

A HEURISTIC APPROACH FOR LOCATING EMS FACILITIES AND VEHICLES

Constance A. Lightner
Department of Management
Fayetteville State University
Fayetteville, North Carolina

Jonathan M. Graham
Computer Science Department
Norfolk State University
Norfolk, Virginia

ABSTRACT

Many industries thrive on their ability to provide rapid and efficient services to customers. Attaining this goal is particularly critical for life sensitive service providers like fire departments and emergency medical services (EMS). This issue is complicated by the possibility that a vehicle may be unavailable when called for service. In this paper we analyze current techniques for locating EMS and highlight some of the problems associated with these methods. We propose a genetic algorithm to locate emergency facilities and vehicles to minimize the number of locations which cannot be covered in a specified time. Our proposed genetic algorithm is driven by the evaluation of the fitness of chromosomes through a discrete event simulation model. This method of evaluating the fitness of the chromosomes is used because of the difficulty of evaluating the required probabilities. We believe that our approach will lead to a general cost efficient technique for locating EMS.

Keywords: EMS, allocation, genetic algorithm

1. INTRODUCTION

The emergency medical service (EMS) provides preliminary assistance and transportation for patients in need of emergency care. EMS ultimately aims to reduce mortality resulting from acute illnesses and trauma. Attainment of this goal depends upon the rapid response of qualified individuals, Ling and Benton [16]. The number of ambulances employed by a system, and their respective locations are two factors that EMS officials can plan, in order to minimize response times and contribute to achieving this goal. Several research studies have been conducted to explore the relationship between response time and mortality, Persse, *et. al.* [17]. All of these studies found a direct relationship between these variables. The importance of responding in a minimal time is therefore very important.

The research presented in this paper analyzes the state of the art techniques for locating emergency medical facilities and vehicles to minimize the number of calls not serviced within a critical time period. A review of literature revealed that many linear programming models for emergency vehicle allocation depend upon accurately modeling the probabilistic event of a vehicle being unavailable for service when called. Subtle changes in estimating the likelihood of this event can drastically affect the optimal allocations rendered by the model. The lack of an available model to provide a closed form, optimal, solution to our problem motivated the development of our heuristic technique. To overcome these problems we propose a genetic algorithm to obtain the optimal location/allocation. Our technique models vehicle availability using a discrete event simulation of an actual EMS system, to evaluate potential allocations.

This paper makes the following contributions to the literature:

- An analysis of current locating techniques for EMS. This will permit researchers in the field to determine if these techniques are applicable for their particular problem.
- A proposed genetic algorithm (GA) approach which is flexible enough to solve problems of this type.

2. RELATED WORK

Early literature on location problems was based upon the traditional set covering problem. Set covering describes the minimum number of facilities needed to cover all system demand. A population is considered served, or covered, when a facility is sited within the maximal service distance. Church & Reville [4] presented the Maximal Covering Location Problem (MCLP) to maximize the total amount of population served within a maximal service distance, given a fixed number of facilities.

Hakimi [13] developed a model similar to the MCLP termed the P-median problem. The P-median problem weights the distance between demand nodes and facilities by the associated demand quantity and calculates the total weighted travel distance between demand nodes and facilities. The model then seeks to find the location of P facilities so as to minimize the total demand weighted travel distance between demands and facilities. Additional research has been conducted to expand upon the P-median approach to the emergency vehicle location problem and analyze sources of modeling errors (Erkut & Bozkaya [6], Francis *et. al.* [7]). These models offer techniques to improve the efficiency of the traditional P-median problem.

A review of the EMS system in Austin, Texas that relied upon MCLP based models exposed the necessity to construct alternative models that addressed the issue of vehicle availability. Daskin's [5] Maximal Expected Coverage Location Problem (MEXCLP) examines equipment availability within the context of location models. The objective of the MEXCLP is to locate emergency vehicles so as to maximize the expected coverage area, even in the event that multiple vehicles are in use. The model is based on the assumption that the probability that a randomly chosen vehicle is busy is independent of any other vehicle being in use. Bernardo and Repede [1] later modified this model to incorporate time varying demand and multiple states of vehicle availability.

Schilling *et. al.* [19] developed the Tandem Equipment Allocation Model (TEAM) and Facility-Location Equipment-Emplacement Technique (FLEET) model to allocate equipment with varying capabilities and demands, and possibly allocate the equipment and facilities simultaneously. The FLEET model has proven to be effective in the location-allocation of fire stations and equipment.

Bianchi and Church's [2] Multiple coverage, One-unit FLEET problem (MOFLEET) consolidates the MEXCLP and FLEET models by examining factors that must be addressed in ambulance location problems. This model minimizes the chances that an ambulance will not be available when demanded, by simultaneously allowing for coverage by alternative units. In multiple coverage models, a demand is defined as being completely covered when the demand area can be reached by at least M vehicles within the specified time period. Gendreau, Laporte, and Semet [9] also addressed the multiple vehicle coverage issue by using a tabu search heuristic to determine optimal allocations.

The models created by Daskin [5], Schilling, *et. al.* [19], and Bianchi and Church [2] all address the possibility of an ambulance being busy when called, using a system wide busy probability. This system-wide probability simplifies the model greatly; however, it is highly unlikely that a system will actually have a uniform system busy probability. Instead, it is more realistic to expect vehicles in the inner city to have different levels of activity from vehicles located in remote, rural areas. This realization motivated the incorporation of local busy estimates (Reville and Hogan [18], Lightner and Tavakoli [15]).

Brotcorne, *et. al.* [3] and Goldberg [12] compiled a comprehensive review of the 30 year history of ambulance location and relocation models. They categorized past models as follows: 1) Deterministic models which ignore stochastic considerations regarding the availability of ambulances; 2) Probabilistic models which reflect the fact that ambulances operate as servers in a queuing system and can sometimes be unavailable for calls; and 3) Dynamic models for relocating and deploying ambulances based upon real time data (Gendreau, *et. al.* [9]). The first two categories directly address the issue of initially designing an EMS system to optimally locate facilities and distribute vehicles to these sites. The third is focused on relocating and deploying vehicles as demand patterns change. This research is concerned with addressing the initial system design.

The literature on deterministic approaches fail to address the possibility that a vehicle could be busy (responding to a previous call) when emergency services are requested. This inaccurately assumes that a single vehicle could immediately cover all calls that are accessible within the target response time. Realistically the potential for vehicle unavailability necessitates that multiple vehicles be available to some areas in order to ensure that all calls are immediately answered. The available probabilistic techniques rely upon accurately estimating either global or local busy probability estimates.

Lightner and Tavakoli [15] conducted a study comparing Daskin's [5] traditional MOFLEET model and a modified MOFLEET model which incorporated local busy estimates. The allocations rendered by both models, for a given number of vehicles and facilities, were compared using a discrete event simulation. Simulation results showed that the model which rendered the maximum level of coverage was not consistent, as the number of

facilities and vehicles to be allocated varied. Furthermore, they revealed that subtle changes in busy probability estimates significantly altered the optimal allocations rendered by the models. Current literature on emergency facility/vehicle allocation does not provide an analytic model for optimally locating resources. This void, coupled with the need to model probabilistic behavior, motivated the creation of the heuristic technique presented in this paper.

3. MODEL DESCRIPTION AND PROBLEM FORMULATION

3.1. Notation

- I = the set of all demand grids
- J = the set of potential facility sites ($J \in I$)
- x_j = the number of vehicles located at facility j for ($j \in J$)
- a_{it} = the number of calls generated from grid i at time t
- E = the number of vehicles that need to be sited
- S = the maximum time period to answer a call
- F = the number of facilities to be sited
- C = the maximum number of vehicles to site at any facility
- $N_i = \{j \in I \mid m_{ij} \leq S\}$
- $y_{it} = \begin{cases} 1 & \text{if there is no vehicle available within } N_i \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$
- $z_j = \begin{cases} 1 & \text{if there is a facility located at grid } j \\ 0 & \text{otherwise} \end{cases}$

3.2. Model

Our model consists of a grid of g areas, each corresponding to a geographical location in a city. A call arrives to the central 911 number from an area in the grid and is routed to a facility located within the grid. The 911 center aims to have EMS vehicles respond to calls within a target of s minutes. Thus a 911 call is routed to a facility with an available vehicle that can respond within s minutes. If there are no vehicles available to service a call within s minutes, for the purpose of this research, we will consider the call lost. Calls arrive with a specified inter-arrival time and are drawn from a determined probability distribution. The probability of calls arriving from each grid is known in advance.

3.3. Problem formulation

Our goal is to find an allocation of facilities to grid areas and vehicles to facilities so that the fraction of lost calls is minimized. We are constrained by the number of facilities, the number of available vehicles and the capacity of each facility. More formally, we can state our optimization problem as follows:

$$\text{Minimize } \sum_{i \in I} \sum_t a_{it} y_{it} / \sum_{i \in I} \sum_t a_{it}$$

Subject to:

$$\sum_{j \in J} x_j = E$$

$$\sum_{j \in J} z_j = F$$

$$x_j \leq Cz_j \quad \text{For all } j \in J$$

$$x_j \geq 0 \quad \text{Integer for all } j \in J$$

$$z_j \quad \text{and} \quad y_{it} = \{0,1\} \quad \text{For all } i, j, \text{ and } t$$

In the next section we will discuss the relationships between the variables in the above model, and provide an overview of our how our approach will manipulate the variables to solve this problem.

4. VEHICLE LOCATION

Our approach to the problem is to design a genetic algorithm (GA) that selects the location of F facilities (from set J) and then allocates E vehicles to the facility sites. The fitness of a potential facility/vehicle allocation will be evaluated using a discrete event simulation. The simulation will be given the allocation of vehicles, x , and will generate a_{it} (for each grid in the system), according to known probability distributions. Each time a call is generated, a vehicle will be dispatched from a site within s minutes (if available). As vehicles are dispatched, x will be updated to reflect unavailable vehicles that are currently providing emergency services. Moreover, y_{it} will be modified based upon the adjusted x , to reflect whether grid i has an available vehicle within its service area at time t to respond to a call. Ultimately our genetic algorithm will seek to find the allocation of facilities and vehicles which yields the lowest fraction of lost calls.

4.1. Genetic algorithm

We represent the chromosome as a string of digits with values ranging from zero to the number of vehicles located at that facility. An allele value of zero means that there is no facility at that grid area. The length of the chromosome corresponds to the number of grid area.

4.1.1. Chromosome

The fitness function is the fraction of calls that were lost. This fitness is determined by passing the chromosome data to a simulation package. The allocation specified by the chromosome is simulated to determine the expected fraction of lost calls.

4.1.2. Fitness function

The selection operator selects two parents for mating based on their fitness, relative to the population. The selection operator we use is the tournament selection [11] with a size of two.

4.1.3. Selection

The selection operator selects two parents for mating based on their fitness, relative to the population. The selection operator we use is the tournament selection [11] with a size of two.

4.1.4. Crossover

Single and multi-point crossover defines crossover points as places between loci where an individual can be split. Uniform crossover [10] generalizes the scheme to make every locus a potential crossover point. In our adaptation of uniform crossover we generated a random binary number. If the value was 0 then child one received an allele from parent 1 and child 2 its allele from parent 2. If the value was 1 then child one received an allele from parent 2 and child 2 its allele from parent 1.

4.1.5. Mutation

In mutation, we moved a vehicle from one location and placed it at another location.

4.1.6. Ensuring valid chromosomes

The versions of crossover and mutation used may sometimes result in an invalid chromosome. We adjust these invalid chromosomes in such a way as to keep the chromosome as close to the invalid chromosome as possible.

4.2. Model input data

The data required for our heuristic are compiled from the Cumberland County emergency dispatch database (managed by Cumberland County Information Services) and EMS officials. Based on calls received, the county was divided into grid zones. The structure of these zones is based on travel-time, population density, traffic flow, and other pertinent considerations. Following analysis of these factors, a network consisting of 119 grid zones was established. In effort to meet the national goal, our model was designed to minimize the amount of unanswered calls within eight minutes. Selecting eight as our target response time sets the standard for determining the expected coverage that an optimal allocation will yield. The response time for EMS services is measured from the time a call comes in through the system's dispatching center, until an EMS vehicle reaches the emergency site. Increasing or

decreasing this value will respectively alter the expected coverage rendered by our technique for a single vehicle distribution.

We analyzed the system to determine the eight-minute coverage for each zone. The accuracy of the travel distances was verified through several processes. A sample of 14,130 EMS calls was used to determine the volume of calls in each node. The average daily volume of calls was used to generate calls within grids.

4.3. Simulation

We developed a discrete event simulation model in Awesim to evaluate the fitness of each chromosome. The GA algorithm which is implemented in C++, copies the chromosome information, as well as information about the distance between facilities, to a data file. The GA program then invokes the AWESIM environment, which reads the data file and performs the simulation of calls and dispatching of vehicles. The result of the simulation is an estimate of the fraction of calls lost. The model simulates network calls for three years (in a single simulation run) and the results are averaged over 20 simulation runs.

5. DISCUSSION AND FUTURE WORK

We looked at several approaches for locating EMS facilities and vehicles. The deterministic approaches fail to consider the possibility of vehicle unavailability, while the probabilistic approaches suffer from heavy dependence upon accurately estimating the probability of this occurrence. A slight error in its estimation could drastically change the results. Our approach is promising because it addresses vehicle unavailability by simulating the actual EMS system for a given vehicle allocation. Ultimately, it selects the vehicle allocation which provides the lowest fraction of lost calls.

In the future we intend to compare the genetic algorithm and several other heuristic techniques to those obtained in the literature. In addition we will explore various ways of choosing the available vehicle from among several choices. Currently, when a call arrives, we choose any available vehicle.

6. REFERENCES

- [1] Bernardo, J., and J. Repede (1988). Two-tier Emergency Medical Service System: Analysis and Extension of Current Location Methodologies. *Presented at the 1988 Annual Meeting of the Decision Sciences Institute, Las Vegas, NV.*
- [2] G. Bianchi and R.L. Church (1988). A Hybrid Fleet Model for Emergency Medical Design. *Social Sci. Med*, 26, 163-171.
- [3] Brotcorne, L. G. Laporte, and F. Semet (2003). Ambulance Location And Relocation Models. *European Journal of Operational Research*, 147 (3), 451-463.
- [4] Church, R. and C. Reville (1974). The Maximal Covering Location Problem. *Papers Regional Sci.*, 32, 101-118.
- [5] Daskin, M.S. (1983). A Maximal Expected Set Covering Location Model: Formulation, Properties, and Heuristic Solution. *Transportation Science*, 17, 48-69.
- [6] Erkut, E., and B. Bozkaya (1999). Analysis of Aggregation Errors for the P-Median Problem. *Computers and Operations Research*, 26, 1075-1096.
- [7] Francis, R.L., T.J. Lowe, and Arie Tamir (2000). Aggregation Error Bounds for a Class of Location Models. *Operations Research*, 48/2, 294-307.
- [8] Gendreau, M., G. Laporte, and F. Semet (2000). Solving an Ambulance Location Model by Tabu Search. *Location Science*, 5/7, 75-88.
- [9] Gendreau, M., G. Laporte, and F. Semet (2001). A Dynamic Model And Parallel Tabu Search Heuristic For Real-Time Ambulance Relocation. *Parallel Computing*, 27, 1641-1653.

- [10] Goldberg, D. (1989). Genetic Algorithms in Search, Optimization and Machine. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA.
- [11] Goldberg, D. E. and Deb, K (1991). A comparative analysis of selection schemes used in genetic algorithms. *Foundations of Genetic Algorithms*, 69-93.
- [12] Goldberg, J.B. (2004). Operations Research Models for the Deployment of Emergency Services Vehicles. *EMS Management Journal*, 1 (1), 20 -39
- [13] Hakimi, S.L. (1964). Optimum Locations of Switching Centers and the Absolute Centers and Medians of a Graph. *Operations Research*, 12: 450-459.
- [14] Jayaraman, V. and Srivastava-R. (1995). A service logistics model for simultaneous siting of facilities and multiple levels of equipment. *Computers and Operations-Research*, 91-204.
- [15] Lightner, C. and Tavakoli, A (2005). Implementing A Mathematical Model For Locating Facilities And Vehicles To Minimize Response Time,” *Proceedings of the 2005 International Academy of Business and Public Administration*, 2005
- [16] Ling, L. and W.C. Benton (2003). Hospital Capacity Management Decisions. *European Journal of Operational Research*, 146/3, 596-614.
- [17] Persse, D.E., Key, C.B., Bradley, R.N., Miller, C.C. and A. Dhingra, (2003). Cardiac Arrest Survival as a Function of Ambulance Deployment Strategy in a Large Urban Emergency Medical Services System. *Resuscitation*, 59/1, 97-104.
- [18] Revelle, C. and K. Hogan (1986). A Reliability Constrained Siting Model with Local Estimates of Busy Fractions. *Environ. Planning*, B 15, 143-152.
- [19] Schilling, D., Elzinga, D.J., Cohon, J., Church, R. and C. Revelle (1979). The TEAM/FLEET Models for Simultaneous Facility and Equipment Siting. *Transportation Sci_*, 13, 163- 175.