Transportation Research Part C 69 (2016) 120–133

Contents lists available at ScienceDirect

Transportation Research Part C

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

A double standard model for allocating limited emergency medical service vehicle resources ensuring service reliability

Yi Liu [a], Zongzhi Li [b,*], Jingxian Liu [a], Harshingar Patel [b]

[a] *College of Navigation, Wuhan University of Technology, Wuhan 430070, China*
[b] *Department of Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, Chicago, IL 60616, United States*

ARTICLE INFO

Article history:
Received 23 December 2015
Received in revised form 23 April 2016
Accepted 30 May 2016
Available online 7 June 2016

Keywords:
Emergency
Ambulance
Reliability
Optimization
Resource allocation

ABSTRACT

This paper introduces a new double standard model (DSM), along with a genetic algorithm (GA), for solving the emergency medical service (EMS) vehicle allocation problem that ensures acceptable service reliability with limited vehicle resources. Without loss of generality, the model is formulated to address emergency services to human injuries caused by vehicle crashes at intersections within an urban street network. The EMS fleet consists of basic life support (BLS) and advanced life support (ALS) vehicles suited for treating crashes with different severity levels within primary and secondary service coverage standards corresponding to extended response times. The model ensures that all demand sites are covered by at least one EMS vehicle within the secondary standard and a portion of which also meets the service reliability requirement. In addition, a portion of demand sites can be covered by at least one of each type of EMS vehicles within the primary standard. Meanwhile, it aims to achieve maximized coverage of demand sites within the primary standard that complies with the required service reliability. A computational experiment is conducted using 2004–2010 data on top two hundred high crash intersections in the city of Chicago as demand sites for model application. With an EMS fleet size of 15 BLS and 60 ALS ambulances maintained by the Chicago Fire Department, at best 92.4–95.5% of demand could be covered within the secondary standard at 90% of service reliability; and 65.5–68.4% of high severity demand and 50.2–54.5 low severity demand could be covered within the primary standard at 90% of service reliability. The model can help optimize EMS vehicle allocation in urban areas.

1. Introduction

In an environment of ever-increasing urban travel demand, traffic safety becomes a major concern in urban areas around the globe. An incident caused by vehicle crashes imposes adverse impacts on both traffic safety and mobility at and around the incident site. It may lead to severe vehicle damages, property losses, and personal injuries and fatalities. In order to mitigate losses of a traffic incident particularly related to the loss of human lives, maintaining effective responses is critical and immediately providing emergency medical services (EMS) is an essential part of such actions (Dobson, 2003; Wells, 2007). The average response time, which is greatly affected by the distribution of EMS vehicle depots and allocation of available EMS vehicles to incident sites, becomes a key measure to assess the effectiveness of emergency responses. However,

* Corresponding author.
E-mail address: lizz@iit.edu (Z. Li).

http://dx.doi.org/10.1016/j.trc.2016.05.023
0968-090X/

Nomenclature

Symbols

q_i^k	local busy fraction of one type k vehicle deployed at any $j \in S_{i,r'}$
r, r'	service coverage standards
r_1	primary service coverage standard
r_2	secondary service coverage standard
$\bar{t}^k$	average service time of type k vehicle, hours per service
t_{ij}	vehicle travel time on the shortest path from vehicle location j to demand site i
f_l	frequency of requests for emergency service by competing demands around demand site i, demand calls per day
$S_{i,r'}$	$\{\forall j \mid t_{ij} \leqslant r'\}$, the set of depot locations within r' of demand site i
$C_{i,r}$	$\{\forall m \mid t_{mj} \leqslant r,\ j \in S_{i,r'}\}$, the set of m competing demand sites around demand site i, which are located within r of any depot location $j \in S_{i,r'}$
y_j^k	number of type k vehicles deployed at vehicle location j
ρ_i^k	utilization ratio
$e_{i,r'}^k$	smallest number of type k vehicles assigned around demand site i at depot location $j \in S_{i,r1}$, that satisfies $1 - \left(\frac{\rho_i^k}{e_{i,r'}^k}\right)^{e_{i,r'}^k} \geqslant \gamma$
e_{i,r_1}	smallest number of vehicles assigned around demand site i at depot location $j \in S_{i,r1}$, that can satisfy $1 - \left(\frac{\rho_i}{e_{i,r1}}\right)^{e_{i,r1}} \geqslant \gamma$
e_{i,r_2}	smallest number of vehicles assigned around demand site i at depot location $j \in S_{i,r2}$, that can satisfy $1 - \left(\frac{\rho_i}{e_{i,r2}}\right)^{e_{i,r2}} \geqslant \gamma$
D	set of demand sites
D_1	set of high crash severity demand sites
D_2	set of low crash severity demand sites
N_d	total number of demand sites
d_i	demands at demand site $i \in D$
p_j	maximum number of vehicles that can be deployed to vehicle location j
p^k	fleet size of type k vehicles
S	vehicle depot locations
$S_{i,r1}$	set of depot locations that can reach demand site i within primary standard r_1
$S_{i,r2}$	set of depot locations that can reach demand site i within secondary standard r_2
z_{i,r_1}^k	1, if demand site i is reachable by at least one type k vehicle within primary standard r_1; 0, otherwise
γ	service reliability level
x_{i,r_2}^k	1, if demand site i is covered by type k vehicle within secondary standard r_2 at service reliability level γ; 0, otherwise
x_i	1, if demand site i is covered within primary standard r_1 at reliability level γ; 0, otherwise
α	the portion of demand sites covered within the primary standard r_1
β	the portion of demand sites covered within the secondary standard r_2
m_i	indicator of covered demand calls at demand site i
M	fitness measure
i	demand site i
j	vehicle depot location j
k	vehicle type k, 1 for ALS, 2 for BLS

Abbreviations

DSM	double standard model
GA	genetic algorithm
EMS	emergency medical service
BLS	basic life support
ALS	advanced life support
SCLP	set covering location problem
MCLP	maximal coverage location problem
MEXCLP	maximum expected covering location problem
NP	nondeterministically polynomial
PDO	property damage only
FB	frequency-based scenario
SB	severity-based scenario
CFD	Chicago Fire Department

considering the extensive spatial range in which traffic incidents may occur, resources of emergency service providers in major urban and metropolitan areas are always limited by available budget, manpower, vehicles, and equipments. Optimal deployment of limited emergency resources in a large area is vital in maintaining cost-effective EMS and efficient emergency responses.

The EMS vehicle allocation problem falls into the general category of the facility location problem, which has been studied extensively. In this context, mathematical models have been developed over the years to locate EMS, fire stations, bus garages, airline hubs, and so forth (Current et al., 2002). They can be differentiated into covering, *P*-median, and *P*-center facility location models. The covering location models attempt to locate servers so that either maximum coverage of demands or minimum cost of facility location with pre-obtained covering level is achieved. The *P*-median location models aim to minimize average or total cost of servers and demands by optimizing server locations. While *P*-center location models focus on minimizing the highest travel costs between any demand and its nearest facility and simultaneously determine the service reliability (Huang et al., 2007).

For the EMS allocation problem, which is addressed in this paper, the covering location model formulation is more applicable in that: (i) acceptable service coverage standards measured by response time are typically pre-specified by the emergency management authorities; and (ii) full coverage is a prerequisite of *P*-median and *P*-center location models. However, it is not always guaranteed to maintain full demand coverage in the EMS allocation problem due to resource scarcity. Conversely, providing adequate service coverage from the EMS vehicle location to the demand site within a predefined service time becomes essential. As one of the early contributors in this research area, Toregas et al. (1971) introduced the Set Covering Location Problem (SCLP), in which a full demand coverage was still mandatory. Church and Revelle (1974) were the first to investigate the maximum deterministic covering model, named the Maximal Coverage Location Problem (MCLP). Subsequently, Daskin (1982, 1983) explored a stochastic version of the MLCP, termed as the Maximum Expected Covering Location Problem (MEXCLP), by employing a uniform probability distribution for any EMS vehicle being busy within the network to capture service uncertainty. The objective was to maximize the expected demand coverage. Later efforts on the covering type models mainly dealt with improving the basic models with the consideration of multiple server types and multiple demand coverages (Hogan and ReVelle, 1986; Marianov and ReVelle, 1992). Gendreau et al. (1997, 2006) introduced the double standard model (DSM), along with a Tabu search algorithm, that considered two response standards with different coverage requirements as constraints to maximize the portion of the demand being doubly covered using a more restrictive standard (primary standard) and also ensure the demand being fully covered by a less strict standard (secondary standard). Doerner et al. (2005) solved the ambulance location problem for eight provinces in Austria. A variant of the ant colony optimization model with a modified double-coverage objective function and single-coverage standard constraints was formulated and a Tabu Search algorithm was used to find a near optimal solution. As an extension of the MCLP, Erkut et al. (2007) introduced the Maximum Survival Location Problem (MSLP) in which a monotonic decreasing survival function was incorporated into the formulation to relax the relation between response time and survival rate. Subsequently, Laporte et al. (2009) developed a dynamic version of DSM and tested the model on the Island of Montreal. Meanwhile, the static version of the DSM was applied to data from three different countries. Schmid and Doerner (2010) formulated a mixed integer model for the EMS allocation problem that accounts for time-varying travel time and deployment by allowing vehicles to be repositioned to ensure a certain coverage level. The problem was solved meta-heuristically using the variable neighborhood search technique. Kepaptsoglou et al. (2012) proposed a DSM that maximizes the vehicle crash site coverage with a predefined number of ambulances dedicated to provide EMS services. The model relied on a genetic algorithm (GA) to derive solution and was applied in the city of Thessaloniki, Greece. A most recent study conducted by Liu et al. (2014) further improved the DSM formulation, along with a solution algorithm, that explicitly considers multiple EMS vehicle types and variant coverage requirements for demand sites with different priority levels. The refined model was applied for allocating EMS vehicles in the Chicago urban area. For the comprehensive review of facility location problems and optimization techniques employed, addition references can be found in Brotcorne et al. (2003), Goldberg (2004), Arostegui et al. (2006), and Li et al. (2011).

The majority of the EMS allocation models found in the literature are deterministic in nature. In fact, the uncertainty of EMS vehicle availability and demand site accessibility in emergency management operations is common and should be adequately addressed (Huang and Fan, 2011). Even with EMS vehicles made available to specific demand sites, vehicle accessibility can be affected by background traffic and the travel time may fluctuate due to spatial and temporal traffic dynamics or other nonrecurring traffic spikes. An empirical study by Kolesar et al. (1975) revealed that some variations in travel time were encountered by firefighters to arrive at emergency demand sites in the city of New York. More recently, Budge et al. (2010) reached a similar conclusion in a study of assessing changes in the average travel time experienced by EMS ambulances in the city of Calgary, Canada. In both studies, the magnitude of travel time changes was found to be marginal. As such, treating the average travel time or speed of EMS vehicles as a deterministic factor is potentially acceptable. However, it is less so for the issue of uncertainty concerning service availability of EMS vehicles. This is especially true when limited EMS fleets are used to serve extensive emergency calls across large geographical areas.

Service reliability is defined as the availability of EMS service when it is so desired. So far, the majority of existing covering location models for solving the EMS allocation problem have not handled the issue of uncertainty associated with service reliability of EMS vehicles. For the very few models that have attempted to do so, approaches used generally include adopting: (i) multiple service coverage (Hogan and ReVelle, 1986; Marianov and ReVelle, 1992; Gendreau et al., 1997, 2006; Doerner et al., 2005; Laporte et al., 2009; Kepaptsoglou et al., 2012; Liu et al., 2014); and (ii) chance constraints to ensure a certain level of service reliability (ReVelle and Hogan, 1989; Ball and Lin, 1993; Huang and Fan, 2011).

The DSM is a typical example of the multiple service coverage approach. It aims to allocate facilities to maximize the portion of demands being doubly covered within a tight response standard (primary standard) and also ensures a full coverage within a less tight response standard (secondary standard). Both service coverage standards are measured as the travel time in minutes required by an EMS vehicle to travel from the vehicle depot location to the EMS demand site. In the DSM, the double service coverage treats each demand site separately and uniformly, which may generate results of service redundancy for some demand sites and shortage for other sites. The chance constraint approach establishes a performance-based reliability level by simultaneously considering the EMS vehicles and competing demand sites, which could avoid the above limitations of imbalanced EMS vehicle allocation from the multiple service coverage approach. However, it still lacks concurrent considerations of achieving maximized number of the demand sites to be reliably covered in the DSM.

Motivated by weaknesses identified in the existing models, this study introduces a new model that stems from the DSM developed by Liu et al. (2014). The previous DSM explicitly considers different types of EMS vehicles and two service coverage standards, while the proposed model goes one step beyond by combining the double-standard concept with chance constraints instead of double coverage to minimize possible service redundancy. Since this new formulation simultaneously handles availability of EMS vehicles and service reliability requirements from competing demand sites, it helps maximize the coverage of important demand sites with guaranteed service reliability, leading to a more efficient demand-responsive EMS vehicle allocation. Here, important demands include all critical sources of EMS calls that are mainly considered when assigning EMS vehicles in practice. For demands associated with all other non-critical sources of demand calls, EMS vehicles will still try to provide service as possible.

Without loss of generality, the new model is formulated to address emergency services to human injuries caused by vehicle crashes at high crash frequency intersections within an urban street network. The high crash frequency intersections are selected as demand sites and are grouped into high and low crash severity sites. The EMS vehicles are classified into two types of vehicles to provide services for the two groups of demand sites according to two service coverage standards. The model aims to maximize the portion of demand sites with reliable service coverage. In this study, if a demand site can be timely responded by an EMS for most of the times, say 90% of the times, it is a reliable service coverage. A heuristic algorithm is introduced to produce an efficient solution. The model and solution algorithm are applied in a computational experiment.

The remainder of this paper is organized as follows: Section 2 presents details of the proposed model and its solution algorithm. Section 3 elaborates on the computational study for model application. Finally, Section 4 presents a summary and draws conclusion.

2. Proposed methodology

2.1. Consideration of resource uncertainty

When an emergency service call is received from a demand site i, the EMS responder may consider to dispatch a type k EMS vehicle that could potentially arrive from its depot location to the demand site within the arrival time of service coverage standard r. However, the EMS vehicle may be busy serving a competing demand at that time. As such, one or more EMS vehicles in the neighborhood of the demand site that could meet the service coverage standard should be targeted for deployments to provide reliable service. In order to estimate the lowest number of EMS vehicles needed to be contacted for deployments around this demand site i to ensure a certain level of service reliability, the concept of local busy fraction introduced by ReVelle and Hogan (1989) and ReVelle and Marianov (1991) can be employed. The local busy fraction is the probability that an EMS vehicle is not able to serve demand site i when it is busy in serving another emergency call. It is estimated as the total service time required by competing demand sites around demand site i divided by the total available service time of the EMS vehicles deployed around demand site i. In this study, in order to embed the concept of local busy fraction into the proposed model, a modified definition of local busy fraction is proposed as follows:

$$q_i^k = \frac{\bar{t}^k \cdot \sum_{l \in C_{i,r}} f_l}{24 \cdot \sum_{j \in S_{i,r'}} y_j^k} = \frac{\rho_i^k}{\sum_{j \in S_{i,r'}} y_j^k} \tag{1}$$

where q_i^k = local busy fraction of one type k EMS vehicle deployed at any $j \in S_{i,r'}$; r, r' = service coverage standards; $\bar{t}^k$ = average service time of type k EMS vehicle (hours per service); f_l = frequency of requests for EMS by competing demands around demand site i (service calls per day); $S_{i,r'} = \{\forall j | t_{ij} \leqslant r'\}$, the set of depot locations within r' of demand site i; $C_{i,r} = \{\forall m | t_{mj} \leqslant r, j \in S_{i,r'}\}$, the set of m competing demand sites around demand site i, which are located within r of any depot location $j \in S_{i,r'}$; y_j^k = number of type k EMS vehicles deployed at depot location j; ρ_i^k = utilization ratio.

Differentiated from the original definition of local busy fraction mentioned above, the proposed definition of local busy fraction considers the involvement of multiple service coverage standards. In the new formulation, r is service coverage standard corresponding to the m competing demand sites and r' is the service coverage standard relevant to the demand site i that requests for EMS. Considering multiple service coverage standards used in the analysis, r and r' might be different and, generally, r should always be the secondary standard in the double standard model formulation. The reason is that it is safe to assume that all types of EMS vehicles should respond to any service request that called within any service coverage standard. In this paper, r is always denoted the secondary standard.

Assuming that the requests for services from different demand sites are independent, for demand site i, all types of EMS vehicles deployed at any vehicle depot location $j \in S_{i,r'}$ should have the same local busy fraction value. Therefore, the probability of one or more EMS vehicles not available follows a binomial distribution. The probability of having at least one of the type k EMS vehicles available can be computed in the following:

$$Prob(at\ least\ one\ type\ k\ EMS\ vehicle\ is\ available) = 1 - Prob(all\ type\ k\ EMS\ vehicles\ are\ busy)$$

$$= 1 - (q_i^k)^{\sum_{j \in S_{i,r'}} y_j^k} = 1 - \left(\frac{\rho_i^k}{\sum_{j \in S_{i,r'}} y_j^k} \right)^{\sum_{j \in S_{i,r'}} y_j^k} \tag{2}$$

With a predefined service reliability level, $\gamma \in [0, 1]$, for demand site i, in order to satisfy the requirement that the probability of having at least one type k EMS vehicle available within service coverage standard r', the following condition holds:

$$1 - \left(\frac{\rho_i^k}{\sum_{j \in S_{i,r'}} y_j^k} \right)^{\sum_{j \in S_{i,r'}} y_j^k} \geqslant \gamma \tag{3}$$

which can be replaced by the following expression:

$$\sum_{j \in S_{i,r'}} y_j^k \geqslant e_{i,r'}^k \tag{4}$$

where $e_{i,r'}^k$ is the smallest integer that satisfies

$$1 - \left(\frac{\rho_i^k}{e_{i,r'}^k} \right)^{e_{i,r'}^k} \geqslant \gamma \tag{5}$$

In this study, the utilization ratio, ρ_i^k, which is defined as the average service time of type k EMS vehicle, $\bar{t}^k$, multiplied by the hourly frequency of requests for EMS by competing demands around demand site i, f_l, which can be estimated from historical data using Eq. (1). Thus, with known ρ_i^k, $e_{i,r'}^k$ is also predetermined, which means Eq. (4) needs to be treated as a constraint in the proposed model, while all other equations are to be used as inputs at the pre-processing stage.

2.2. Double standard model

The original DSM formulation introduced by Gendreau et al. (1997, 2006) aims to maximize the doubly covered demand sites according to two service coverage standards for a given ambulance fleet. In particular, it ensures that a portion of demand sites is covered by two EMS vehicles within the primary standard and all demand sites must be covered by at least one EMS vehicle within the secondary standard where the primary standard maintains a narrower time window for site arrival.

With EMS vehicle allocation problem centered on rescuing human injuries caused by vehicle crashes, Liu et al. (2014) extended the original DSM formulation by splitting the EMS vehicles into two groups, termed as BLS and ALS vehicles. The BLS vehicles are equipped with basic rescue equipment, but can be quickly dispatched to crash sites demanding for emergency services. Conversely, the ALS vehicles are augmented from BLS vehicles with advanced equipment, which are in need of a longer time to be dispatched. When providing emergency services to human injuries caused by vehicle crashes, time of essence. For high severity crash sites, it is advantageous to quickly dispatch BLS vehicles. If the crashes can be handled on sites, medical treatments will be immediately implemented. Otherwise, the injured travelers could be sent to the nearest hospitals as early as possible. Whereas high or low crash severity sites involving less emergent EMS calls, ALS vehicles could be assigned for treatments of injured travelers at crash scenes. In fact, a recent study (Sanghavi et al., 2015) found that patients with life-threatening emergencies, including trauma, stroke, and respiratory failure have higher survival rate for 90 days if they were transported in a BLS ambulance rather than a ALS one. With the requirements of having each demand site covered by at least one EMS vehicle within the secondary standard with a portion of such sites meeting the required service reliability and ensuring a portion of demand sites covered by each type of EMS vehicles within the primary standard, the extended model aims to maximize the coverage of demand sites according to the primary standard with guaranteed service reliability. The new model proposed hereinafter further refines the previous two DSM models to explicitly ensure service reliability for the coverage of demand sites.

2.3. Model description

Let D be the set of demand sites where travelers injured from vehicle crashes at intersections within an urban street network are in need of emergency medical services, with d_i representing the demand for service calls at demand site $i \in D$. The set of potential vehicle locations is denoted as S, and for each depot location $j \in S$, the maximum number of vehicles that can be deployed at this location is predefined as p_j. All demand sites/depot locations are located on an undirected graph

$G = (N, E)$, where N and E are the sets of nodes and edges that correspond to intersections and street segments of an urban street network. An undirected graph is used to indicate the permission for EMS vehicles to move in any direction of the street regardless of the vehicular traffic direction.

Consistent with typical double standard models, the current model also contains two service coverage standards, r_1, primary standard, and r_2, secondary standard. The demand sites are high crash frequency intersections within an urban street network. An intersection is identified as a demand site if (i) the total number of fatal and injury crashes of Type A (major injury), Type B (moderate injury), and Type C (minor injury) without inclusion of property damage only (PDO) crashes; or (ii) equivalent injury Type B crashes by excluding PDO crashes occurred within its range on average per year over a multi-year period is significantly higher than the grand average value per year per intersection for the entire network. In both site selection criteria, PDO crashes are excluded. Also, intersection-related crashes refer to those occurred within the range measured from its center along each approach by 250 ft or 75 m. For each demand site, the crash severity index is calculated as the ratio between the total number of equivalent injury Type C crashes per year and the total number of fatal and all types of injury crashes per year. A demand site with high crash frequency is regarded as a high crash severity site if the crash severity index exceeds a certain threshold value. As such, the demand sites are further grouped into low and high crash severity sites.

Two types of EMS vehicles, BLS and ALS, are used to provide emergency services. The BLS vehicles can be quickly dispatched that are suitable to provide emergency services for high crash severity sites that often require immediate emergency medical services. Compared with BLS vehicles, the ALS vehicles carry advanced equipment that can properly treat all severity level crashes on-site or on-the-way to the emergency center, but require a longer time for dispatching due to medical equipment preparation. EMS service providers normally maintain a larger fleet size of ALS vehicles because of their better functionality. In all, BLS vehicles can be assigned for high crash severity sites to quickly respond to service calls and ALS vehicles can be allocated to both high severity and low severity sites to initiate on-site treatments immediately after site arrival. The fleet size is controlled and an upper limit is set for total number of vehicles that can be deployed at a vehicle depot location.

The objective of the proposed model is to optimally deploy EMS vehicles at potential depot locations within an urban street network to demand sites according to the double service coverage standards. The coverage of demand sites within the primary standard r_1 obeys the following rules: (1) demands of high crash severity sites will be covered if a certain reliability level (γ) is ensured, implying that there is always at least one EMS vehicle, either a BLS or an ALS vehicle, is available within the primary standard r_1; (2) demands of low crash severity sites will be covered only if a certain reliability level (γ) is ensured, indicating that there is always at least one ALS vehicle is available within the primary standard r_1. This demand coverage rules encourage BLS vehicles, which have a shorter response time and smaller fleet size, to concentrate around high crash severity demand sites that need to immediate emergency medical services.

Similar to the typical double standard models, this model also imposes constraints of having at least one EMS vehicle arrive at any demand site within the secondary standard r_2 and at least a portion (α) of demand sites is guaranteed to be reachable by each type of the EMS vehicles within the primary coverage standard r_1. In practice, the use of either a BLS or an ALS vehicle to serve a demand site depends on the severity level of crashes occurred at the site and the vehicle availability. Unlike the typical double standard models, one more constraint is added in the proposed model. This will ensure that the coverage of at least a portion (β) of demand sites will be guaranteed by a certain level of service reliability (γ) when they request for emergency services according to the secondary standard r_2.

2.4. Model formulation

The model formulation is of the following specification:

$$\text{Maximize} \sum_{i\in D} d_i \cdot x_i \tag{6}$$

Subject to :

$$\sum_{j\in S_{i,r2}} y_j^k \geqslant 1 \quad \forall i \in D,\ k \in K \tag{7}$$

$$\sum_{i\in D} d_i \cdot z_{i,r_1}^k \geqslant \alpha \cdot \sum_{i\in D} d_i \quad \forall k \in K \tag{8}$$

$$\sum_{i\in D} d_i \cdot x_{i,r_2}^k \geqslant \beta \cdot \sum_{i\in D} d_i \tag{9}$$

$$\sum_{j\in S_{i,r1}} y_j^k \geqslant z_{i,r_1}^k \quad \forall i \in D,\ k \in K \tag{10}$$

$$\sum_{j\in S_{i,r2}} \sum_{k\in K} y_j^k \geqslant e_{i,r_2} \cdot x_{i,r_2}^k \quad \forall i \in D \tag{11}$$

$$\sum_{j\in S_{i,r1}} y_j^1 \geqslant e_{i,r_1} \cdot x_i \quad \forall i \in D_2 \tag{12}$$

$$\sum_{j\in S_{i,r1}}\sum_{k\in K} y_j^k \geqslant e_{i,r_1}\cdot x_i \quad \forall i\in D_1 \tag{13}$$

$$\sum_{j\in S} y_j^k \leqslant p^k \quad \forall k\in K \tag{14}$$

$$\sum_{k\in K} y_j^k \leqslant p_j \quad \forall j\in S \tag{15}$$

$$D = D_1 \cup D_2 \tag{16}$$

$$z_{i,r_1}^k, x_{i,r_2}^k, x_i \in \{0,1\} \quad \forall i\in D,\ k\in K \tag{17}$$

$$y_j^k \geqslant 0, integer \quad \forall j\in S,\ k\in K \tag{18}$$

$$S_{i,r1} = \{j\in S : t_{ij} \leqslant r_1\} \quad \forall i\in D \tag{19}$$

$$S_{i,r2} = \{j\in S : t_{ij} \leqslant r_2\} \quad \forall i\in D \tag{20}$$

where **input parameters**: D = Set of demand sites; D_1 = Set of high crash severity demand sites; D_2 = Set of low crash severity demand sites; d_i = Demands at demand site $i \in D$; $k \in K$, k = Vehicle type, 1 for ALS, 2 for BLS; r_1 = Primary service coverage standard; r_2 = Secondary service coverage standard; e_{i,r_1} = Smallest number of EMS vehicles assigned around demand site i at depot location $j \in S_{i,r1}$, that can satisfy Eq. (5) with $r' = r_1$ in Eq. (1); e_{i,r_2} = Smallest number of EMS vehicles assigned around demand site i at depot location $j \in S_{i,r2}$, that can satisfy Eq. (5) with $r' = r_2$ in Eq. (1); p_j = Maximum number of vehicles that can be deployed to vehicle location j; p^k = Fleet size of type k vehicles; S = Vehicle depot locations; $S_{i,r1}$ = Set of depot locations that can reach demand site i within primary standard r_1; $S_{i,r2}$ = Set of depot locations that can reach demand site i within secondary standard r_2; t_{ij} = Vehicle travel time on the shortest path from vehicle location j to demand site i; **decision variables**: y_j^k = Number of type k vehicles deployed at vehicle location j; z_{i,r_1}^k = 1 if demand site i is reachable by at least one type k vehicle within primary standard r_1; x_{i,r_2}^k = 1 if demand site i is covered by type k vehicle within secondary standard r_2 at service reliability level γ; x_i = 1 if demand site i is covered within primary standard r_1 at reliability level γ.

The objective function (6) aims to maximize the number of demand sites D covered within the primary standard r_1 at required service reliability level γ, where BLS vehicles will be only dispatched to demand sites involved with high severity crashes. The constraints as shown in Eqs. (7), (8), and (10) ensure that each demand site is covered by at least one EMS vehicle as appropriate according to the secondary standard r_2, and at least a portion (α) of demand sites are covered by each type of EMS vehicles as per the primary standard r_1. Constraints (9), (11), (12), and (13) are service reliability constraints. Constraints (9) and (11) jointly ensure that at least a portion (β) of demand sites are covered by EMS vehicles with service reliability level γ according to the secondary standard r_2. Eq. (12) ensures that, for low crash severity demand sites, if the number of ALS vehicles deployed within primary standard r_1 is greater than or equal to the smallest integer that obtains service reliability level γ, then those demand can be covered. For high crash severity demand sites, Eq. (13) ensures that if the total number of BLS and ALS vehicles deployed within primary standard r_1 is greater than or equal to the smallest integer that obtains service reliability level γ, those demand sites can be covered.

Eqs. (14) and (15) ensure that the total number of vehicles equals to the given fleet size, p^k, for both types, and the number of vehicles deployed at potential vehicle location j does not exceed its capacity, p_j. As shown in Eq. (16), any demand site is either in demand set D_1 or D_2. Eqs. (17) and (18) indicate that variables x_{i,r_2}^k and x_i take 0/1 integer values, and the number of vehicles of both types located at each potential vehicle location is a non-negative integer. Finally, Eqs. (19) and (20) define depot location sets for demand site i according to the primary and secondary standards r_1 and r_2, $S_{i,r1}$ and $S_{i,r2}$.

2.5. Solution algorithm

The current double standard model formulation belongs to the general class of mixed integer optimization models. This type of problems is often *NP*-Complete and/or *NP*-Hard, which is hard to solve by an exact solution algorithm in polynomial time or to verify a possible solution in polynomial time. Conversely, heuristic algorithms are sought to obtain near optimal solutions in reasonable time. For instance, the original double standard model introduced by Gendreau et al. (1997, 2006) was solved by a Tabu Search algorithm. For the current model, the solution technique of the genetic algorithm (GA) is employed to reach a near optimal solution (Aickelin, 2002; Arostegui et al., 2006). Specifically, the model is implemented and solved by a Microsoft Excel-based solver, EVOLVER 6. The first step for the iterative computation of the GA is to represent the decision variables by a population of random strings, which is then operated by reproduction (selection), crossover, and mutation to create a new population of points. At each of the three steps of operators' process, if a bad string is produced it is removed from the population and those with good features are kept as a part of the new population. The resultant collection of a better population is used to generate the next population, and the fitness of the newly generated population is obtained as the value of the objective function. At each step, the solution obtained is stored as best, if the solution is an improvement from the last process.

3. Applying the methodology for EMS vehicle allocation in city of Chicago

3.1. Data collection

Detailed crash records of one thousand major signalized intersections in the city of Chicago are collected for 2004–2010 and are categorized by crash severity level, including fatal, injury Types A, B, and C for major, moderate, and minor injury levels, and property damage only (PDO), as well as categorization by major crash types.

In this study, two traffic safety metric scenarios are utilized for identifying the top two hundred high crash frequency intersections as demand sites and to calculate the demand for emergency service at each identified intersection, d_i: (i) total number of fatal and all types of injury crashes as the frequency-based (FB) scenario; and (ii) total number of equivalent injury Type B crashes as the severity-based (SB) scenario. For demand sites on each list, they are further split into high and low crash severity demand sites. Specifically, a demand site is considered as a high crash severity site if its crash severity index, calculated as the ratio of equivalent injury Type C crashes to the total of fatal and all types of injury crashes, is greater than a threshold value of 2. Otherwise, it is treated as a low crash severity site. The crash data analysis has yielded 91 and 132 intersections as high crash severity sites, and 109 and 68 intersections as low crash severity sites for the FB scenario and SB scenario, respectively.

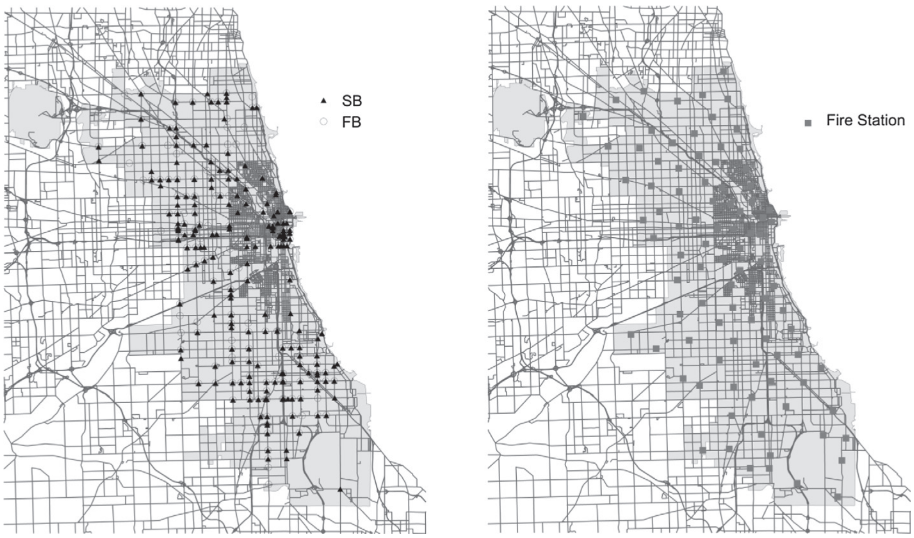
On the supply side, 92 Chicago Fire Department (CFD) operated fire stations are considered as potential EMS vehicle depot locations. All 75 EMS vehicles could be assigned to these fire stations with a capacity of up to 2 vehicles per station. Fig. 1 illustrates the demand sites classified for the two scenarios and the EMS vehicle depot locations within the study area.

3.2. Model settings

Several important factors, including response priorities, emergency service types, and EMS vehicle characteristics, that highly affect an efficient and quality emergency service for roadway safety, are considered in the proposed model. Parameters are defined based mainly on the requirements and realities in practical EMS operations.

A 4–5 min response time for BLS and a 8–10 min response time for ALS standard are recommended for use as the EMS effectiveness measurement by several fire departments (Sa'adah, 2004); and based on this a 5-min primary standard, r_1, and a 10-min secondary standard, r_2, are adopted in this study. The average service time for EMS vehicles is set to be 1.5 h, $\bar{t}^k = 1.5\forall k$, which is adopted from a previous work by ReVelle and Marianov (1991). The service reliability level γ is set at 90%.

The shortest travel time for an individual dispatch is defined by the shortest route distance in the network divided by the average speed of an ambulance running at 30 mph that is consistent with the speed limit for most of the Chicago urban streets. Based on the recommended standard response times for both BLS and ALS ambulances, all demand sites should be covered within the secondary standard r_2. The minimum coverage level of demand sites by at least one ALS ambulance and one BLS ambulance within the primary standard r_1 is set at 40%, namely, $\alpha \geqslant 40\%$. Furthermore, at least 50% of the demands are set to be served by the two types of ambulances within secondary standard r_2 at the pre-specified service



(a) Demand sites (b) EMS vehicle depot locations

Fig. 1. Demand sites and potential EMS vehicle depot locations in the city of Chicago.

reliability level of 90%, which means $\beta \geqslant 50\%$ and $\gamma = 90\%$. The objective is to achieve maximized coverage of demand sites within the primary standard r_1 and guaranteed by 90% of the service reliability.

3.3. Utilization ratio calculation

Since e_{i,r_1} and e_{i,r_2} in the proposed model are determined by the utilization ratio ρ_i^k and service reliability level γ using Eq. (5), it is imperative to compute ρ_i^k by using Eq. (1) prior to estimate e_{i,r_1} and e_{i,r_2}. The service call frequency at each demand site, f_l, is estimated by using the historical data reported in 2012. Specifically, a total of 333,842 EMS calls were received and handled by CFD in 2011 (Roche and Friszell-Neroulas, 2012). It is assumed that all service call frequencies are normally distributed with the same mean and variance and the value of f_l at all demand sites is randomly generated based on this assumption. With the value of estimated service call frequency per day at all demand sites, the steps to compute utilization ratios for a demand site i are as follows: (1) defining $S_{i,r'}$, $r' \in \{r_1, r_2\}$, to locate all depot locations at which EMS vehicles assigned are capable of responding to the service call from i; (2) determining all demand sites, including i, that are within secondary service coverage standard of depot locations in $S_{i,r'}$, $l \in C_{i,r} = \{\forall i | t_{ij} \leqslant r, j \in S_{i,r'}\}$, then sum their competing demands f_l ensuring no-duplication; and (3) computing utilization ratio by using Eq. (1). With known utilization ratio, the values of e_{i,r_1} and e_{i,r_2} are determined by executing a MATLAB script coded for this process. Since the range of utilization ratios for all the demand sites in this computational study is relatively small, a loop statement is used in executing MATLAB script to test integers from one to a relatively large number in Eq. (5) to derive their values.

3.4. GA application

3.4.1. Chromosome design and criteria representation

For the purpose of this study, the definition of Chromosome is adopted from the existing literature, as a set of EMS vehicles deployed in every depot location (Liu et al., 2014). Below is a possible solution of the model represented by an array:

$$Y = \begin{bmatrix} y_1^1 & y_2^1 & y_3^1 & \cdots & y_j^1 & \cdots & y_{N_s}^1 \\ y_1^2 & y_2^2 & y_3^2 & \cdots & y_j^2 & \cdots & y_{N_s}^2 \end{bmatrix} \tag{21}$$

where N_s = the total number of depot locations, which is 92 for this study.

Given the known shortest travel time between each potential vehicle location j and every demand site i, t_{ij}, and coverage indicators z_{i,r_1}^k, $x_{i,r_2}^{1,2}$, x_i, $\forall k \in K$, a valid solution Y is represented as follows.

$$z_{i,r_1}^k = \begin{Bmatrix} 1, & \exists\, j \in S_{i,r_1} : t_{ij} \leqslant r_1 \cap y_j^k \geqslant 1 \\ 0, & otherwise \end{Bmatrix} \quad \forall i \in D \tag{22}$$

$$x_{i,r_2}^k = \begin{Bmatrix} 1 & \sum_{j \in S_{i,r2}} \sum_{k \in K} y_j^k \geqslant e_{i,r_2} \\ 0 & otherwise \end{Bmatrix} \quad \forall i \in D \tag{23}$$

$$x_i = \begin{Bmatrix} 1 & \forall i \in D_1 : \sum_{j \in S_{i,r1}} \sum_{k \in K} y_j^k \geqslant e_{i,r1};\ \forall i \in D_2 : \sum_{j \in S_{i,r1}} y_j^1 \geqslant e_{i,r1} \\ 0 & otherwise \end{Bmatrix} \tag{24}$$

In GA representation, z_{i,r_1}^k, $x_{i,r_2}^{1,2}$, and x_i can be directly determined based on definitions using conditional statements in programming. By using the structure of GA representation shown above in Eqs. (21)–(24), constraints (10)–(13) and (16)–(17) in the proposed model are handled internally. Meanwhile, Eqs. (8) and (9) are transplanted into GA. To represent Eq. (7) in GA, two more coverage indicators, z_{i,r_2}^k, $k \in K = \{1, 2\}$, are used with the following condition holds:

$$z_{i,r_2}^k = \begin{Bmatrix} 1, & \exists\, j \in S_{i,r_2} : t_{ij} \leqslant r_2 \cap y_j^k \geqslant 1 \\ 0, & otherwise \end{Bmatrix} \quad \forall i \in D \tag{25}$$

where $z_{i,r_2}^k = 1$ if demand site i is reachable by at least one type k vehicle within secondary standard r_2; 0, otherwise. Then, representation of Eq. (7) is with the following inequalities:

$$\sum_{i \in D} z_{i,r_2}^k \geqslant N_d \quad \forall k \in K \tag{26}$$

where N_d = the total number of demand sites, which is 200 in this study.

3.4.2. Fitness measure

Each component of the objective function is treated as the product of demand at demand site i, d_i, and its coverage indicators, x_i. In this study, m_i is introduced in the GA with definition provided as follows:

$$m_i = \left\{ \begin{array}{ll} d_i & \text{if } x_i = 1 \\ 0 & \text{otherwise} \end{array} \right\} \quad \forall i \in D \tag{27}$$

Then, the fitness measure of the GA for proposed model is presented as M, where

$$M = \sum_{i \in D} m_i \tag{28}$$

3.4.3. Termination condition

The algorithm is terminated after 50,000 successive trials when the current best solution could not be improved by more than 0.1%. The justified number of successive trials in this condition is experimentally obtained by conducting several test runs that terminate when the solution could not be improved for a sufficient long time, say 24 h.

3.4.4. Initial feasible solution

The execution of the GA requires a feasible solution to begin with. For the purpose of this study, a pre-optimization is conducted specifically to help obtain an initial feasible solution that all model constraints are satisfied. In the pre-optimization process, a special form of the proposed GA is run, in which all constraints are retained while a new objective that maximize the number of constraints satisfied is employed. This special GA optimization terminates when all model constraints are satisfied. In other words, an initial feasible solution of the proposed GA is obtained.

3.4.5. GA calibration

The FB scenario is used with different population sizes, crossover rates, and mutation rates for calibrating the GA. Candidate population sizes are 25 and 50, crossover rates are considered as 0.2, 0.4, and 0.6, while the mutation rates are considered as 0.05, 0.15, and 0.25. Average fitness measure of 5 runs for each combination of GA settings are used to make the decision as shown in Table 1. The best value is obtained for a population size of 25, crossover value of 0.2, and mutation rate of 0.05. Fig. 2 depicts the evolution of GA fitness measure over trials.

3.5. Results

3.5.1. EMS vehicle allocation results

Table 2 summaries results of assigning EMS vehicles by executing the proposed model to cover the top 200 high crash frequency intersections as demand sites selected using the FB and SB scenarios. In both cases, 60 ALS and 15 BLS are allocated among the 92 candidate local fire stations as EMS vehicle depot locations consistent with the current practice of the city of Chicago to maximize reliable coverage for the demand sites. Eight runs are conducted for each scenario. When using an Intel i7-2600k CPU@3.40 GHz, the time needed to reach final solution varies from 100 to 140 min.

Following the requirement of the 10-min secondary service coverage standard, any demand site can be adequately covered by at least one EMS vehicle without considering the service reliability. Further imposing 90% of required service reliability, the results show that at best 92.4–95.5% of the demand sites can be covered and on average about 83.4–91.3% can be covered. The coverage levels for both cases are significantly higher than the β value of at least 50%. The differences in the best and average levels of service coverage are within 5–9%. This reveals that the existing EMS fleet operated by Chicago Fire Department is sufficient in guaranteeing a high level of service coverage for vehicle crash-related emergency services within a 10-min service coverage standard at a reliability level of 90%.

According to the 5-min primary service coverage standard, the best results show that 87.6–95.9% of the demands can be singly covered by ALS ambulances and 71.9–75.5% can be singly covered by BLS ambulances, which are significantly greater than the α value of at least 40%. After imposing an additional constraint of guaranteed service reliability γ at 90%, at best 65.5–68.4% of the high severity demands and 50.2–54.5% of the low severity demands are properly covered. On average, the service coverage levels are 83.9–90.1% and 68.6–70.1% by ALS and BLS ambulances without the service reliability requirement; and are 59.7–62.2% and 48.5–49.9% for high and low severity demands with guaranteed service reliability, respectively. The differences between the best and average coverage levels for all cases are well within 9%.

With requirements of the 5-min primary standard and 90% of the required service reliability, the coverage level for high severity demand sites is higher than low severity demand sites by 11–18.2%. This may be explained by the fact of assigning both BLS and ALS ambulances to high crash severity demand sites to handle more severe crashes at high severity demand sites that need immediate services, but only allocating ALS ambulances to low severity demand sites to treat less severe

Table 1
GA calibration.

Average fitness measure	Crossover rate (%)		20			40			60		
	Mutation rate (%)		5	15	25	5	15	25	5	15	25
FB	Population size	25	3190	3105	2958	2874	2898	2762	2743	2927	2616
		50	3015	2869	3049	2955	3038	2876	3007	2748	2868

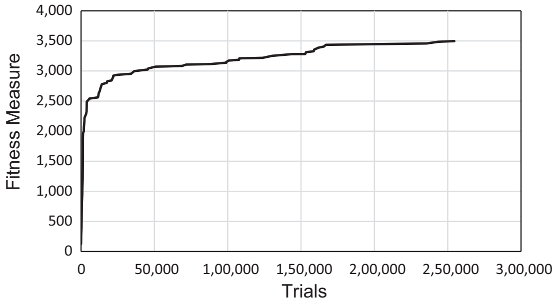


Fig. 2. GA fitness measure progress over trials.

crashes. Having a smaller number of ALS ambulances for low severity demand sites than that of both types of ambulances for high severity demand sites, it leads to a lower level of demand coverage for low severity demand sites accordingly.

As shown in Fig. 3, EMS vehicles tend to be assigned in pairs (larger circles) to depot locations, which is evident by the larger nodes outnumbering the smaller nodes for both FB and SB scenarios. This is not unexpected because the proposed model aims to provide more reliable service coverage. Consequently, EMS vehicles are allocated more intensively in clusters to achieve the service objective. This can also be observed in Fig. 3. In the FB scenario, ambulances concentrate around central and southern Chicago areas, while nodes that denote EMS vehicle deployments can be hardly found in the western area. In the SB scenario, southern Chicago area is again well covered. Conversely, in the middle part of the city, only the central business district is assigned with EMS vehicles. In addition, Fig. 4 shows that BLS ambulances (circles) mainly deployed to rapidly treat high severe crashes occurred at the sites classified as high severity demand sites (triangles).

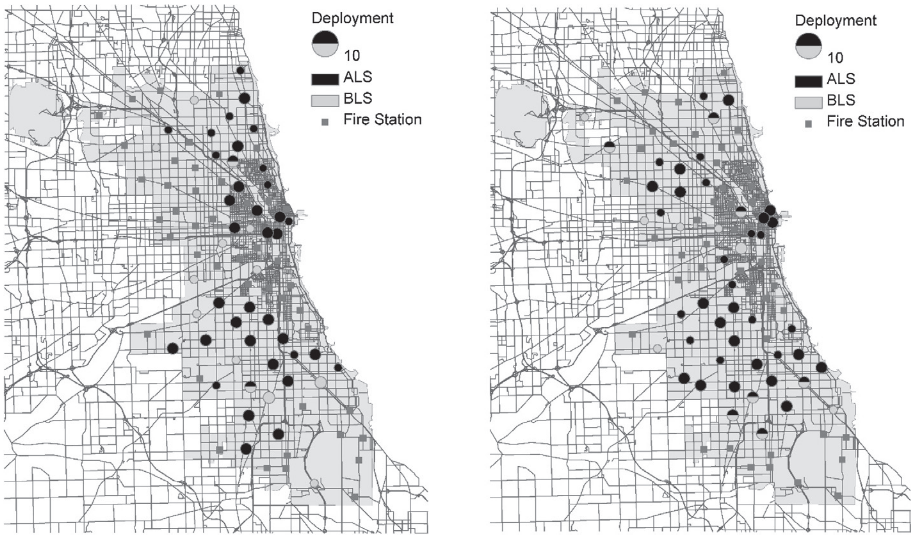
3.5.2. Model comparisons

In the proposed model formulation that incorporates the constraint of guaranteed service reliability, it aims to find the optimal deployments of EMS vehicles to achieve maximized coverage of demand sites to ensure that for each covered demand site at least one EMS vehicle is available in 90% of the time for site dispatching within the 5-min primary standard. In order to meet this high reliability requirement, EMS vehicles are deployed intensively as clusters. As a result, a certain extent of losses for singly covered demand within the primary standard without guaranteed service reliability may occur. In order to examine this effect, the proposed model and an earlier double standard model developed by Liu et al. (2014) are slightly modified by removing the reliability constraint in the proposed model (namely, the current model) and replacing the double coverage of demand sites in the objective function of the earlier model by single coverage (i.e., the previous model). Such modifications enable cross comparisons of EMS vehicle allocation results on an equal basis in that both models concentrates on maximizing the number of singly covered demand sites within the primary standard without imposing the service reliability requirement as a model constraint. Next, the two models are applied to the current data set for result comparisons. As shown in Table 3, for the FB and SB scenarios, demand calls singly covered within the primary standard without constraining service reliability are 95.2–98.1% for the current model and 99.2–100.0% for the previous model, respectively. Both models are able to ensure a high level of demand coverage at over 95%. Compared with the previous model, the reductions in the single coverage of demand from the current model are by 1.9–4%.

In order to assess the impacts of guaranteed service reliability on demand coverage, the proposed model (aims to achieve maximized coverage of demand sites within the primary standard with guaranteed service reliability) and the previous model (focuses on reaching maximized double coverage of demand sites within the primary standard) are further executed

Table 2
EMS vehicle allocation results.

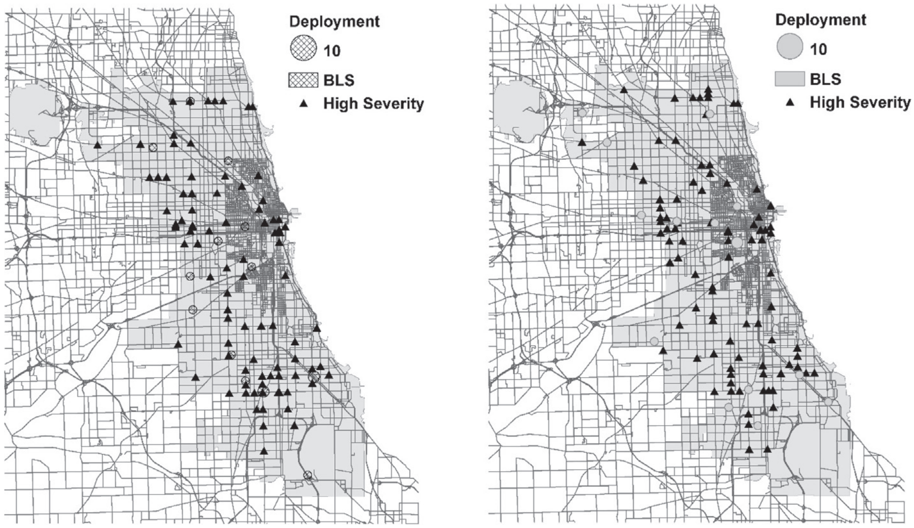
Coverage of demand (%)		Fitness measure M	Secondary standard: $r_2 = 10$ min		Primary standard: $r_1 = 5$ min		Trials to optimal solution
			Required coverage: $\beta \geqslant 50\%$ Reliability: $\gamma = 90\%$	Required coverage: $\alpha \geqslant 40\%$	Reliability: $\gamma = 90\%$		
					High severity sites (BLS and ALS)	Low severity sites (ALS)	
Best	FB	3496	92.4	ALS 87.6 BLS 71.9	65.5	54.5	254,000
	SB	4343	95.5	ALS 95.9 BLS 75.5	68.4	50.2	284,000
Avg.	FB	3217	91.3	ALS 90.1 BLS 70.1	62.2	48.5	211,000
	SB	4022	83.4	ALS 83.9 BLS 68.6	59.7	49.9	190,000



(a) FB scenario

(b) SB scenario

Fig. 3. EMS vehicle allocation within the Chicago street network.



(a) FB scenario

(b) SB scenario

Fig. 4. BLS vehicles quickly dispatched to high severity demand sites.

using the current data set for pre-specified levels of service reliability using increments of 10% from 10% to 90%. As shown in Table 4, increases in the required service reliability will lead to decreases in demand coverage. For the current model, the demand coverage levels are maintained at 59.5–64.8% and 58.6–64.3% for the FB and SB scenarios by reducing the required service reliability from 90% to 10%. The reductions in demand coverage caused by enhancing the service reliability from 10% to 90% are well within 6%. For the previous model, the demand coverage levels are 22.6–51.4% and 9.5–40.7% for the FB and SB scenarios by relegating the required service reliability from 90% to 10%. Unlike the current model, reductions in demand coverage are more than 30% due to increases in the required service reliability from 10% to 90%. Regardless of the level of required service reliability, the current model outperforms the previous model significantly in terms of the percentage of demand sites being covered and the growing gaps of demand coverage for higher levels of required service reliability. This may be attributable to the use of different model formulation concepts. In the current model, a demand site is covered within the 5-min primary standard as long as it is reachable by at least one EMS vehicle that is not busy in 90% of the time.

Table 3
Comparisons of single coverage of demand within the primary standard without considering required service reliability.

Single coverage of demand (%)	Primary standard: r_1 = 5-min, without service reliability requirement Scenario	
	FB	SB
The current model	95.2	98.1
The previous model	99.2	100.0

Table 4
Comparisons of demand coverage within the primary standard for different levels of required service reliability.

Coverage of demand (%)	Primary standard: r_1 = 5-min									
	Scenario	Required service reliability levels: γ								
		10%	20%	30%	40%	50%	60%	70%	80%	90%
The current model	FB	64.8	64.0	64.0	64.0	62.9	61.5	60.8	59.5	59.5
	SB	64.3	62.7	62.3	62.3	62.3	61.7	59.5	59.5	58.6
The previous model	FB	51.4	48.4	44.0	42.2	40.1	36.4	32.8	27.2	22.6
	SB	40.7	36.5	36.2	31.7	30.1	23.1	19.5	15.3	9.5
Changes in demand coverage	FB	13.4	15.6	20	21.8	22.8	25.1	28.0	32.3	36.9
	SB	23.6	26.2	26.1	30.6	32.2	38.6	40.0	44.2	49.1

Contrariwise, the previous model does not consider the required service reliability. This limitation is compensated by requesting for at least two EMS vehicles reachable to the demand site for its coverage, which may generate results of service redundancy for some demand sites and shortage for other sites. With the total number of EMS vehicles being fixed, this implies that fewer demand sites could be simultaneously served. If it so happens that multiple vehicles reachable to a specific demand site within 5 min are all busy in serving other concurrent competing demand calls, the demand site is virtually not covered. As such, EMS vehicle allocation relying on the current model is more efficient (by assigning just no less than one EMS vehicle to a demand site) and reliable (by ensuring the EMS vehicle could reliably reach the demand site) than that of the previous model.

4. Conclusions

This study introduces a new model for optimal emergency service vehicle management in the urban area that achieves a certain level of service reliability under limited vehicle resources. The model requires that any demand site is covered by at least one EMS vehicle within the secondary standard with a portion of the covered sites achieving the service reliability requirement, ensures that a portion of the sites is covered by each type of EMS vehicles within the primary standard, and aims to maximize the demand sites covered within the primary standard meeting the required service reliability.

The model application provides valuable insight into reliable urban emergency service vehicle management. For the top two hundred high crash intersections in the city of Chicago with vehicle crashes involving approximately 333,842 emergency service calls per year, the city's EMS vehicle fleet that consists of 15 BLS and 60 ALS ambulances could provide service coverage for about 92.4–95.5% of demand within the 10-min secondary service coverage standard at 90% of service reliability; and 65.5–68.4% of high severity demand and 50.2–54.5 low severity demand within the 5-min primary standard at 90% of service reliability.

Cross comparisons of the proposed model and an earlier model for emergency service vehicle management are made by model applications to the same dataset. Both models are able to ensure a high level of demand coverage at over 95% and the demand coverage achieved by the previous model is marginally higher by 1.9–4%. Further analysis is performed to assess the impacts of guaranteed service reliability on demand coverage. The demand coverage levels within the 5-min primary standard are calculated for pre-specified levels of service reliability using increments of 10% from 10% to 90%. It is discovered that the demand coverage levels obtained from the current model are consistently higher than those of the previous model. Also, reduction in demand coverage as a result of increases in the required service reliability associated with the current model (within 6%) is significantly smaller than that relevant to the previous model (exceeds 30%). Model comparison results indicate that the proposed model is advantageous over the earlier model in efficient allocation of limited EMS vehicle resources by simultaneously considering the extent of demand coverage and the reliability of service coverage. In this context, the proposed model is more effective to assist in urban emergency service vehicle management.

One major contribution of the proposed model is to combine the double service coverage standards and multi-vehicle assignments to each demand site to ensure service coverage with the service reliability requirement using chance constraints. This treatment helps maximize the coverage of demand sites with guaranteed service reliability, leading to a more efficient demand-responsive emergency service vehicle management. The model rigorously connects required emergency service coverage standard, service reliability, and available vehicle resources (as inputs) with the service coverage level (as the outcome) that could be achieved. The proposed model offers new research directions in urban emergency service vehicle management, particularly by considering adjusted travel time from the vehicle location to the demand site responsive to traffic dynamics, interactions of EMS vehicles, and EMS service reliability across the large urban area and vehicle delays that could be experienced at intersections owing to inefficient traffic signal coordination in the urban context.

Acknowledgments

The authors acknowledge the transportation agencies in the Chicago metropolitan area for data collection, processing, and computational analysis as part of model application.

References

Aickelin, U., 2002. An indirect genetic algorithm for set covering problems. J. Oper. Res. Soc. 53, 1118–1126.
Arostegui, M.A., Kadipasaoglu, S.N., Khumawala, B.M., 2006. An empirical comparison of tabu search, simulated annealing, and genetic algorithms for facilities location problems. Int. J. Prod. Econ. 103, 742–754.
Ball, M.O., Lin, L.F., 1993. A reliability model applied to emergency service vehicle location. Oper. Res. 41, 18–36.
Brotcorne, L., Laporte, G., Semet, F., 2003. Ambulance location and relocation models. Eur. J. Oper. Res. 147 (4), 451–463.
Budge, S., Ingolfsson, A., Zerom, D., 2010. Empirical analysis of ambulance travel times: the case of calgary emergency medical services. Manage. Sci. 56 (4), 716–723.
Church, R.L., ReVelle, C.S., 1974. The maximal covering location problem. Pap. Reg. Sci. 32 (1), 101–118.
Current, J., Daskin, M., Schilling, D., 2002. Discrete network location models. In: Drezner, Z., Hamacher, H.W. (Eds.), Facility Location: Application and Theory. Springer-Verlag, NY.
Daskin, M.S., 1982. Application of an expected covering model to emergency medical service system design. Decis. Sci. 13 (3), 416–439.
Daskin, M.S., 1983. A maximum expected location model: formulation, properties and heuristic solution. Transport. Sci. 7, 48–70.
Dobson, R., 2003. Advances in medical technology account for a third of the reduction in road traffic death. BMJ 326 (7397), 1004.
Doerner, K.F., Gutjahr, W.J., Hartl, R.F., Karall, M., Reimann, M., 2005. Heuristic solution of an extended double-coverage ambulance location problem for Austria. Cent. Eur. J. Oper. Res. 13 (4), 325–340.
Erkut, E., Ingolfsson, A., Erdogan, G., 2007. Ambulance location for maximum survival. Naval Res. Logist. 55, 42–58.
Gendreau, M., Laporte, G., Semet, F., 1997. Solving an ambulance location model by tabu search. Location Sci. 5 (2), 75–88.
Gendreau, M., Laporte, G., Semet, F., 2006. The maximal expected coverage relocation problem for emergency vehicles. J. Oper. Res. Soc. 57, 22–28.
Goldberg, J., 2004. Operations research models for the deployment of emergency services vehicles. EMS Manage. J. 1, 20–39.
Hogan, K., ReVelle, C.S., 1986. Concepts and applications of backup coverage. Manage. Sci. 34 (1986), 1434–1444.
Huang, Y., Fan, Y., 2011. Modeling uncertainties in emergency service resource allocation. J. Infrastruct. Syst. 17 (1), 35–41.
Huang, Y., Fan, Y., Cheu, R.L., 2007. Optimal allocation of multiple emergency service resources for protection of critical transportation infrastructure. J. Transport. Res. Rec. 2022 (1), 1–8.
Kepaptsoglou, K., Karlaftis, M.G., Mintsis, G., 2012. Model for planning emergency response services in road safety. J. Urban Plann. Dev. 138 (1), 18–25.
Kolesar, P., Walker, W., Hausner, J., 1975. Determining the relation between fire engine travel times and travel distances in New York City. Oper. Res. 23 (4), 614–627.
Laporte, G., Louveaux, F.V., Semet, F., Thirion, A., 2009. Application of the double standard model for ambulance location. Innov. Distrib. Logist., 235–249
Li, X., Zhao, Z., Zhu, X., Wyatt, T., 2011. Covering models and optimization techniques for emergency response facility location and planning: a review. Math. Methods Oper. Res. 74, 281–310.
Liu, Y., Arash, M.R., Li, Z., Kepaptsoglou, K., Patel, H., Lu, X., 2014. Heuristic approach for optimizing emergency medical services in road safety within large urban networks. J. Transport. Eng., 04014043
Marianov, V., ReVelle, C.S., 1992. The capacitated standard response fire protection sitting problem: deterministic and probabilistic models. Ann. Oper. Res. 40 (1992), 303–322.
ReVelle, C., Marianov, V., 1991. A probabilistic FLEET model with individual vehicle reliability requirement. Eur. J. Oper. Res. 53 (1), 93–105.
ReVelle, C., Hogan, K., 1989. The maximal covering location problem and α-reliable P-center problem: derivatives of the probabilistic location set covering problem. Ann. Oper. Res. 18 (1), 155–174.
Roche, K., Friszell-Neroulas, E., 2012. 2011 National Run Survey. Retrieved from <www.emsworld.com>.
Sa'adah, R., 2004. The Eight-Minute ALS Response Time Standard: A Review and Discussion of its Use as a Strategic Result Goal by the District of Columbia. District of Columbia Fire and Emergency Medical Services Department.
Sanghavi, P., Jena, A.B., Newhouse, J.P., Zaslavsky, A.M., 2015. Outcomes of basic versus advanced life support for out-of-hospital medical emergencies. Ann. Intern. Med. 163, 681–690.
Schmid, V., Doerner, K.F., 2010. Ambulance location and relocation problems with time-dependent travel times. Eur. J. Oper. Res. 207 (3), 1293–1303.
Toregas, C., Swain, R., ReVelle, C.S., Bergman, L., 1971. The location of emergency service facilities. Oper. Res. 19 (6), 1363–1373.
Wells, P., 2007. Deaths and injuries from car accidents: an intractable problem? J. Clean. Prod. 15 (11–12), 1116–1121.