set at \$0.30 per mile. The cost of diverting for the urgent demand is set at five levels in Table 1 to show the potentially higher diversion cost of the urgent

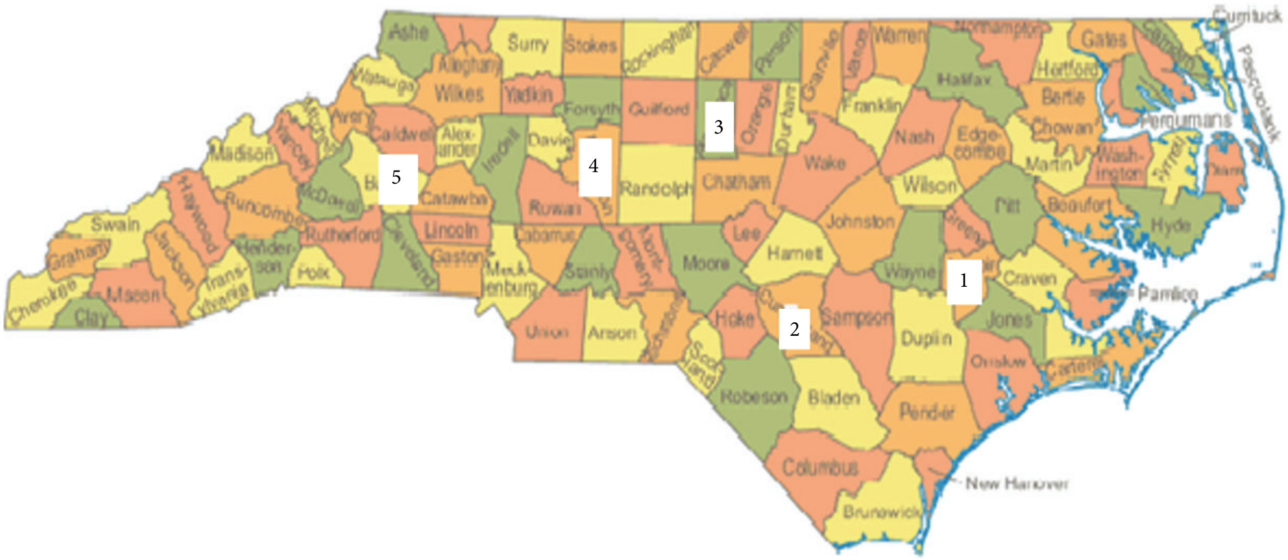


FIGURE 1: The geographic locations for the hospital network in North Carolina in the United States, problem set 1.

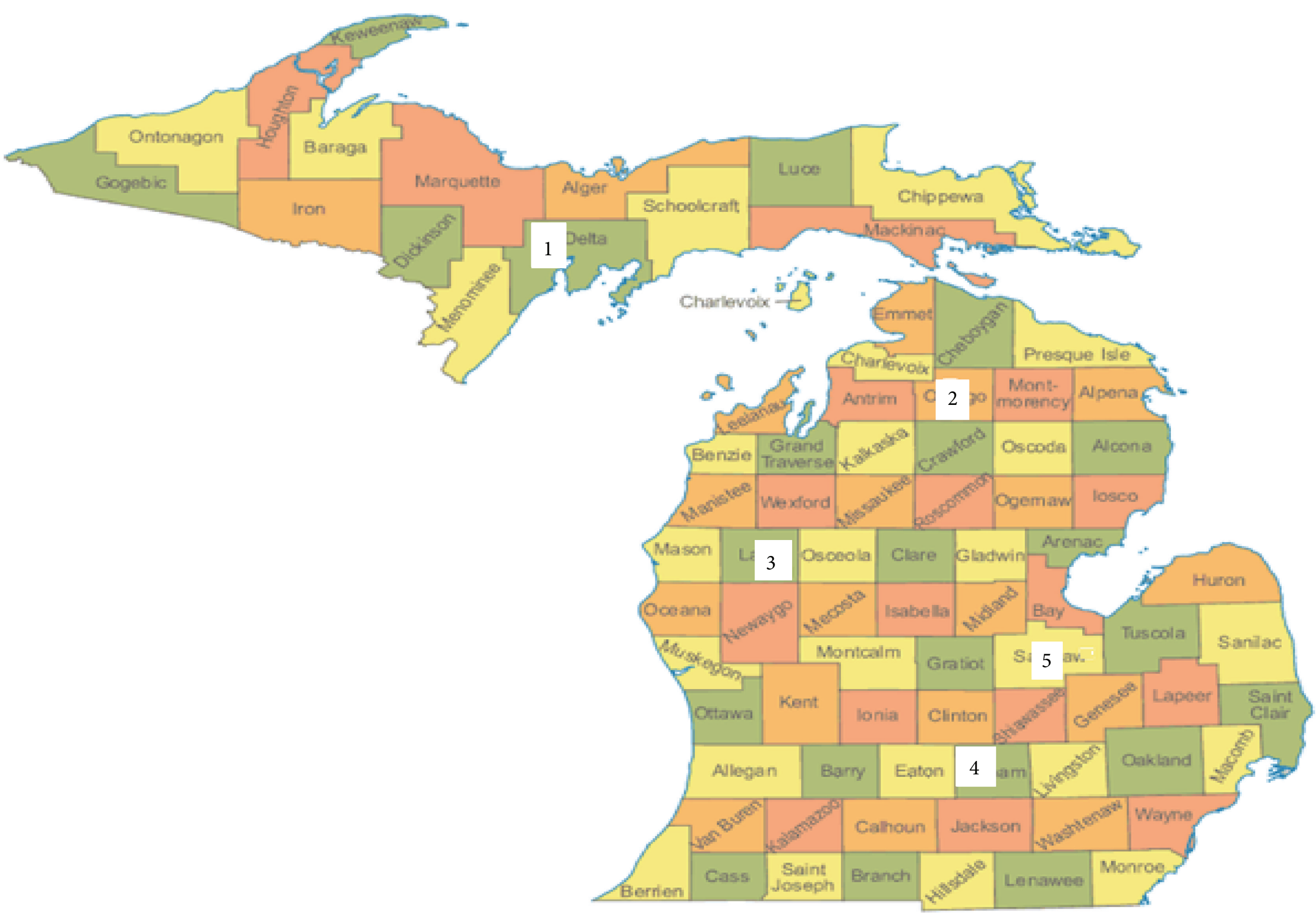


FIGURE 2: The geographic locations for the hospital network in Michigan in the United States, problem set 2.

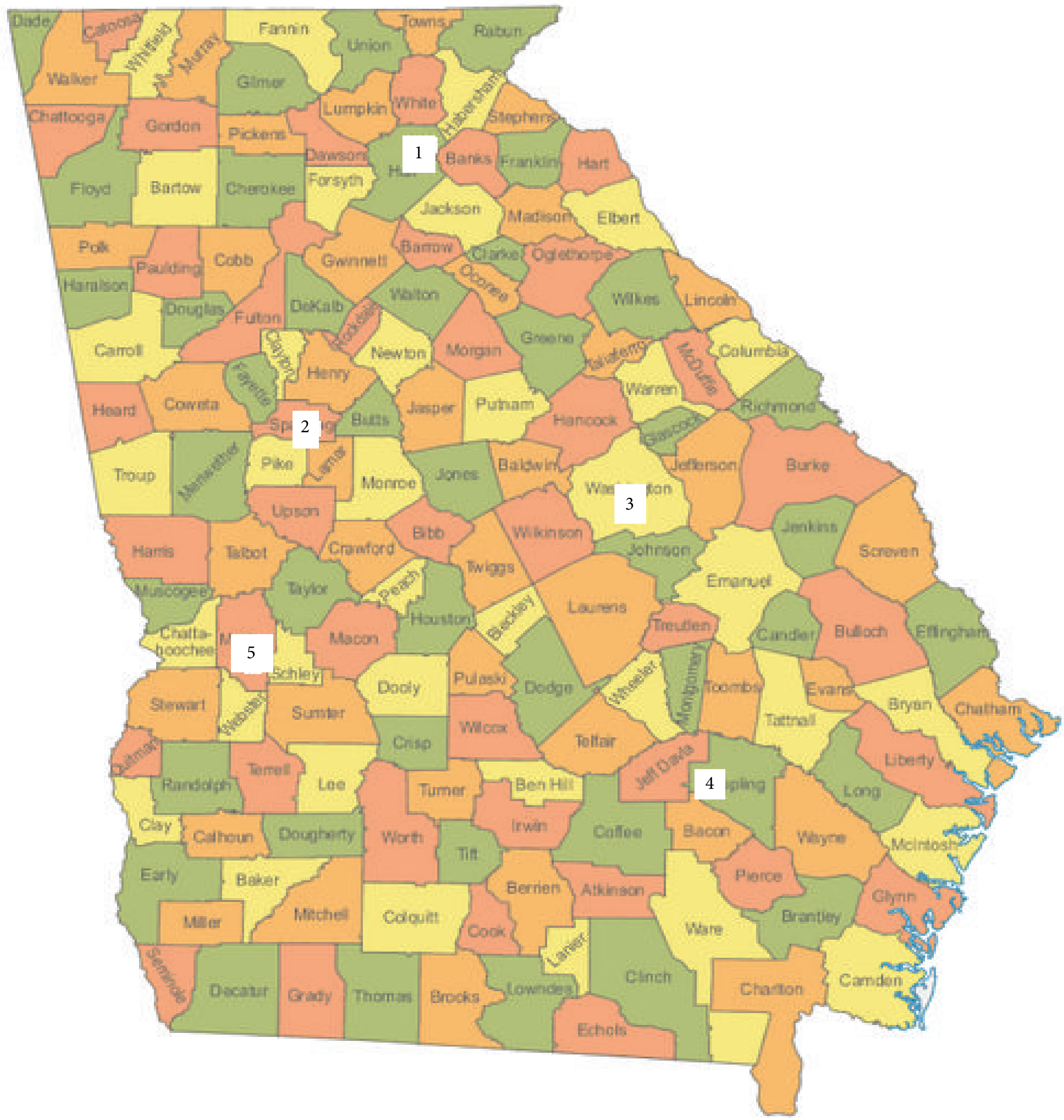


Figure 3: The geographic locations for the hospital network in Georgia in the United States, problem set 3.

demand. The penalty cost for the unmet demand is set at \$220 per unit. The average lodging cost per day and average length of stay (in days) for admissions are estimated to be \$50 per day and $N(3, 1)$, respectively. Note that these data are simulated based on four real-world studies [31–34] related to our work and data provided by Health Economic Resource Center (HERC) (http://www.herc.research.med.va.gov/). Since this model may be applied to a specific period of time, all the costs are projected to that period. For example, consider a MRI machine, if its cost is considered 1.2 Million\$ and its lifetime is 10 years and during the planning period is 1 year, the considered MRI cost in the planning period (1 year) will be 0.12 Million\$. Moreover, the costs borne by the patient are related to costs incurred in the specific planning period. Therefore, the demand for each service is equal to all the patients who tend to come in hospital system in the planning period. The patient retention rates for urgent and nonurgent demand based on slightly gradual loss of retention (retention that decreases slowly at 7% for every 80 miles of distance for urgent demand and 10% for every 120 miles of distance for nonurgent demand) according to the data provided in [33] are set as follows.

The patient retention rates for urgent demand are

0–80 miles: 80%,

81–160 miles: 77%,

Table 1: Considered factors and their levels for problem set 1.

Factor	Level	Urgent demand	Nonurgent demand
Diversion cost (Ds) ($/mile)	1	0.3	0.3
	2	0.3	0.9
	3	0.3	3
	4	0.3	18
	5	0.3	Infinite
Required service level at each site (α_h)	1	≥70%	≥70%
	2	≥75%	≥75%
	3	≥80%	≥80%
	4	≥85%	≥85%
	5	≥90%	≥90%
	6	≥95%	≥95%

161–500 miles: 65%.

The patient retention rates for nonurgent demand are

0–120 miles: 85%,

121–240 miles: 75%,

241–500 miles: 65%.

We provide the values of experimental factors for problem sets 1, 2, and 3 in Tables 1, 2, and 3, respectively. Furthermore, we consider 3 policies (MP, CD, and FC) for each of the problem sets.

The factors shown in Tables 1, 2, and 3 along with three mentioned policies result in the experiment with 90 + 648 + 540 = 1278 test problems across the three problem sets. The aims of the next section are to conduct computational analysis and offer results of the experiments by solving MINLP with the GAMS solver. Specifically, we use BARON and AlphaECP solvers under the 30 h time limit for each of the problems and introduced suitable upper and lower limits on the variables and expressions [43]. We also set to resolve a problem by BARON if by AlphaECP after 30 h the problem is not solved to within 0.4% of the optimal solution. The solvers provide global optimality of 99% of the 1278 test problems within 0.01% of the time limit while the remaining test problems are solved within 0.4% of the time limit and 45 h. The test problems of the all policies (MP, CD, and FC) are solved to obtain global optimum solution. Generally, the results represented in the next section denote that MP policy outperforms both CD and FC policies in terms of the majority of the factors.

Table 2: Considered factors and their levels for problem set 2.

Factor	Level	Urgent demand	Nonurgent demand
Percent urgent demand (pU)	1	0%	100%
	2	20%	80%
	3	40%	60%
	4	60%	40%
	5	80%	20%
	6	100%	0%
Required service level at each site (α_h)	1	≥70%	≥70%
	2	≥75%	≥75%
	3	≥80%	≥80%
	4	≥85%	≥85%
	5	≥90%	≥90%
	6	≥95%	≥95%
Treatment cost (T) ($/unit)	1	100	100
	2	120	120
	3	140	140
	4	160	160
	5	180	180
	6	200	200

Table 3: Considered factors and their levels for problem set 3.

Factor	Level	Urgent demand	Nonurgent demand
Required service level at overall system (α_{system})	1	≥70%	≥70%
	2	≥75%	≥75%
	3	≥80%	≥80%
	4	≥85%	≥85%
	5	≥90%	≥90%
	6	≥95%	≥95%
Required service level at each site (α_h)	1	≥70%	≥70%
	2	≥75%	≥75%
	3	≥80%	≥80%
	4	≥85%	≥85%
	5	≥90%	≥90%
	6	≥95%	≥95%
Penalty cost (p) ($/unit)	1	220	220
	2	450	450
	3	650	650
	4	850	850
	5	1050	1050

5. Discussion of Results

For each of the 1278 combinations of the factors earlier mentioned, we obtain the total cost. Based on the results in Table 4 and Figures 4–9, optimally sharing resources across a MHS can enhance performance of the system in terms of service levels, financial aspects, and accessibility. For problem sets 1, 2, and 3, the influence of the combinations of the factors (the service levels, percent of urgent demand and diversion cost, penalty cost, and treatment cost) on the performance of the different policies (MP, CD, and FC) is discussed as shown below. When we want to evaluate the impact of one factor on the MHSs, the value of service level and percentage of urgent demand are set at 0.95 and 40%, respectively, and the value of other factors is set at level 1 of those factors, unless it is mentioned in the experiment.

Table 4: Average distance patients travel to receive the service under the different policies (MP, CD, and FC).

Required service level at each site (α_h)	(CD) Nonurgent/urgent demand	(FC) Nonurgent/urgent demand	(MP) Nonurgent demand	(MP) Urgent demand	
0.7	6.3	80.1	31.6	16.9	Problem set 1
	7.9	120.3	81.4	48.8	Problem set 2
	11.1	169.2	100.1	68.0	Problem set 3
0.9	6.9	92.6	29.5	15.0	Problem set 1
	9.1	127.0	72.8	42.3	Problem set 2
	13.9	184.5	93.9	60.7	Problem set 3

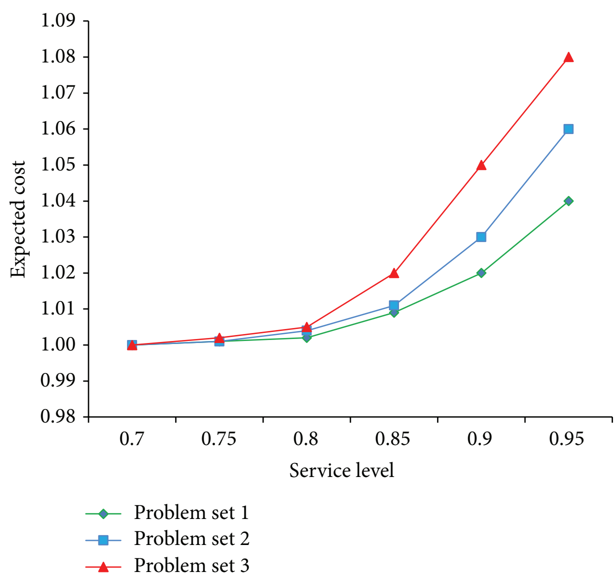


Figure 4: Relative cost against service level at each site (α_h) under MP policy, problem sets 1, 2, and 3. Note: the total cost is represented relative to the cost at service level equal to 75%.

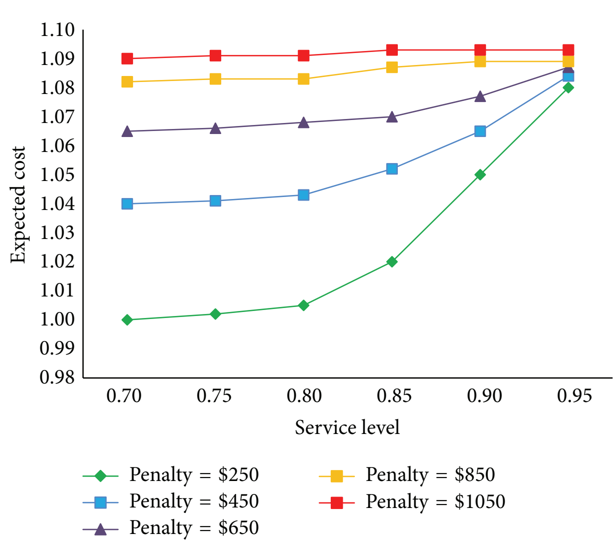


Figure 5: Relative cost against service level at all sites (α_{system}) as the penalty cost increases under MP policy, problem set 3.

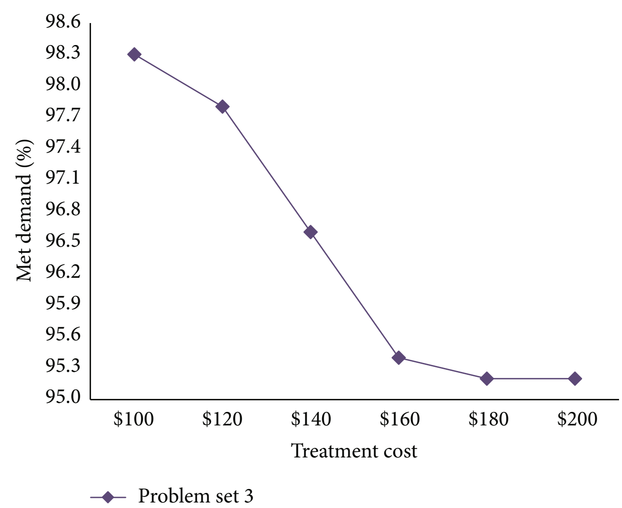


Figure 6: Met demand against treatment cost under MP policy, problem set 2.

5.1. Trade-Off between Service Levels and Total Cost under MP Policy. At first, we consider the trade-off between the service level at each site and total cost. Figure 4, comparing the 6 levels for service at each site (α_h) under MP policy for each of the three problem sets, indicates that the objective value (i.e., total cost) increases slowly as the service level rises from 0.7 to 0.8. The increase becomes much faster as the service level rises from 0.8 to 0.95. Another fact is that as the service level rises, the jump in the total cost is larger in problem set 3 than problem set 2 and greater in problem set 2 than problem set 1. It means that in MHSs with hospital locations very near to each other, we can improve the service level without considerable increase in the cost. Note that the retention rates are primarily a function of distance; therefore, the relatively shorter travel distances in locations close to each other correspond to higher retention for each level of the service.

A slight increase in the total cost by raising the service levels from 0.7 to 0.8 in all the three networks results from the fact that when the service level rises to 0.8, a considerable amount of network demand (larger than required proportion) is met and therefore the total penalty cost, as a part of total cost, decreases. This decrease partially compensates for

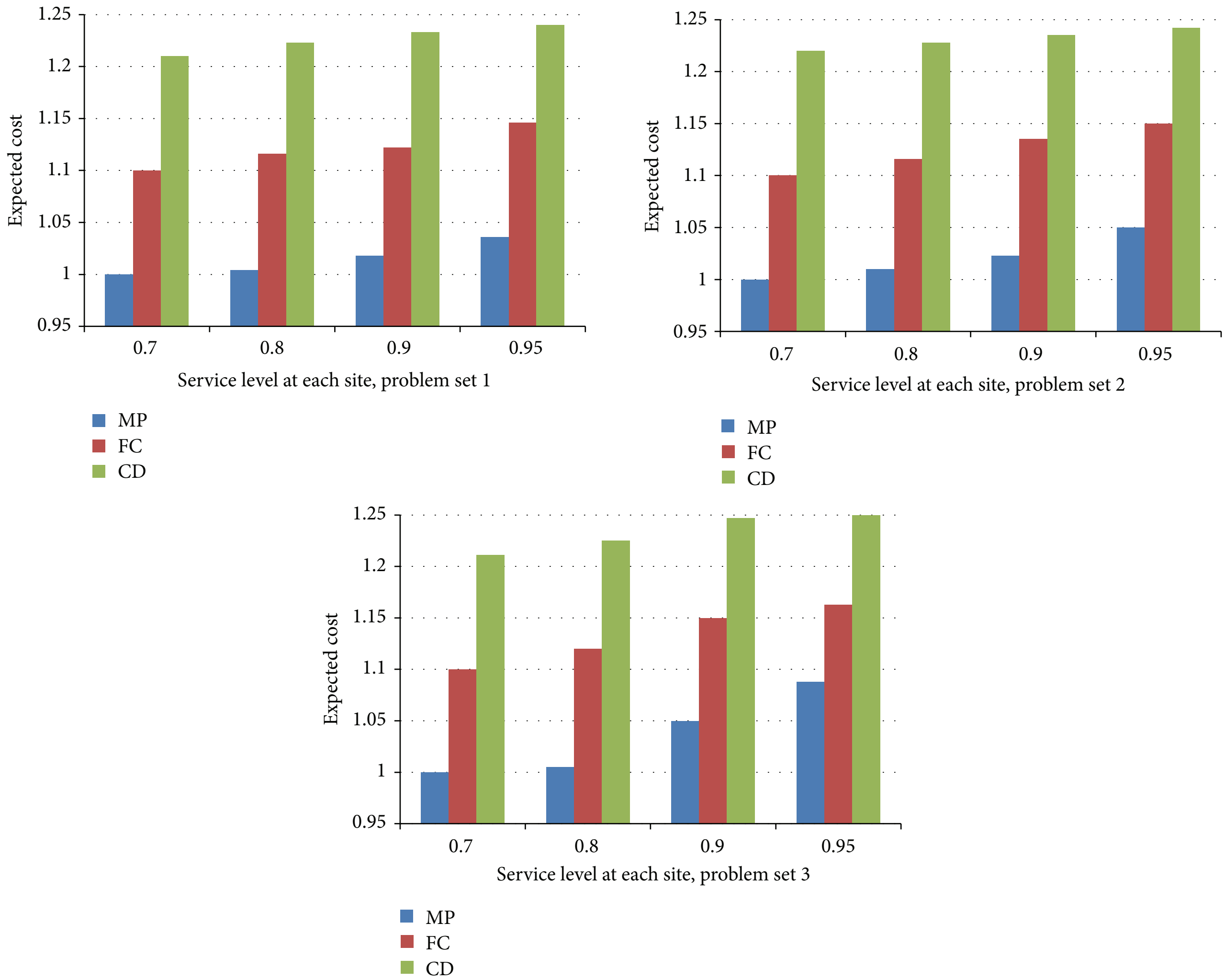


FIGURE 7: Relative total cost against service level in each site under various policies (MP, CD, and FC), problem sets 1, 2, and 3. Note: total costs are represented relative to the cost at service level equal to 70% under MP policy.

the effects of increases in other costs (associated with additional met demand) on the total cost. It means that raising the service level up to a specific level (in this experiment from 0.7 to 0.8) does not result in significant changes in the cost. Therefore, with a slight increase in the cost we would be able to achieve the higher service levels and increase met demand in MHSs. As mentioned above, raising the service levels more than this specific value (0.8) causes a significant increase in the total cost.

These results also indicate that the numbers of hospitals set up in all problem sets are more than one. Because the required service level for the overall system is constant (0.95) and the maximum value for the required service level at each site is equal to 0.95 (α_h is not larger than α_{system}), if only one hospital is set up, changing the service level at each site would have no effect on the total cost. It is interesting to note that in problem set 1, the numbers of hospitals that need to be set up are equal to 2 for all the service levels. In the case of problem set 2, this value is equal to 2 for the service level from 0.7 to 0.85, but it increases to 3 for the service level from 0.85 to 0.95. In the last case (problem set 3), the numbers of hospitals that need to be set up with specialized service are equal to 2 for the service level from 0.7 to 0.8, but it increases to 3 for the service level from 0.8 to 0.95. It means that in MHSs with hospital locations relatively far from each other, for providing higher service level, it is necessary to increase the number of hospitals that should be set up to deliver specialized service.

The results of the realistic instances of a location-allocation model with 50 admission districts, 20 candidate medical center locations, and 5 open treatment units also indicated that as the minimum service level at each hospital increases, the optimal cost increases at an increasing rate [33].

5.2. Trade-Offs between the Penalty Cost, Service Level at Overall System, and Total Cost under MP Policy. Figure 5 indicates the effect of the different service levels of the overall system (α_{system}) on the total expected cost under MP policy as penalty cost increases. From this figure, we conclude that as the overall service level in problem set 3 increases, the total expected cost increases too. In analyzing the results of the service levels of the overall system factor, α_{system}, an interaction with the penalty cost factor becomes apparent.

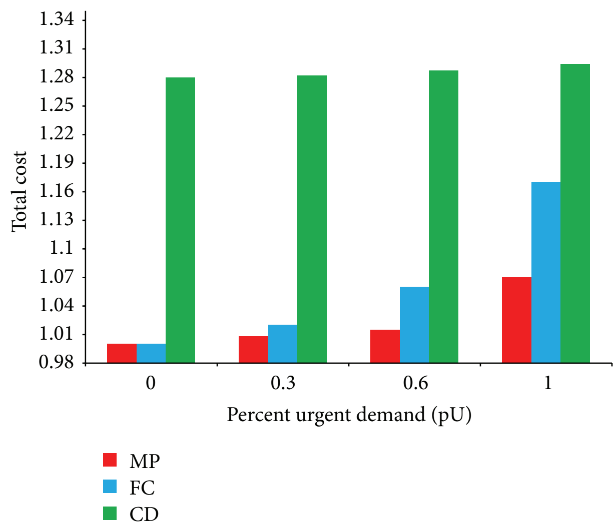


Figure 8: Relative total cost against percent urgent demand (pU) under various approaches (MP, CD, and FC), problem set 2. Note: total costs are represented relative to the cost at percent urgent demand equal to 0% under MP policy.

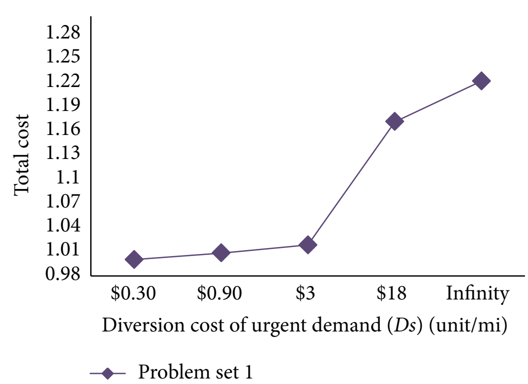


Figure 9: Relative total cost against different diversion cost per mile for urgent demand under MP policy, problem set 1. Note: total costs are represented relative to the cost where diversion cost per mile for urgent demand is equal to 0.3.

The results illustrate that under $1050 per unit penalty cost, the effect of increasing the overall service level on the total cost is not significant. It stems from the fact that when per unit penalty cost is so high compared to other costs, the system tends to handle more network demand (often larger than required) and therefore the overall service level rises automatically, even if we require a mediocre service level. In other words, at higher per unit penalty costs, slight change in the total cost by increasing the overall service level results from the fact that the penalty cost and service level affect each other. It means that a relatively higher penalty cost results in a higher service level.

The results obtained by locating specialized service capacity in MHSs located in Indiana and Pennsylvania [34] showed that when the penalty cost associated with unmet demand increases, the service level increases too.

5.3. Trade-Off between the Treatment Cost and Met Demand under MP Policy. Another interesting result is the impact of the treatment cost on satisfied demand. Figure 6 shows how met demand in the problem set 2 differs under the different treatment costs. We set penalty cost at $450 per unit in this experiment. Results illustrate that as the treatment cost associated with satisfied demand increases, the proportion of met demand decreases under MP policy. On average 98.1% of the network demand is satisfied under $100 per unit treatment (versus 94.8% satisfied demand under $200 per unit treatment). This happens because when the treatment cost is lower in comparison to the penalty cost, the system prefers to meet more proportion of the network demand to decrease the total cost. Note that this trend recurs until satisfying more demand does not require setting up a new hospital because setting up a new hospital imposes considerable amount of cost on the network. In other words, the model firstly evaluates trade-offs between costs associated with additional met demand such as treatment, higher capacity level, and setting up a new hospital and cost of unsatisfying demand, then decides to handle network demand more than mandated proportion or not. As the treatment cost increases and reaches $180 per unit, the network tries to handle only the required proportion of the network demand not beyond.

The relationship between treatment cost and met demand has not been studied directly yet; however, the closest study to this is in [33]. The study was about the integrated service networks of Veterans Affairs Department and demonstrated that by increasing the lost admission cost, the average treatment cost increases too. The lost admission cost is federal funding not achieved if patients are not admitted (served).

5.4. Trade-Off between Total Cost and Service Levels under Each of the Policies (MP, CD, and FC). Figure 7, comparing the 4 levels for service at each site (α_h) under various policies (MP, CD, and FC) for each of the problem sets, indicates that as the service level in each site increases, the total cost increases too in all three policies. For CD and FC policies it happens because of increasing the treatment and variable costs associated with maintaining higher capacity level. This relationship is consistent with MP policy too, but increasing cost in this case may also be related to the additional cost of setting up new facility for some service levels. Results also reveal that MP policy provides some 16.8% reduction in the cost of CD approach and some 8.1% reduction in the cost of FC one. This improvement is moderately consistent across the three levels for the service factor. However, for all of the problem sets, we may see that as the service level increases, the benefit of MP policy decreases a little with respect to CD approach and increases a small degree with respect to FC one. This happens because the cost of setting up the new facility in MP policy for higher service levels is considerable.

Another fact is that as the service level rises, the improvement in the total cost in MP policy compared to CD and FC policies is larger in problem set 3 than problem set 2 and greater in problem set 2 than problem set 1. It means that in MHSs with hospital locations slightly far from each other we can better utilize the benefits of pooling resources.

The results of a location-allocation model with 25 potential service locations, 40 districts, and 5 open treatment units [32] also indicated that as the service level in each site increases, the total cost increases too in all three levels of treatment unit centralization, and partial centralization has better performance (lower total cost) in comparison with full centralization and complete decentralization.

5.5. Distances a Patient Must Travel to Receive the Service under Each of the Policies (MP, CD, and FC). Distance traveled by a customer to receive the service is another customer service measure in designing a network. Table 4 shows the average distance a patient traveled to receive the service under the MP, CD, and FC policies. As it is predicted, CD policy has the best performance with regard to this criterion. Because when all hospitals are set up with the specialized capacity, all customers receive service locally. However, as showed in Figure 7, offering services at all sites may be prohibitively expensive and even impossible. MP policy also has good performance with regard to this measure; it does not need long-distance travel for urgent demand. Table 4 shows that the average distances traveled by customers with urgent demand under MP policy are 16.9, 48.8, and 68.0 miles for service level 0.7 in problem sets 1, 2, and 3, respectively, while these values are 15.0, 42.3, and 60.7 miles for service level 0.9. Since the number of hospitals that should be set up increases in higher service levels, and more customers receive service locally, increasing the service level in MP policy results in a decrease in the average distance traveled by the patients for urgent demand. The similar results can be obtained for nonurgent demand (see Table 4). Moreover, CD policy's good performance in average traveled distance criterion is moderately consistent across the various service levels versus to MP policy. However, for all the problem sets, as the service level increases, the benefit of CD policy decreases a little with respect to MP one. Under FC policy, solutions require long-distance travel for both urgent and nonurgent demand in various service levels.

The results of locating traumatic brain injury treatment units in the Veterans Affairs Department and allocating of admissions to these units [31] also showed that the traveled distance and total travel cost in complete decentralization and partial centralization approaches are moderately similar (complete decentralization has better performance) and lower than full centralization one.

5.6. Percent Urgent Demand (pU) under Different Policies (MP, CD, and FC). Figure 8 investigates the impact of percentage of urgent demand on the total cost in problem set 2 under different policies (MP, CD, and FC). Results show that as the percentage of urgent demand increases, the total cost increases at an increasing rate under the MP policy. When the percentage of urgent demand is low, more patients enjoy flexibility in time or location of receiving the service. In such circumstances, scheduling procedures at alternate facilities in advance becomes easier. Intuitively this requires setting up fewer sites and therefore the solution to the MP policy moves toward FC solution. Nevertheless, when the percentage of urgent demand is high, because diversion cost of urgent demand is expensive, more sites tend to be set up to offer the specialized service and, therefore, the solution to the MP moves toward CD solution. On the other hand, as the percentage of urgent demand increases, the benefit of the MP policy or optimally placement of the specialized capacity increases relative to an FC policy but decreases relative to a CD one.

The results of using panel data that comprises information for 43 Portuguese National Health Service (NHS) hospitals for the period 2007 to 2009 [44] indicated that for most of their samples, hospitals that face higher urgent demand have higher cost. In other words, as the percentage of urgent demand increases, the total cost increases at an increasing rate.

5.7. Trade-Off between Total Cost and Diversion Cost under MP Policy. Division cost for CD and FC policies is zero. In the case of a CD policy, based on retention rates and other factors, patients refer to a hospital and each hospital has commitment to handle its own demand and in the case of a CD policy, all of the network demands must be handled by one hospital and that hospital does not have permission of rerouting the demand to another hospital in the network. In this experiment, we examine the advantage of permitting hospitals in a MHS to reroute urgent demand to another location rather than making a hospital to handle its own demand. For this purpose, we compare the cost obtained for the MP policy when percentage of urgent demand in the network is equal to 0.5 and the five levels for diversion cost of urgent demand factor (Ds) are considered.

Figure 9 shows how the total cost for problem set 1 under MP policy varies under the different diversion costs for the urgent demand. The diversion cost for the nonurgent demand is constant and set at 0.3 but the diversion cost for the urgent demand is set at five levels (0.3, 0.9, 3, 18, and infinite) to represent the potentially higher cost of diverting a unit of the urgent demand. One of the diversion costs for urgent demand factor values is equal to infinite which indicates in this case that the demand cannot be diverted. Results indicate that, as the diversion cost of the urgent demand increases, the total cost increases too. Note that when Ds = infinity, the optimal solution to the MP positions all the hospitals with specialized capacity (as the CD policy). It means that, as the diversion cost of urgent demand increases, the benefit of the MP policy or optimally placement of the specialized capacity decreases relative to a CD approach. Figure 9 shows that even if diversion cost of the urgent demand is so high, permitting a hospital in the MHS for diverting urgent demand cuts the total cost considerably. For example, when $Ds = 18$ (60 times of Dn), network's cost saving is equal to 5.3%. This value for $Ds = 3$ is equal to 20.6%.

The results of a retrospective review of administrative data from one academic medical center for the period 2003 to 2006 [45] indicated that the decreasing in overall cost is significant for periods of severe divert compared to no divert. On the other hand, permitting a hospital in the MHS for diverting the urgent demand cuts the total cost considerably.

These results especially from a managerial view are valuable. Providing flexibility (even a little amount) in handling urgent demand through diversions of patient, we can achieve considerable cost savings even if the diversion cost of urgent demand is quite high.

A number of potential extensions exist for this research. First, in the model we consider two types of demand: urgent and nonurgent. We can consider different categories of patients based on their acuity levels and differentiate the facilities of the network by their capability to handle and offer specialized services to the one, some, or entire of the acuity levels. Second, our model takes into account many key factors of the specialized services location-allocation, but even so it has some limitations. For example, it does not directly consider influence of number of patients on the quality of service. If increasing in the number of patients for handling in a particular facility enhances or decreases the quality of service in that medical facility, then exogenously such impacts must be studied. Third, the model developed concentrates on cost minimization, accessibility, and service levels, and an explicit budget limitation is not directly considered to dictate the number and maximum service level of facilities that may be set up in the MHS. Last, based on [46] patient length of stay has effect on emergency department diversions.

6. Conclusion

This paper considered a stochastic optimization model and determined the number, capacity and location of hospitals in a MHS for offering specialized services while the hospitals leverage benefits of sharing resources. Importantly, the model minimizes the total cost borne by the MHS and its patients and incorporates patient service levels at each site and overall system, patient retention rates, and types of demand. Computational results denoted that MHSs can better use resources by delivering specialized services across fewer locations and the MP policy outperformed both the CD and FC approaches for the majority of the considered factors.

Sharing specialized hospital operations in a MHS may not only leverage economies of scale to decrease costs but also result in an increase in actual and perceived quality of service. Increasing the number of patients for a specialized operation leads medical staffs to build their cumulative experience in administering the service and this can enhance the quality of service. While applying a shared resources policy has some advantages, it may result in long-distance travel for a patient to access the service. Therefore, the research described in this paper takes into account accessibility and service levels as well as the financial aspect. Our results indicate that average distance that patients (especially with urgent demand) must travel to receive specialized service in the MHSs is not extensive. Our results also show that, in MHSs with hospital locations very near to each other, we can improve the service level without considerable increase in the cost. It also can be concluded that providing flexibility (even a little amount) in handling the urgent demand through diversion of the demand can achieve considerable cost savings even if the diversion cost of the urgent demand is quite high.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] J. Yonek, S. Hines, and M. Joshi, *A Guide to Achieving High Performance in Multi-Hospital Health Systems*, Health Research & Educational Trust, 2010.

[2] V. L. Smith-Daniels, S. B. Schweikhart, and D. E. Smith-Daniels, "Capacity management in health care services: review and future research directions," *Decision Sciences*, vol. 19, no. 4, pp. 889–919, 1988.

[3] R. McClellan, *Improving Health While Reducing Cost Growth: What is Possible*, Engelberg Center for Health Care Reform at Brookings, 2014.

[4] D. Straus, "Managing healthcare spending in the United States," *The Journal of Global Health Care Systems*, vol. 2, no. 1, pp. 1–11, 2012.

[5] Centers for Medicare & Medical Services, "National health expenditures 2012 highlights," 2013, https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/index.html.

[6] Bipartisan Policy Center, *What is Driving U.S. Healthcare Spending?*, 2012, http://bipartisanpolicy.org/wp-content/uploads/sites/default/files/BPC%20Health%20Care%20Cost%20Drivers%20Brief%20Sept%202012.pdf.

[7] The Organization for Economic Cooperation and Development (OECD), "OECD Health Statistics 2014: How Does the United States Compare?" 2014, http://www.oecd.org/unitedstates/Briefing-Note-UNITED-STATES-2014.pdf.

[8] M. J. Warshawsky, "An enhanced macroeconomic approach to long-range projections of health care and social security expenditures as a share of GDP," *Journal of Policy Modeling*, vol. 21, no. 4, pp. 413–426, 1999.

[9] United States Department of Health and Human Services, "The effect of healthcare cost growth on the U.S. Economy," 2008, http://aspe.hhs.gov/health/reports/08/healthcarecost/report.

[10] Pricewaterhouse Coopers' Health Research Institute, "The price of excess identifying waste in healthcare spending," 2010, http://pwchealth.com/cgi-local/hregister.cgi/reg/waste.

[11] Association of Medical Colleges (AAMC), "Help Wanted: More U.S. Doctors Projections Indicate America Will Face Shortage of M.D.s by 2020," 2006.

[12] The Organization for Economic Cooperation and Development (OECD), "Health at a Glance: 2013," 2013, http://www.oecd.org/els/health-systems/Health-at-a-Glance-2013.pdf.

[13] U.S. Census Bureau, *The Next Four Decades the Older Population in the United States: 2010 to 2050*, 2010, http://www.census.gov/prod/2010pubs/p25-1138.pdf.

[14] M. Larkin, "Shortage strategies: retaining the experienced nurse," *The Journal of Nursing Administration*, vol. 37, no. 4, pp. 162–163, 2007.

[15] Healthcare Financial Management Association (HFMA)'s Value Project, "The value journey Organizational Road Maps for Value-Driven Health Care," 2012, http://bipartisanpolicy.org/wp-content/uploads/sites/default/files/BPC%20Health%20Care%20Cost%20Drivers%20Brief%20Sept%202012.pdf.

[16] C. M. Christensen, J. H. Grossman, and J. Hwang, *The Innovator's Prescription: A Disruptive Solution for Health Care*, McGraw-Hill Education, 2009.

[17] Center for Health Market Innovations, "Narayana Hrudayalaya Hospital (NH)," 2014, http://healthmarketinnovations.org/program/narayana-hrudayalaya-hospital-nh.

[18] A. Shih, K. Davis, S. Schoenbaum et al., "Organizing the U.S. Health Care Delivery System for High Performance," The Commonwealth Fund, 2008, http://www.commonwealthfund.org/usr_doc/Shih_organizingushltcaredeliverysys_1155.pdf.

[19] M. S. Daskin and L. K. Dean, *Location of Health Care Facilities*, Operations Research and Health Care, Springer, New York, NY, USA, 2004.

[20] M. Panwar and K. Rathi, "Social sustainability: contextual facility location planning model for multi-facility hierarchical healthcare system in India," *International Journal of Applied Engineering Research*, vol. 9, no. 3, pp. 275–284, 2014.

[21] G. Buzai, "Location–allocation models applied to urban public services: spatial analysis of primary health care centers in the city of Luján, Argentina," *Hungarian Geographical Bulletin*, vol. 62, no. 4, pp. 387–408, 2013.

[22] US Agency for International Development (USAID), *Improving Resource Allocation in Kenya's Public Health Sector*, US Agency for International Development (USAID), Washington, DC, USA, 2010.

[23] N. Coskun and R. Erol, "An optimization model for locating and sizing emergency medical service stations," *Journal of Medical Systems*, vol. 34, no. 1, pp. 43–49, 2010.

[24] P. Mitropoulos, I. Mitropoulos, I. Giannikos, and A. Sissouras, "A biobjective model for the locational planning of hospitals and health centers," *Health Care Management Science*, vol. 9, no. 2, pp. 171–179, 2006.

[25] H. Shavandi and H. Mahlooji, "Fuzzy hierarchical location-allocation models for congested systems," *Journal of Industrial and Systems Engineering*, vol. 1, no. 2, pp. 171–189, 2007.

[26] H. Afshari and Q. Peng, "Challenges and solutions for location of healthcare facilities," *Journal of Industrial Engineering and Management*, vol. 3, article 127, 2014.

[27] E. Cetin and L. S. Sarul, "A blood bank location model: a multiobjective approach," *European Journal of Pure and Applied Mathematics*, vol. 2, no. 1, pp. 112–124, 2009.

[28] H. K. Smith, P. R. Harper, and C. N. Potts, "Bicriteria efficiency/equity hierarchical location models for public service application," *Journal of the Operational Research Society*, vol. 64, no. 4, pp. 500–512, 2013.

[29] C. S. Phibbs, L. C. Baker, A. B. Caughey, B. Danielsen, S. K. Schmitt, and R. H. Phibbs, "Level and volume of neonatal intensive care and mortality in very-low-birth-weight infants," *The New England Journal of Medicine*, vol. 356, no. 21, pp. 2165–2175, 2007.

[30] S. Ketabi, "Efficiency measurement of cardiac care units of Isfahan hospitals in Iran," *Journal of Medical Systems*, vol. 35, no. 2, pp. 143–150, 2011.

[31] M. J. Côté, S. S. Syam, W. B. Vogel, and D. C. Cowper, "A mixed integer programming model to locate traumatic brain injury treatment units in the Department of Veterans Affairs: a case study," *Health Care Management Science*, vol. 10, no. 3, pp. 253–267, 2007.

[32] S. S. Syam and M. J. Côté, "A location-allocation model for service providers with application to not-for-profit health care organizations," *Omega*, vol. 38, no. 3-4, pp. 157–166, 2010.

[33] S. S. Syam and M. J. Côté, "A comprehensive location-allocation method for specialized healthcare services," *Operations Research for Health Care*, vol. 1, no. 4, pp. 73–83, 2012.

[34] S. Mahar, K. M. Bretthauer, and P. A. Salzarulo, "Locating specialized service capacity in a multi-hospital network," *European Journal of Operational Research*, vol. 212, no. 3, pp. 596–605, 2011.

[35] P. R. Harper and A. K. Shahani, "Modelling for the planning and management of bed capacities in hospitals," *Journal of the Operational Research Society*, vol. 53, no. 1, pp. 11–18, 2002.

[36] P. R. Harper, A. K. Shahani, J. E. Gallagher, and C. Bowie, "Planning health services with explicit geographical considerations: a stochastic location-allocation approach," *Omega*, vol. 33, no. 2, pp. 141–152, 2005.

[37] M. Ndiaye and H. Alfares, "Modeling health care facility location for moving population groups," *Computers & Operations Research*, vol. 35, no. 7, pp. 2154–2161, 2008.

[38] G. Melnick and E. Keeler, "The effects of multi-hospital systems on hospital prices," *Journal of Health Economics*, vol. 26, no. 2, pp. 400–413, 2007.

[39] C. Blanda and K. Gould, "Financing options for large non-profit hospitals and multi-hospital systems," White Paper III, 2013.

[40] C. W. Burt, L. F. McCaig, and R. H. Valverde, "Analysis of ambulance transports and diversions among US emergency departments," *Annals of Emergency Medicine*, vol. 47, no. 4, pp. 317–326, 2006.

[41] S. Winitzki, "A handy approximation for the error function and its inverse," A lecture note obtained through private communication, 2008.

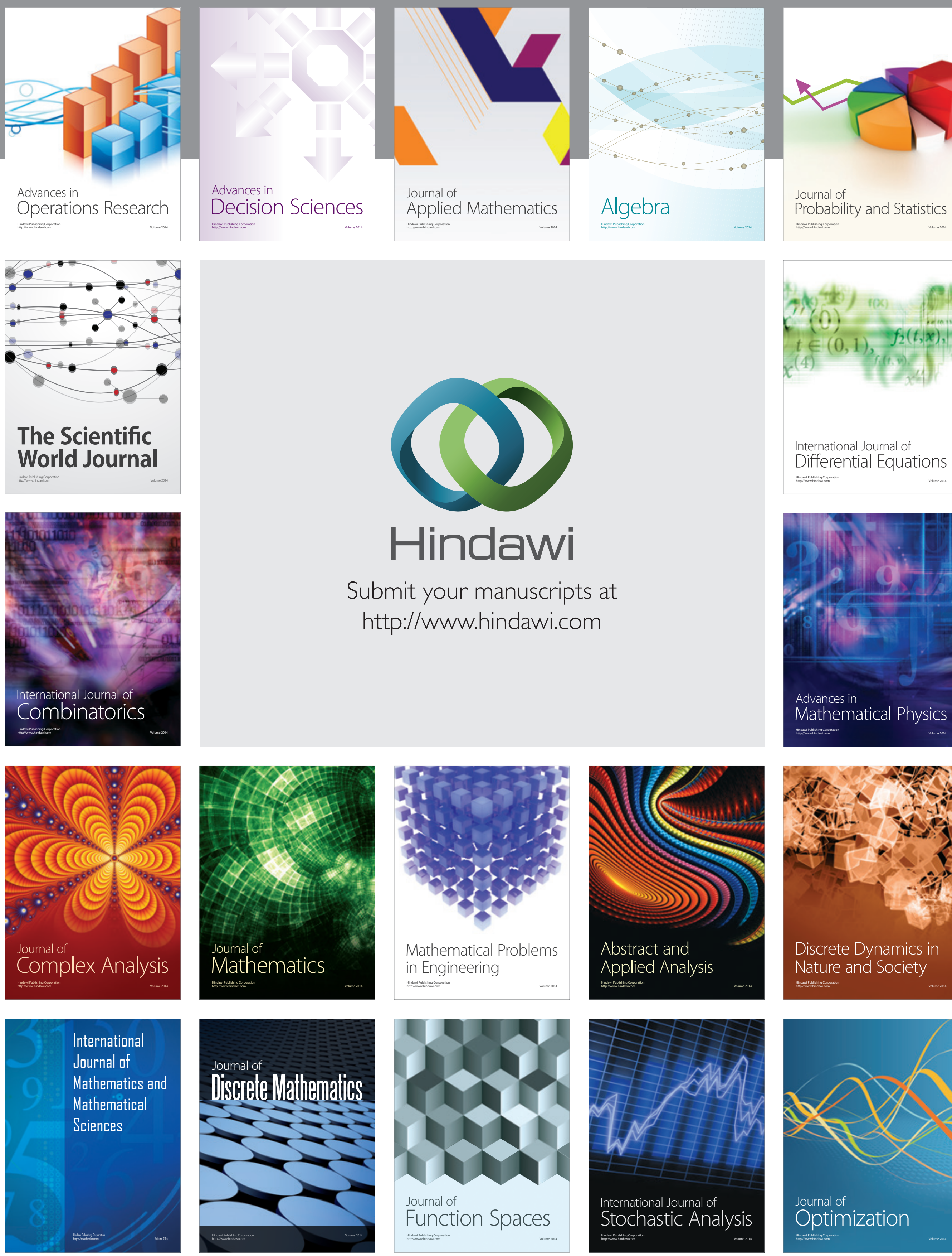
[42] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, San Francisco, Calif, USA, 1979.

[43] The Solver Manuals, http://www.gams.com/help/topic/gams.doc/solvers/allsolvers.pdf.

[44] A. Almeida and J. Cima, "Demand uncertainty and hospital costs: an application to Portuguese NHS hospitals," FEP Working Papaer 499, University of Porto, 2013.

[45] D. A. Handel and K. J. McConnell, "The financial impact of ambulance diversion on inpatient hospital revenues and profits," *Academic Emergency Medicine*, vol. 16, no. 1, pp. 29–33, 2009.

[46] A. Kolker, "Process modeling of emergency department patient flow: effect of patient length of stay on ED diversion," *Journal of Medical Systems*, vol. 32, no. 5, pp. 389–401, 2008.


Advances in
Operations Research
Advances in
Decision Sciences
Journal of
Applied Mathematics
Algebra
Journal of
Probability and Statistics
The Scientific
World Journal
International Journal of
Differential Equations
Hindawi
Submit your manuscripts at
http://www.hindawi.com
International Journal of
Combinatorics
Advances in
Mathematical Physics
Journal of
Complex Analysis
Journal of
Mathematics
Mathematical Problems
in Engineering
Abstract and
Applied Analysis
Discrete Dynamics in
Nature and Society
International
Journal of
Mathematics and
Mathematical
Sciences
Journal of
Discrete Mathematics
Journal of
Function Spaces
International Journal of
Stochastic Analysis
Journal of
Optimization