Computers & Industrial Engineering 72 (2014) 12–23

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

A simulation optimization framework for ambulance deployment and relocation problems

Lu Zhen [a,*], Kai Wang [a], Hongtao Hu [b], Daofang Chang [b]

[a] School of Management, Shanghai University, 99 Shangda Road, Shanghai 200444, China
[b] Logistics Engineering College, Shanghai Maritime University, 1550 Haigang Avenue, Shanghai 201306, China

ARTICLE INFO

Article history:
Received 7 August 2013
Received in revised form 4 March 2014
Accepted 6 March 2014
Available online 15 March 2014

Keywords:
Ambulance deployment
Ambulance relocation
Simulation optimization
Genetic algorithm

ABSTRACT

This paper studies ambulance deployment and relocation problems, which are two of the core decisions faced by emergency medical service control centers in metropolis. The challenge in the problems is to estimate the operational performance of a deployment plan under stochastic environments. More specifically, the stochastic and dynamic nature of request arrivals, fulfillment processes, and complex traffic conditions as well as the time-dependent spatial patterns of some parameters complicate the decisions in the problems. This paper proposes a simulation optimization method that enables evaluating the operational performance of deployment plans through a detailed simulation model. For guiding the search process in the simulation optimization method, the genetic algorithm is employed in this study. On the basis of the deployment decisions, a mathematical model on ambulance relocation is also proposed for adapting to the dynamic changing environments along the time. To illustrate the proposed method's usage in practice, a demo example about its application in Shanghai is given. Some numerical experiments are also performed to validate the effectiveness and the efficiency of the proposed methods.

1. Introduction

Rising costs of medical equipments, increasing call volumes, and worsening traffic conditions in metropolis make emergency medical service control centers face increasing pressure so as to meet performance targets. The service control centers are supposed to locate a proper number of ambulances in some bases (waiting locations) so that medical service requesters can be reached in a time efficient manner. Uneven distribution of population in the city makes the ambulances should not be evenly deployed in the bases. A medical service control center needs to decide how many ambulances should be deployed in all the waiting locations, respectively. This decision making process is in a dynamic environment where the spatial distribution of potential requesters are changing along the time, and the spatial patterns of traffic situations in a city are also different in peak hours and off-peak hours. The ambulance deployment decision is also in a stochastic environment where the request calls arrive at the control center in a random manner; the travel time for a certain journey may contain randomness; the service time at the request calls' scenes and hospitals is also uncertain. The above mentioned dynamic and stochastic nature of the request arrivals and ambulance fulfillment processes as well as the environments complicates the ambulance deployment decision.

The challenge in ambulance deployment decision is to estimate the operational performance of a deployment plan. Simulation optimization method is a proper way that can enable assessing the operational performance of deployment plans through a detailed simulation model, which can capture multiple sources of uncertainties. Therefore this paper makes an explorative study on ambulance deployment problem by using a simulation optimization methodology.

Moreover, the ambulance deployment decision is not constant and static; it should be dynamically changing along the time because the input data for deployment decisions at different time intervals are varying. Thus, there are ambulance relocation processes between two consecutive time intervals. The relocation decision is based on the optimal deployment decisions in two consecutive time intervals. Therefore, besides the ambulance deployment problem, this paper also investigates the ambulance relocation problem, so that the results of this study can supply a potentially useful decision support tool on intelligent ambulance scheduling for emergency medical service providers.

The remainder of this paper is organized as follows. Section 2 is the literature review. Problem backgrounds of the ambulance deployment and relocation problems are elaborated in Section 3. Then the detailed introductions about the proposed simulation

* Corresponding author. Tel.: +86 21 66134237.
E-mail address: lzhen.sh@gmail.com (L. Zhen).

http://dx.doi.org/10.1016/j.cie.2014.03.008
0360-8352/

optimization framework and some key components are given in Section 4. Section 5 illustrates a demo example of the ambulance deployment and relocation problem applied in Shanghai. Some numerical experiments are performed in Section 6 for a further investigation on the proposed method. Closing remarks and conclusions are outlined in the last section.

2. Related works

Most studies on ambulance deployment or location problems are based on a minimal covering model (Toregas, Swain, ReVelle, & Bergman, 1971), which tries to minimize the number of ambulances necessary so as to cover all demand points, and a maximal covering model (Church & ReVelle, 1974), which tries to maximize the total demand covered given a feet of fixed size. Then based on the above models, some new models were proposed to consider the possibility that ambulances may be unavailable and some demand points may not be covered. For example, the double standard model (DSM) was proposed by Gendreau, Laporte, and Semet (1997); the probability model for unavailable ambulance was developed by Daskin (1983). The DSM model was applied to the data coming from the eight rural provinces in Austria (Doerner, Gutjahr, Hartl, Karall, & Reimann, 2005). Gendreau, Laporte, and Semet (2001) also extended their model into a dynamic environment to take advantage of the available time between consecutive calls by anticipating future decisions on the deployment of the fleet. The DSM was extended from single period to multiple periods (Schmid & Doerner, 2010). Some scholars used integer programming models to study the real-time ambulance redeployment problems (Brotcorne, Laporte, & Semet, 2003; Gendreau, Laporte, & Semet, 2006; Kolesar & Walker, 1974). The objectives of these integer programming models are mainly formulated from two perspectives, i.e., the backup coverage for future calls, and the relocation cost of ambulances. Solving these models is usually time-consuming as they need to solve an optimization sub-problem every time a decision is made. Therefore, decision makers usually resort to a parallel computing environment for implementing a real-time system. Shariat-Mohaymany, Babaei, Moadi, and Amiripour (2012) proposed two reliability-based linear models for optimal location of ambulances. Some studies considered the randomness in the system explicitly, either through a dynamic programming formulation or through heuristic approaches. Berman (1981a, 1981b, 1981c) proposed some frameworks by using the dynamic programming approaches for the ambulance redeployment problem. These papers follow an exact dynamic programming formulation, thus the formulation is tractable only in oversimplified versions of the problem with few vehicles and small transportation networks. Andersson and Vaerband (2007) investigated the ambulance deployment decision by using a preparedness function that essentially measures the capability of a certain ambulance configuration to cover future calls. The preparedness function is similar with the value function in a dynamic program, which assesses the impact of a current decision on the future evolution of the system. However, the way of the preparedness function belongs to some sort of heuristic methods in nature. The simulation optimization method was used in the personnel deployment at an emergency department healthcare unit (Ahmed & Alkhamis, 2009). Iannoni, Morabito, and Saydam (2009) studied how to optimize the locations of ambulance bases on highways. Zhang, Puterman, Nelson, and Atkins (2012) integrated the simulation optimization method with the demographic and survival analysis. A decision support system was developed for setting long-term care capacity planning. Underwood, Zhang, Denton, Shah, and Inman (2012) proposed a genetic algorithm based simulation optimization method to design PSA screening policies. Some interesting findings and policy recommendations were obtained from their study.

When compared with the above described models, our method provides some merits. In contrast to the models that are based on either integer programming or dynamic programming methodologies, our method can captures the more complex and random evolution of the system over time, and the stochastic nature of request arrivals, fulfillment processes, and complex traffic conditions as well as the time-dependent spatial patterns of some parameters, all of which complicate the decisions in the problem of this study. This paper proposes a simulation optimization method by involving the GA meta-heuristic and a simulator. Moreover, an ambulance relocation model is also proposed for adapting to the dynamic changing environments along the time.

3. Problem background

This paper studies the ambulance deployment problem and the ambulance relocation problem. (1) The ambulance deployment problem origins from the uneven spatial distribution of potential medical service requests. The problem is concerned with how to deploy a given amount of ambulances among the waiting stations in the city so as to optimize the service performance, which can be quantified by some indicators. (2) The ambulance relocation problem exists because the information about the request arrivals and traffic conditions is time dependant. This relocation problem is on the basis of the previous deployment problem. The optimal ambulance deployment decision depends on the parameter setting about the request distributions and traffic conditions, which are influenced by the time. Given two optimal deployment plans that belong to two consecutive time intervals, the ambulance relocation problem is concerned with how to relocate ambulances so as to minimize the total cost (distance) of relocation routes between the ambulance waiting stations in the city.

A request for medical service usually arrives by phone and is answered by a dispatcher, who inputs the information by asking some predefined questions, and determines the priority of the request calls. When a request becomes known to the dispatching system, the dispatcher checks with available ambulances and assigns the request to an ambulance. In this process, there are some possible criteria for the decision. For example, the nearest available ambulance may be chosen; or some other criteria can also be embedded in the decision support system for dispatching. The time span between the request arrival and the setting off of a chosen ambulance should be the shorter the better. This time span usually contains the time necessary for querying some information about the actual incident, negotiation with ambulance drivers, and the setup time for the crew to get ready. When the ambulance arrives at the patient location (i.e., the call's scene), the crew of the ambulance may need to take some first-aid measures for the patient. When the service is completed, the ambulance takes the patient to a hospital. Sometimes, the patient may appoint a hospital in case the incident is not very urgent; otherwise, the ambulance transport the patient to a hospital, which has the available resources for treating the patient and is the nearest from the call's scene. When the ambulance arrives at the hospital, the crew starts to unload the patient and deliver him (or her) to the corresponding department in the hospital. Then the whole service for the request is finally completed. The ambulance becomes idle (available) again and goes to a waiting location if there are not new requests assigned to it; otherwise, the ambulance needs to set off to the next call's scene immediately. Fig. 1 shows the above process.

The above elaborates the operational level process of ambulance dispatching. On the tactical level, the control center of ambulances needs to make a decision about how to deploy a given amount of ambulances among the waiting stations in the city