Operations Research for Health Care 1 (2012) 73–83

Contents lists available at SciVerse ScienceDirect

Operations Research for Health Care

journal homepage: www.elsevier.com/locate/orhc

A comprehensive location-allocation method for specialized healthcare services

Siddhartha S. Syam [a,1], Murray J. Côté [b,*]

[a] *Department of Management, Marquette University, Milwaukee, WI 53201-1881, United States*

[b] *Department of Health Policy and Management, Texas A&M Health Science Center, College Station, TX 77843-1266, United States*

ARTICLE INFO

Article history:
Received 9 January 2012
Accepted 11 September 2012
Available online 5 October 2012

Keywords:
Design of experiments
Integer programming
Location-allocation
Simulated annealing

ABSTRACT

This paper focuses on the development, solution, and application of a location-allocation model for specialized health care services such as the treatment and rehabilitation necessary for strokes or traumatic brain injuries. The model is based on our experience with the Department of Veterans Affairs' integrated service networks. The model minimizes the total cost borne by the health system and its patients and incorporates admission acuity levels, service proportion requirements, and admission retention rates. A common resource constraint is introduced at the facility level since treatment of multiple acuity levels involves the pooling of common resources. Realistic instances of the model with 20 potential service locations, 50 admission districts and up to five open treatment units for three levels of severity are solved in about 300 seconds. The applicability of the model is tested by an extensive managerial experiment using data derived from one of the Department of Veterans Affairs specialized healthcare services. We investigate the effects of five critical factors: (1) the degree of service centralization, (2) service level mandates by acuity, (3) lost admission cost by acuity, (4) facility overload penalty cost by acuity and (5) target utilization level by acuity and treatment unit. We examine the countervailing forces present in making healthcare service location decisions and the resulting tradeoffs from the implicitly multiobjective nature of the system. The experiment and analysis demonstrate that the major factors of the experiment have a significant bearing on the optimal assignment of admission districts to treatment units.

1. Introduction and motivation

In the United States, healthcare remains an area of crucial concern for millions as evidenced by the current ongoing debate over the federal healthcare reform bill. Healthcare providers are pressured by two conflicting dimensions: ever-increasing healthcare costs and the public demand for access to cutting-edge treatment. As a result, healthcare providers have virtually no choice but to constantly seek to become as efficient as possible in all aspects of their operations.

The original study that this paper extends is based on a perceived need to improve the delivery of specialized health services at the Department of Veterans Affairs (VA) in terms of effectiveness and efficiency [1]. The VA is the primary organization charged with providing healthcare to veterans and the original study was based upon a funded research project that aimed to determine the optimal location of traumatic brain injury (TBI) treatment units for VA medical centers. As a not-for-profit service organization, the VA has to define optimality using a multi-objective approach where the cost of providing care (i.e., efficiency) and the level and extent of healthcare provided in terms of access and availability (i.e., effectiveness) are generally viewed as equally critical objectives. The research described in this paper extends the previous study and model in meaningful and important ways that we expect have applicability to nonprofit healthcare providers beyond the VA.

The specific improvements to efficiency and effectiveness included in this paper are as follows: (a) Multiple levels of severity with capacity limits by facility and severity (sometimes called acuity). The justification for this is that treatment costs (including fixed costs of specialized equipment) are much higher for high levels of acuity. In turn, the high costs have to be offset by scale economies which make it impossible to replicate the same capacity to treat severe cases in every facility. (b) Common resource constraints at each facility. The need for this stems from the fact that many kinds of resources such as physicians, supplies, storage facilities, and operating theaters are common across severity levels but used at varying rates at different levels. Common resource availability constrains treatment at all severity levels in a facility. (c) Service level mandates, overload penalties, lost admission costs, and target utilization by severity level are factors introduced in the optimization model, whose impact is investigated in a managerial experiment described later. In this context, for each level of acuity (i) a service level mandate is a requirement to serve a particular proportion of

* Corresponding author. Tel.: +1 979 458 3031; fax: +1 979 458 0656.
E-mail addresses: siddhartha.syam@marquette.edu (S.S. Syam), mjcote@srph.tamhsc.edu (M.J. Côté).
1 Tel.: +1 414 288 5462; fax: +1 414 288 5754.

2211-6923/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.orhc.2012.09.001

eligible patients, (ii) a target utilization percentage is a proportion of capacity that management wishes to utilize in order to balance costly load on a facility versus the need to serve as many patients as possible (iii) an overload penalty is a cost associated with exceeding a facility's utilization target, and (iv) lost admission costs are federal funding not obtained if patients are not served (admitted).

In the model developed in this paper, the service level mandate is an explicit constraint while the other factors are included in the objective function so that violations of the corresponding managerial targets are penalized in accordance with the levels of the factors. The factors other than service level mandate are associated with targets that often trade off against each other rather than explicit requirements and the managerial experiment described later tracks the costs of not achieving the targets as the factor levels are varied. Instances of tradeoffs are service level versus overload and lost admissions versus target capacity utilization at a facility. Lost admissions and service level are related to effectiveness and overload and capacity utilization are related to efficiency (cost).

This paper is organized as follows. The following section provides the literature review germane to our research. Section 3 presents the optimization model and defines the relevant decision variables and model coefficients. Following the model, Section 4 describes an extensive managerial experiment meant to evaluate the impact of important managerial parameters and discusses the results that we obtained from that experiment. Last, concluding remarks and future research directions are provided in Section 5.

2. Literature review

The purpose of location-allocation models is to concurrently determine optimal facility locations and the assignment of customers to open facilities. Since the research literature in facility location is vast, no attempt is made here to provide a comprehensive review. Rather, we direct the reader to a full review of general facility location models and the methods used to solve them found in Love et al. [2] and Cornuéjols et al. [3]. Instances of the application of location-allocation models to healthcare issues include hospital location in rural regions [4], geographical considerations in healthcare planning [5], locating blood banks [6], service mix and location in managed healthcare [7], trauma care involving hospitals and ambulances [8], the reorganization of liver transplant regions [9], and optimizing the location of specialized treatment facilities [10]. An extensive survey of the research regarding the location of healthcare facilities may be found in Daskin and Dean [11].

The location-allocation model developed in this paper involves mixed integer programming and is rooted in the classic uncapacitated facility location (UFL) model [12]. In practice, the UFL model is often modified to provide feasible templates of frequently occurring service or business scenarios which correspond to more complicated models. For example, the introduction of facility capacity limits leads to the capacitated facility location problem, and limiting the permitted number of open facilities leads to the p-median problem. The different versions of the UFL model such as one in which every location can be a facility and demand node and another in which facilities are limited to a subset of nodes and variants of the UFL model are known to be difficult to solve (particularly for large instances) since they belong to the NP-complete class of problems [13]. However, it may be noted that the variants may be significantly harder to solve than the UFL model since they contain constraints not found in the UFL model. Consequently, optimal solutions to our type of problem are difficult to obtain and very large instances may necessitate specialized solution approaches or heuristics such as Lagrangian relaxation [14], simulated annealing [15], and dual ascent [16]. While these sophisticated heuristics offer the advantage of reduced computational times, they generally provide near-optimal rather than optimal solutions to complex problems. After some experimentation, we found that the commercial-grade general-purpose optimization software CPLEX-OPL [17] solved most instances of problems in our research environment in reasonable computing time. Consequently, we adopted CPLEX-OPL as our solver engine for the model developed here. This model is described in the next section.

3. The optimization model

The primary goal of the optimization model developed in this paper is to provide a mathematical framework that incorporates the primary criteria of the VA as it seeks to serve veterans: (1) the cost of providing service and (2) the service level provided to the VA's patients. The costs included in the model include fixed costs, treatment costs, travel costs, lodging costs, lost service costs, and overloading penalty costs. We note that patients or their families bear the cost of travel and hotel lodging of family members who accompany the patient to a given facility and other costs are borne by the VA. The service level, for each level of acuity, is defined as the proportion of eligible admissions served by the VA for a given geographical area. The model incorporates retention rates by distance traveled and these are incorporated according to multiple levels of acuity (i.e., reflecting the observation that patients are willing to travel relatively longer distances for higher levels of acuity). For the purposes of our model and its generalized application beyond the VA, we refer to acuity as representing the general medical condition of a potential patient where higher acuity patients will require longer lengths of stay and more resources than lower acuity patients. Our purpose behind this is simply to demarcate differing patient classes. A similar definition is applied to service levels in that we recognize the healthcare organization may need to maintain a certain volume to justify (either economically or by the decision-maker's prerogative) why a service is provided at a given facility. Last, we treat the eligibility of veterans as potential patients as detailed in [1].

An important application of our model is to analyze the tradeoff between a centralized capacity policy with a relatively small number of treatment units generally located in large metropolitan areas versus a decentralized capacity policy with a relatively large number of geographically disperse treatment units, some of which may be located in rural or low-population density areas. While the primary analysis is in terms of cost, our model also assists decision-makers in fulfilling the service mission of the VA, including examining secondary objectives such as patient travel and lodging costs. A by-product of the degree of centralization adopted by the organization that is captured by our model is the level of employment corresponding to various policies. In government and/or unionized work environments, the staffing level may be important enough to be a criterion in its own right. While we do not view staffing in that manner in our research, our model allows the decision-maker to analyze the consequences of alternative staffing level restrictions.

In addition to retention rates by acuity, the model includes constraints that ensure that the capacity of each potential treatment unit by acuity level is not exceeded and that mandatory service levels by acuity level are met. Parameter definitions ensure that the fixed cost of a treatment unit is a piecewise linear function of that unit's capacity. It also includes a common resource limit for each medical center where the treatment unit may be located such that the common resources are those that are used by all acuity levels. There are also restrictions on the number of open treatment units by acuity level. It should be noted that the service mission of the VA is maintained through constraints that enforce service level mandates for each level of acuity (e.g., at least 70% of a target