Proceedings of the 2012 Winter Simulation Conference
C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A.M. Uhrmacher, eds

A SIMULATION-BASED ITERATIVE METHOD FOR A TRAUMA CENTER – AIR AMBULANCE LOCATION PROBLEM

Taesik Lee

KAIST
335 Gwahangno, Yuseong-gu
Daejeon, 305-701, REPUBLIC OF KOREA

Hoon Jang

KAIST
335 Gwahangno, Yuseong-gu
Daejeon, 305-701, REPUBLIC OF KOREA

Soo-Haeng Cho

Tepper School of Business, Carnegie Mellon Univ.
5000 Forbes Avenue, Pittsburgh
PA, 15213, USA

John G. Turner

The Paul Merage School of Business, UC Irvine
SB 338, Irvine
CA, 92697-3125, USA

ABSTRACT

Timely transport of a patient to a capable medical facility is a key factor in providing quality care for trauma patients. This paper presents a mathematical model and a related solution method to search for optimal locations of trauma centers and air ambulances. The complicatedness of this problem stems from the characteristic that optimal locations for the two resources are coupled with each other. Specifically, this coupling makes it difficult to develop *a priori* estimates for the air ambulance's busy fraction, which are required to construct a probabilistic location model. We propose a method that uses integer programming and simulation to iteratively update busy fraction parameters in the model. Experimental results show that the proposed method is valid and improves the solution quality compared to alternative methods. We use real data on Korean trauma cases, and apply the method to the design of a trauma care system in Korea.

1 INTRODUCTION

Trauma refers to major injury that can lead to a loss of life if the appropriate level of care is not administered in a timely fashion. Trauma care requires highly skilled medical workers and expensive resources that are often beyond the capability of a regular ER. Many countries have accordingly developed systems dedicated to treating trauma patients, referred to as trauma care systems. In designing a trauma care system it is necessary to address a variety of aspects including facilities (trauma centers), patient transport, governance, human resources, financing, and quality control. This paper addresses a location problem where optimal locations of trauma centers and helicopter ambulances are sought.

Timely transport to appropriate medical facilities is one of the critical performance measures for a trauma care system. Delays in providing clinical intervention affect the survivability of trauma patients. In medicine, it is generally accepted that a patient should be given appropriate medical care within an hour from the moment of an incident. From this timeliness requirement arises the need to optimally locate trauma centers and transport systems across the country.

Studies on healthcare facility location problems have a long history in operations research. Among many types of location problems, our problem falls under the category of a covering problem. In a covering problem, one tries to find a set of optimal locations of facilities, e.g. ambulances, that can *cover* all or

978-1-4673-4782-2/12/$31.00 ©2012 IEEE

a maximum number of demands, where coverage of a facility is commonly defined by the distance or time from the facility location.

Covering problems are classified into two groups. One is to minimize the number of facilities to cover all demand points. Toregas et al. (1971) first introduced this type of problem for ambulance locations, calling it the Location Set Covering Problem (LSCP). Alternatively, one might be interested in locating a limited number of resources that maximize the number of covered demand points. This type of problem is referred to as the Maximal Covering Location Problem (MCLP), and was initially proposed by Church and ReVelle (1974). Since the introduction of these two basic model formulations, a number of extensions and modifications have been proposed with more realistic features being considered.

Two extensions of the basic models are particularly relevant to our study. One is a capacitated covering problem incorporating capacity constraints for each facility (Chung et al. 1983; Current and Storbeck 1988; Pirkul and Schilling 1991; Haghani 1996). A trauma center cannot serve an infinite number of patients, and each center has a limit for the number of patients it can cover. A capacitated covering problem is computationally much harder to solve than an uncapacitated version.

The other extension is to consider probabilistic availability of resources such as ambulances. A patient at demand location i, even when it is within the coverage of an ambulance at location j, may not receive the service if the ambulance is busy serving another patient. Many probabilistic models have been developed to incorporate ambulances' probabilistic availability (Daskin 1983; ReVelle and Hogan 1989; Ball and Lin 1993; Borras and Pastor 2002). These models typically use an approximate estimate for the probability of a vehicle being unavailable, which is known as the busy fraction. The busy fraction is often defined for a patient i as the ratio of the total workload of nearby ambulances to the operating hours of those vehicles. One problem with this simplifying approximation is that the servers are assumed to be independent. In an attempt to obtain a more accurate estimate, Larson (1974) proposed a hypercube queuing model where the dependency among ambulances is considered, and later developed an approximate hypercube queuing model to reduce computational loads of the original model (Larson 1975). Larson's work was followed by many studies aimed at incorporating more realistic model features and integrating them into a probabilistic location model (Jarvis 1975; Ingolfsson et al. 2008).

Apart from the mathematical programming approaches, simulations have been used as an important tool for dealing with such problems. Simulations offer advantages over mathematical models in their modeling flexibility. They allow easier incorporation of the details and the stochastic nature of real world operations than mathematical models. The most common mode of using simulations in the context of medical facility/resource location problems is to use a simulation model to evaluate the performance of solutions obtained by a mathematical model (Goldberg et al. 1990; Henderson et al. 2004; Gunes et al. 2005; Aringhieri et al. 2007). Some of the rather recent simulation studies in this area use simulations as a major part of optimization (Silva et al. 2010). These simulation optimization approaches are largely based on meta-heuristic methods to find near-optimal solutions. However, there are many practical challenges in simulation optimization approaches, including the difficulty of assessing optimality of obtained solutions and possibly long computation time.

In this study, we address a problem of simultaneously locating a limited number of trauma centers and helicopter ambulances to maximize service coverage. A similar problem has been studied by Branas and ReVelle (2001), but their model is simpler in that it does not consider probabilistic availability or capacity constraints. Our problem presents an interesting challenge: locating two different types of resources – trauma centers and helicopter ambulances – that are coupled with each other. This means that location decisions for these resources mutually affect demands for each other. This coupling is particularly problematic when we build a probabilistic location problem formulation. Solving a probabilistic location problem generally requires that a parameter, called a busy fraction, be known. However, for our problem, the coupling makes it difficult to derive a reasonable estimate of the busy fraction *a priori*. To tackle this difficulty, we develop a scheme that uses a mathematical program model and discrete event simulation iteratively. Experimental results suggest that the proposed iterative method is valid and effective.

This paper is organized as follows. In section 2, we describe the problem in detail, and present a probabilistic covering model. A proposed iterative approach using a simulation model to overcome the difficulties in estimating the busy fraction parameter is discussed in Section 3. Section 4 describes experimental settings. Results of the experiments are discussed in section 5, followed by concluding remarks in section 6.

2 PROBLEM DESCRIPTION

2.1 Problem Description

We consider a location problem for trauma centers in conjunction with helicopter ambulances. The objective is to maximize the expected number of trauma patients that can be transported to a trauma center within 60 minutes from the moment of an incident. Our model includes a capacity constraint for each trauma center: each center has a capacity limit for the maximum permissible number of patients per given time period. For the helicopter ambulance component of the model, the availability of helicopter ambulances is modeled by incorporating a busy fraction. The following is a list of additional operational features and modeling assumptions:

- The number of trauma centers m and the number of helicopter ambulances q are finite and given.
- Each of the m trauma centers will be established in one of the $|\boldsymbol{W}|$ candidate locations ($|\boldsymbol{W}| \geq m$).
- q helicopter ambulances will be allocated to heliport bases in $\boldsymbol{H}$; more than one helicopter ambulance can be allocated to one heliport base.
- The total expected number of trauma patients who are transported within the 60-min requirement is the sum of patients transported by ground ambulances N_G and those transported by helicopter ambulances N_A.
- A patient whose location is within 30 km from an available trauma center will always be transported by a ground ambulance.
- A helicopter ambulance dispatched from heliport base h to serve a demand region i by bringing a patient to trauma center j performs the following sequence of events:
  - Take-off at heliport base h (6 min).
  - Travel h to patient i (at a constant velocity of 180km/h).
  - Field management at the scene (8 min).
  - Transport i to j (at a constant velocity of 180km/h).
  - Landing and hand-off the patient at j (6 min).
  - Return to base h from j (at a constant velocity of 180km/h).
  - Land and refuel at h (5 min).
- There are ample ground ambulances, and thus there is no shortage of ground ambulances for any service request.
- A patient whose location is outside the 30km boundary will be transported by the nearest available helicopter ambulance.
- To transport a patient within 60-min, the sum of distances d_{hij} from a heliport base h to the demand region i and from the demand region i to trauma center j should be no greater than 120 km; that is, for patient i to be covered by a heliport base h and a trauma center j, $d_{hij} \leq 120$ km.
- Patients whose location is outside the 30 km boundary and is also outside of helicopter ambulance coverage ($d_{hij} > 120$ km) are not covered; that is, these patients are *discarded*.

2.2 Optimization Model

We first define the indices i, j, and h to denote demand regions, trauma center sites, and heliport base sites. Unlike typical ambulance location problems where we work with a demand region and ambulance location pair (i, j) – j being an index for the ambulance location, our problem works with a triplet of heliport base, demand region, and trauma center, (h, i, j). Demands from region i can then be grouped by their destination trauma center j, and we designate each subgroup as a demand pair (i, j).

Let $\boldsymbol{W}$ and $\boldsymbol{V}$ represent the set of all candidate sites of trauma centers and all demand regions to be served, respectively. For demand region i, $\boldsymbol{W}_i^G$ and $\boldsymbol{W}_i^A$ are defined. $\boldsymbol{W}_i^G$ is a set of trauma center candidate sites within the 30-km boundary from demand region i. $\boldsymbol{W}_i^A$ is a set of trauma center candidate sites for which there exists a heliport base h with $d_{hij} \leq 120$ km. In other words, if a trauma center is established at one of $\boldsymbol{W}_i^G$, patients from demand region i are transported by a ground ambulance within an hour to the center. Similarly, if a center is located at one of $\boldsymbol{W}_i^A$, they can be transported by an helicopter ambulance within an hour to the center. $\boldsymbol{V}_j^G$ and $\boldsymbol{V}_j^A$ are defined for trauma center j in the same manner. $\boldsymbol{V}_j^G$ is the set of demand regions within a 30km boundary from the center j, and $\boldsymbol{V}_j^A$ is the set of demand regions for which there exists a heliport base h with $d_{hij} \leq 120$ km. $\boldsymbol{H}_{ij}$ is a set of heliport bases that can serve demand pair (i, j) within an hour. A heliport base h belongs to $\boldsymbol{H}_{ij}$ if $d_{hij} \leq 120$ km. f_i is the number of patients at demand region i, and c_j denotes the capacity limit for trauma center j.

In the optimization model, we use four main decision variables:

- x_{ij}^G is an integer variable to indicate the size – number of patients – of demand pair (i, j) served by ground ambulances.
- x_{ij}^A is an integer variable to indicate the size – number of patients – of demand pair (i, j) served by helicopter ambulances.
- z_h is an integer variable to denote the number of helicopter ambulances allocated to heliport base h.
- y_j is a binary variable, and is 1 if a trauma center is established at candidate site j, and 0 otherwise.

There are two auxiliary variables used in the model:

- $x_{ij,k}^A$ is an integer variable, and it takes the value of x_{ij}^A if served by k helicopter ambulances, and zero otherwise.
- $u_{ij,k}^A$ is a binary variable, and is 1 if $x_{ij,k}^A$ has a positive value, and 0 otherwise.

A full formulation of the location problem for trauma centers and helicopter ambulances is provided below. The objective function (1) maximizes the expected number of successfully transported patients by ground ambulances or helicopters.

$$\textit{maximize} \quad \sum_{i\in I}\sum_{j\in W_i^G} x_{ij}^G + \sum_{i\in I}\sum_{j\in W_i^A}\sum_{k=1}^{q}(1-(\rho_{ij,k})^k)x_{ij,k}^A \tag{1}$$

The first term is the number of patients transported by ground ambulances, and the second term represents the expected number of patients transported by helicopter ambulances. $\rho_{ij,k}$ in (1) is the average busy fraction of k helicopter in $\boldsymbol{H}_{ij}$, and thus $(1-(\rho_{ij,k})^k)$ is the probability that at least one of the k helicopters is available at the moment. $\rho_{ij,k}$ is discussed in detail in the next section.

The following set of constraints is applied:

$$\sum_{j\in W_i^G} x_{ij}^G + \sum_{j\in W_i^A} x_{ij}^A \leq f_i \qquad \text{for all } i \tag{2}$$

$$\sum_{i\in V_j^G} x_{ij}^G + \sum_{i\in V_j^A} x_{ij}^A \leq c_j y_j \qquad \text{for all } j \tag{3}$$

$$\sum_{k=1}^{q} u_{ij,k}^A \leq 1 \qquad \text{for all } (i,j) \text{ where } \boldsymbol{H}_{ij}\neq\phi \tag{4}$$

$$x_{ij,k}^A \leq f_i u_{ij,k}^A \qquad \text{for all } (i,j) \text{ where } \boldsymbol{H}_{ij}\neq\phi;\ k=1,\ldots,q \tag{5}$$

$$\sum_{k=1}^{q} x_{ij,k}^A = x_{ij}^A \qquad \text{for all } (i,j) \text{ where } \boldsymbol{H}_{ij}\neq\phi \tag{6}$$

$\sum_{k=1}^{q} k u_{ij,k}^{A} \leq \sum_{h \in H_{ij}} z_h$ for all (i, j) where $\boldsymbol{H}_{ij} \neq \phi$ (7)

$\sum_{j \in J} y_j \leq m$ for all j (8)

$\sum_{h \in H} z_h \leq q$ for all h (9)

Constraint (2) requires that the number of patients from demand region i served by ground or helicopter ambulances does not exceed the total number of patients at i. Constraint (3) limits the capacity for each trauma center. Constraints (4) – (6) relate $x_{ij,k}^{A}$ to a dummy variable $u_{ij,k}^{A}$ and the main decision variable x_{ij}^{A}. Constraint (7) requires that at least k helicopters be allocated to heliport bases in $\boldsymbol{H}_{ij}$. Constraints (8) and (9) limit the number of trauma centers and helicopter ambulances at no more than m and q, respectively.

3 PROPOSED SOLUTION METHOD

3.1 Iterative Method to Estimate Busy Fraction

In probabilistic ambulance location models, the busy fraction ρ_i for demand region i is defined as the duration for which nearby ambulances are busy serving a patient as a fraction of the total operating hours of those ambulances (Daskin 1983). In our model, we use refined indices for the busy fraction, denoted as $\rho_{ij,k}$. Index ij is used to indicate that it is defined for a demand region and trauma center pair (i, j). It is also indexed by the number of helicopter ambulances k located at heliport bases in $\boldsymbol{H}_{ij}$. $\rho_{ij,k}$ is written as

$$\rho_{ij,k} = \frac{\sum_{h \in H_{ij}} L_h}{24k} \qquad \text{for all } (i, j) \text{ for which } \boldsymbol{H_{ij}} \neq \emptyset \qquad (9)$$

where L_h is the amount of service time worked by all helicopters at heliport base $h \in \boldsymbol{H_{ij}}$. The constant, 24, in the denominator indicates that we assume 24-hr operation for helicopter ambulances.

Unlike the prior studies on probabilistic ambulance location models, it is not possible to presume a reasonable busy fraction value for our problem. This is due to the dependence of the total workload – numerator in (9) –on the location solutions for trauma centers. That is, locations of trauma centers determine the demands for helicopter ambulances. This is referred to in this paper as coupling between the two resource locations. Thus, we have a dilemma where we need $\rho_{ij,k}$ to solve the optimization problem, but it can be properly estimated only after the optimization problem is solved. This dilemma is our motivation to use a simulation model as a means to iteratively update $\sum_{h \in H_{ij}} L_h$.

The essence of the proposed method is to estimate $\sum_{h \in H_{ij}} L_h$ in (9) by iteratively using the integer program to obtain an optimal location solution for given $\sum_{h \in H_{ij}} L_h$ and then running a simulation to recompute $\sum_{h \in H_{ij}} L_h$ for the obtained location solution. The procedure begins with solving the math model with an initial guess, $\rho_{ij,k}^{0}$. As an initial guess, we set $\rho_{ij,k}^{0} = 0$ for all (i, j) pairs. This is equivalent to solving a deterministic version of the location problem. With this initial guess, we obtain an initial solution for locations of trauma centers (y_j^0) and helicopter ambulances (z_h^0). This initial solution is fed into the simulation model, and we run a simulation. From the simulation runs, we collect information on the busy hours for each helicopter ambulance and summate them by each heliport base to compute L_h ($h \in \boldsymbol{H}$). We can then update the estimate for $\sum_{h \in H_{ij}} L_h$. With the updated estimates, we solve the integer program for the next round of iteration. We repeat this process until the location solution at the $(n+1)^{\text{st}}$ step $\{ y_j^{n+1}: z_h^{n+1} \}$ is the same as the solution from the previous step $\{ y_j^{n}: z_h^{n} \}$.

During the iteration, we may obtain a location solution where there is no helicopter allocated around some demand pair (i, j). That is, for all $h \in \boldsymbol{H_{ij}}$, $z_h = 0$. When this happens, we have no data on L_h from the simulation run, and thereby no information on $\sum_{h \in H_{ij}} L_h$. If we arbitrarily choose to set $\rho_{ij,k}$ to be zero, this would result in an illogical solution: locating one helicopter is equivalent to locating an infinite number of helicopters, leading to a static coverage solution for that region. To avoid this situation, we exclude the demand pair (i,j) for the next iteration. This rule leads to fast convergence in the proposed algorithm, because it reduces the set of demand points to be covered as the iteration proceeds.

The proposed procedure described above is summarized in Algorithm 1 below.

Algorithm 1

Step 0. Set $n = 0$; Set the initial value of the busy fraction to zero, i.e. $\rho_{ij,k}{}^{n} = 0$ for all i, j, k.
Step 1. Solve the IP (Integer Program) to obtain the location solution $\{y_j^{0}: z_h{}^{0}\}$.
Step 2. Set $n = n+1$; Simulate with $\{y_j^{n-1}: z_h{}^{n-1}\}$; update using the simulation results.
Step 3. Solve the IP with $\rho_{ij,k}{}^{n}$ with to obtain $\{y_j^{n}: z_h{}^{n}\}$
Step 4. IF $\{y_j^{n}: z_h{}^{n}\} = \{ y_j^{n-1}: z_h{}^{n-1}\}$ OR $n = n_{\max}$
– $\{y_j^{*}: z_h{}^{*}\} = \{y_j^{n}: z_h{}^{n}\}$; Stop
ELSE
– Go to Step 2

It should be noted that we can in fact make estimates without using a simulation in Algorithm 1. This can be done by replacing Step 2 in Algorithm 1 with a reverse-counting procedure. Given a location solution (from the previous iteration step), we can identify coverage by a heliport h and trauma center j. This coverage is an elliptical region with h and j as foci (Figure 1). We can then simply count the number of helicopter transports at each demand region i ($x_{ij}{}^{A}$) in the coverage. This gives an estimate for L_h. While this reverse-counting procedure is simple and we do not need to construct a simulation model, it involves some arbitrariness in allocating demands to heliport bases. This is discussed further in the results section.

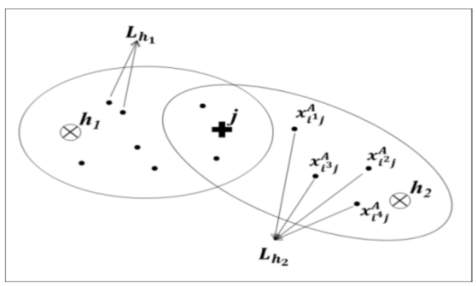

Figure 1: Workload of a heliport, L_h, is estimated by counting the number of patients at each demand region ($x_{ij}{}^{A}$) within the coverage of heliport base h and trauma center j (h, j), indicated as an elliptical area.

3.2 Simulation Model

We build a discrete event simulation model to assess the workload of each helicopter so that we can make estimates for $\sum_{h\in H_{ij}} L_h$. Operational logic of the simulation model is carefully designed to reflect what has been modeled in the math program formulation. The model is implemented in C# programming language. Simulation logic implemented in the model is briefly described here.

Upon an arrival of an ambulance request from demand region i, a ground ambulance or helicopter ambulance is dispatched. Dispatch decision determines 1) which trauma center the patient is sent to, and 2) which ambulance transports the patient. This decision is made based on a nearest-available resource protocol – i.e. a patient is transported to the nearest trauma center j that has not yet reached its capacity limit, and the nearest ambulance among currently available ones is dispatched. If $d_{ij} \leq 30$km, a ground ambulance is dispatched. If $d_{ij} > 30$km and $d_{hij} \leq 120$km, a helicopter ambulance is dispatched. This means that a helicopter ambulance does not fly to a patient if the route h-i-j is too long to transport the patient within the 60-min threshold. Though rather unrealistic, this last assumption is implemented to align the simulation model behavior to the math model.

Number of patients currently present in the trauma center is monitored throughout a given day, and if it reaches the daily capacity limit, no more patients can be transported to the center unless all other trauma centers in the region are also fully occupied. At the beginning of each day, trauma centers are assumed to operate with a zero census level.

In addition to estimating $\sum_{h \in H_{ij}} L_h$ for the iterative solution approach, we also use the simulation model to evaluate the performance metric for location solutions. We use the fraction of successful transport r as the performance metric: $r = \frac{N_G + N_A}{N}$ where N is the total number of trauma cases, and N_G and N_A are the number of successfully transported patient by ground and helicopter ambulances, respectively. At the end of each transport, we determine whether the transport was successful. We measure the time of arrival at a trauma center to compare with the 60-min threshold. We also check the census of the trauma center. A transport is considered a success if a patient arrives at a trauma center within 60-min and the trauma center is under its capacity limit. Note that all ground transports are successful by definition. Also note that, for a given location solution, $\{y_j : z_h\}$, there can be demand regions that are so remotely located that they are outside of geographic coverage. Patients at these demand regions are discarded and considered failed transport cases.

4 EXPERIMENTAL SETTINGS

4.1 Data Collection & Analysis

This research was conducted as part of a preliminary feasibility study for Korea's national trauma care system design initiative. In the study, we draw trauma patient data from the following national trauma databases: the National Health Insurance database, National Emergency Department Information System (NEDIS), and EMS database. All data have been anonymized to protect patient privacy. The total number of trauma cases is approximately 190,000 for the year 2008. We use the time of EMS calls and their locations. There are 38 candidate sites for trauma centers, most of which are general and university hospitals. The number of candidate sites for heliport bases is 16 in total. Location information has been geo-coded to generate their coordinates using a commercial GIS software, ArcGIS™.

For simplicity, Euclidean distance is used for measuring distance between patient location, trauma center candidate site, and the heliport base candidate site. Certainly, this is not the correct representation of true distances, especially for ground ambulance cases, as Euclidean distance tends to underestimate the real distance travelled by ground vehicles. However, given the size of the geographic areas and the total number of cases, this simplifying approximation should not present a major concern.

4.2 Experimental Settings

We test these solution methods for four problem instances. We arbitrarily choose trauma patient data of single days from year 2008 data – Jan. 1, Jan. 9, Feb. 14, Jul. 8, and Nov. 17, and we use half year data from Jan. 1 to June 30. The number of trauma cases in each instance is 361, 370, 395, 412, 551, and 90,265. For the half year data case, we aggregate the original data into 25km x 25km grids to make it computationally tractable. We locate 10 trauma centers and 15 helicopter ambulances.

We compare the quality of the solution obtained from the proposed approach to the solutions yielded by other alternatives. As a reference for comparison, we develop solution approaches using well-known location problem formulations. Both approaches deploy a two-step procedure that decouples the trauma center location problem and helicopter ambulance allocation problem. In the first step, we assume that the number of helicopter ambulances is infinite so that the availability of helicopter ambulances is always guaranteed. This simplifying assumption renders the original problem as a capacitated maximal covering location problem, known as cMCLP (Chung et al. 1983). The optimal locations of trauma centers from the first step are then used as inputs to the second step to locate helicopter ambulances. This part is formulated in two ways, using the maximal covering location problem (MCLP), which does not account for probabilistic availability of helicopter ambulances, and employing a standard maximal expected covering location model, called MEXCLP (Daskin 1983), a probabilistic version of MCLP.

5 RESULTS & DISCUSSION

We first examine how solutions evolve through the iteration steps. Figure 2 shows plots of the fraction of successful transports r as a function of the iteration step n for the six test instances. Note that r is an indirect indicator of the convergence of a location solution. In all test instances, the location solution converges within a few iterations. The iteration step at which the location solution converges – that is, $\{y_j^n : z_h^n\} = \{y_j^{n-1} : z_h^{n-1}\}$ – is indicated by a circle in each plot. This can be interpreted as the point that the prior estimates on the busy fraction from the previous iteration step match the posterior estimates. In our test, not only does the location solution converge to $\{y_j^* : z_h^*\}$, the fraction of successful transports r monotonically increases as well. Recall in Algorithm 1 that we begin the iteration by setting the initial guess, $\rho_{ij,k}^0 = 0$. This gives a deterministic coverage solution, and its performance (r) is the first data point in the plots. As the iteration proceeds toward the converging point, r increases monotonically to the maximum value. While we do not have a guarantee that $\{y_j^* : z_h^*\}$ is the optimal solution, at least within the solution space explored through the iterations, the method does find the best solution.

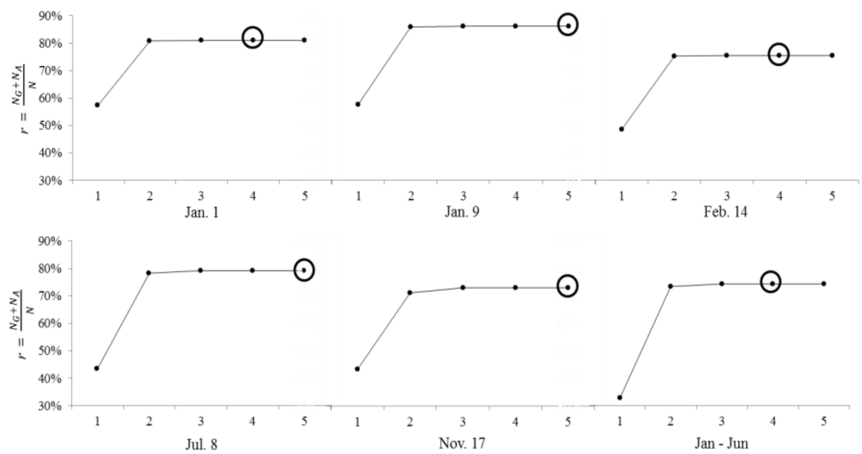

Figure 2: Using Algorithm 1, the location solution converges within a few iterations in all four tested problem instances. Circles in each plot indicate that the location solutions have converged.

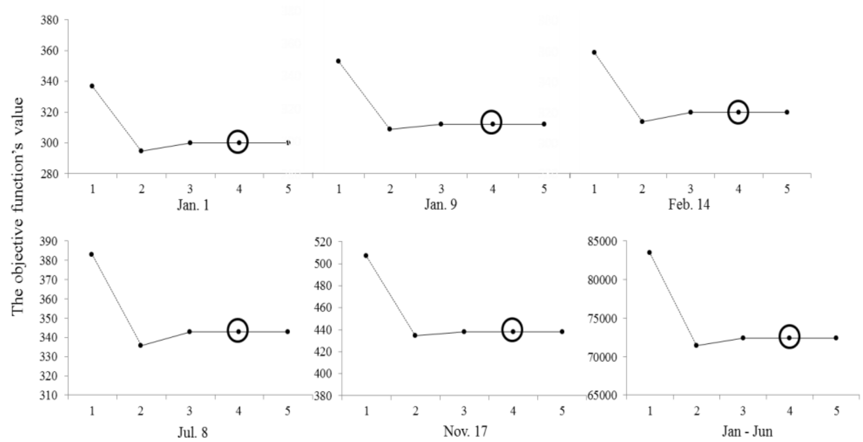

Figure 3: Using the reverse-counting procedure, the location solution converges within a few iterations in all four tested problem instances. Circles in each plot indicate that the location solutions have converged.

We observe the same behavior when we use the reverse-counting procedure in place of simulation. Figure 3 shows the same plot, but the objective function's value instead of *r* from simulation is used as the *y*-axis. This is because the solution we obtain from this procedure is not directly relevant to the simulation as it only works with the mathematical model. The results show an almost identical pattern in which the location solution converges within a few iteration steps and the objective function's value increases toward the maximum. The high objective value at the first iteration in Figure 3 is due to the fact that the initial guess $\rho_{ij,k}^{0} = 0$ makes the second term in (1) artificially small, and thus irrelevant to the procedure. Results from both with- and without-simulation iteration appear to support the use of an iterative method to estimate the busy fraction for our problem.

Now we examine the degree of improvement in solution quality that is obtained when we solve the coupled problem as a whole, by comparing to the alternative, two-step approaches: cMCLP-MCLP and cMCLP-MEXCLP. For a metric of comparison, we measure *r* from the simulation by inputting location solutions from each solution method. Figure 4 shows that solving the problem as a whole by using the iterative method (shaded bars) clearly outperforms the two-step approaches (hashed bars). This can be explained by the assumption made in the first step of the two-step approaches. In decomposing the original problem into two steps, we assume that there is an unlimited number of helicopter ambulances available at every heliport base. This assumption was introduced to eliminate the probabilistic factor in the first step, cMCLP, thereby avoiding the difficulty of estimating the busy fraction. However, a consequence of this assumption is that optimizing the cMCLP may yield a trauma center location solution that is possibly sub-optimal for the overall problem. The iterative method approach removes this arbitrary assumption and attempts to find the location solution for both types of resources simultaneously. The large gap in *r* between the two approaches suggests the importance of finding a way to work out the busy fraction estimation for our coupled location problem.

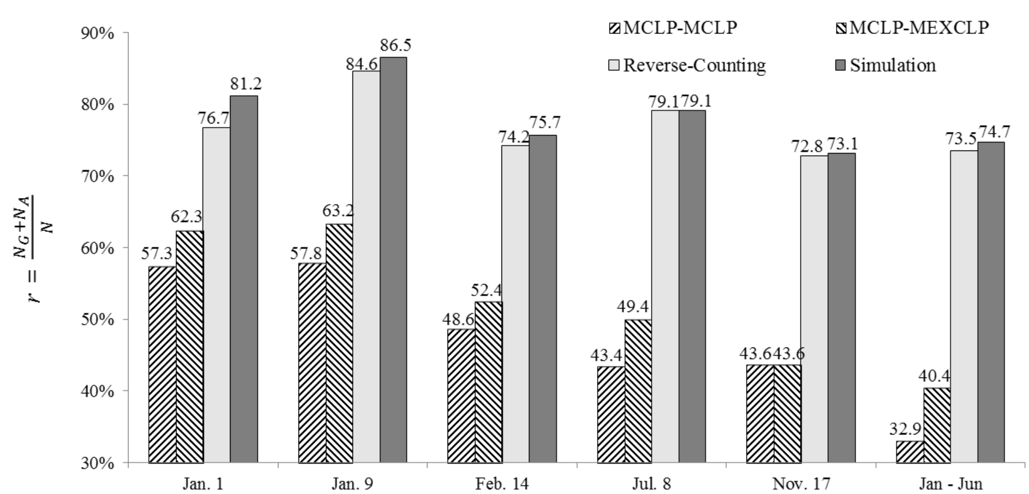

Figure 4: The simulation based iterative method outperforms the three alternatives.

Between the two iterative methods, Algorithm 1 vs. reverse-counting, Algorithm 1 produces a slightly better outcome. This seems to derive from two factors. First, there is some arbitrariness in the reverse-counting procedure. When a demand region *i* and trauma center *j*, (*i*, *j*), can be served by more than one heliport base, it is arbitrarily assigned to one of those heliport bases. This affects the time for helicopter route $h - i - j - h$. This is the service time used for the numerator in (9). Second, it ignores the temporal dimension in the demands. Even when there are only a few service requests in a day, there can be a shortage in helicopter resources if those requests are concentrated in a small time window. In the mathematical model and the reverse counting procedure, the total number of demands are looked at in a unit time, whereas the simulation takes the actual temporal pattern of call arrivals from the historical data. Thus, the busy fraction estimated by the simulation model can be said to be more relevant.

6 CONCLUDING REMARKS

In this paper, we discuss a location problem where trauma centers and helicopter ambulances are located simultaneously. These two location problems cannot be decoupled due to their interdependency. From a facility location problem perspective, this leads to an interesting and challenging problem in a few aspects. Two different types of resources must be considered simultaneously as they mutually affect the overall coverage and workload. In particular, this causes difficulty when we formulate the problem as a probabilistic coverage model. The busy fraction, $\rho_{ij,k}$ needs to be estimated to solve the coupled problem, but the interdependency between trauma centers and helicopter ambulance locations renders this estimation unattainable.

We propose an iterative method where we use an optimization model and a discrete event simulation in an iterative fashion. Location solutions obtained from the optimization model at the first step, with an estimated value of the busy fraction for individual heliport bases, are fed to the simulation model, which in turn provides updated estimates for the busy fraction to be used in the next step of optimization. Our experimental results suggest that the proposed method is a viable solution approach. The location solution converges through the iterations, and also results in the best performance value of the measure. The proposed method outperforms the other alternatives including two-step heuristic approaches.

There are a few potential areas of improvement for our work. One shortcoming of the proposed method is that there exists a better rule for updating optimization results. In this study, we exclude demand points that have no reachable helicopters. This rule leads to quick convergence, but it may produce a poor solution. Considering that the updating optimization results rule has an important role in the proposed method, finding a better updating rule is an attractive research direction to improve the solution quality. The other area where there is room for improvement is in the initial solution for starting the proposed method. Related to the updating optimization results rule, the solution quality from the proposed method depends on the initial solution. Thus, developing a strategy to choose a better initial solution for obtaining a better final solution is also a worthwhile direction for study.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(N01110571).

REFERENCES

Aringhieri, R., G. Carello, and D. Morale. 2007. "Ambulance Location through Optimization and Simulation: The Case of Milano Urban Area." *XXXVIII Annual Conference of the Italian Operations Research Society Optimization and Decision Science*: 1-29.

Ball, M.O., and F.L. Lin. 1993. "A Reliability Model applied to Emergency Service Vehicle Location." *Operations Research* 41: 18-36.

Borras, F., and J.T. Pastor. 2002. "The Ex-post Evaluation of The Minimum Local Reliability Level: An Enhanced Probabilistic Location Set Covering Model." *Annals of Operations Research* 111: 51-74.

Branas, C.C., and C. ReVelle. 2001. "An Iterative Switching Heuristic to Locate Hospitals and Helicopters." *Socio-Economic Planning Sciences* 35: 11-30.

Ingolfsson, A., S. Budge, E. Erkut. 2008. "Optimal Ambulance Location with Random Delays and Travel Times." *Healthcare Management Science* 11: 262-274.

Chung, C.H., D.A. Schilling and R. Carbone. 1983. " The Capacitated Maximal Covering Problem: A Heuristic." *In Proceedings of Fourteenth Annual Pittsburgh Conference on Modeling and Simulation*: 1423-1428.

Church, R., and C. ReVelle. 1974. "The Maximal Covering Location Problem." *Papers of the Regional Science Association* 32: 101-118.

Current, J., and J. Storbeck. 1988. "Capacitated Covering Models." *Environment and Planning B: Planning and Design* 15: 153-164.

Daskin, M.S. 1983. "A Maximum Expected Covering Location Model: Formulation, Properties, and Heuristic Solution." *Transportation Science* 17: 48-70.

Goldberg J., R. Dietrich, J. Chen, M. Mitwasi, T. Valenzuela, and E. Criss. 1990. "A Simulation Model for Evaluating A Set of Emergency Vehicle Base Locations: Development, Validation, and Usage." *Socio-Economic Planning Sciences* 24: 125-141.

Gunes, E., and R. Szechtman. 2005. "A Simulation Model of A Helicopter Ambulance Service." In *Proceedings of the 2005 Winter Simulation Conference.* Edited by M.E. Kuhl, N.M Steiger, F.B. Armstrong, and J.A. Joines, 951-957. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Haghani, A., and S.C. Oh. 1996. "Formulation and Solution of A Multi-commodity, Multi-modal Network Flow Model for Disaster Relief Operations." *Transport. Res. Part A: Policy and Practice* 30: 231-250.

Henderson, S., and A. Mason 2004. "Ambulance Service Planning: Simulation and Data Visualisation." In *Operations Research and Helath Care*: *A Handbook of Methods and Applications* edited by M.L. Brandeau, F. Sainfort, and W.P. Pierskalla, pp 77-102. Boston: Kluwer Academic.

Jarvis, J.P. 1975. "Optimization in Stochastic Service Systems with Distinguishable Servers." Ph.D. dissertation, Massachusetts Institute of Technology.

Larson, R.C. 1974. "A Hypercube Queuing Model for Facility Location and Redistricting in Urban Emergency Services." *Computers and Operations Research* 1: 67-95.

Larson, R.C. 1975. "Approximating the Performance of Urban Emergency Service Systems." *Operations Research* 23: 845-868.

Pirkul, H., and D. Schilling. 1991. "The Maximal Covering Location Problem with Capacities on Total Workload." *Management Science* 37: 233-248

ReVelle, C., and K. Hogan. 1989. "The Maximal Availability Location Problem." *Transportation Science* 23: 192-200.

Silva, P.M.S., and L.R Pinto. 2010. "Emergency Medical Systems Analysis by Simulation and Optimization." In *Proceedings of the 2010 Winter Simulation Conference*, Edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yucesan, 2422-2432, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Toregas, C., R. Swain, C. ReVelle, and L. Bergman. 1971. "The Location of Emergency Service Facilities." *Operations Research* 19: 1363-1373.

AUTHOR BIOGRAPHIES

Dr. TAESIK LEE is an assistant professor of Industrial & Systems Engineering Department at KAIST, Korea. His research interests are primarily in the area of system design, modeling and simulation, especially for large-scale, complex systems. For the past few years, he has been working on designing various aspects of healthcare delivery systems. Dr. Lee obtained his MS and PhD degree from MIT in the Dept. of Mechanical Engineering. His email address is <taesik.lee@kaist.edu>.

HOON JANG is a Ph.D candidate in Industrial and Systems Engineering Department at KAIST. He received his undergraduate and master degree in Industrial and Systems Engineering Department at KAIST. Korea. His research interests included modeling and simulation, optimization in healthcare delivery systems. His email is <hoon.jang@kaist.ac.kr>.

Dr. SOOHAENG CHO is an assistant professor of Tepper School of Business at Carnegie Mellon University, USA. His major research areas include supply chain management, healthcare operations, operations strategy and new product development. He received his MS from MIT and PhD from UCLA. His email address is <soohaeng@andrew.cmu.edu >.

Dr. JOHN G. TURNER is an assistant professor of Paul Merage School of Business at UC Irvine, USA. His research interests include media management, revenue management, applied optimization problems. Dr. Turner obtained his MS and PhD degree from Tepper School of Business at Carnegie Mellon University. Dr. Turner's email address is <john.turner@uci.edu>.