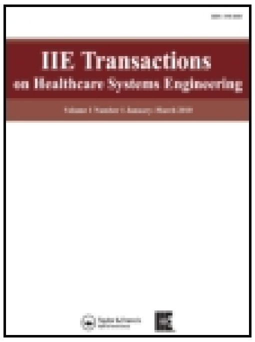

Taylor & Francis
Taylor & Francis Group

IIE Transactions on Healthcare Systems Engineering

ISSN: 1948-8300 (Print) 1948-8319 (Online) Journal homepage: https://www.tandfonline.com/loi/uhse20

The minimum *p*-envy location problem: a new model for equitable distribution of emergency resources

Sunarin Chanta , Maria E. Mayorga , Mary E. Kurz & Laura A. McLay

To cite this article: Sunarin Chanta , Maria E. Mayorga , Mary E. Kurz & Laura A. McLay (2011) The minimum *p*-envy location problem: a new model for equitable distribution of emergency resources, IIE Transactions on Healthcare Systems Engineering, 1:2, 101-115, DOI: 10.1080/19488300.2011.609522

To link to this article: https://doi.org/10.1080/19488300.2011.609522

IIE Transactions on Healthcare Systems Engineering (2011) **1**, 101–115

ISSN: 1948-8300 print / 1948-8319 online
DOI: 10.1080/19488300.2011.609522

The minimum *p*-envy location problem: a new model for equitable distribution of emergency resources

SUNARIN CHANTA[1], MARIA E. MAYORGA[1,*], MARY E. KURZ[1] and LAURA A. McLAY[2]

[1]*Department of Industrial Engineering Clemson University, Clemson, SC, 29634, USA*
E-mail: mayorga@clemson.edu
[2]*Department of Statistical Sciences & Operations Research Virginia Commonwealth University, Richmond, VA, USA*

Received February 2011 and accepted July 2011

Equity is an important consideration in public services such as Emergency Medical Service (EMS) systems. In such systems not only equitability but also performance depends on the spatial distribution of facilities and resources. This paper proposes the minimum *p*-envy facility location model which aims to find optimal locations for facilities in order to balance customers' perceptions of equity in receiving service. The model is developed and evaluated through the lens of EMS systems, where ambulances are located at facilities (stations) with the objective of minimizing the sum of "envy" among all demand zones (customer points) with respect to an ordered set of *p* operating stations weighted by the proportion of demand in each zone. The problem is formulated as an integer program, with priority weights assigned according the probability that an ambulance is available, which is estimated using the hypercube model. Because of the computational effort required to obtain solutions using commercially available software, a tabu search is developed to solve the problem. A case study using real-world data is presented. The performance of the proposed model is tested and compared to other location models such as the *p*-center and maximal-covering-location problems (MCLP).

Keywords: Emergency services, equity, facility location, tabu search

1. Introduction

Emergency medical service (EMS) systems are public service systems that provide emergency medical service to patients within a service area. The services provided vary depending on the calls such as providing emergency medical care via a technician or paramedic, or providing transportation. An important factor in determining EMS performance is not only the quality of emergency medical care provided but also the timeliness or *response time* in which care is provided (McGinnis, 2004). In urban areas, the most widely used ambulance response time standard is to respond to 90% of calls within 8 minutes and 59 seconds as compared to responding to 90% of calls within 14 minutes and 59 seconds in rural areas (Fitch, 2005). In practice, however, it may not be possible to meet this standard depending on the geographical area, the EMS resources available, and the location of EMS resources at the time of a call. In addition, response times may be much longer than the standard, especially in rural or remote areas. Even within a contained geographic area, guaranteeing the same (or similar) response times to all customers in the system may be infeasible.

Unlike private services, such as supermarkets or banks, which are free to locate their facilities in densely populated areas in order to maximize profits, public services such as EMS systems provided by governmental or non-profit agencies need to locate their facilities in a way that serves all residents (customers) fairly as they provide essential life-saving services (Savas, 1978; Stone, 2002). Locating ambulances in EMS systems is an important resource allocation problem that has many implications for equity.

We briefly provide a review of facility locations models that have been applied to public service problems and consider equity. Two well-known facility location models often used to locate ambulances are the *p*-median and *p*-center problems. In the facility location problem with *p* facilities, the *p*-median objective minimizes the total distance from demand points (customers) to their closest facility. Suppose a facility is to be located on a line between two demand points at the ends of the line, moving the facility from one end to another end does not change the total distance between the two demand points and the facility location. Thus, the *p*-median problem is reflective of aggregate level outcomes rather than individual level outcomes, meaning that in the previous example, it does not matter where the facility is located along the line. On the other hand, the *p*-center problem minimizes the maximum

*Corresponding author

distance from demand points to their closest facility. As with the previous example, if again the facility is moved along the line between two demand points, the distance to one demand point is reduced while the distance to the other demand point is increased. Thus, the optimal solution of the p-center problem locates the facility equidistant to both demand points, which reflects one concept of equity (Leclerc et al., 2010). Although the p-center problem belongs to a family of equitable location design problems, its objective improves only the "worst" customer instead of explicitly reflecting the outcomes of all individuals. For a review of the p-center and p-median problems see Daskin (1995).

Public services such as EMS systems have an expectation of fairness for their customers (Stone, 2002). The facility locations directly affect how customers access services. In order for all customers to have an equal chance to obtain services, inequity among all customers must be reduced. Several measures have been proposed to capture inequity of the system or the effect of distribution of the facilities to customers. The most common inequity measure is the maximum distance between customers and the closest facility, assuming that all customers are only serviced by their closest facility. Such a measure is reflected in the p-center problem. Other inequity measures suggested in the literature include range (see e.g., Brill *et al.*, 1976; Erkut and Neuman, 1992), variance (see e.g., Maimon, 1986; Kincaid and Maimon, 1989; Berman, 1990), and mean absolute deviation (see e.g., Berman and Kaplan, 1990; Mulligan, 1991) in the distances between customers and their closest facility. Marsh and Schilling (1994) provide a comprehensive review of equity measures.

Range is a measure that considers the difference between the closest and the farthest customers, while variance and mean absolute deviation are two measures that consider minimizing the difference between individual outcomes and some system standard. However, these measures do not consider that even though a customer may receive better access to care than a given standard, he may still feel dissatisfied if he is "worse off" than other customers. Another equity measure that considers the difference in the outcomes between individual customers is the sum of absolute differences in distance between customers and their closest facility (Keeney, 1980; Lopez-de-los-Mozos and Mesa, 2001; Lopez-de-los-Mozos, 2003). Similarly, the Gini coefficient and Lorenz curves are popular metrics that have been developed for evaluating inequity in economic and social welfare literature, and have been applied to deal with inequity in facility location problems (Maimon, 1988; Erkut, 1993; Drezner *et al.*, 2009). These measures are functions of the absolute value between individual differences, such that penalties are incurred for any differences in individual outcomes (whether a customer is worse off or better off). Since people feel no dissatisfaction when they are better off than others, only negative effects are considered in the minimum envy location problem (MELP) introduced by Espejo *et al.* (2009). They propose several ways to formulate the minimum envy problem; however, their formulations do not necessarily fit well with EMS models. In particular, the formulations provided by Espejo *et al*. (2009) assume that there is strict preference ordering information about customers' dissatisfaction. This is not practical for application to EMS systems because a customer is able to have two stations at the same preference ordering (equidistant). Furthermore, ordinal preferences lack information about distance which is an important metric when assessing quality of service. In our model we are able to relax the strict and ordinal preference order assumptions.

Furthermore, most inequity measures, including all equity location models mentioned above, consider customers' dissatisfaction based only on the closest facility. These inequity measures are appropriate for some public services, such as post office or school locations where the customer travels to the facility, but not necessarily for EMS systems, where open facilities indicate the location where EMS ambulances are stationed. In an EMS system, the ambulance stationed at the closest facility is not always available to serve customers, and in that case the ambulance stationed at the next closest facility might instead be dispatched. To resolve this, many researchers account for the probability that a particular ambulance is available or busy at the time a call for service arrives. For probabilistic location models, see (Larson, 1974, 1975; Daskin, 1983; ReVelle and Hogan, 1989; Batta *et al.*, 1989; Galvao *et al.*, 2005; Iannoni and Morabito, 2007). Other proposed location models explicitly consider backup or multiple coverage (Hogan and ReVelle, 1986; Daskin *et al.*, 1988; Araz *et al.*, 2007; Iannoni and Morabito, 2007). Since EMS systems are an important public service that affects wellness of the service population, we are interested in developing a practical equitable location model that represents the inequity of all customers in the system, and more realistically represents the operations and performance criteria of EMS systems. To the best of our knowledge, this is the first equitable location model that integrates the concept of envy while taking into account the degree of importance of the different servers and incorporates the probability of servers being available to respond calls.

In particular, we propose the Minimum p-Envy Location Problem (MpELP) for locating EMS ambulances at possible station locations in order to increase equity of receiving service among all demand zones. Envy is selected as a way to measure equity, where envy is defined as a function of the distance from a demand zone to its closest EMS station and the distance from a demand zone to its backup EMS stations weighted by priority of the serving stations and weighted by proportion of demand. The performance of our model is investigated by comparing it with two popular equity measures, p-center and Gini coefficient, and the well-known maximal covering location problem (MCLP). Because of its complexity, this problem cannot be solved efficiently to optimality, even for small test cases, using commercially available optimization software; thus a tabu search is developed which yields near-optimal solutions with little computational effort.

The rest of the article is organized as follows. In Section 2, we describe the concept of envy, introduce notation, and formulate the model. An illustrative example is presented in Section 3. Section 4 details how to assign the station weights using the hypercube model. Section 5 presents the procedure of the solution method that we developed for solving the problem using a tabu search (TS). We conduct computational experiments for tuning the tabu search parameters in Section 6. In Section 7 a case study is selected to test the proposed approach using real-world data and computational results are reported in Section 8. Section 9 shows the performance of the minimum p-envy location model in comparison to other location models. Finally, conclusions and discussion are provided in Section 10.

2. Minimum *p*-envy location model

In this section, we modify the concept of envy to create an objective which is meaningful for the ambulance location problem. From Longman's English dictionary, envy is "the feeling of wanting something that someone else has." Therefore, customers in demand zone i feel envy when they receive worse service than others, but when they receive better service than others they have no feeling of envy. These concepts reflect definitional notions of equity in the social science domain (Stone, 2002) in that they clarify the recipients (the potential patients), what is being distributed (delivery of ambulances to patients according to the patients' relative dissatisfaction) and the process for equitably allocating resources (ambulance location). In our model, we define "envy" of demand zone i as a level of customers' dissatisfaction in demand zone i as compared to other demand zones, where a demand zone is a demand point where customers are located. The dissatisfaction of customers in demand zone i is an ordered vector of the distance from demand zone i to its serving stations (facility locations) in decreasing order. That is, the distance to the station closest to demand zone i, which is the primary station, is the first element in the dissatisfaction vector, followed by the distance to the next closest station or the secondary station, and so on. The serving stations, except for the primary stations, are called *backup stations*, of which we can have one or more for each demand zone. Envy is defined as the difference in dissatisfaction between demand zones. Since different demand zones have different total number of customers (demand or call density), we weight the total envy in each demand zone by the proportion of demand in that zone. An illustrative example of how envy is calculated is presented in the next section. We use the following notation:

n = the number of demand zones
m = the number of potential stations
p = the number of ambulances to be located (stations to be opened)
q = the number of serving stations being considered which consists of one primary station and q-1 backup stations where $q \leq p$
λ_l = weight assigned to the l^{th}-priority station, $l = 1, \ldots, q$
H_i = demand (call volume) in zone i
h_i = weight (proportion of demand) of zone $i = \frac{H_i}{\sum_{i=1}^{n} H_i}$
d_{ij} = the distance between zone i and station j

The objective of our equitable location model is to minimize the sum of weighted envy among all demand zones, as shown in Equation (1). Note that the proportion of demand at node i is the weight (h_i) that we assign to differentiate between call volume at different demand zones. As mentioned earlier, customers in each demand zone may have dissatisfaction with respect to all serving stations; first priority station, second priority station, and so on. Thus, we can differentiate the envy with respect to different serving stations by adding the different weights (λ_1) to each level of priority; $l = 1, \ldots, q$, where q is the number of serving stations that are restricted to respond to a particular zone. Note that $q \leq p$ where p is the number of stations that will be opened. We introduce $q \leq p$ here because it may be that only a certain number of back-up stations are allowed or that the decision maker only wants to consider envy with respect to some subset of stations; however, all stations need to be located and thus p cannot simply be replaced by q. Note also that $p \leq m$ where m is the number of potential station locations. A station is said to be opened when there is at least one ambulance stationed for serving customers. We note from (1) that since there is no contribution to the objective of locating more than one ambulance at the same station, the number of open stations and the number of ambulances are the same.[1] We specify the effect that each station has on a demand zone through the vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_q)$, $\lambda_l \geq 0 \, \forall l$. Without loss of generality, we assume that $\sum_{l=1}^{q} \lambda_l = 1$ and $\lambda_1 \geq \lambda_2 \geq \ldots$. Station priority weights can be assigned in various ways, depending on how the system administrator values backup service. For example, if a system only utilizes one backup station, we can set $q = 2$, so $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ where $\lambda_1 \geq \lambda_2$, and $\lambda_1 + \lambda_2 = 1$. How the weights $\boldsymbol{\lambda}$ may be assigned is further discussed in Section 4.

The minimum p-envy location problem is introduced as an integer programming model. The objective function captures the total weighted envy among all demand

[1]This assumption may be relaxed by incorporating constraints on the number of ambulances per station and modifying the envy calculation. For example, if up to two ambulances are allowed at each station, the first backup station is considered the same as the primary station when there are two ambulances at the primary station.

zones as shown in Equation (1). The decision variable e_{ik}^{l} represents the envy of demand zone i compared with demand zone k based on their serving stations at the l^{th} priority level. Note that the l^{th} priority station serving demand zone i is not necessarily the same as the l^{th} priority station serving demand zone k. Equations (2)–(3) work together to calculate the envy between all possible pairs of customers. The variable e_{ik}^{l} takes on value 0 when zone i is served by a closer facility than zone k compared with the same priority station, otherwise it is equal to the difference between the distance from zone i to its serving station and the distance from zone k to its serving station, that is $e_{ik}^{l} = \max\{0, (\sum_{j=1}^{m} d_{ij} y_{ij}^{l} - \sum_{j=1}^{m} d_{kj} y_{kj}^{l})\}$, where y_{ij}^{l} are binary variables that indicate whether an ambulance at station j is assigned to serve zone i as the l^{th} priority station. These variables help to ensure that the service order for the bases is by increasing distance. Equation (4) limits the number of ambulances that are available to be located, or equivalently, number of stations to be opened. Equation (5) ensures that a demand zone be served by exactly one facility at each l^{th} priority station. Equation (6) ensures that a station can either serve as a 1st or 2nd or l^{th} priority of zone i. Equation (7) requires that a demand zone i can be served by facility j if station j is open. Equation (8) assigns a station to serve zone i by considering the distance from an open station to the zone; the closer station receives the higher priority to serve zone i.

The minimum p-Envy Location Model (p-Envy):

$$\text{Minimize } Z = \sum_{l=1}^{q} \sum_{i=1}^{n} \sum_{k=1}^{n} \lambda_l h_i e_{ik}^{l} \tag{1}$$

Subject to:

$$e_{ik}^{l} \geq \sum_{j=1}^{m} d_{ij} y_{ij}^{l} - \sum_{j=1}^{m} d_{kj} y_{kj}^{l} \quad \text{for } i, k = 1, \ldots, n : k \neq i;\ l = 1, \ldots, q \tag{2}$$

$$e_{ik}^{l} \geq 0 \quad \text{for } i, k = 1, \ldots, n; l = 1, \ldots, q \tag{3}$$

$$\sum_{j=1}^{m} x_j = p \tag{4}$$

$$\sum_{j=1}^{m} y_{ij}^{l} = 1 \quad \text{for } i = 1, \ldots, n; l = 1, \ldots, p \tag{5}$$

$$\sum_{l=1}^{p} y_{ij}^{l} \leq 1 \quad \text{for } i = 1, \ldots, n; j = 1, \ldots, m \tag{6}$$

$$y_{ij}^{l} \leq x_j \quad \text{for } i = 1, \ldots, n; j = 1, \ldots, m;\ l = 1, \ldots, p \tag{7}$$

$$d_{ij} y_{ij}^{l} \leq d_{ij} y_{ij}^{l+1} \quad \text{for } i = 1, \ldots, n; j = 1, \ldots, m;\ l = 1, \ldots, p-1 \tag{8}$$

Where:

$$x_j = \begin{cases} 1 & \text{if a facility is located at station } j \\ 0 & \text{otherwise} \end{cases}$$

$$y_{ij}^{l} = \begin{cases} 1 & \text{if a facility at station } j \text{ assigned to serve zone i as the } lth \text{ priority station} \\ 0 & \text{otherwise} \end{cases}$$

e_{ik}^{l} = envy of customers at demand zone i compared to zone k with respect totheir facilities at the l^{th} priority.

3. Illustrative example

In this section, a small example is provided to illustrate the concept of p-envy and how the objective function is calculated. Suppose there are three demand zones, three potential stations for locating EMS ambulances, and two ambulances. In this case, $n = 3$, $m = 3$, and $p = 2$. Assume that one backup station is considered, $q = 2$. The number of rows in the distance matrix (d_{ij}) represents the number of demand zones (n) while the number of columns represents the number of potential stations (m), where d_{ij} represents the distance from demand zone i to station j. Other inputs include the proportion of demand in each demand zone i (h_i), and the weights assigned each priority open station (λ_l). The inputs to this small example are given below in matrix form.

$$\mathbf{d} = \begin{bmatrix} 2 & 2 & 10 \\ 8 & 4 & 6 \\ 10 & 5 & 2 \end{bmatrix} \quad \mathbf{h} = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix} \quad \lambda = \begin{bmatrix} 0.60.4 \end{bmatrix}$$

The vector $\mathbf{h}$ denotes that 20%, 30%, and 50% of customer calls originate in demand zones 1, 2, and 3, respectively. The vector $\boldsymbol{\lambda}$ indicates that a consumer's envy will be comprised of 60% resulting from envy regarding their primary serving station and 40% envy regarding their secondary serving station. Suppose ambulances are located at station 1 and station 2, demand zone 1 is 2 units away from its 1st priority or primary station, and also 2 units away far from its 2nd priority or secondary station. Demand zone 2 is located closer to station 2, so station 2 serves as a 1st priority station of zone 2, and station 1 serves as a 2nd priority station of zone 2. The same with demand zone 3, it is served by station 2, and station 1 as 1st and 2nd priority stations respectively. Then, the envy of demand zone i with respect to demand zone j, in regards to their 1st priority station is calculated from the difference of the distance from demand zone i to its 1st priority station and the distance from demand zone j to its 1st priority station whereas if demand zone i is closer to its 1st priority station demand zone j, the envy of demand zone i with respect to j is equal to 0, because demand zone j does not have better access than demand zone i. If demand zone i is farther from its 1st priority station than demand zone *j* is to theirs, demand zone *i* envies demand zone *j* which we quantify as the difference in dissatisfaction between demand zone *i* and demand zone *j*. The envy matrix corresponding to locating ambulances at station 1 and 2 (e^l_{ik}) is calculated from the summation of $\max\{0, \sum_{j=1}^{m} d_{ij} y^l_{ij} - \sum_{j=1}^{m} d_{kj} y^l_{kj}\}$ where $l = 1,2$; $i,k = 1,2,3$; $k \neq i$. For example, $e^1_{12} = \max\{0, (2\text{-}4)\} = 0$, $e^1_{13} = \max\{0, (2\text{-}5)\} = 0$, and $e^1_{23} = \max\{0, (4\text{-}5)\} = 0$, $e^1_{21} = \max\{0, (4\text{-}2)\} = 2$. The total envy of all demand zones with respect to all serving stations is equal to the summation of all elements in the envy matrix multiplied by the demand zone weight (h_i) and the station weight (λ_l). If we locate ambulances at station 1 and 2, the total envy of all demand zones is equal to 4.28. Our goal is to find the station locations that give the minimum total of envy. With this small example, one can easily enumerate all possible solutions, and the optimal solution is opening stations at locations {2, 3} with a total envy value of 1.56. Using the integer programming formulation of the minimum *p*-envy model, developed in the previous section, a solver found an optimal solution at

$$\begin{aligned}
\boldsymbol{x} &= \{x_1, x_2, x_3\} = \{0, 1, 1\},\\
\boldsymbol{y} &= \{[(y^1_{11}, y^2_{11}), (y^1_{12}, y^2_{12}), (y^1_{13}, y^2_{13})], [(y^1_{21}, y^2_{21}), (y^1_{22}, y^2_{22}),\\
&\qquad (y^1_{23}, y^2_{23})], [(y^1_{31}, y^2_{31}), (y^1_{32}, y^2_{32}), (y^1_{33}, y^2_{33})]\}\\
&= \{[(0,0), (1,0), (0,1)], [(0,0), (1,0), (0,1)],\\
&\qquad [(0,0), (0,1), (1,0)]\},\\
\boldsymbol{e} &= \{[(e^1_{11}, e^2_{11}), (e^1_{12}, e^2_{12}), (e^1_{13}, e^2_{13})], [(e^1_{21}, e^2_{21}), (e^1_{22}, e^2_{22}),\\
&\qquad (e^1_{23}, e^2_{23})], [(e^1_{31}, e^2_{31}), (e^1_{32}, e^2_{32}), (e^1_{33}, e^2_{33})]\}\\
&= \{[(0,0), (0,4), (0,5)], [(2,0), (0,0), (0,1)], [(0,0),\\
&\qquad (0,0), (0,0)]\}.
\end{aligned}$$

4. Determining appropriate station priority weights

The station weights should be assigned according to how a system administrator views the importance of the resources, or according to how they believe customers feel envy. The minimum *p*-envy problem is specifically designed to consider backup stations; thus, the number of backup stations should affect the values of the weights that are assigned. Suppose the system has no backup station, in other words only one station has 100% responsibility to serve a particular zone, the station weight should be set to 1 and $\lambda = (\lambda_1, 0, 0, \ldots, 0)$ where $\lambda_1 = 1$; in that case the minimum *p*-envy location problem becomes the original minimum envy problem except that envy is measured nominally rather than with strict preference ordering. The only restriction on the weights assigned is that $\sum_{l=1}^{q} \lambda_l = 1$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q$. For example, these values could be assigned to be linearly decreasing; that is, $\lambda_l = \frac{q+1-l}{K}$ where $K = 1 + 2 + \ldots + q$, l = station priority, $l \in \{1, \ldots, q\}$. For example, if $q = 5$, $\lambda_1 = 5/15$, $\lambda_2 = 4/15$, …, $\lambda_5 = 1/15$, respectively. Next we provide a recommendation for how these weights might be assigned to reflect the actual performance of EMS systems.

In real EMS systems the closest vehicle may not be available to answer a call. Thus we suggest that the probability of vehicle being available be assigned as a station weight. Daskin (1983) developed the earliest probabilistic location model, the maximum expected coverage location problem (MEXCLP), which assumed that servers operate independently and have the same busy probability which is independent of their locations. Later, Batta *et al.* (1989) developed an adjusted MEXCLP (AMEXCLP) which relaxes some assumptions of the MEXCLP by embedding the hypercube queuing model into MEXCLP. The hypercube model developed by Larson (1974, 1975) considers a correction factor that accounts for busy probabilities depending on server locations. The model has several underlying assumptions: 1) calls for service arrive according to a Poisson process, 2) if a call arrives while all servers are busy; it enters at the end of a queue and will be served in a FIFO manner. One of the drawbacks of this model is that an implicit assumption is that all stations have equal workload.

In this paper, the busy probability of vehicles is estimated by the hypercube queuing model. Let P_b denote the probability that a randomly selected vehicle will be busy which depends on the number of ambulances that are deployed (assuming *q*-1 backup stations). Using actual system data, we estimate the probability P_b by $P_b = \Lambda/q\mu$ where, Λ is the average number of calls per hour, $1/\mu$ is the average service time per call (hours), and q is number of ambulances that are deployed. Constructing an M/M/q queuing model operating at steady state, we get the probability that all servers are available, P_0, as given in Equation (9). The correction factors Q are calculated as in Equation (10). If there are l ambulances that may respond to a call, the probability that the l^{th} vehicle will be dispatched or is available is calculated from the probability that l-1 ambulances are busy and the l^{th} vehicle is available. The probability that the l^{th} vehicle is available (λ_l) is shown in Equation (11) where $Q(q, P_b, l\text{-}1)$ is the correction factor and $Q(q, P_b, 0) = 1$.

$$P_0 = \left(\frac{(p P_b)^q}{q!(1 - P_b)} + \sum_{j=0}^{q-1} \frac{q^j P_b^j}{j!} \right)^{-1} \tag{9}$$

$$Q(q, P_b, j) = \sum_{k=j}^{q-1} \frac{(q-j-1)!(q-k)q^k P_p^{k-j} P_0}{(k-j)!q!(1-P_b)}, \quad j = 0, \ldots, q-1 \tag{10}$$

$$\lambda_l = Q(q, P_b, l-1)(1 - P_b)(P_b^{l-1}), \quad l = 1, \ldots, q \tag{11}$$

5. Tabu search

The complexity of the minimum *p*-envy location problem makes finding the optimal solution via a commercial optimization sotfware impractical due to the

Table 1. Results of solving *p*-envy location problem via an optimization solver, ILOG OPL 5.5

n	*M*	*p*	Time (sec.)	Gap(%)
5	5	3	1.06	0
		2	0.81	0
10	10	5	25.53	0
		3	5.57	0
		2	3.09	0
20	10	5	647.60	0
		3	121.26	0
		2	47.37	0
30	15	10	>1H	0
		5	>1H	0
		3	426.56	0
50	20	10	>1H	NA
	15	5	>1H	NA
		3	>1H	80.24
100	30	10	>1H	NA
	20	5	>1H	NA
	10	5	>1H	NA

computational effort required to solve real-world size problems. Although the model has been linearized to reduce computational effort, it requires a large number of additional variables and constraints to remove the nonlinear (maximization) terms involved in calcualting envy. The number of variables and constraints make the problem size grow exponentially as the number of demand zones and potential stations increase, which directly leads to increased computational costs. To illustrate the complexity, we generated 17 test problem sets with different combinations of parameters; $n \in \{5,10,20,30,50,100\}$, $m \in \{5,10,15,20,30\}$, $p \in \{2,3,5,10\}$, and assuming that $q = p$ in all test cases. The integer programing formulation of the M*p*ELP was implemented in the commercial optimization solver ILOG OPL 5.5., running on a Dell Latitude D410 machine with Intel Pentium processor 1.73 GHz, 1 GB of RAM. With traditional branch and cut methods, the solver was able to find the optimal solutions in some cases, as shown in Table 1. The running time limit was fixed to 1 hour. Based on this experiment, for problem sizes equal to or larger than 30 demand nodes, it is not practical to obtain the optimal solution via the optimization solver. We also observed that in the case of $n = 30$, $m = 15$, $p = q = 5$, it took about 6 hours to get the optimal solution. The notation >1H states that running time exceeded 1 hour and NA states that no feasible solutions have been found after running the solver for 1 hour.

To overcome this problem, one might try to reduce the number of variables by improving the formulation. For a discussion of developing efficient minimum envy formulations see Espejo *et al*. (2009). However these integer programming formulations tend to have limitations depending on the problem structure, and they still suffer from dimensionality issues. Espejo *et al*. (2009) developed several formulations for the minimum envy location problem with the underlying assumption that a demand zone must have predefined strict preferences for all potential stations, and the computational running time for the problem size $n = 40$ was reported to be longer than 1 hour. In this paper, we are interested in providing a practical approach that will enable us to solve the real-world size problem, which tend to have a large number of demand zones ($n \geq 100$). Therefore, we developed a tabu search that enables us to find near-optimal solutions efficiently.

Tabu search (TS), a metaheuristic algorithm, was formalized in 1986 by Glover (1986). The characteristics of TS are based on the mechanism of human memory. During the search process, TS keeps a memory of a predetermined number of solutions that have already been evaluated and records them on a tabu list. These solutions that have been evaluated are protected for a limited period of time using short-term memory in an attempt to escape local optima. If the new solution yields a better objective than the best solution found so far, a move is performed regardless of the tabu list, otherwise moving to the new solution will only occur when the new solution is not in the tabu list. The TS algorithm is composed of the following procedures 1) initializing a feasible solution, 2) improving upon the current solution, 3) managing the tabu list, 4) checking the stopping criteria. The algorithm repeats steps 2 through 4 until the stopping criteria is satisfied.

5.1. *Representation and initialization*

We choose a permutation representation for our solution. That is, suppose we have two ambulances to be located among five potential stations, and consider the solution of locating ambulance 1 at station 3 and ambulance 2 at station 5; the permutation representation string will be {3,5}. The initial solution is randomly generated using the concept of random keys as introduced by Bean (1994). We start with one feasible solution at the initial stage.

5.2. *Improving process*

To improve a current solution, we consider all the solutions in the neighborhood of the current solution and replace it with its best neighbor. The swap neighborhood used in Ghosh (2003) is applied in which each neighbor is found by replacing one located ambulance with one non-located ambulance. In other words, an open station is replaced by a closed station. Suppose we have two ambulances to be located among five potential stations and the current solution is {3,5}. If we chose station 3 to be replaced, the possible neighbors are {1,5}, {2,5}, and {4,5}. The total number of possible neighbors to each solution are $(m\text{-}p)p$. Because we only replace one station at each iteration, the number of neighbors at each iteration are (m-p), and the best neighbor is the solution that yields the lowest total weighted envy.

5.3. *Tabu list*

To avoid selecting an old solution that has been recently evaluated, we create a tabu list to record the old moves or old solutions. We propose two types of tabu lists and apply each one to the TS algorithm we developed. These are swap record and solution record.

5.3.1. *Swap record*

As described in section 5.2, a new solution is obtained by swapping an open station with a closed station. This tabu list consists of pairs of recent stations that have been replaced and the stations that replaced them. For example, if we have a current solution, {3,5}, and we want to move ambulance 1 from station 3 to station 2, our new solution is {2,5}. In this case, we record the move {3,2}. Thus the swap record tabu list is an mxm matrix where m is the number of the potential stations. We record the swap {3,2} by updating the value of element (3,2) and (2,3) in the swap record tabu list. This tabu list structure has the advantage of being convenient to manage; however, the size of the list grows as the number of the candidate stations increases.

5.3.2. *Solution record*

Instead of recording the swap move, we can alternatively record the solutions that have been evaluated. Note that the swap record tabu list cannot protect some solutions that have recently been evalutated in the case that the order of the stations is different. For example, if the current solution is {1,2,3}, and the next solution is {1,2,5}; the swap record {3,5} would be added to the swap record list. However, this does not rule out the possibility that in two moves we would see solution {3,2,5} then {3,2,1} which is the same as the previous solution {1,2,3}. This problem is solved when using the solution record tabu list. However, this type of list structure requires more steps to create the list. With this record we use binary encoding to store the solution such that each distinct set of ambulance locations yields the same value, despite the order. For example, the solution {1,2,3} or {3, 2, 1} will be recorded as a value of $2^1+2^2+2^3 = 14$. In this case, the length of the list is fixed at the number of the candidate stations at each iteration (m).

5.4. *Short-term memory*

Independent of which tabu list structure is used, the solutions in the tabu list will be protected for the next solution, which means we never have the same solution in the following iteration. This protection is set to be active for a limited time, called the tenure time. The tenure time works as a short-term memory of the TS algorithm which is one of the parameters that might effect the performance of the TS algorithm. There are three possible ways to manage the tenure time: fixed, dynamic, and random. In this study, we used fixed tenure time, and considered list lengths of 7, 10, 15, 20 as suggested by Glover (1990). One advantage of using a dynamic tenure is that the size of the neighborhood can be increased early in the search (in a manner similar to simulated annealing) by allowing a longer tabu list when the incumbent solution is poor relative to the best solution seen so far. On the other hand, a random tabu length tenure allows the search to diversify its neighborhood during the search. In this particular case, we find the fixed tenure time option provides reasonable results, and so we present results based on that setting.

5.5. *Aspiration critera*

An aspiration criteria is applied when the better move is tabu. In other words, a tabu move (solution that is in the tabu list) is allowed when this solution yields a better objective than the best found so far. While a tabu search may be performed without accepting improving tabu moves, such a tabu search would inhibit the quality of the method.

5.6. *Stopping criteria*

Several potential stopping criteria have been proposed such as maximum CPU time, a maximum number of solutions, a maximum number of iterations, or a maximum number of iterations with no improvement. Based on preliminary experiments, we terminate the program after a fixed number of iterations which depends on the problem size and is dicussed further in later sections. For each scenario the TS is run for 30 replications (with different initial solutions randomly generated at each replication). The steps of the tabu search at each iteration are shown below:

```
Step 1: Initialize solution
Step 2: Best : = Initial Solution
  Current : = Initial Solution
Step 3: While (Stopping criterion not met) do
  Select a station to swap
  Evaluate all possible neighbors
  Best_nb : = Best neighbor
  If Best_nb is better than Best
    Then Go to Step 5
  Else   Go to Step 4
Step 4: If Best_nb is not in the tabu list
```

```
        Then Go to Step 5
      Else Best_nb : = Next best neighbor
      If Best_nb is the last neighbor
        Then Go to Step 5
      Else Go to Step 4
    Step 5: Current : = Best_nb
      Update tabu list
      If Current is better than Best
        Then Best : = Current
      End while
```

While we realize that the proposed TS algorithm is quite simple, we will demonstrate below that it is both quite effective and efficient at finding solutions. Furthermore, the algorithm is robust in the sense that it works with any location model objective. Lastly, while we do perform parameter tuning for the TS, we do not test using alternate heuristic methods. The focus of the article is the development and analysis of the MpELP. The TS is developed here to allow us to analyze real-world size problems.

6. Parameter tuning experiments

In this section, we conducted experiments to find the best combination of two parameters: the type of tabu list structure and the choice of tenure time length. These parameters were identified as influential factors based on initial testing. Two data sets have been used. The first is a real-world data set consisting of 122 demand zones and 16 potential station locations (details regarding this data set are provided in Section 7). The second is a publicly available data set with 30 nodes and 30 stations, taken from Lorena (http://www.lac.inpe.br/~lorena/correa/Q_MCLP_30.txt). For each data set, we create 6 instances by varying the number of stations that can be opened, i.e., p varies from 5 to 10. Each case is tested under the two types of tabu list structures (as discussed in Section 5.3) and four tenure time lengths of 7, 10, 15 and 20, respectively. The resulting 96 test cases were run on a Dell Latitude D410 machine with Intel Pentium processor 1.73 GHz, 1 GB of RAM. The TS algorithm was coded using C. We also obtain the optimal solution to each problem by enumerating all possible solutions. We note that full enumeration takes anywhere from 1 hour to 2 days depending on the problem size and is only used to evaluate the performance of our algorithm, not recommended as an approach to solving the p-envy problem. The results are represented as the median and range of the solution gap (%gap = the relative difference between the best tabu search solution value and the optimal solution value) over the 30 replications, which are reported in Table 2.

We report the median rather than mean because the solution gaps are not normally distributed, as will be later discussed. In this experiment we terminated each run after 50 iterations. We performed statistical analysis to identify if the tabu list structure and tenure time length significantly affect the performance of the TS. Because our results do not satisfy the assumptions required to use traditional ANOVA analysis (the solution gaps are not normally distributed and the variance in solution gaps is non-homogeneous), the Friedman test, a non-parametric statistical test, is selected to assess if differences in performance exist due to choice of list structure and tenure time length. At a significance level of 0.05, the Friedman test indicated that there is a statistically significant difference between using different types of tabu lists and among all levels of tenure time length. The swap record yielded the lowest overall median solution gap of 0.8%. We also observed that a tenure time equal to 15 yielded the best solutions with the smallest median solution gaps among all test cases regardless of the type of tabu list used. Therefore, the swap record tabu list structure with tenure time length of 15 is suggested as the best parameters for our TS.

7. Case study

Our case study uses real-world data from the Hanover Fire and EMS department, which is located in Hanover County, Virginia. The Hanover EMS department responds to 911 calls 24 hours a day and serves a county of 474 square miles, with a population of approximately 97,000 individuals. The data were collected from the Fire and EMS department, and captures the life-threatening calls received during 2007. We divided the coverage area into 175 distinct demand zones made up of approximately 2 by 2 mile areas. In this way, we ensure that originating demand is represented realistically. Currently, there are $m = 16$ existing potential stations for locating EMS ambulances. All station locations are shown in Fig. 1. Based on the data, requested calls did not originate from all 175 zones. Therefore, we ignore the zones that have no demand and only considered the $n = 122$ zones in which demand existed in 2007.

The input data to the model are the number of the requested calls (or number of customers) in each demand zone, the geographical coordinates of the 122 demand zones and 16 potential stations, and the weights assigned to different priority stations. To set up the locations of the stations and demand zones, we drew grid lines over the area of interest, with one block representing four square miles. The coordinates (a, b) of the stations and center point of demand zone blocks are used to calculate the distance between each demand zone and each station. Distance between two points can be measured in many ways (see Drezner and Hamacher, 2004). The most familiar two are

Table 2. Median solution gaps and solution gap range among the 30 replications for the parameter tuning experiments, expressed as Median (Min, Max)

			%Gap[a] [Median (Min, Max)]			
			Tenure time			
Data set ($n \times m$)	List	p	7	10	15	20
30QMCLP (30×30)	Swap	5	0.000 (0.000, 0.700)	0.000 (0.000, 3.397)	0.000 (0.000, 2.007)	0.000 (0.000, 1.188)
		6	0.716 (0.201, 1.833)	0.873 (0.000, 1.629)	0.5445 (0.000, 1.983)	1.045 (0.201, 1.886)
		7	2.170 (0.000, 3.022)	1.980 (0.720, 4.566)	2.101 (0.000, 3.199)	2.761 (0.747, 4.573)
		8	1.462 (0.000, 3.453)	1.596 (0.269, 2.877)	2.021 (0.827, 4.236)	1.596 (0.000, 3.926)
		9	1.712 (0.700, 4.764)	2.615 (0.778, 5.584)	2.159 (0.700, 3.940)	2.858 (0.489, 4.711)
		10	1.553 (0.000, 3.360)	1.263 (0.000, 3.300)	0.991 (0.061, 3.189)	1.869 (0.000, 3.302)
	Solution	5	0.000 (0.000, 1.003)	0.956 (0.000, 4.455)	0.694 (0.000, 1.310)	0.694 (0.000, 4.596)
		6	0.870 (0.201, 2.253)	0.876 (0.000, 2.240)	0.873 (0.000, 1.629)	1.056 (0.544, 2.936)
		7	2.514 (0.000, 5.238)	2.101 (0.000, 5.539)	2.723 (0.000, 4.116)	1.604 (0.720,3.877)
		8	1.955 (1.041, 3.995)	1.633 (0.973, 4.204)	1.966 (0.827, 4.546)	1.495 (0.269, 4.236)
		9	2.248 (0.700, 4.764)	2.194 (0.778, 5.584)	2.703 (0.700, 3.940)	2.896 (0.489, 4.711)
		10	2.135 (0.000, 3.300)	1.265 (0.000, 3.300)	0.757 (0.061, 3.1895)	1.265 (0.000, 3.3021)
Hanover County (122×16)	Swap	5	0.000 (0.000, 8.254)	0.000 (0.000, 10.088)	0.000 (0.000, 8.254)	0.000 (0.000, 5.507)
		6	0.000 (0.000, 3.801)	0.000 (0.000, 5.253)	0.000 (0.000, 3.752)	0.000 (0.000, 3.943)
		7	0.000 (0.000, 7.479)	0.000 (0.000, 3.556)	0.000 (0.000, 4.216)	0.569 (0.000, 2.772)
		8	1.022 (0.000, 1.617)	1.022 (0.000, 3.493)	0.000 (0.000, 1.767)	0.000 (0.000, 2.914)
		9	0.434 (0.000, 4.579)	0.433 (0.000, 5.039)	0.433 (0.000, 6.005)	0.433 (0.000, 4.878)
		10	1.492 (0.002, 6.944)	0.894 (0.002, 5.841)	0.894 (0.002, 5.039)	0.894 (0.002, 5.648)
	Solution	5	0.000 (0.000, 8.254)	0.000 (0.000, 5.004)	0.000 (0.000, 5.507)	0.000 (0.000, 8.254)
		6	0.000 (0.000, 3.943)	0.000 (0.000, 8.900)	0.000 (0.000, 3.943)	0.000 (0.000, 3.943)
		7	0.000 (0.000, 2.786)	0.284 (0.000, 3.673)	0.569 (0.000, 3.673)	0.000 (0.000, 6.491)
		8	1.022 (0.000, 1.617)	1.022 (0.000, 3.493)	0.000 (0.000, 1.767)	0.000 (0.000, 2.914)
		9	0.433 (0.000, 5.039)	0.433 (0.000, 3.374)	0.433 (0.000, 3.374)	0.000 (0.000, 3.374)
		10	0.894 (0.002, 4.218)	0.894 (0.002, 4.218)	0.894 (0.002, 5.039)	0.894 (0.002, 6.561)
Overall median	Swap		0.869 (0.000, 2.170)	0.883 (0.000, 2.615)	0.489 (0.000, 2.159)	0.731 (0.000, 2.858)
						0.800 (0.489, 0.883)
	Solution		0.882 (0.000, 2.514)	0.925 (0.000, 2.194)	0.725 (0.000, 2.723)	0.794 (0.000, 2.896)
						0.838 (0.725, 0.925)

[a]%Gap = [(Best known of TS - Optimal solution) *100] / Optimal solution.

rectilinear distance and Euclidean distance. In this case we use the Euclidean metric because approximately 70% of the Hanover County area is rural, and can thus be reached via highways or county roads that do not conform to a grid. Given a demand zone i at (a_i, b_i) and a station location j at (a_j, b_j), the distance (d_{ij}) between demand zone i and station j is calculated using the Euclidian metric.

8. Computational results

In this section we test the performance of our tabu search heuristic using the same two data sets. Based on the parameter tuning experiments in Section 5, the swap record tabu list with a tenure time of 15 is used with both data sets. Since the numbers of neighborhoods in both cases are different we used different termination criteria for each data set. We terminated the program after 500 iterations for the 30QMCLP data set, and 100 iterations for the Hanover County data set. The solution gaps over 30 replications of both cases are shown in Table 3. We can see that the median and average solution gap is less than 1% for all cases and that, within a few seconds the TS obtained the optimal solution for all instances of the Hanover data set and for 2 out of 6 instances of the 30QMCLP data set (recall that a commercial solver was not able to obtain solutions to problems with $n = 30$ in 1 hour).

9. Performance of the minimum *p*-envy location problem model

While our model seeks to reduce inequity through the *p*-envy objective, we must be careful not to sacrifice efficiency of the current EMS system. Hanover County EMS measures efficiency in terms of coverage, where the coverage level is the total proportion of demand that can be reached within a response time threshold (RTT). Following

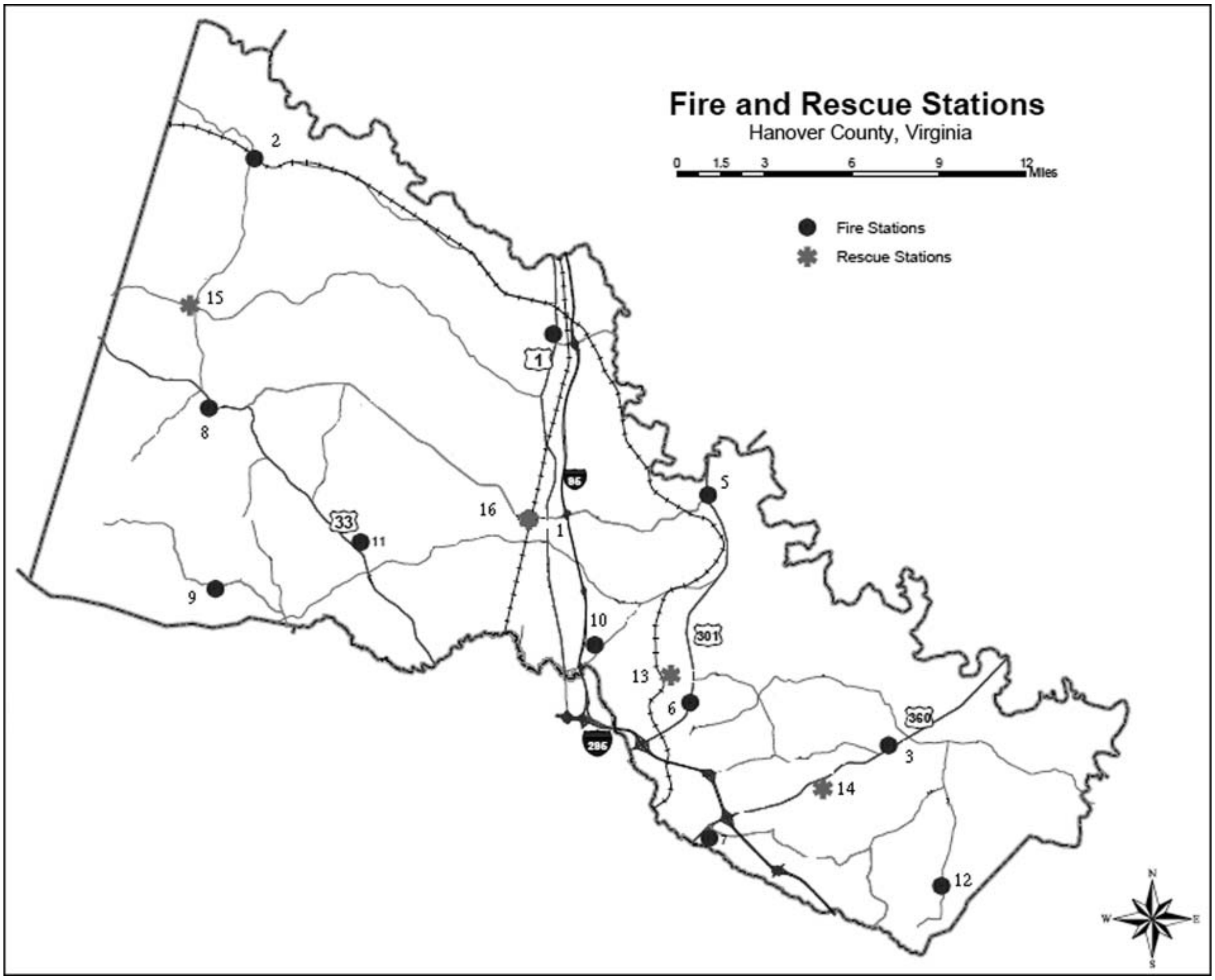

Fig. 1. Map of fire and rescue stations in Hanover County, Virginia.

Table 3. Experimental results of TS using tuned parameters

Data set ($n \times m$)	P	%Gap[a]					CPU time (sec)				
		Median	Avg	SD	Min	Max	Median	Avg	SD	Min	Max
30QMCLP (30×30)	5	0	0	0	0	0	2.656	2.661	0.035	2.625	2.781
	6	0.544	0.351	0.274	0	0.873	3.742	3.778	0.081	3.703	4.079
	7	0	0.293	0.365	0	0.747	3.734	3.734	0.038	3.703	3.922
	8	0.269	0.408	0.382	0	1.254	3.703	3.706	0.020	3.672	3.750
	9	0.572	0.545	0.534	0	1.844	3.984	4.025	0.097	3.953	4.375
	10	0.061	0.098	0.148	0	0.757	5.078	5.085	0.031	5.062	5.234
Hanover County (122×16)	5	0	0	0	0	0	3.953	3.966	0.035	3.937	4.078
	6	0	0.080	0.304	0	1.201	3.531	3.569	0.132	3.406	3.922
	7	0	0.154	0.294	0	0.896	3.578	3.586	0.046	3.547	3.750
	8	0	0.102	0.312	0	1.022	4.172	4.188	0.039	4.156	4.344
	9	0	0.300	0.639	0	3.374	4.562	4.577	0.061	4.515	4.813
	10	0	0.446	0.681	0	2.087	3.453	3.477	0.085	3.406	3.828

[a]%Gap = [(Best known of TS - Optimal solution) *100] / Optimal solution.

current Hanover County standards, we use a response-time threshold of 9 minutes. Thus, a demand zone is said to be covered when there exists an EMS ambulance that is able to respond to a call in that demand zone within 9 minutes. In particular, we assume based on distance, average ambulance speed, and road conditions that for a call to be responded to within 9 minutes, at least one station should be open within 4 miles of the demand zone. In this case, there are 1711 calls spread over 122 demand zones; given the set of possible station locations, there are four zones that cannot be covered, since they are more than 4 miles from the closest possible station. Therefore, the maximum percentage of coverage for Hanover County is 98.8%.

We compare our model to a traditional covering location model, which maximizes efficiency, and to other equity models. Two standard measures of equity are selected for comparison, *p*-center and Gini coefficient. The *p*-center is a classic equity model that intends to improve the worst customer (minimizes the distance of the customer located the furthest away from their closest station). The Gini coefficient is an equity measure that considers the average dissatisfaction among all customers. The traditional maximal covering location (MCLP) model is selected as a baseline to measure coverage. The formulations of the models are provided below.

- Minimum *p*-envy location problem (M*p*ELP): Objective is to minimize sum of envy weighted by proportion of demand:

$$\min Z = \sum_{l=1}^{q} \sum_{i=1}^{n} \sum_{k=1}^{n} \lambda_l h_i e_{ik}^l$$

Subject to (2) – (8).

- Maximal covering location problem (MCLP), see original version in Church and ReVelle (1974): Objective is to maximize proportion of demand that can be covered (reached within a given response time threshold):

$$\max Z = \sum_{i=1}^{n} y_i H_i$$

Subject to

$$\sum_{j=1}^{m} a_{ij} x_j \geq y_i \quad \text{for all } i = 1, 2, \ldots, n \tag{12}$$

and (4)

$$\text{Where } y_i = \begin{cases} 1 & \text{if demand zone } i \text{ is covered by an open station} \\ 0 & \text{otherwise} \end{cases}$$

$$a_{ij} = \begin{cases} 1 & \text{if station } j \text{ can cover demand at zone } i \\ 0 & \text{otherwise} \end{cases}$$

- *p*-center, see details in Daskin (1995): Objective is to minimize the maximum distance from customers to their closest station:

$$\min Z$$

Subject to

$$\sum_{j=1}^{m} d_{ij} y_{ij} \leq Z \quad \text{for all } i = 1, 2, \ldots, n \tag{13}$$

$$\sum_{j=1}^{m} y_{ij} = 1 \quad \text{for all } i = 1, 2, \ldots, n \tag{14}$$

$$y_{ij} \leq x_j \quad \text{for all } i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, m \tag{15}$$

and (4)

$$\text{Where } y_{ij} = \begin{cases} 1 & \text{if a demand zone } i \text{ is served by facility at station } j \\ 0 & \text{otherwise} \end{cases}$$

- Gini coefficient measure (Gini), see details in Drezner *et al*. (2009): Objective is to minimize Gini coefficient (a weighted measure of absolute differences):

$$\min \frac{\sum_{i=1}^{n} \sum_{k=1}^{n} \left| \sum_{j=1}^{m} d_{ij} y_{ij} - \sum_{j=1}^{m} d_{kj} y_{kj} \right|}{2n \sum_{i=1}^{n} \sum_{j=1}^{m} d_{kj} y_{ij}}$$

Which is equivalent to minimizing the numerator:

$$\min \sum_{i=1}^{n} \sum_{k=1}^{n} \left| \sum_{j=1}^{m} d_{ij} y_{ij} - \sum_{j=1}^{m} d_{kj} y_{kj} \right|$$

Subject to (4), (14) – (15)

We use the Hanover County data, which contains 122 demand zones and 16 potential station locations. We vary the total number of ambulances to be located from 5 to 10. Thus, in this case $n = 122$, $m = 16$, and $p = q$ varies from 5 to 10. h_i is the proportion of demand at location i; $i = 1, \ldots, 122$ and all λ_l values are assigned according to probability of vehicles being busy as described in Section 4. In this study we use a 9-minute response time threshold to evaluate coverage. The goal here is to gauge how much improving equity compromises typical EMS performance measures, such as coverage. We solved all four facility location models to optimality (optimal solution to the *p*-envy model was confirmed via full enumeration) and then compared the resulting equity measures and coverage. These results are shown in Tables 4 to 7. In these tables we present several metrics for evaluating the quality of a solution. We measure equity as the sum of weighted total envy, and we measure efficiency by the coverage of demand (this is the traditional

Table 4. Results of *p*-envy

p	Opened stations	Maxdist	Gini coefficient	Total weighted envy	Covered demand
5	{1 4 6 7 8}	12	0.3139	63.7672	1524
6	{1 4 7 8 13 14}	10	0.3120	54.4391	1572
7	{1 3 4 7 9 13 15}	7	0.2810	47.5000	1628
8	{1 4 7 9 10 13 14 15}	8	0.2997	43.7513	1618
9	{1 2 4 7 8 9 10 13 14}	8	0.2995	38.9498	1637
10	{1 2 3 4 5 6 7 8 9 10}	6	0.2881	35.2600	1661

Table 5. Results of MCLP

p	Opened stations	Maxdist	Gini coefficient	Total weighted envy	Covered demand
5	{1 4 6 14 15}	12	0.3157	78.8367	1559
6	{1 4 6 11 14 15}	10	0.2960	75.9275	1604
7	{1 4 5 6 11 14 15}	10	0.3022	71.3908	1636
8	{1 4 5 6 9 11 14 15}	8	0.2925	71.8066	1657
9	{1 2 4 5 6 8 9 11 14}	8	0.2903	72.6818	1674
10	{1 2 4 5 6 8 9 11 12 14}	6	0.2735	72.5320	1688

Table 6. Results of *p*-center

p	Opened stations	Maxdist	Gini coefficient	Total weighted envy	Covered demand
5	{1 2 3 6 8}	8	0.2772	130.4130	1173
6	{1 3 4 9 13 15}	7	0.2623	125.9371	1153
7	{2 3 4 8 9 11 13}	6	0.2658	170.9044	978
8	{2 3 4 5 6 8 9 11}	6	0.2736	137.3588	1208
9	{1 2 3 4 5 6 8 9 10}	6	0.2790	109.6269	1397
10	{1 2 3 4 5 6 7 8 9 10}	6	0.2840	90.9555	1510

Table 7. Results of Gini

p	Opened stations	Maxdist	Gini coefficient	Total weighted envy	Covered demand
5	{3 4 9 10 15}	8	0.2533	146.4598	776
6	{3 4 9 11 13 15}	7	0.2588	154.1969	961
7	{1 3 4 9 13 15 16}	7	0.2640	119.1073	1249
8	{1 2 3 4 8 9 13 16}	7	0.2657	126.2046	1268
9	{2 4 5 8 9 10 11 12 14}	6	0.2677	137.6112	1293
10	{1 2 3 4 5 6 8 9 11 12}	6	0.2695	85.5668	1674

measure of efficiency for EMS systems). We also report the maximum distance (Maxdist) between a demand zone and its closest open station (or the *p*-center objective) and the total covered demand. For all models we calculate the other objective values ex-post, after finding the optimal solution. In these tables, larger values of covered demand are desirable and smaller values of inequity measures (Maxdist, the Gini coefficient, and total weighted envy) are desirable. The *p*-envy, Gini coefficient, and MCLP models produce unique optimal solutions while the *p*-center model often produces multiple solutions. In the case that the *p*-center object produces multiple optimal solutions, we report the average values of the covered demand, and equity measures from all optimal solutions. The four models are compared in Figs. 2 to 4 in terms of the resulting equity and efficiency measures.

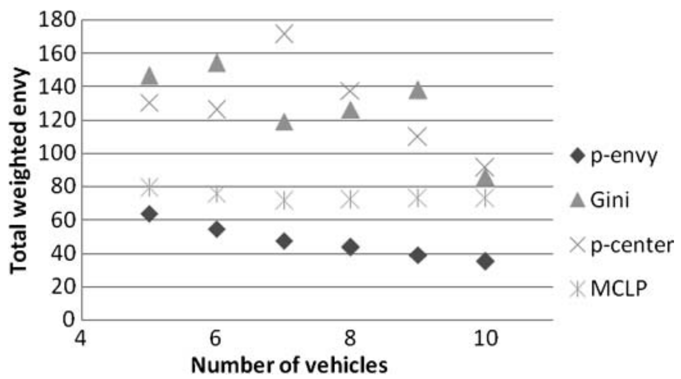

Fig. 2. Equity comparison—p-envy measure for each model.

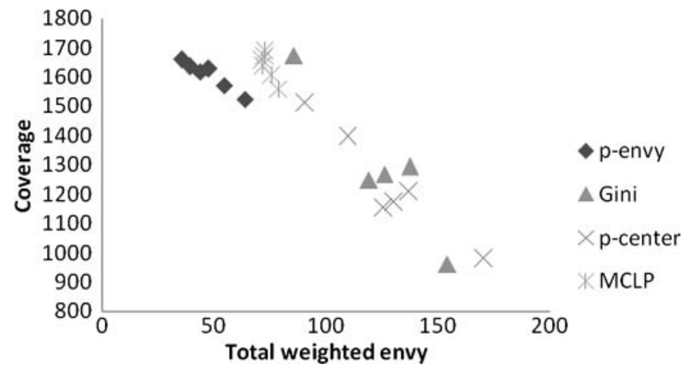

Fig. 4. Coverage-equity trade off comparison (with available probability station weights).

Figure 2 shows the total weighted p-envy for each model for $p = 5$ to 10. As expected, the minimum p-envy model has the lowest sum of total weighted envy among these four models. Interestingly, the p-center and Gini coefficient models, that also try to reduce inequity, do not dominate the MCLP model (in terms of p-envy). A possible explanation for this is that neither the Gini or p-center models weight the demand zones by demand density, such that each zone is treated equally, which may be impractical in real systems, where demand density may vary widely by geographic location. Furthermore, the performance of the p-envy model is robust to the number of ambulances. For all models the resulting Gini coefficient is stable, ranging only from 0.2533 to 0.3157, while the maximum distance from a zone to its closest station (Maxdist) is quite variable, ranging from 6 to 12 miles.

Figure 3 compares the four models in terms of coverage. Here we see that the Gini model performed much worse compared with the other models while the p-envy model performed very close to the MCLP model, whose objective is to maximize coverage. The performance of the p-center model largely depends on the number of ambulances. This is an undesired trait of the p-center model solutions because one would expect that coverage should increase as the number of ambulances increase. However, the p-center model does not weigh demand zones, and it sacrifices the coverage of densely populated areas in order to ensure better service to the demand zone that is "worse off."

To illustrate the tradeoff between equity and coverage, we plot the performance of all four models with respect to these metrics. Figure 4 shows the results of all four models where the p-envy model uses the station availability probabilities (see Section 4) for station weights. Note that the solution points in Fig. 4 are generated by solving all four models with different number of vehicles; the resulting solutions are shown in Tables 4 to 7. Interestingly, we see that the minimum p-envy location model not only yields the lowest total envy, but attains almost the same coverage as MCLP. Therefore, the p-envy model allows us to reduce inequity without sacrificing coverage, for this data set. This is an unexpected outcome for the equity model presented, as equity and coverage tend to be conflicting objectives which necessitate a multi-objective approach, such as the one undertaken by Chanta *et al.* (2011). The results depend on the weights assigned to the priority of the stations (vector $\boldsymbol{\lambda}$). For example, we note that Maxdist could be reduced in the p-envy model by giving more weight to the closest station (increasing λ_1), because Maxdist is only concerned with customers' distance to the closest open station. Figure

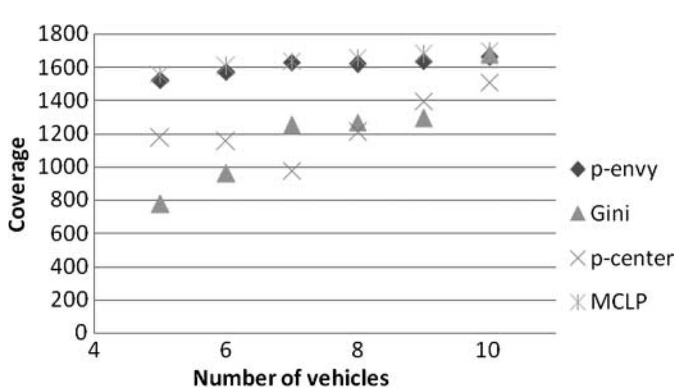

Fig. 3. Efficiency comparison—resulting coverage for each model.

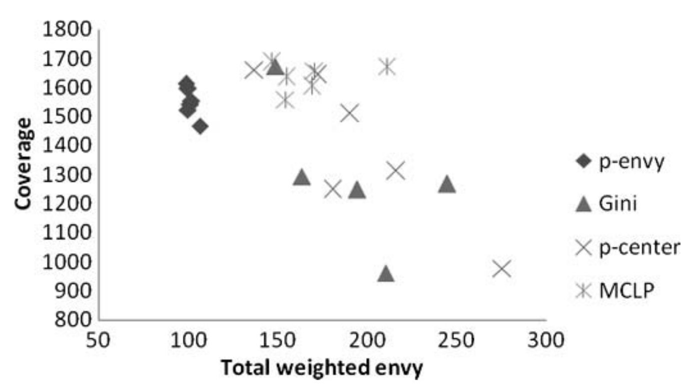

Fig. 5. Coverage-equity trade off comparison (with equal station weights).

5 shows the results when we have equal weight of station priorities ($\lambda_l = 1/q$ for all l). We see that solutions of the minimum p-envy model are more efficient than solutions of other equity models in the sense that for a fixed amount of coverage they achieve lower envy. This difference increases as the weights assigned to the backup stations become increasingly important.

10. Conclusion and discussion

In this paper, we have proposed the minimum p-envy location problem (MpELP) for EMS systems using the concept of envy which minimizes the inequity of access to service among all zones between all serving facilities (stations). While most equity measures only consider the effect of the closest serving station, our model is different in that we consider the effect that all serving stations have on all customers. Because this objective is complex it results in a problem that cannot practically be solved with commercial optimizers. Thus, a tabu search is developed to solve the problem. Solutions are obtained in a few seconds and the performance of the heuristic is very effective with respect to both computational time and quality of solutions. We also compare the minimum p-envy location model with other equity models such as p-center and Gini coefficient. The results show that the proposed model not only yields the lowest total weighted envy compared with other equity models, but also yields highly efficient solutions in terms of coverage. In fact the coverage of the minimum p-envy location model is very close to the coverage resulting from the standard maximal covering location model (MCLP). These results are unexpected, as equity and coverage are usually conflicting objectives (Chanta *et al.*, 2011). The proposed model is helpful for facility location planners, especially in the realm of public service where reducing inequity is of high importance, though not at the expense of efficiency. We believe the contribution of this model goes beyond presenting a way of more realistically capturing customer dissatisfaction, as this new metric outperforms other commonly used "equity" measures both in terms of fairness and efficiency. The MpELP could be especially relevant in locating public health services where the goal is to locate facilities in such a way that all customers are treated fairly while not sacrificing the performance of potentially life-saving services.

Acknowledgments

It is a pleasure to acknowledge Battalion Chief Henri Moore, Jr., Chief Fred C. Crosby II, Mr. Lawrence Roakes, and Mr. Edward Buchanan of the Hanover County Fire and EMS Department in Hanover County, VA, for the knowledge, experience and data they provided to support this research effort.

References

Araz, C., Selim, H. and Ozkarahan, I. (2007) A fuzzy multi-objective covering-based vehicle location model for emergency services. *Computers & Operations Research*, **34**(3), 705–726.

Batta, R., Dolan, J. M. and Krishnamurthy, N. N. (1989) The maximal expected covering location problem: revisited. *Transportation Science*, **23**(4), 277–287.

Bean, J. C. (1994) Genetic algorithms and random keys for sequencing and optimization. *ORSA Journal on Computing*, **6**(2), 154–160.

Berman, O. (1990) Mean-Variance location problems. *Transportation Science*, **24**(4), 287–293.

Berman, O. and Kaplan, E. H. (1990) Equity maximizing facility location schemes. *Transportation Science*, **24**(2), 137–144.

Brill, E. D., Jr., Liebman, J. C. and ReVelle, C. S. (1976) Equity measures of exploring water quality management alternatives. *Water Resources Research*, **12**(3), 845–851.

Chanta, S., Mayorga, M. E. and McLay, L. M. (2011) Improving emergency service in rural areas: A bi-objective covering location model for EMS systems. In press, *Annals of Operations Research*, June.

Church, R. L. and ReVelle, C. (1974) The maximal covering location problem. *Papers of Regional Science Association*, **32**, 101–118.

Daskin, M. S. (1983) A maximal expected covering location model: formulation, properties, and heuristic solution. *Transportation Science*, **17**, 48–70.

Daskin, M. S. (1995) *Network and Discrete Location: Models, Algorithms, and Applications*, John Wiley & Sons, Inc, New York, NY.

Daskin, M. S., Hogan, K. and ReVelle, C. (1988) Integration of multiple, excess, backup and expected covering models. *Environment and Planning B: Planning and Design*, **15**, 15–35.

Drezner, T., Drezner, Z. and Guyse, J. (2009) Equitable service by a facility: Minimizing the Gini coefficient. *Computers & Operations Research*, **36**(12), 3240–3246.

Drezner, Z. and Hamacher, H. W. (2004) *Facility Location: Applications and Theory*, Springer, New York, NY.

Erkut, E. (1993) Inequality measures for location problems. *Location Science*, **1**, 199–217.

Erkut, E. and Neuman, S. (1992) A multiobjective model for location of undesirable facilities. *Annals of Operations Research*, **40**, 209–227.

Espejo I., Marin, A., Puerto, J. and Rodriguez-Chia, A. M. (2009) A Comparison of formulations and solution methods for the minimum-envy location problem. *Computers & Operations Research*, **36**, 1966–1981.

Fitch, J. (2005) Response times: Myths, measurement and management. *Journal of Emergency Medical Services*, **30**(9), 46–56.

Galvao, R. D., Chiyoshi, F. Y. and Morabito, R. (2005) Towards unified formulations and extensions of two classical probabilistic location models. *Computers & Operations Research*, **32**(1), 15–33.

Ghosh D. (2003) Neighborhood search heuristics for the uncapacitated facility location problem. *European Journal of Operational Research*, **150**, 150–162.

Glover, F. (1986) Future paths for integer programming and links to artificial intelligence. *Computers & Operations Research*, **13**, 533–549.

Glover, F. (1990) Tabu search: A tutorial. *Interfaces*, **20**(4), 74–94.

Hogan, K. and ReVelle, C. (1986) Concepts and applications of backup coverage. *Management Science*, **32**, 1434–1444.

Iannoni, A. P. and Morabito, R. (2007) A multiple dispatch and partial backup hypercube queuing model to analyze emergency medical systems on highways. *Transportation Research, Part E*, **43**, 755–771.

Keeney, R. L. (1980) Equity and public risk. *Operations Research*, **28**(3), 527–534.

Kincaid, R. K. and Maimon O. (1989) Locating a point of minimum variance on triangular graphs. *Transportation Science*, **23**, 216–219.

Larson, R. C. (1974) A hypercube queuing model for facility location and redistricting in urban emergency services. *Computers & Operations Research*, **1**, 67–95.

Larson, R. C. (1975) Approximating the performance of urban emergency service systems. *Computers & Operations Research*, **23**, 845–868.

Leclerc, P. D., McLay, L. A. and Mayorga, M. E. (2010) Modeling equity for allocation in public resources. Forthcoming, *Community-Based Operations Research: Decision Modeling for Local Impact and Diverse Populations*, Springer, NY.

Lopez-de-los-Mozos, M. (2003) The sum of absolute differences on a network: Algorithm and comparison with other equity measures. *Information Systems & Operational Research*, **41**(2), 195–210.

Lopez-de-los-Mozos, M. and Mesa, J. A. (2001) The maximum absolute deviation measure in location problems on networks. *European Journal of Operational Research*, **135**, 184–194.

Lorena, L. A. N. Problem Instances: Queuing Maximal Covering Location-Allocation Problem, available at http://www.lac.inpe.br/~lorena/correa/Q_MCLP_30.txt, last accessed February 5, 2011.

Maimon, O. (1986) The variance equity measure in locational decision theory. *Annals of Operations Research*, **6**, 147–160.

Maimon, O. (1988) An algorithm for the Lorenz measure in locational decisions on trees. *Journal of Algorithms*, **9**(4), 583–596.

Marsh, M. and Schilling, D. (1994) Equity measurement in facility location analysis: A review and framework. *European Journal of Operational Research*, **74**(1), 1–17.

McGinnis, K. K. (2004) *State EMS Rural Needs Survey*, National Association of State EMS Directors, Falls Church, VA.

Mulligan, G. F. (1991) Equity measures and facility location. *Papers in Regional Science*, **7**(4), 345–365.

ReVelle, C. and Hogan, K. (1989) The maximum availability location problem. *Transportation Science*, **23**, 192–200.

Savas, E. S. (1978) On equity in providing public services. *Management Science*, **24**(8), 800–808.

Stone, D. (2002) *Policy Paradox: The Art of Political Decision Making*, W. W. Norton & Company, New York, NY.

Biographies

Sunarin Chanta is a faculty of King Mongkut's University of Technology North Bangkok, Thailand. She received her BS and MS degrees in Statistics from Kasetsart University, Thailand, in 2002; and MS and PhD degrees in Industrial Engineering from Clemson University in 2011. She teaches courses in operations research. She has received the best paper award in operations research track for proceedings of the 2009 IIE Annual Conference and Expo. Her research interests are mathematical modeling for emergency response service systems, OR applications on healthcare systems, multiple objective decision making, optimization, and metaheuristics. Chanta is a member of INFORMS and IIE.

Maria E. Mayorga is an Assistant Professor of Industrial Engineering at Clemson University. She received her PhD in Industrial Engineering & Operations Research from the University of California, Berkeley. Dr. Mayorga has research and teaching interests in probability models and stochastic processes and optimization, with applications to healthcare systems engineering. Dr. Mayorga has authored over two dozen publications in archival journals and refereed proceedings. Her research has been supported by NIH and BMW Manufacturing among others. She is a member of INFORMS and is a board member for the IIE Operations Research Subdivision.

Mary E. Kurz, Associate Professor of Industrial Engineering at Clemson, has research interests in deterministic operations research and metaheuristics with over 55 archival journal and refereed conference proceedings articles. Her PhD is in Systems and Industrial Engineering and her MS and BS degrees are in Systems Engineering, all from The University of Arizona. She teaches graduate courses in analytics and metaheuristics and undergraduate courses in deterministic operations research and decision support systems. She received IIE's Operations Research Division's 2005 Award for Excellence in the Teaching of Operations Research. Kurz is a member of INFORMS and a senior member of IIE.

Laura A. McLay is an Assistant Professor of Statistical Sciences and Operations Research at Virginia Commonwealth University. She received her PhD from the University of Illinois at Urbana-Champaign. Her research interests are in the field of operations research, with a particular focus on discrete optimization with application to homeland security and emergency response problems. Dr. McLay's research has been supported by NSF, DHS, and the Department of the Army. She has authored or coauthored more than twenty publications in archival journals and refereed proceedings. She is a member of INFORMS, IIE, Tau Beta Pi, Alpha Pi Mu, and the Virginia Academy of Science.