THE EFFECT OF FIRETRUCK RELOCATIONS ON THE PERFORMANCE OF FIREFIGHTER SERVICE PROVIDERS

MASTER THESIS

MATHEMATICS DEPARTMENT

FACULTY OF SCIENCES

VU UNIVERSITY AMSTERDAM

AMSTERDAM, 2017

BY

MOUMNA RAHOU

This page intentionally left blank

MASTER THESIS MATHEMATICS

THE EFFECT OF FIRETRUCK RELOCATIONS ON THE PERFORMANCE OF FIREFIGHTER SERVICE PROVIDERS

A research about the relationship between executing relocations and the performance of the firefighter service provider

BY

MOUMNA RAHOU

Centrum Wiskunde & Informatica
Research Group Stochastics
Science Park 123
1098 XG Amsterdam

Fire Department Amsterdam Amstelland
Business Intelligence
Karspeldreef 16
1101 CK Amsterdam-Zuidoost

VU University Amsterdam
Faculty of Sciences
Mathematics Department
De Boelelaan 1081a
1081 HV Amsterdam

Supervisors:
Drs. G. A. G. Legemaate
Dr. ir. P.M van de Ven
Prof. dr. R. D. van der Mei
Prof. dr. S. Bhulai

DECEMBER 2017

This page intentionally left blank

Preface

The program for the Master of Science in Mathematics on the VU University Amsterdam is concluded by a research project or an external project to be carried out within a business, industry or research facility other than the department of Mathematics. I opted for an external project at the research facility the Centrum Wiskunde & Informatica (CWI) in close collaboration with the Fire Department Amsterdam Amstelland (FDAA). This report describes the process and the results of my research.

The research area of my Master Mathematics is Stochastics. The head of the Stochastic Group at CWI, prof. dr. Rob van der Mei, proposed several subjects for my master project. One of them was about the effect of relocations of firetrucks in order to improve the response time, which I decided upon because of its societal relevance.

I would like to thank Arjen ter Heide for providing the data which I needed to perform this research, Guido Legemaate and dr. ir. Peter van de Ven for tracking my progress and providing me with positive feedback, prof. dr. Rob van der Mei and prof. dr. Sandjai Bhulai for being my supervisors. Because of their suggestions and remarks I was able to finish this long term project. My thanks also go to all my colleagues from the CWI and the FDAA for the positive work environment. Above all, I would like to thank my lovely parents for their interest, patience and supporting words.

Moumna Rahou

December 2017

Abstract

In life-threatening situations where every second counts, the timely arrival of firefighter services at the emergency scene can make the difference between survival and death. The location of firetrucks has a huge impact on the response time to an emergency scene, i.e., the total time between an incoming emergency call and the moment that a firetruck arrives at the emergency scene. The potential for improving performance of firefighter services is directly related to reducing response time. To realize short response times, it is crucial to plan firefighter services efficiently.

Motivated by this, an algorithm has been developed which leads to additional movements of firetrucks compared to the reactive paradigm, where firetrucks depart from the base station when an emergency is reported. We study the effect of the relocations on the response time performance. We formulate the relocations from one configuration to a target configuration by the Linear Bottleneck Assignment Problem, so as to provide the quickest way to transition to the target configuration. Moreover, the performance is measured by a general penalty function, assigning to each possible response time a certain penalty.

The purpose of this project is to develop the model Dynamic Firefighter Management (DFM). The results consistently show that DFM mainly gives a large potential for areas in which the coverage is rather low. When relocating is permitted, the coverage increases with approximately 59.9% in a normal situation and with 91.0% in busy situations. This is important for the implementation of DFM in practice. This model should serve as a basis for further research of this topic.

List of Abbreviations

Abbreviation	Explanation
AMEXCLP	Adjusted Maximum Expected Covering Location Problem
CBS	Centraal Bureau Statistiek
CWI	Centrum Wiskunde & Informatica
DAM	Dynamic Ambulance Management
DFM	Dynamic Firefighter Management
FDAA	Fire Department Amsterdam Amstelland
MCLP	Maximum Coverage Location Problem
MECRP	Maximum Expected Covering Relocation Problem
MEXCLP	Maximum Expected Coverage Location Problem
MEXPREP	Minimum Expected Penalty Relocation Problem
RDW	Rijksdienst Wegverkeer

Contents

1. Introduction

In this chapter, we will start with some words about the CWI and about the FDAA where this internship took place. Then we will provide the context and a motivation of the research done during this internship. The chapter will be concluded with the approach, and the structure of this master thesis.

1.1 About CWI

Founded in 1946, CWI is the national research center for mathematics and computer science in the Netherlands. The vision of CWI is twofold: to perform cutting-edge fundamental research in Mathematics and Computer Science, and to transfer knowledge to academia and to Dutch and European industry. This results in importance for our economy, from payment systems and cryptography to telecommunication and the stock markets, from public transport and internet to water management and meteorology.

Within CWI the research group *Stochastic Group* has a long-standing tradition in the field of performance modeling and solution techniques for stochastic evaluation and optimization problems. Examples can be found in areas like communication and information systems, biology, economics and logistics. This group develops and studies stochastic and statistical models that yield fundamental understanding and enable control and optimization of such systems. Analysis of these models relies on techniques from fundamental probability theory, queueing theory, stochastic scheduling, spatial stochastics and stochastic geometry. Besides its focus on methodological aspects of stochastic models, the group also has a strong focus on the applicability of the results. The group has a broad national and international network of collaborations with industrial partners, governmental and academic institutions.

1.2 About FDAA

The Dutch capital Amsterdam was the first city in the Netherlands with a professional fire service, which was established in 1874. With 144 personnel and 9 fire stations covering 30 square kilometers, it ensured fire protection for approximately 285,000 inhabitants. Today, it is the regionally organized FDAA that, with 1150 personnel and 19 fire stations covering 354 square kilometers, is responsible over 1,000,000 inhabitants.

Within FDAA the research group Business Intelligence carries out investigations to pursue the vision, which is maximizing safety in the region. To achieve this, the following activities are carried out: collecting and analyzing historical quantitative and qualitative datasets, translate this into valuable insights and attempts are made to apply new data analyzes to improve the safety in the region. This research group plays a crucial role in guaranteeing quality of both the preparation of as the repressive actions of the firefighters.

1.3 Context

The response time to an emergency scene depends on three variables which are (1) the travel time from the base to a location, (2) the pre-trip delay, i.e., the time elapsed before the firetruck starts driving to the emergency scene, and (3) the dispatch time, i.e., the total time between an incoming emergency call and the moment that a firetruck gets the order to drive to the emergency scene.

The evaluation of firefighter services providers, judged by the authorities, heavily relies on their performance regarding the response times. The most common measure on which firefighter service providers base their performance, is the fraction of calls reached within some response time or coverage radius. An emergency scene is covered if an idle firetruck is present within this coverage radius. This coverage radius is expressed in time, e.g., 5 minutes.

In emergency situations, the location of firetruck has a huge impact on the response time to an emergency scene. To realize short response times, it is crucial to plan firefighter services well. This encompasses a variety of planning problems at the strategic, tactical, and operational level. At the strategic level, the locations of the fire stations are determined. Then, at the tactical level, the number of firetrucks and thus the number of firefighters per fire station is specified. At the operational level, real-time dispatching of firetrucks to emergency scenes and real-time relocation of firetrucks is considered. At the strategic level, at the tactical level as well as at the operational level, the focus is on the search for the best possible coverage, based on the coverage radius. In this thesis, we focus on the last part of the operational level: the relocation of firetrucks.

Firetrucks are relocated in real-time, using dynamic and proactive relocation strategies, in order to achieve shorter response times to emergency scenes. These relocation decisions are typically made when an emergency call happens, e.g., when a firetruck is dispatched or when a firetruck is newly free after the service at the scene. At certain moments in time, crews may be required to park up at a waiting site away from their base station, to increase coverage of the region. Such a relocation decision is usually made when an emergency scene happens, i.e., a change of the system occurs. Examples of emergency scenes are, for instance, a change in availability of firetrucks (when a firetruck is dispatched to an emergency scene or when a firetruck finishes service) or the arrival of a firetruck at the emergency scene.

The relationship between performance and the number of relocations is complex. The consequences of moving a firetruck to a different fire station are not known a priori, due to uncertainty that plays an important role in the process. It is usually not the case that ‘more’ is ‘better’, i.e., the more relocations are made, the better the performance of the firefighter service provider. But even if this was the case, there is still a trade-off: would one carry out extra relocations for only a small gain in performance? Opinions of different firefighter providers differ on this question and it is hard to set a standard concerning the execution of relocations. Therefore, useful insights about the relationship between performance and the number of relocations are desirable.

1.4 Objective

In this thesis, we study the relationship between the execution of relocations and the performance of the firefighter service provider. Therefore, we present a firefighter redeployment model, in which we are able to incorporate different performance criteria. We use a heuristic method that computes an action concerning the relocation of firetrucks in such a way that the expected performance is maximized. This computation is done at decision moments: the time of occurrence of a new emergency scene or the time of the idle report of a firetruck. We use a heuristic policy instead of the optimal one because computation of the optimal policy is very complex, if not impossible. Besides, even if it was possible to compute, the optimal policy is probably a complex one: it is not easy to understand and to execute by the dispatcher. Instead, we use a heuristic method that is not too farfetched, while it is highly likely that this heuristic policy contains the same characteristics as the optimal one.

1.5 Approach

In Chapter 2 the literature related to this research is reviewed. This chapter provides a literature study about the studies that are already done in this field and the surveys on relocation models. In Chapter 3 the historical data which is provided by the FDAA will be analyzed. The number of firetrucks requests will be observed for the highest priority. Based on the observations a non-homogeneous Poisson process describing the number of firetrucks rides per day and an intra-day pattern of the firetrucks requests are provided, which will be validated. In Chapter 4 the description of the model will be presented. Hereby we will have a view at the system dynamics and we will discuss how this system can be controlled in the best way. In Chapter 5 the heuristic method will be explained extensively. This method is used to calculate the relocations that must be performed in order to maximize the coverage. The results of the research will be discussed in Chapter 6. The main topic in this chapter will be the improvement of the response time as a consequence of the implantation of the model. In Chapter 7 the conclusion is given followed by a discussion about the strengths of the model and the extensions which can be made to improve this model. Other topics for further research will also be provided in the same chapter.

2. Literature Study

In this chapter, we provide a literature on emergency dispatching and the surveys on relocation models.

2.1 Performance

The potential for improving performance of EMS systems is directly related to reducing response time, i.e., the total time between an incoming emergency call and the moment that a firetruck arrives at the emergency scene, see e.g., [41]. Studies in [24] and [17] have shown that shorter response times lead to an increased probability of patient survival. Furthermore, the study in [38] has shown that shorter response times are associated with reduced complications, especially for the most critical emergency scenes, such as people whom suffering from heart attack or a large fire in a building.

2.1.1 Performance Measure

The proportion of calls reached within some response time or coverage radius, is the most widely used measure in practice and many EMS systems base their performance on this metric as stated in [14]. An area is *covered* if an idle firetruck is present within a certain coverage radius. Note that this coverage radius is expressed in time units, e.g., 8 minutes.

Coverage models have been used frequently by researchers and practitioners for several reasons:

- The concept is simple to communicate to decision makers and the public; a call is either covered or it is not.
- Many EMS systems use the percentage of calls covered as a performance measure. A common standard is to respond to 90% of all urgent calls within 8 minutes [12].
- Deterministic coverage models typically result in integer programs that are easy to solve using standard optimization software.

Despite these advantages, it is stated in [13] that the binary nature of the coverage concept is an important limitation, and standard coverage models should not be used for EMS vehicle location. Coverage has the limited ability to discriminate between different target response times. Another limited ability is the fact that coverage cannot make a distinguishing in the extent to which a firetruck arrives late. When the coverage radius is passed, it does not seem to matter how long it takes before the firetruck finally arrives at the emergency scene. This could result in large optimality errors when one uses covering models to locate emergency facilities instead of a model that takes survival probabilities into account. In [3] a penalty function is used to model general performance measures based on response times. It is used to incorporate different performance measures, such as minimizing the number of incidents for which the maximum allowed response time is exceeded, minimizing the average response time or measures related to survival probabilities, as studied in [13]. Other performance measures could be: the probability of all vehicles being occupied, the capability of a certain vehicle configuration to cover future calls and cost effectiveness, see, e.g., [22], [37].

2.1.2 Operational Level

For the realization of the increase in performance, an efficient planning of the firetrucks is crucial. This encompasses a variety of planning problems at the *strategic, tactical, and operational level*. At the strategic level, the locations of the fire stations are determined, see e.g., [6]. At the tactical level, the number of firetrucks per fire station is specified and, as a direct consequence, the number of firefighters per fire station. At the operational level, the real-time dispatching of firetrucks to emergency scenes and the real-time relocation of firetrucks is controlled, see e.g. [17]. At the strategic level, at the tactical level as well as at the operational level, the focus is on the search for the best possible coverage, based on the coverage radius. In this thesis, we focus on the operational level.

2.1.3 Environments

When one large fire, or several small fires, being fought in a single area of a city, the fire stations of the dispatched firetrucks are left empty, resulting in a sharp degradation in the fire protection afforded to the surrounding area. At the operational level it is common practice in many cities to spread out the available firetrucks by relocating some firetrucks into selected empty fire stations, see e.g., [17]. Much research has been focused on solving various kinds of vehicle relocation problems in *static, probabilistic and dynamic* environments to improve the performance of EMS systems.

In a static environment relocations are not allowed. In [20] it is stated that static models are inadequate in the sense that their solution is unlikely to provide sufficient coverage once a vehicle is dispatched to a call. One way around this difficulty has been to develop probabilistic models which reflect the fact that emergency vehicles are busy only a fraction of the time and become unavailable once they respond to a call. They must in fact be considered as servers that operate within a queueing system. This line of research was pioneered in [25], [26] and [27]. For well-known probabilistic covering models, we refer to [2], [5], [29], [33] and [34]. These models make different assumptions on travel time distribution, on the spatial dependency of vehicle availability, and on desired probabilistic covering thresholds.

In a dynamic environment, instead of seeking a single solution to a static or probabilistic model, the idea is to dynamically relocate vehicles in real-time as vehicles are dispatched to calls. Each relocation amounts to solving a static model subject to side constraints on vehicle moves. For example, one should avoid relocating too many vehicles at once or moving the same vehicle too often over a short period. An early dynamic model was proposed in [17] for the relocation of fire companies. More recently a dynamic firetrucks relocation model is developed in [16] which can be applied in real-time through the use of parallel computing. For surveys of emergency vehicle location and relocation problems, see [30], [7].

Decisions regarding firetruck location strategies can be used to improve coverage radius. The firetruck location problem refers to the assignment of a limited number of firetrucks to maximize coverage, given that the system has a fixed number of fire stations, and a demand location is considered to be covered when a firetruck is located within a predetermined time standard. However, in reality the arrival of calls is stochastic and dynamic. Dynamic vehicle relocation can improve the performance of systems in situations with fluctuating arrivals, and as a result, in [1] it shows a drastic increase in the number of dynamic strategies used.

2.1.4 A posteriori & A priori

There are two main ways of solving dynamic optimization problems. The *a posteriori* approach is most common. It consists of computing and implementing a new solution whenever an emergency scene occurs. This is the strategy used in several dynamic vehicle routing algorithms which is described in [15]. The main drawback of this approach is the need to compute a new solution whenever a vehicle is dispatched to a call. This can be time consuming or even infeasible when two or more successive emergency scenes occur in quick succession. This was the case in the Island of Montreal firetrucks relocation problem studied in [16]. There is also the *a priori approach*. In [16] an a priori methodology is proposed in which several solutions are precomputed in anticipation of future emergency scenes and the appropriate solution, if available, is implemented whenever an emergency scene occurs. This methodology was successfully applied to the firetrucks relocation problem in Montreal. It was shown that a precomputed solution was available in more than 95% of emergency scenes. Many researchers think that the a posteriori approach is not always a practical option in emergency vehicle dispatching centers, an alternative is to precompute a series of location scenarios for $k = 1, \ldots, n$ vehicles which can readily be applied whenever a call is made. Therefore, the relocation strategies that we will discuss are based on the *a priori* methodology.

2.2 Relocation Strategies

Existing methods to perform relocations use preplanned assignments. One type of dynamic strategy for relocating firetrucks is to use a compliance table. Another relocation strategy, the Dynamic Ambulance Management, is described in [3].

2.2.1 Compliance Tables

The aim of the compliance tables is to provide a dynamic relocation strategy for emergency vehicles in such a way that the expected covered demand is maximized and the number of relocations is controlled. The compliance tables are modeled as an integer linear programming.

A compliance table indicates the possible empty fire stations in relation to the number of available firetrucks. That is, a compliance table shows where firetrucks should be located when there are a certain number of firetrucks available. Each row in a compliance table indicates, for a given number of available firetrucks, the desired fire stations for these firetrucks. If these firetrucks are at their desired fire stations, the system is in compliance. To understand what a compliance table policy entails, consider the example in Table 2.1.

Table 2.1. An example of a compliance table.

No. of available firetrucks	Fire Stations											
	1	2	3	4	5	6	7	8	9	10	11	12
13	1	1	1	1	1	1	1	1	2	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1
11	1		1	1	1	1	1	1	1	1	1	1
10	1		1	1		1	1	1	1	1	1	1
9	1		1	1			1	1	1	1	1	1
8	1		1	1			1		1	1	1	1
7	1		1	1			1		1	1	1	
6			1	1			1		1	1	1	
5			1	1					1	1	1	
4			1						1	1	1	
3			1						1	1		
2			1						1			
1									1			

In Table 2.1, each row represents the number of available firetrucks and the column represents specific fire stations. To read the table, given a current number of available firetrucks, the goal is to have firetrucks located at stations with a ‘1’ or a ‘2’ in that row. For example, based on Table 2.1, with one available firetruck, it should be located at fire station 9. Now suppose one firetruck just changed status from busy to available, so that two firetrucks are available. Based on Table 2.1, the available firetruck should go to station 3.

There are different kinds of models that compute compliance tables for EMS. In [39] a nested-compliance table policy is discussed which is based on the adjusted maximum expected covering location problem (AMEXCLP) with the addition of the admission of relocations. The model requires steady state probabilities of a Markov chain model to be input parameters for the integer programming model. The benefit of nested policies is that only one vehicle, which is already on the move, is relocated, thus avoiding unnecessarily moving other vehicles that can be disruptive to service providers. The results showed that the model provided improvement of solutions over the results of the non-relocation (AMEXCLP) model with average improvement of 2.7% based on an original data set from a real-world problem and an improvement of 6.1% based on a generated data set in which demand was randomly assigned to demand zones according to the Poisson distribution.

In [20] the Maximum Expected Covering Relocation Problem (MECRP) is proposed. The objective is to maximize the expected coverage over time subject to an upper bound on the number of waiting sites that can be changed at each event. This model needs the assumption that there are at least as many waiting site locations as vehicles. This is a rather strong assumption and not generally true in practice for EMS. This however does hold for the FDAA. The Minimum Expected Penalty Relocation Problem (MEXPREP) in [4] is an extension of the MECRP. The MEXPREP has the ability to control the number of waiting site relocations.

Different performance measures can be incorporated. Other performance measures than coverage and average response times are considered in [13]. In this thesis, performance is based on the survival probability of a patient suffering from a cardiac arrest. Survival functions proposed in [24], [12], [42], and [43] are incorporated in the Maximum Coverage Location Problem (MCLP) in [10] and the Maximum Expected Coverage Location Problem (MEXCLP) in [11]. In some of these models, probabilistic response times are incorporated, based on the work by [18]. Moreover, [28] proposes a methodology for evaluating the performance of response time thresholds in terms of resulting patient survival rates.

A strength of the compliance table policies is that it allows for EMS systems to respond to the dynamic nature of the problem, but can be calculated in advance, thus not requiring optimization in real time. Furthermore, it is simple to explain to and to use by dispatchers: the state of the EMS system is only described by the number of available vehicles. A weakness of the compliance table is that it is adequate when the number of idle firetrucks is large but breaks down at high alarm rates when the firetrucks preassigned to relocate are not available.

2.2.2 Dynamic Ambulance Management

A similar problem which we want to research for the FDAA is investigated by the ambulance service provider. The relationship between the number of ambulances relocations and the performance of the ambulances service provider is discussed. In this paper, a heuristic method is described which computes an action concerning the relocation of vehicles in such a way that the expected performance is maximized. This computation is done at decision moments: The time of occurrence of a new emergency scene or the time of the idle report of an ambulance. The dispatcher decides at these moments whether to relocate or not. This heuristic policy is easy to understand and for the dispatcher it is easy to use in practice. The model and the heuristic policy governed by DAM algorithms will extensively be discussed in Chapter 5.

The difference between this thesis and most of the papers in the literature, e.g. [20], is the assumption that the computed action is always carried out. However, it may be the case that the expected gain in performance by taking this action is very small and that this benefit does not outweigh the disadvantages regarding the number of additional ambulances relocations to achieve this gain. In [3] it is determined whether or not a relocation is really necessary. Another important difference between the mainstream literature and this thesis is the way in which a redeployment action is carried out. As assumed in most of the papers, it is not necessarily one particular vehicle has to move from origin to destination. Instead, other idle vehicles may also be used in this relocation process in order to decrease the time required to attain a new ambulances configuration.

In September 2015, the region Flevoland started with the DAM algorithms as a trial. The implementation of DAM algorithms leads to additional movements of ambulances compared to the reactive paradigm, where ambulances depart from the base station when an incident is reported. The results of DAM show that there is a significant improvement if ambulances are relocated, compared to the static policy in which always the static motion if performed.

2.3 Dependent Travel Times

In areas with high traffic volume during peak hours travel speed and the resulting travel time can vary significantly throughout the day. Due to variations in speed and the resulting travel times it is not sufficient to solve the static relocation problem once using fixed average travel times as the coverage areas themselves change throughout the day. In [36] a multi period version is developed where it is shown that it is essential to consider time-dependent variations in travel times and coverage, taking into account time-varying coverage areas, where vehicles are allowed to be relocated in order to maintain a certain coverage standard throughout the planning horizon. A mixed integer program is formulated which optimizes the coverage at various points in time is simultaneously optimized. Just using (averaged) fixed travel times is no longer sufficient and will lead to sub-optimal or infeasible solutions. Taking time-dependent travel times into account significantly influences the quality of the solution obtained. The performances increase by more than 10%.

The New York City RAND fire project, lasting from 1968 to 1975, was a successful research project, which made use of analytical and statistical modeling, which led to key changes at the New York City Fire Department, [22]. A part of this research project was about the travel times of the firetrucks. The travel time is divided into an acceleration phase, a cruising speed phase, and a deceleration phase. They use these different phases to model the travel time depending on the distance, and they take the mean of the travel times as the expectation of the travel time. This model is improved in [18] by taking the median as the expectation of the travel time. It is argued that since travel times are non-negative, the distribution of the travel times is probably skewed to the right, and that therefore it is better to take the median rather than the mean as the expectation of the travel time.

In [44] a different model is introduced where they look at the different street segments separately, and take differences in travel speed into account instead of looking at the total route as one stochastic variable. They use a Bayesian method to estimate the travel time distribution, and they compare that model to the model of [8]. They conclude that the Bayesian method seems to give more realistic results than the method in [8], since that method does not take into account the different speeds of different roads. For the case study which is described in [44], GPS-data is used and Markov Chain Monte Carlo simulation is applied to find the true route, since their GPS-data is not accurate enough. One could, of course, refine this model by not only looking at different roads separately, but also taking the transitions of one road to another into account. In [21] a trip is seen as a combination of running travel times along links, and delays at intersections and traffic signals. They treat this delay as a deterministic penalty. They do not consider emergency services in particular, but give a model for general services.

Often, researchers assume travel times to be deterministic, see e.g., [3], [9] and [32]. In this case, fire engines are either always on time or always too late, which is not realistic, since there is always some variability, because traffic itself is variable, but also because of weather, the time of day, heavy traffic etcetera. In [19] it is shown that, by considering firetrucks travel times between a particular station and demand point pair, for a total of 352 trips, stochastic travel times lead to a more realistic model.

2.4 Thesis Outline

Our research is based on the knowledge gained from literature which is summarized in the previous section. The methods and techniques which are described in this literature study are mainly used to find answers to our research questions. The goal is to develop a dynamic algorithm which determines when relocations should be made, which fire stations should be filled and which available firetrucks should be moved. The processes are carried out in the statistical program R, see [40] and the model is built in Matlab, see [31].

First, we explain why we have taken this particular approach since an understanding of our objectives is crucial to an understanding of the finished product. If the objective were to provide equal first unit response time to all areas of the city, i.e., fairness, the available firetrucks should be spread out rather uniformly. If the goal were to minimize the region wide average response time to alarms, i.e., efficiency, the companies should be highly concentrated in the areas where the expected emergency scenes is greatest. Different objectives about balancing fairness and efficiency are discussed in [22].

Fire stations are not uniformly distributed over the city but are concentrated in some areas and spread out in others. This distribution is the result of positioning the fire stations efficiently within the constraints of politics. In working with this distribution, the FDAA has implicitly decided how it wishes to balance fairness against efficiency in a short run. In the long run, of course the FDAA may modify the distribution by building new fire stations or relocating the old fire stations. The latter case is very expensive as the distribution of the fires changes with time, the city is expanding, flammable old buildings are replaced with new buildings that are more resilient etc. which makes it for the long run difficult to keep up with the changes of the region. This however is a problem which can be solved at the strategic level. Since our focus is on the operational level, we will look at a cheaper alternative which is the use of a relocation method. The relocation method consists of four stages each of which is solved by the application of heuristics. The four interrelated stages are:

Stage 1: Determination of the need for a relocation.
A call for relocations will be made whenever the fire protection being provided to any area of the city falls below a given minimum level.

Stage 2: Determination of the demand location which needs to be covered.
The fire stations to be filled are chosen to bring fire protection in all areas above minimum levels while moving as few firetrucks as possible.

Stage 3: Determination of the available firetrucks which relocate.
A penalty function will be used to compare alternative relocations which expresses the penalty of relocation in term of response time to future fires. The function takes into account such factors as relocation distance and expected response times. The firetruck which produces the lowest penalty is selected for the relocation.

Stage 4: Specify relocation assignments
The set of relocating firetrucks is assigned to the set of fire stations to be filled so that the total distance travelled by the firetrucks is minimized.

2.4.1 Contribution

By ignoring time-dependent variations in travel times one severely misestimates the resulting coverage, which will lead to inferior solutions. By taking into account those variations during the optimization phase and allowing vehicles to be relocated, the quality of the solutions obtained can be increased significantly. The goal is to provide the decision maker with time-dependent location plans, such that the resulting coverage can be kept at the required level throughout the planning horizon.

In this thesis, we will develop a relocation model whereby the time-dependent travel time which was proposed in [36] will be taken into account. First, we take advantage of the availability of time-dependent data in order to get a clearer view of the traffic situation and the resulting changes in coverage throughout the day. Next, we introduce a model, which incorporates this information and allows vehicles to be repositioned to optimize the coverage at several points in time.

There are different ways to consider time-dependent data. First, we solve our model in a myopic way at various points in time using the prevailing travel times respectively. All resulting relocations will be calculated ex post. Rather than solving the model independently several times, we try to solve the model simultaneously for various points in time. We will explicitly take into account time-dependent variations in speeds and the resulting changes with respect to the corresponding coverage. All resulting relocations will be considered implicitly during the optimization phase

Another extension of the model is taking the variation of response times into account. The FDAA has categorized the emergency scenes into four groups. Each group has its own response time. These response times depend on the type of danger at the emergency, where the emergency takes place and whether people are involved. For the region Amsterdam Amstelland, the firetrucks need to be at the location of the emergency within 6 to 10 minutes after a call has been placed.

3. Data Description

A historical dataset is provided by the FDAA of emergencies records from January 1, 2008 until December 31, 2015. We use the raw data for a detailed study with R software and Matlab. The data included 137647 incoming emergency calls.

The data contained times stamps of the incoming emergency calls, of alarming a fire station, of the departure of the firefighters from the fire station to the emergency scene, of the arrival at the emergency scene, the location of the emergency, the object(s) that is (are) involved, the kind of priority each emergency has and more information about the handling of the emergencies. The FDAA distinguishes three levels of priority of the emergencies:

Priority 1: This is a fire to which there has to be sent a firetruck as soon as possible;
Priority 2: This is a small fire which doesn't spread out, like a trash can in the street, such that there has to be sent a firetruck, but not with a great hurry;
Priority 3: This is a small incident for which there is not a great hurry, like a cat in a tree.

Of the emergency calls 66% were labeled as priority 1, 21% as priority 2 and 13% as priority 3. For the emergencies that are labeled as priority 2 or priority 3, the arrival of the firetruck within a given time standard does not hold. Since we want to improve the response time for emergency scenes in which the rapid arrival of the firetrucks does matter, we will only consider in this thesis emergencies that are labeled as priority 1.

There are four different emergencies, which are labeled as priority 1. Each kind of emergency has its own maximum allowed response time (MART). These times depend on the kind of object(s) that is (are) involved in the emergency and are set by the government of the Netherlands, see Table 3.1.

Table 3.1. The Maximum Allowed Response Time (MART) and the Maximum Allowed Travel Time (MATT) for different objects. MART and MATT are given in seconds.

Type	MART	MATT	Objects
A	300	120	Building with a closed construction: retail, cells, residential function above retail
B	360	180	Blocks, flats, residential function for reduced self-reliant.
C	480	300	Building with the functions: health care, education, accommodation.
D	600	420	Office, industry terrains, places where people sport and general gatherings.

Of the emergencies that occurred, 10.1% of them had a MART of 300 seconds, 8.6% of them had a MART of 360 seconds, 60.5% of them had a MART of 480 seconds and 20.8% of them had a MART of 600 seconds.

Although the data is gathered automatically, some cleaning was necessary. The preliminary analysis revealed some peaks in the emergency scene handling times.

According to the policy officer of the FDAA the peaks originate from cases where the matter still continues to determine the cause of the emergency. However, the firefighters have already handled the emergency. An appropriate replacement of these data was not possible, considering we have no knowledge of the actual duration. For that reason we removed these outliers. Furthermore, the preliminary analysis did not reveal any obvious trend but seasonality in the daily and weekly cycle of the call arrival rate. In addition, we discovered an obviously different pattern of the hourly call arrival rate for public holidays. For this reason public holidays, e.g., Christmas, New Year's Eve, King's Day (24 days in total) were omitted. Apart from this, we used all collected call and operation data for our analysis. We partitioned the remaining data into 24 hourly intervals. For the assignment to a certain interval we used the time of call. Therefore, the interval 16 represents the time window between 3 p.m. and 4 p.m. of all days.

3.1 Traffic Intensity

The data of the Centraal Bureau Statistiek (CBS) shows that the number of vehicles increased in the recent years. In Table 3.2 we see that the total number of vehicles has increased over the past few years.

A logical consequence of the increased number of vehicles is the increasing of the traffic density and intensity on the road. In the execution of the relocation the firetrucks participate in normal traffic in terms of speed and the use of the same roads, i.e. not using the flashing lights or sirens and no driving on the tram tracks. It is therefore logical to think that the greater the traffic intensity is on the road, the longer the relocation will take to complete. Taking this into account is an important addition to the model which we discussed in section 2.2.2.

Table 3.2. The size and composition is based on the registration of the license plates which is performed by the RDW.

	2011	2012	2013	2014	2015
All Vehicle Types	321993	328488	333019	336153	336859
Passenger Cars	221620	225297	228764	230677	228691
Company Cars	34966	34766	33957	33493	33906
Company Vehicles	24832	24523	23749	23199	23482
Vans & Trucks	22379	22127	21438	20886	21105
Special Vehicles	1473	1436	1416	1447	1413
Buses	377	328	269	273	280
Motors	16420	16964	17344	17635	18009
Mopeds	48987	51461	52954	54348	56253

The Nationale Databank Wegverkeersgegevens (NDW) provided us with information relating to the traffic situation in the region Amsterdam Amstelland within the ring and excluding the highways. The data can be obtained on request. We received data related to the *traffic intensity* of the month June 2016. We define the traffic intensity as the average number of vehicles on the road network during a certain hour in the region Amsterdam Amstelland. In Figure 3.1 two peaks can be seen, which corresponds with the rush hours. The rush hours are between 8:00 – 10:00 AM and between 5:00 – 7:00 PM. At these moments, the traffic intensity and traffic density are the highest. Note that we do not make a difference in working days and weekend days.

The number of vehicles on the road is much lower during the night than during rush hours We will take the behavior of the traffic into account by introducing a traffic intensity factor. For simplicity, we use four different traffic intensity factors. We have associated the intensity factor with the corresponding hours, see Table 3.3.

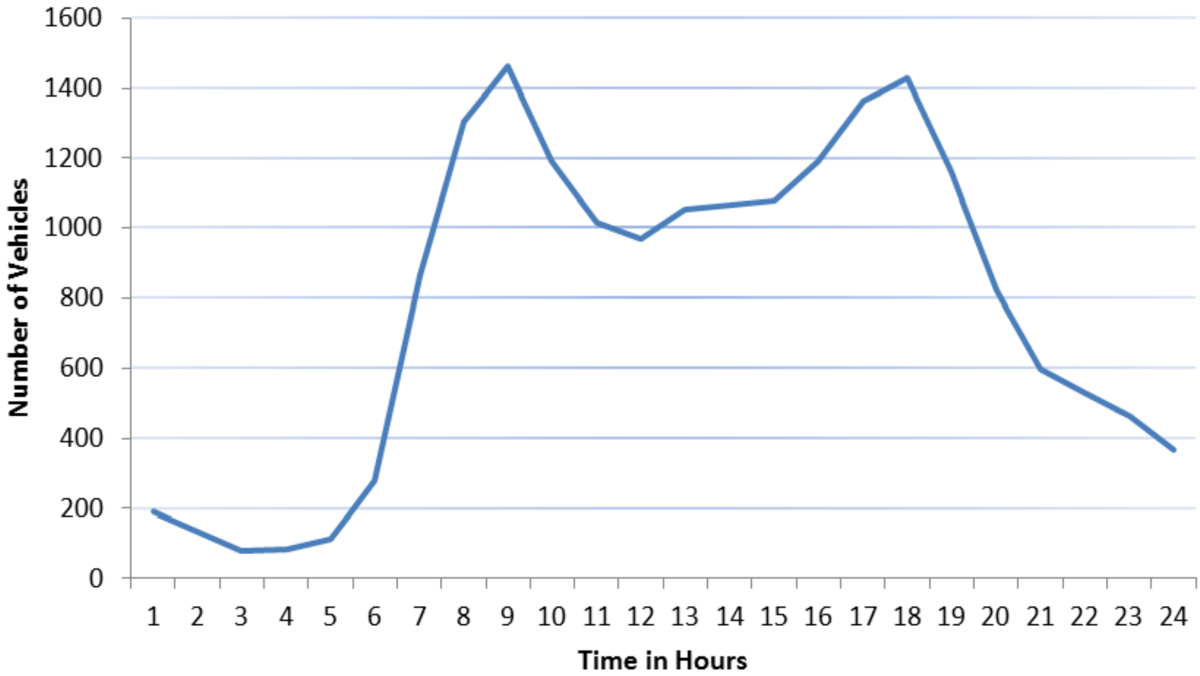


Figure 3.1. The number of vehicles on the road network in the region Amsterdam Amstelland per hour.

Table 3.3. The intensity factors.

Intensity factor	0.8	1.0	1.2	1.5
Hours	1, 2, 3, 4, 5, 6, 21, 22, 23, 24	7, 11, 12, 13, 14, 20	8, 9, 15, 19	10, 16, 17, 18

4. Model Description

As discussed in Chapter 2, we will develop a model for the setting of the FDAA. Then we add the time dependent travel times. This chapter includes the description of the demand-model, the performance matrix used by the FDAA in practice and the description of our model.

4.1 Setting FDAA

4.1.1 Demand Locations

The geographical regions which are used by the FDAA are called the *demand locations* from which calls can rise. For operational purposes, FDAA divides its service region into $N = 2648$ demand locations, each of which contains a number of buildings. The sizes of the demand locations are determined by the intensity of new fires. In each demand location, the FDAA wants to have approximately the same intensity of new fires. In the center of the region Amsterdam Amstelland the intensity of new fires is higher and therefore, the demand locations near and in the center are small. As we approach the edge of the region Amsterdam Amstelland, the size of the demand location gets larger as the intensity of new fires gets smaller, see figure 4.1.

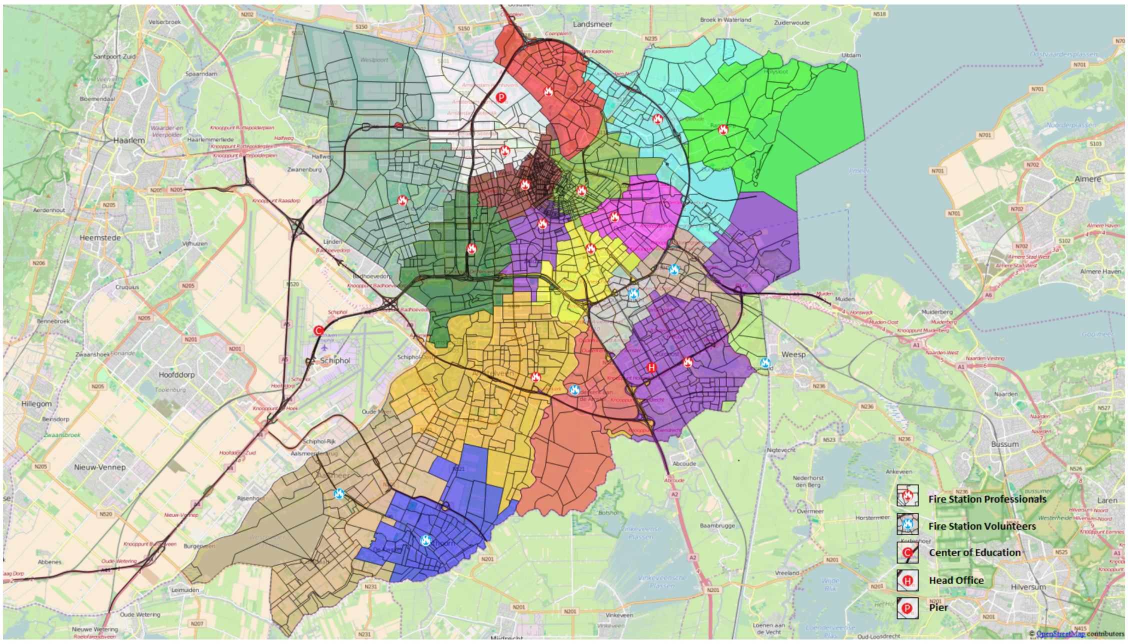


Figure 4.1: The division of Amsterdam Amstelland into service areas (the colored areas) and demand locations (the smaller boxes with black borders).

Each demand location is associated with a *demand probability*. We define the demand probability as the probability that an emergency will occur in a specific demand location. Denote the demand probabilities by $p = (p(1), p(2), \dots, p(N))$, where N is the number of demand locations.

4.1.2 Service Areas

We define a *service area* as a subset of demand locations. In the current configuration, the region Amsterdam Amstelland is partitioned into 19 service areas. In each service area is a fire station that is responsible for all the emergencies that occur in that specific service area. Each fire station is provided with at least one fire apparatus. This is the well-known firetruck with a water tank and fire hoses. The other vehicles like the aerial apparatus, the rescue apparatus and the marine rescue units are not located at each fire station. In this analysis, we only include the most common type of vehicles used at the FDAA which is the fire apparatus. We refer to it by firetruck.

There are 7 fire stations which rely on volunteers and 12 fire stations which rely on professionals, see Figure 4.1. The response time of a fire station which relies on volunteers is much higher. From different places, the volunteers need to gather first at the fire station then they will get ready for the emergency. This process takes too much time. This causes that the volunteers are not able to get at the emergency scene within the MART. Therefore, the fire stations which rely on volunteers are disregarded in this thesis.

4.1.3 Maximum Allowed Response Time

The maximum allowed response time depends on the kind of emergency scene. Each type can have different response times targets, see Table 3.1. These times are set by the government of the Netherlands.

The response time consists of the time between the emergency request comes in and the arrival of the firetruck at the emergency scene. It can be divided in three different parts: the triage and dispatch time, the turnout time and the travel time. The triage and dispatch time is the time spent in the call center to assess the importance of the call and assign a firetruck. The turnout time is the amount of time that elapses between the assignment of a call and the firefighters departure from the fire station. The travel time is the amount of time between the departure from the fire station and the arrival at the emergency scene.

4.2 Model Dynamics

The dynamics of our system closely mimics realistic situations. Incidents are assumed to occur according to a time-dependent Poisson point process as in [46] and [47]. We only consider incoming incidents of the highest urgency and ignore low-priority calls. This is justified by the fact that firefighter service providers are mostly judged on their performance regarding the emergency scenes with the highest priority and high priority incidents may preempt low-priority calls.

4.2.1 Maximum Allowed Travel Time

The dispatch and the turnout time are stochastic. The dispatch time is set on two minutes and the turnout time on one minute. It is only possible for us to have an influence on the travel time of the firetruck to the emergency scene. For simplicity, we assume that both the dispatch and turnout time are deterministic. We refer to Figure 4.2 for a graphical representation of this process.

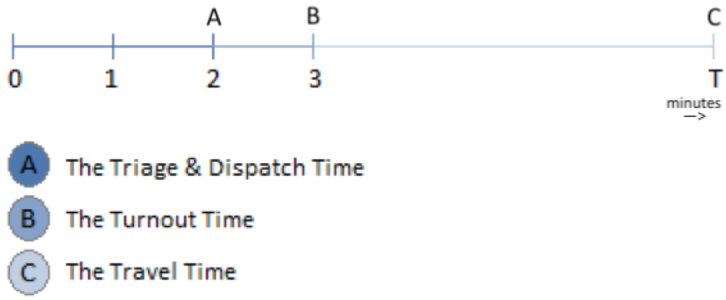


Figure 4.2: The graphical representation of the response time with time in minutes. The dispatch time (A) is set on two minutes and the turnout time (B) is set on one minute. These times are assumed deterministic. The travel time (C) is dynamic and is expressed in T minutes.

We subtract the dispatch time and the turnout time from the maximum allowed response time to obtain a *maximum allowed travel time*, see Table 3.1. However, dispatch times are typically smaller when the dispatcher is able to determine in a short period of time enough information to send a firetruck to the emergency scene. Therefore, the use of a maximum allowed travel time is a simplification of reality. When an emergency occurs, the closest firetruck is dispatched. This could be an idle firetruck at a fire station or a driving firetruck.

4.2.2 Time Dependent Travel Times

The driving times between the demand locations are derived from driving table R and are estimated beforehand and thus assumed to be given, see Table 4.3. This is a N x N matrix with deterministic driving times. These driving times, which are estimated by the FDAA, are based on a regular truck that participates in daily traffic on the road network city at usual speed. In Chapter 3 we have seen that the traffic intensity on the road during the day is time varying. Since the firetrucks drive along with the other traffic at usual speed when performing the relocation, it is fair to say that the driving times are also stochastic but assumed deterministic. We multiply the driving times with the traffic intensity by an *intensity factor* w_h *with* h = {1,…, 24}. This intensity factor depends on the hour of the day. We have deduced this from the movement of the traffic intensity, which we have seen in Chapter 3.

We model the road network in the area Amsterdam Amstelland as a directed complete graph. The demand locations are represented by nodes, see Figure 4.3. The arc connecting two demand locations is weighted according to the driving table. Thus, the length of an arc represents the driving time in seconds which is derived from the driving table.

Table 4.3. The driving times in seconds are given for the demand locations (DL) 1 t/m 4.

From ↓ To →	DL 1	DL 2	DL 3	DL 4	…
DL 1	0	762	1320	1069	
DL 2	783	0	878	1364	
DL 3	1415	876	0	1830	
DL 4	1219	1366	1799	0	
…					⋱

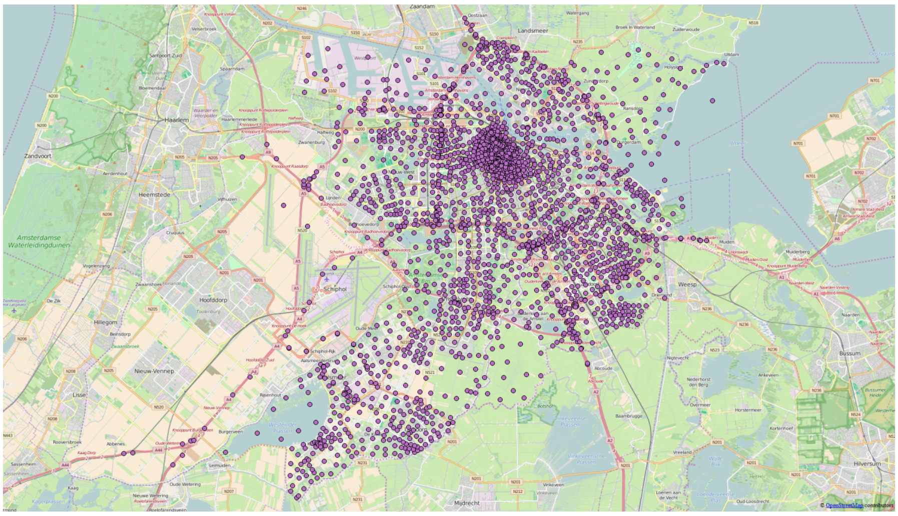


Figure 4.3. Demand Locations of the Region Amsterdam Amstelland represented by nodes.

4.3 Control of the System

Firefighter service providers use as their performance criterion the percentage of priority 1 incidents reached within the maximum allowed response time. As we have already discussed in Chapter 1, this performance matrix may not be ideal because there is no difference between a response time that is slightly below the maximum allowed one, and one that is really short. Something similar holds for the opposite case: a response time that is slightly above the maximum allowed response time and one that is really long are equally poor according to this metric. However, it does matter for the concerned victims. To be able to differentiate between different response times a *penalty function* will be used.

4.3.1 Penalty Functions

A penalty function is a mapping from the response time to positive real numbers. For example, a penalty function can be used to minimize the number of incidents for which the maximum allowed response time is exceeded, or to minimize the average response time or measures related to survival probabilities, as studied in [13]. The amount of penalty generated by an emergency solely depends on the response time to this emergency scene. Hence, penalty functions are non-decreasing functions of the response time.

For our model, we use a penalty function whereby the focus lies on minimizing the number of emergencies for which the response time exceeds the maximum allowed one. In this model, we take four types of emergencies into account. We distinguish emergency type A, B, C and D, see table 3.1. Each type has its own maximal allowed travel time and therefore also its own penalty function.

These functions where composed in consultation with a policy officer of the FDAA. An emergency reached within the maximum allowed travel time induces a penalty between 0 and 0.1, while an incident for which the maximum response time is exceeded, induces a penalty between 0.9 and 1. The penalty functions that we use for each type of incident is described as follows.

Let $f: \mathbb{R}^+ -> \mathbb{R}^+$. The penalty function which is used is:

$$f(t) = \begin{cases} \dfrac{1}{5e^{-0.045(t-\tau)+5}} & \text{if } t \leq \tau \\ \dfrac{4}{5} + \dfrac{1}{5e^{-0.045(t-\tau)+5}} & \text{if } t > \tau \end{cases}$$

whereby τ stands for the maximum allowed travel time, $\tau = \{120, 180, 300, 420\}$.

The graph which belongs to the penalty function is displayed in Figure 4.4 for various values of τ. The standard penalty function as described in the previous chapter is also displayed. The difference between Figure 4.4(a) and 4.4(b), 4.4(c), 4.4(d) and 4.4(e) is the increasing of the penalty as we approach the maximum allowed travel time. By implementing this we differentiate between firetrucks who arrive a long period before the expiration of the maximum allowed travel time and firetrucks who arrive just in time. Simultaneously, we distinguish between firetrucks who arrive late on a close call and firetrucks who arrive way too late. The curve in the graph becomes sharper as the response time increases. With a small response time, the time elapsed is wider.

4.3.2 Firetruck Phases

Once a firetruck is dispatched to an emergency scene, we assume that it immediately starts driving, since the turnout time is part of the pre-travel time we subtracted. We distinguish the following firetrucks phases:

Phase 0: A firetruck is currently not involved in handling of an emergency. It is either at a base station or executing a relocation.

Phase 1: A firetruck travels to the emergency.

Phase 2: A firetruck is busy at the emergency scene. The firefighters execute the handling of the emergency. Hereafter, the firetrucks becomes idle. Then, the dispatcher has to make a decision to which base location that firetruck should be send.

4.3.3 Decision Moments

The dispatcher has some freedom in the way he/she can control the system by making relocation decisions. We allow the dispatcher to make these decisions at the following moments:

1. When a firetruck is dispatched to an incoming emergency.
2. When a firetruck enters phase 0 again and becomes idle.

We refer to these moments as decision moments of the first and second type, respectively. At decision moments of the first type, we restrict the dispatcher to change the firetrucks configuration, i.e., number of idle firetrucks per fire location, at at most one pair of fire locations. That is, the dispatcher may select one fire station from which (origin) and one fire station to which (destination) he/she relocates a firetruck in phase 0. At decision moments of the second type, the origin is given. This is the current location of the firetruck which just finished service. If the selected fire station does not equals the base station of the crew, then we call this a *motion*.

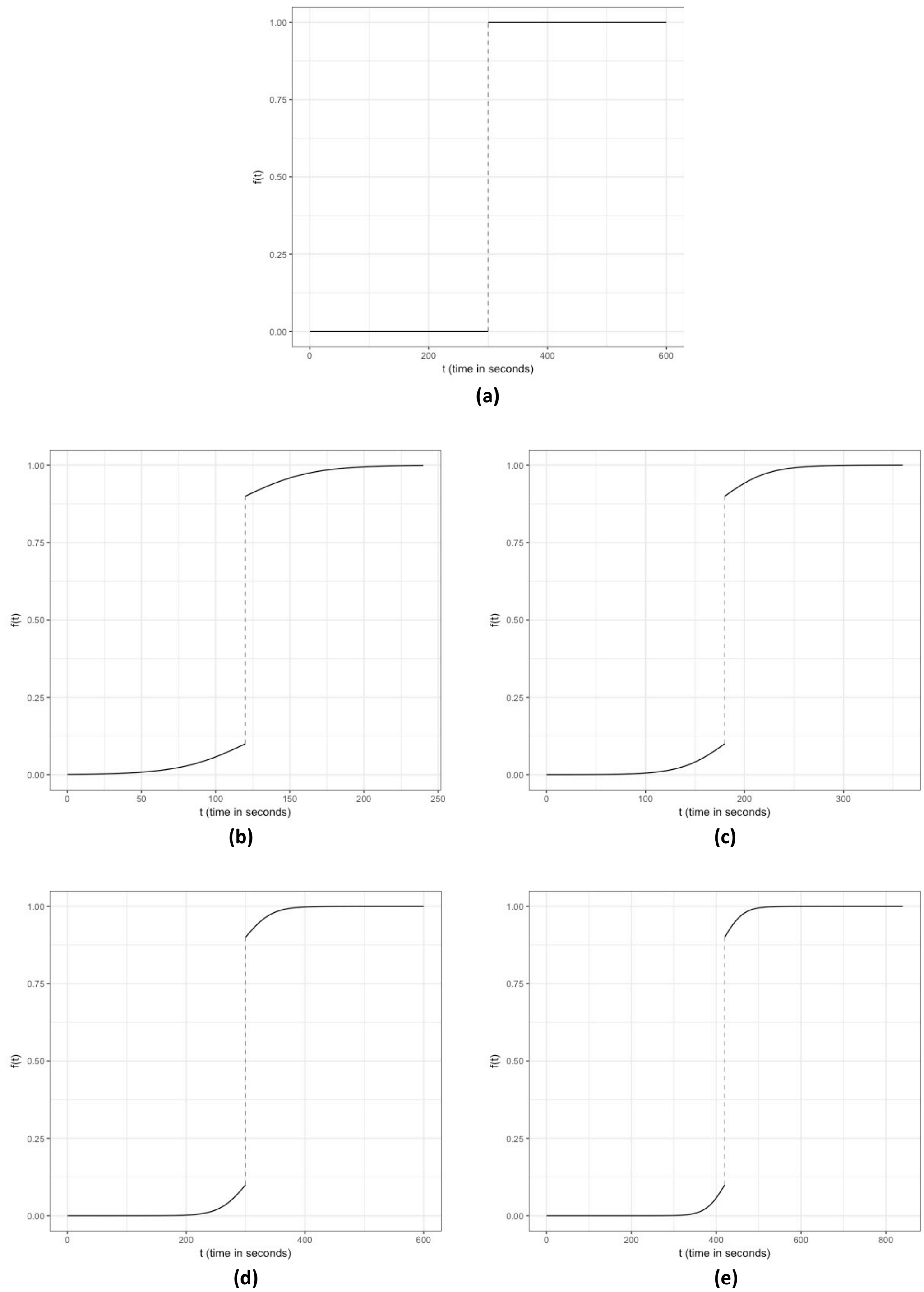


Figure 4.4. The graphical representation of the penalty function. In (a) the standard penalty function is plotted for emergency type 3. The penalty function displayed in (b), (c), (d) and (e) belongs to a maximum allowed travel time of 120, 180, 300 and 420 seconds, respectively.

Note that it is not necessarily one and the same firetruck that leaves the origin and arrives at the destination. After all, the same configuration is attained if a firetruck drives from the origin to a fire station somewhere in the middle, and from there a firetruck departs to the destination. Thus, a motion may consist of multiple relocations. The terminology motion will be further explained in the Chapter 5.

Note the difference in terminology between a firetruck motion and a firetruck relocation. We call a move of a firetruck a relocation if the firetruck crew is idle at a fire station and it is given the task to relocate themselves. When a firetruck becomes available again, this firetruck needs to be relocated to a fire station. If it is relocated to its own fire station, this does not count as a relocation. After all, this does not inconvenience the crew since that fire station is their base.

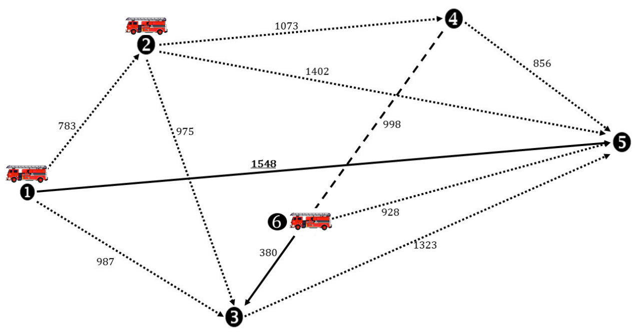
Redirection occurs when a firetruck in phase 0 is driving to its destination but while executing the relocation its destination will be changed. For simplicity, we make the assumption that during the execution of a relocation, the firetruck is for one-thirds of the driving time present at its origin and two-thirds of the time present at its destination. The dispatcher usually faces three issues at decision moments:

1. Is a relocation or a motion necessary?
At decision moments of the first type, it may be the case that the resulting configuration after the dispatch is still satisfactory, in terms of expected response times to future incidents. That is, it may not be beneficial to execute a relocation/motion by reasons mentioned in Chapter 1. This question does not arise at decision moments of the second type, since the dispatcher is always required to perform a firetrucks motion for the firetrucks that just became idle.

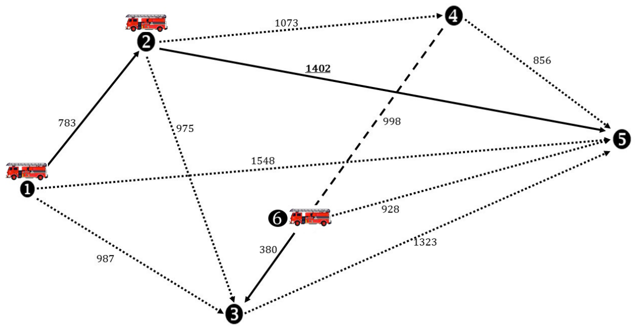
2. Which relocation or motion should be executed?
The dispatcher must select two fire stations: one serving as origin, one as destination. A heuristic method for calculating the best relocation/motion is described in Chapter 5.

3. How to execute this relocation or motion exactly?
When the dispatcher has selected an appropriate relocation/motion, the dispatcher faces a third problem regarding the exact execution of this firetruck relocation/motion. The obvious way is to select a firetruck from the origin and to relocate it to the destination of the relocation/motion. However, the origin and destination are not necessarily close to each other and thus the travel time between them may be long. Such long trips are not desirable, since the new firetruck configuration must be attained as soon as possible.

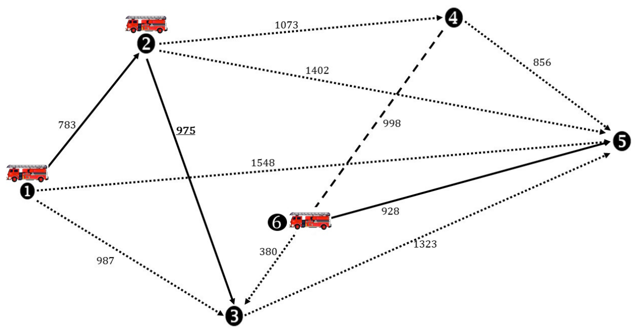
A possibility to avoid long trips is the usage of multiple firetrucks in phase 0, either driving or at a base location, in a motion. Instead of moving just one firetruck, it could be beneficial to break up the firetrucks motion in two or more separate firetrucks relocations to ensure that the new firetrucks configuration is attained earlier. We refer to Example 1 for an illustration of a firetruck motion.



(a)



(b)



(c)

Figure 4.5. Illustration of the usage of multiple firetrucks per motion. The motion is (1; 5) and full arc denote the way in which firetrucks are relocated. The numbers next to the arcs are the driving times in seconds.

Example 1. The firetruck motion we want to perform is (1; 5). There are idle firetrucks in 1 and 2. In addition, one firetruck is traveling from 4 to 3, and is currently in node 6. The obvious way would be to relocate the firetruck from 1 to 5. However, it takes 1548 seconds before the motion is completely performed (Figure 4.5a). If one uses the firetruck at 2, this time can be reduced to 1,402 seconds, at the expense of one extra relocation (Figure 4.5b). In addition, if redirection is allowed, one cause the driving firetruck to decrease the time in which the new firetruck configuration is attained to 975 seconds (Figure 4.5c). We again face a trade-off between the number of relocated firetrucks and the time it takes to attain the new firetruck configuration. In the next section, we will present a heuristic method concerning these three problems.

5. Heuristic Method

For the evaluation of the usefulness of firetruck motions and relocations, we use a heuristic that can easily handle several types of restrictions on the decisions of the dispatcher. First, we describe the heuristic method. Then, we will provide a more detailed explanation regarding the incorporation of these constraints.

5.1 Unpreparedness

The *unpreparedness* at a decision moment plays an important role in the heuristic method. This is a measure for the configuration of firetrucks. The dispatcher observes the current state of the system at a decision moment and executes the motion that minimizes the unpreparedness.

The interpretation of the term unpreparedness generalizes to being an approximation of the expected penalty the next emergency request generates, for a given configuration. We proceed with a formal definition of unpreparedness of a firetruck configuration. To do this we need some additional definitions.

Let s be the current state of the system: the current location or destination of firetrucks and the phases they are in and let F be the set of firetrucks where $F := |F|$. We define $F^k(s)$ as the set of firetrucks in phase k if the state of the system is s.

To define unpreparedness formally, we need some additional definitions. Consider node $i, 1 \leq i \leq N$. Let $des(j, s)$ denote the destination of firetruck j if the state of the system is s, and R is the driving time matrix. Let w be the traffic intensity factor. The driving time between the destination of the closest firetruck in phase 0 and demand location i in timeslot h is defined as

$$\mathrm{r}^{0}_{\mathrm{i,h}}(\mathrm{s}) = w_h \min_{k \in F^0(s)} R(des(j, s), i)$$

The destination equals the current location of the firetruck if the firetruck is not on the road. The reason we use the destination instead of current location, is twofold:

1. If we had used the actual location of the driving phase 0-firetrucks, one might think that one can quickly respond to an incident in the area in which the firetruck is currently driving. However, we are uncertain about the time of the next incident. If the next incident happens in that particular area after some time, it may take long to respond to this incident, since the firetruck has left that area.

2. The opposite case holds as well: for the destination of the firetruck the ability to respond to incidents happening there quickly may be poor, since the firetruck is still far away. Hence, the heuristic may decide to send a firetruck to that destination. However, this is probably useless, since there is already a firetruck traveling to that destination.

Therefore, we use the destinations of driving firetrucks instead of the current location. Moreover, let $p_h(i)$ denote the demand probability in timeslot h: the probability that an incoming incident will occur in node i. Now we have all the ingredients to define the unpreparedness in timeslot h of the configuration of firetrucks, denoted by $U_h(s)$ if the current state of the system is s:

$$U_h(s) := \sum_{i=1}^{N} f\left(\mathrm{r}_{\mathrm{i,h}}^{0}(\mathrm{s})\right) p_h(i),$$

where f is the penalty function.

Example 2. Consider the system in Figure 4.5a. Suppose it is 9 o'clock in the morning. Assume each node has the same demand probability at this hour of the day: $p_9(i) = \frac{1}{5}, i = 1, \ldots, 5$. Moreover, suppose we use the penalty function corresponding to the minimization of the average response time: $f(t) = t, t \geq 0$. We compute $\mathrm{r}_{1,9}^{0}(\mathrm{s}) = \mathrm{r}_{2,9}^{0}(\mathrm{s}) = 0$, since firetrucks are present at nodes 1 and 2. Moreover, $\mathrm{r}_{3,9}^{0}(\mathrm{s}) = 0$ as well, because node 3 is the destination of a driving firetruck. The closest firetruck to node 4 is in node 2, since the firetruck traveling from 4 to 3 is assumed to be at its destination. Therefore, $\mathrm{r}_{4,9}^{0}(\mathrm{s}) = 1.2 \times 1073$ and $\mathrm{r}_{5,9}^{0}(\mathrm{s}) = 1.2 \times 1323$. We have multiplied the driving times with intensity factor w_9= 1.2. At last, the computed unpreparedness is $\frac{3}{5} \times 0 + \frac{1}{5} \times 1287.6 + \frac{1}{5} \times 1587.6 = 575.04$. This is the expected time required to respond to the next incident for the configuration 1, 2, 3.

We did not consider the firetrucks in phase 1 or 2, for specific reasons. The expected remaining busy time of phase 1-firetrucks is probably too large, and thus they are not considered. Expected remaining busy times for phase 2-firetruck are shorter, but highly uncertain since the size of the emergency cannot be determined in advance.

Note that there are several differences between the unpreparedness defined here and the preparedness introduced in [2]. First, unpreparedness has the nice physical interpretation of the expected penalty, e.g., response time, of the next incident. Moreover, no artificial contribution factor is incorporated in the computation. Besides, the definition of preparedness is based on travel times solely, while in the unpreparedness definition a general penalty function is incorporated.

5.2 Evaluation of the Firetruck Motions

At a decision moment of the first type, determining the unpreparedness of the state of the system is the first step in the heuristic. That is, the motion in which none of the firetrucks move except for the ones on the road. We refer to this motion as the static motion, denoted by m_0. For the remainder, we denote the unpreparedness when m_0 is carried out by $U(s_0)$. Subsequently, we evaluate firetruck motions. Denote the remaining possible firetrucks motions by $m_1, m_2, \ldots, m_K$, enumerated by $1, \ldots, K$. Moreover, let s_k denote the state of the system as if m_k was carried out instantaneously and all driving phase 0-firetrucks would be at their destinations. Then, we compute $U(s_k)$ for $1 \leq k \leq K$ to obtain a classification of the firetruck motions. The best motion is the firetruck motion that minimizes the unpreparedness. That is, we select the motion m_l for which

$$s_l = \min_{k = 0,\ldots,K} U(s_k).$$

For decision moments of the second type, we do something similar. However, the firetruck that just finished service of a scene, has to be relocated anyway. This is a consequence of the restriction that each firetruck has to return to a base location. Therefore, we cannot define the static motion as before, in which this firetruck would keep its position. Alternatively, we define our static motion to be equal to the motion in which the just finished firetruck is relocated to the nearest fire station. Moreover, we denote this static motion by m_0.

Note that the number of possible motions is $\mathcal{O}(FB)$, where F and B are the number of firetrucks and base locations, respectively. For decision moments of the second type, the number of firetruck motions is $\mathcal{O}(B)$, since the dispatcher has to decide on a new location only for the firetruck that just finished service. Note that the computation of the unpreparedness can be done in $\mathcal{O}(NF)$ time, since for N demand points we have to determine which of the F firetrucks is the closest phase 0-phase firetruck. Therefore, the total complexity of the algorithm is $\mathcal{O}(NF^2B)$, which is polynomial in the number of demand points, fleet size and number of base locations.

5.3 From Motions to Relocations

Let m_l be the best firetruck motion, and assume $m_l = (b^l{}_1, b^l{}_2)$ is the pair of base stations, where $b^l{}_1$ is the origin and $b^l{}_2$ the destination. Once the firetruck motion is determined, the dispatcher needs to make a decision concerning the exact execution of this motion. To be more specific, the number of additional firetrucks and which ones involved in carrying out this motion need to be determined. We do this by solving a Linear Bottleneck Assignment Problem (LBAP). The formal definition of the LBAP is: Given two sets V and W, together with a weight function $c\colon V \times W \to \mathbb{R}$. Find a bijection $g\colon V \to W$ such that the cost function $\max_{v \in V} c(v, g(v))$ is minimized. The LBAP can be solved to optimality in polynomial time, for instance by methods presented in [4].

In our setting, this is equivalent to the computation of an assignment of phase 0-firetrucks to the base locations that have to be occupied by a firetruck in the new configuration, in such a way that the maximum driving time of a firetruck is minimized. To be more specific, if we denote the set of destinations for phase 0-firetrucks by D_0, we define the set $W = \{D_0 \cup \{b^l{}_2\}\} \setminus \{b^l{}_1\}$. The set V consists of the current locations of the phase 0- firetrucks. When there are multiple firetrucks per location, we specify the elements corresponding to this location with sub-indices in either V or W. Therefore, $|V| = |W|$. Let c be the function describing the driving time between elements of V and elements of W, obtained from the driving time matrix R.

We can interpret the solution to the LBAP in our setting as follows: it is the minimal time required to perform the firetruck motion. Since we base the firetruck motion on the state of the system as it is at the decision moment, apart from the fact that we assume driving phase 0-firetrucks to be at their destination, it is desirable that the new firetruck configuration is attained quickly. There is an obvious relationship between the number of additional firetrucks participating in a firetruck motion, and the completion time of the firetruck motion: the more firetrucks are allowed to be relocated, the faster the new firetruck configuration may be attained. However, it may occur that the number of extra firetruck relocations only has a small impact on the performance, since the gain of participation of additional firetrucks in a motion may be limited. Therefore, in the next chapter we will restrict the dispatcher to relocate a limited number of additional firetrucks. Moreover, we compare the performance and the number of firetruck relocations to the case in which all firetrucks are allowed to take part in the motion.

5.4 Constraints on Decisions

We restrict the dispatcher in two ways: (1) The dispatcher is only allowed to perform the best motion if the gain in unpreparedness with respect to the static motion is substantial, and (2) the dispatcher is not allowed to relocate more than M phase 0-firetrucks in a motion.

In order to get a feeling about the necessity of the best motion, m_l, we compare it to the static motion m_0, defined as above. To be more specific, we compute

$$q := \frac{U(s_0) - U(s_l)}{U(s_0)}$$

where $U(s_0)$ and $U(s_l)$ denote the unpreparedness of the state of the system when, respectively, the static and best motion are performed. Note that $U(s_l) \leq U(s_0)$, since the best motion may equal the static motion. We define Q to be the motion threshold: the dispatcher may carry out the best motion only if $q > Q$. Note that $0 \leq q \leq 1$. If we set $Q = 1$, the dispatcher is restricted to the execution of the static motion solely. In contrast, if $Q = 0$, the dispatcher is always allowed to perform the best motion, even if it results in just a small gain in unpreparedness. Note that we prefer to assess the performance using a relative metric as opposed to an absolute metric. The latter makes sense when a strict $0 - 1$ penalty function is used, however, since we allow for general penalty functions the former is preferable.

The second type of restriction is closely connected to the third question at the end of Section 2.3: the way in which a firetruck motion is carried out, i.e., the number of firetrucks used to perform a firetruck motion. The above-mentioned M is a hard constraint that holds for both types of decision moments and $1 \leq M \leq F$. Remember that a dispatcher may at any time redirect a firetruck if it is already on the road, since this does not count as an extra relocation. Thus, the number of redirected firetrucks is not restricted by M.

In short, the restrictions are parameterized by (Q, M). A summary with the different steps of the method is given at the end of this chapter. In the next chapter, we show some results regarding the performance of the system and the number of relocations as function of Q and M.

Remember that we only consider the closest firetruck. If each base location is the destination of at least one phase-0 firetruck at a decision moment of the second type, all motions are evaluated as equally good. Similarly, for decision moments of the first type, it could occur that the best motion is not unique as well in such a situation. If this is the case, we create scarceness in the number of phase-0 firetrucks by ignoring exactly one firetruck of each base station, and we compute the best motion based on this configuration. If each base location is occupied twice, that is, each base location is the destination of at least two firetrucks, then we always carry out the static motion. However, for the regions and situations we studied, this was hardly the case.

Summary of the approach

1. Consider the system as if each firetruck is at its destination.
2. For each combination of origin and destination:
    (a) Remove one firetruck from the origin.
    (b) Add one firetruck to the destination.
    (c) Compute the unpreparedness of the resulting configuration.
3. Select the best motion and compare it to the static motion.
4. If $q > Q$: Solve LBAP with at most M firetrucks.

6. Results

In this chapter, we show results for the area Amsterdam Amstelland, displayed in Figure 6.1. Amsterdam Amstelland is a region in the Netherlands and covers approximately 282 km^2 and is home to 1.2 million inhabitants, of which 68% lives in Amsterdam itself.

Figure 6.1. Amsterdam Amstelland.

We model Amsterdam Amstelland as a directed complete graph with 2648 nodes, where each arc is weighted according to R. To compute routes between any pair of demand locations, we define a demand location-incidence graph where nodes are only connected by an edge to each other if the corresponding demand locations are adjacent. Using this demand location-incidence graph, we compute the fastest driving routes. We need these routes to keep track of the actual locations of firetrucks when they are not at their base location.

We obtained historical data on the time and place of the incidents and the service time on scene. We used this data for the computation of the demand probabilities per node $p = (p_1, \dots, p_{2648})$, by dividing the number of requests at i by the total number of requests.

The total number of incidents in the data is 90,847. There are 2,896 natural days in our dataset, so on average there are approximately 31 priority 1-incidents during a day. When a day is over, we reset our system to the initial state and proceed with the next day.

We made a distinction between the hours of the day. In Table 6.1 the mean on-scene time with the standard deviations, both in seconds can be found. It can be seen that during the morning rush hour both the mean and the standard deviation are at their highest. Therefore, the greatest gain in relocating can be achieved in the time interval [8,12] given that the mean and standard deviation are the highest in these intervals.

Table 6.1. The mean and the standard deviations of the response time in seconds during a day.

Hours	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Mean	238	244	256	243	247	242	235	238	289	315	267	241
SD	282	396	345	252	250	246	246	365	1178	1234	843	514
Hours	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
Mean	257	247	245	246	233	239	222	226	229	245	233	235
SD	680	469	313	224	273	453	290	254	241	277	253	261

We simulated our system according to historical data, which runs between January 2008 and December 2015. The results are generated by a discrete-event simulation. No randomness is involved, since we use the actual historical data (trace-driven). The simulation evolves according to the system dynamics described in Section 4.2. When a firetruck just got freed from service and there are still requests waiting because no firetrucks were available, the firetruck will immediately respond to the one that is longest in the system.

We consider two different situations: a realistic situation and a busy situation whereby multiple firetrucks are more often occupied e.g. on storm days and holidays like New Year's Eve.

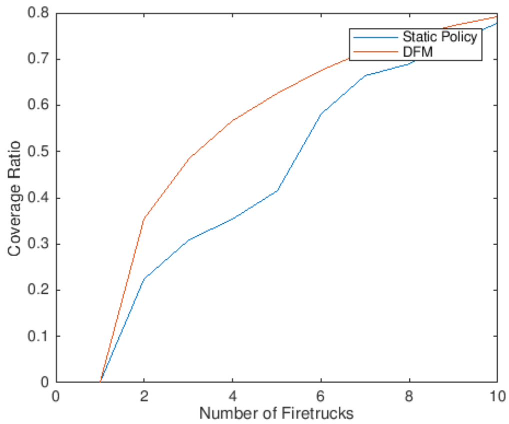
As mentioned before, the redeployment of firetrucks might be beneficial if there is scarceness in the number of available firetrucks. If we apply the heuristic method described in Chapter 5, we implicitly assume available firetrucks are scarce. After all, the contribution of each node to the unpreparedness depends on one firetruck solely, namely the closest one. Therefore, we assume that there is scarceness. To achieve this, we start with one available firetruck. We will increase this number to the maximum of available firetrucks. We assume the same scarceness for both the realistic as for the busy situation.

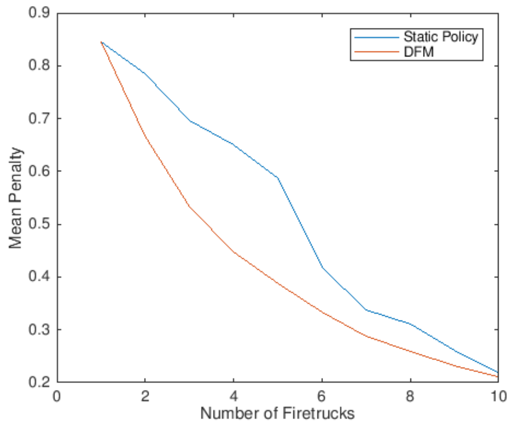
6.1 The Realistic Situation

For the realistic situation, we have displayed the coverage ratio and the mean penalty as function of the number of firetrucks, see Figure 6.2.

Two graphs can be seen in the plots. The red line is the *Static Policy*. No relocations are carried out when this policy is used. For the blue line, the *DFM* is being used. In this case relocating is allowed according to the heuristic method described in Chapter 5.

The first thing to note in Figure 6.2a is that the coverage ratio for both policies increases as the number of available firetrucks grows. The biggest difference between the two policies can be seen with three, four firetrucks. The difference, expressed in





(a) **(b)**

Figure 6.2. The coverage ratio as function of the firetrucks is displayed in Figure 6.2a and the mean penalty as a function of the firetrucks is displayed in Figure 6.2b.

percentages, are approximately 58.8% and 59.9%, respectively. Note that the largest gap between the DFM algorithm and the Static Policy is at $F = 4$. As the availability of more firetrucks increases, the difference between the two graphs is reduced. For example, the difference between the coverage ratio determined by the DFM Algorithm and the Static Policy equals 1.9% when 10 firetrucks are available. Thus, there is a significant gain in performance if more than one firetruck is used in performing a motion. Another striking difference to note is the flow of the graphs. The coverage ratio which is determined by the DFM Algorithm is declining increasing smoothly. While the coverage ratio which is determined by the Static Policy is graph whose line breaks.

In Figure 6.2b the mean penalty is plotted. As the number of firetrucks increases, the penalty decreases. This makes sense since the area is better covered with more firetrucks at our disposal. When the system contains 10 firetrucks the mean penalty with the static policy hardly differs from the penalty. Note that the largest gap between the static policy and the DFM algorithm is at $F = 4$. This gap is approximately 18.8% as observed in Table 6.2. Thus, there is a significant gain in performance if more than one ambulance is used in performing a motion. The biggest gain is achieved when we have 4 firetrucks in the system. In the next section, we will have a closer look at the results for 4 firetrucks in a busier situation. We will also compare these result with the double number of firetrucks, $F = 8$.

Table 6.2: Columns I, II and III represent the gain in performance, the mean response time and the mean penalty for the DFM algorithm compared with the Static Policy, respectively.

	I	II	III
F = 1	0,7%	0%	0.1%
F = 2	56.1%	9.1%	15.2%
F = 3	58.8%	15.2%	23.6%
F = 4	59.9%	18.8%	31.1%
F = 5	50.3%	17.7%	33.9%

	I	II	III
F = 6	16.7%	10.1%	20.6%
F = 7	7.9%	5.7%	14.4%
F = 8	8.3%	5.5%	16.7%
F = 9	4.7%	2.0%	11.3%
F = 10	1.7%	0.4%	4.0%

6.2 The Busy Situation

For the realistic situation we have seen that gain in performance can be achieved by relocating in comparison with the static policy. However, this profit is small. We are now going to look at what happens when we increase the number of emergencies in the same area. We do this by multiplying the arrival rate by 2, 3, 4 and 5. The utilization will also be multiplied in the same way since it is linear to the arrival rate. We call these numbers the utilization factors. The realistic situation is equal to the utilization factor 1.

We have sketched two situations. In the first situation we assume that there are F = 4 firetrucks. In the second situation F = 8 firetrucks are available. We have chosen for these number of firetrucks We will compare these two situations with each other. The results are shown in Figure 6.4.

We have chosen these numbers of firetrucks because we have seen earlier in the realistic situation that there is much to be gained when there are 4 firetrucks in the system. When we double the number of firetrucks, the gain decreases, see Table 6.2. With 4 and 8 firetrucks the difference is indicated between scarcity of firetrucks and sufficient firetrucks.

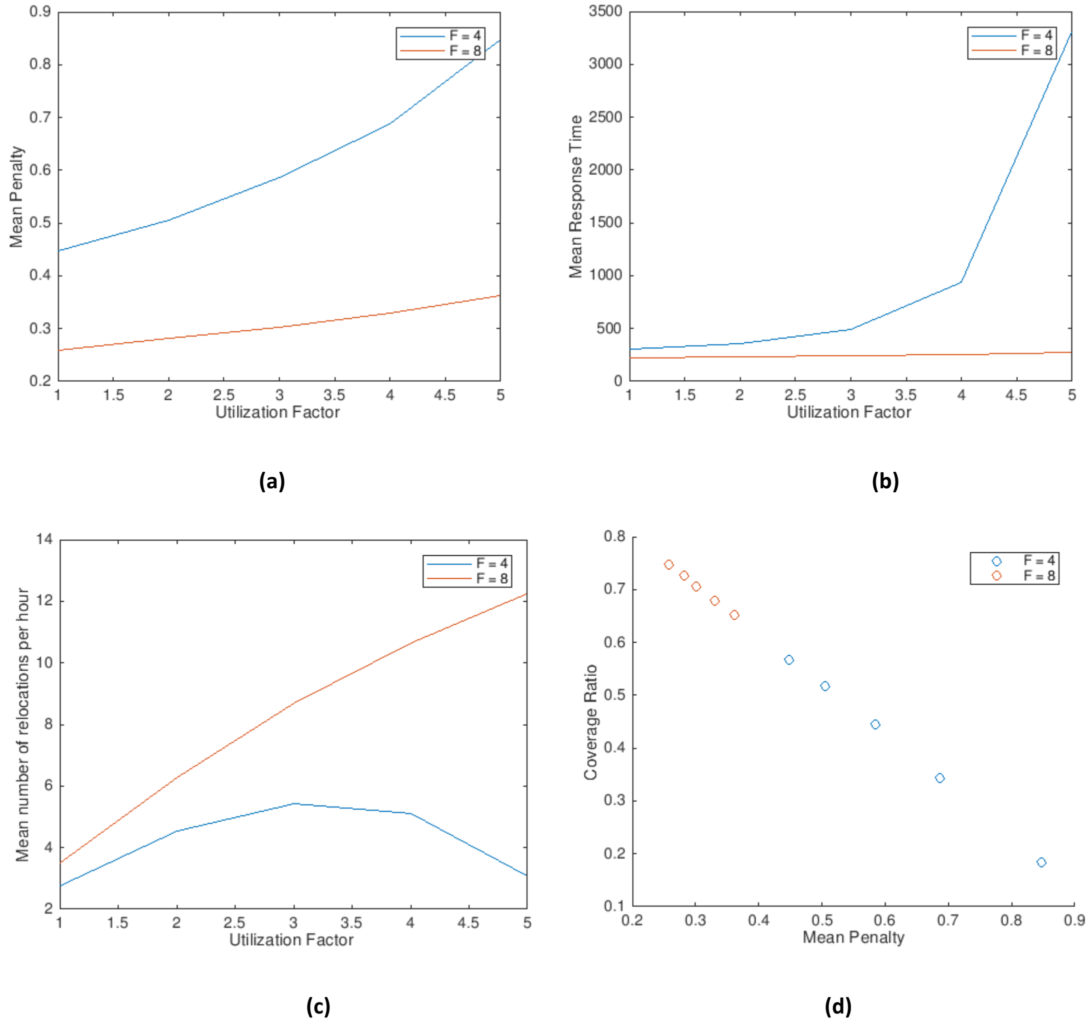


Figure 6.4. The average penalty (Figure 6.4a) and the average response time (Figure 6.4b) as function of the utilization factor. Figure 6.4c displays the number of relocations and Figure 6.4d displays the relation between the average penalty and the coverage ratio.

In Figure 6.4a, we display the penalty as function of the utilization factor, for $F = 4, 8$. The graphs differ a lot from each other. The penalty for $F = 4$ increases with big steps as the number of emergencies increase. While the penalty with $F = 8$ remains relatively small. This outcome makes sense since the area is better covered with 8 firetrucks than with 4 firetrucks. The penalty is closely related to the response time. The penalty increases when the response time increases. In Figure 6.4b the response time as a function of the utilization factor, for $F = 4, 8$ is displayed. For utilization factor 1, the response time increases with 25,9% when $F = 8$ instead of $F = 4$. This increases to 91,8% if there are five time as many emergencies, i.e. utilization factor = 5. With 4 firetrucks at our disposal and five times as many incidents, the response time gets very large because the firetrucks are still busy fighting one emergency when new emergencies occur. We also note that the growth of the response time for $F = 8$ is very marginal. This comes at the price of extra firetruck relocations. This number, as a function of the utilization factor for $F = 4$ and $F = 8$ is displayed in Figure 6.4c. We observe approximately nine additional firetruck relocations per hour. Note that the number of relocations is decreasing for $F = 4$. The cause of this is the occupancy of the firetrucks. When we increase the utilization, the firetrucks are more frequently occupied and cannot be used for performing relocations.

6.2 The Four Emergency Types

As we mentioned in Chapter 4, there are four emergency types. For each of these emergency types we looked at the mean penalty.

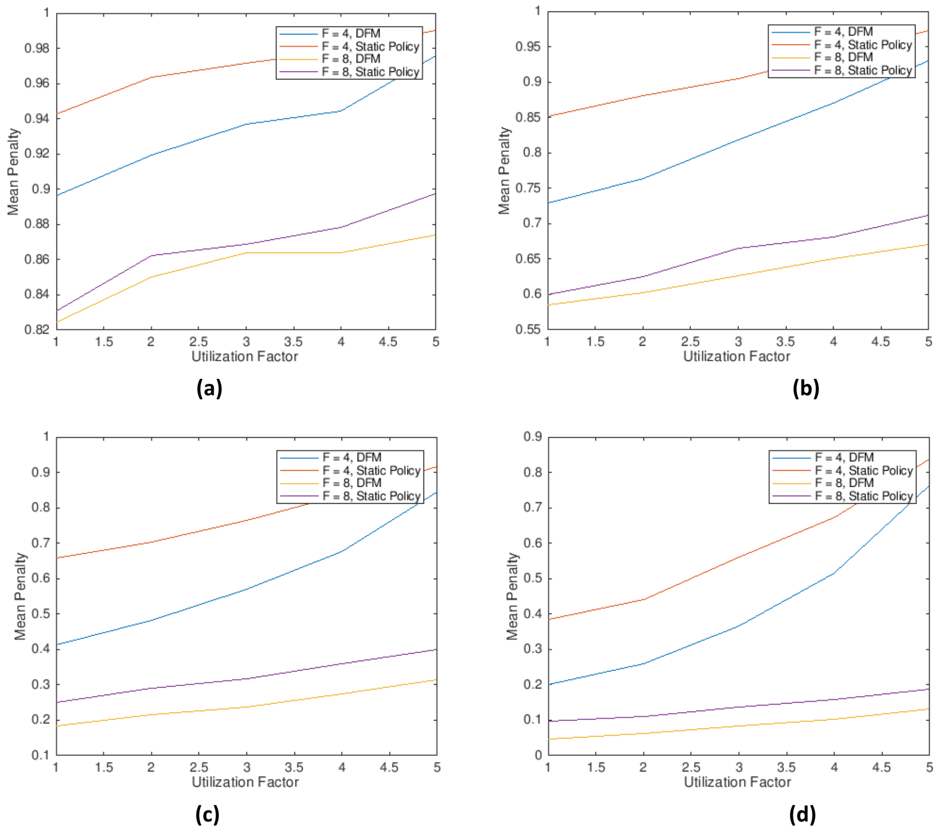


Figure 6.5. The average penalty for the emergency types A (Figure 6.5a), B (Figure 6.5b), C(Figure 6.5c), and D (Figure 6.5d).

In Figure 6.5 we see, as expected, that the penalty for emergency type A in comparison with the other types is the highest since the maximum allowed response time for these type of emergency is the smallest. As the response time increases, the penalty becomes smaller. This applies to both 4 firetrucks and 8 firetrucks and also both policies. We also see in all figures that the graphs which belong to the DFM algorithm a lower penalty compared to the static policy.

In Figure 6.5a we note that the two graphs that belong to 8 vehicles do not differ so much from each other. The gap between the two graphs that belong to 4 vehicles is a lot bigger. For the utilization factor 2 this difference is approximately 4.6% as observed in Table 7. It is remarkable that with utilization factor 5 the penalty suddenly shoots up for 8 firetrucks and that the graph comes closer to that of the Static Policy. As the utilization factor gets larger, the static policy and the FDM are getting closer together. This is in case there are too many emergencies and there is no possibility to relocate the firetrucks because of the unavailability. Regardless of the policy that is used, in both situations all the firetrucks will be occupied all the time. To lower the penalty, more firetrucks are needed. The same applies for Figure 6.5b. In this case, the penalty is lower because the maximum response time for these types of incidents is higher. In Figure 6.5c we see that the graphs for 4 firetrucks already start to look more alike and in Figure 6.5d they are approximately parallel to each other. This means that the achieved gain in performance with the FDM algorithm remains approximately the same be regardless of the number of emergencies that need to be dealt with. For 4 firetrucks, the difference becomes smaller.

In Table 6.3 the gain in performance expressed in percentages can be seen. With 4 firetrucks, the gain in performance for each utilization factor gets larger as the maximum allowed response time increases. This also holds for 8 firetrucks. Note that the difference between emergency type A and B is small. This difference gets larger as the maximum allowed response increases.

Table 6.3. The gain in performance for the mean penalty. The columns represent the utilization factors. The numbers in the cells show the gain in performance when the DFM is used to manage the firetrucks for F = 4 and F = 8.

		I	II	III	IV	V
F = 4	A	4.9%	4.6%	3.6%	3.5%	1.4%
	B	14.4%	13.3%	9.5%	7.2%	4.4%
	C	37.2%	31.2%	25.5%	18.9%	8.0%
	D	47.8%	41.2%	34.7%	23.5%	11.4%
F = 8	A	0.8%	1.4%	0.7%	1.6%	2.6%
	B	2.3%	3.8%	5.8%	4.6%	5.8%
	C	26.3%	25.7%	25.1%	23.8%	21.9%
	D	51.7%	44.8%	38.9%	35.1%	29.7%

7. Conclusion & Discussion

The main objective of this study was to develop a time dependent relocation model for the FDAA such that the response time of the firefighters would be decreased significantly. This objective was reached. In this chapter, we will provide a summary and the conclusion about this objective. However, we think that there is still a lot of potential improvement for this model. The model has to be extended significantly before a meaningful comparison to real operations can be made. The possible extensions will be described in the discussion and we end this thesis with a short note on further research, which can be carried out.

7.1 Conclusion

In this thesis, we analyzed the effect of firetruck relocations on the performance of the firefighter services. Theretofore, we described a firetruck redeployment model, in which a performance measure related to the response time can be chosen by the fire department by defining a corresponding penalty function. We used historical data of the region Amsterdam Amstelland to simulate the system.

The analysis of the dataset of historical emergencies demonstrates that and how the response time can be improved by simply relocating. The results in Chapter 6 showed that DFM mainly gives a great potential for areas in which the coverage is rather low (the busy situation with utilization factor 4 and 5). For these areas, the presented results all imply that there is a significant improvement if firetrucks are relocated compared to the static policy. For the areas that already have a high coverage, the gain in performance by relocating is very small. Likewise, for the realistic situation, the decrease in penalty is largest if only a few firetrucks are allowed in the system. It gets harder and harder to increase the performance by allowing more firetrucks in the system.

We conclude this section by noting that the penalty is mainly caused by the emergency types A and B. This makes sense since the maximum allowed response time for these two types (120 and 130 seconds, respectively) is very small. Even with more firetrucks in the system and the use of DFM, the penalty remains high. For these emergencies, a lot of improvements can be made. If the penalty of these specific emergencies decreases, the total mean penalty will also decrease. The fire department could consider to execute more firetruck relocations for emergency A and B and less for emergency C and D resulting in a better coverage in the area for emergency A and B. This however requires more research.

The graphs presented in this thesis can be very useful for the fire department to gain insights in the relationship between performance and number of relocations. DFM is generally believed to provide means to enhance the response-time performance.

7.2 Discussion

There are various factors that may have an influence on the performance of the firefighters and on the number of relocations as well which we did not have taken into account. Several extensions can be made to refine of the relocation model. One obvious shortcoming of our approach is that we restricted the dispatcher at a decision moment of the first type to change the firetrucks configuration at at most two points: the origin and the destination. However, it could be beneficial for the performance if this restriction would be relaxed, but this probably comes at the expense of more relocations. Moreover, the relation between performance or number of relocations and number of decision moments is an interesting topic as well: what would happen if one decreases (e.g., only when a firetruck is newly free) or increases (e.g., every minute) the number of decision moments?

This model is based on 1 type of vehicle while the FDAA has multiple types of vehicles. In case of fire, two vehicles are needed: water tender and a turntable ladder. These vehicles are in a sense dependent of each other. This dependency should be further examined.

Another extension would be to take the dependence of the weather into account. We believe it is a valuable addition given that the weather affects the number of incidents. For example, the rate of incoming incidents on storm days is much higher than a typical day. The same holds for hot summer days and holidays e.g. New Year's Eve. A logical consequence is that multiple firetrucks are more often occupied and this limits the number of possible relocations that can be performed. These days are outliers compared with the others days of the year. However, they do occur and the timely arrival of firefighter services also holds for these days. Since our model is based on the average rate during a year, it won't work for days wherein the rate is much higher than the average rate.

At last, we also want to emphasize the difference between the weekdays and weekends. The number of the incoming incidents on weekdays differs a lot from the weekend. The flow of the traffic intensity and traffic density during the weekdays differs also from the weekend. This has an influence on the performance and on the number of relocations as well. During the weekdays, we saw two peaks which imply the rush hours during a day. While in the weekend, the flow of the traffic intensity is much smoother. The traffic intensity slowly increases and in the afternoon, it starts decreasing. The graphical representation of the flow of the traffic intensity and density can be seen in Chapter 3. In our model, we took the average over the whole week and did not distinguish the weekdays from the weekend days. This distinction can be a valuable addition to our model.

The relocation model presented in this thesis forms a good basis for these extensions and modifications which includes the most important parts and principles of the operations which are made by the fire department.

Bibliography

[1] Alanis, R., Ingolfsson, A., & Kolfal, B. (2013). A Markov chain model for an EMS system with repositioning. *Production and operations management*, *22*(1), 216-231.

[2] Ball, M. O., & Lin, F. L. (1993). A reliability model applied to emergency service vehicle location. *Operations research*, *41*(1), 18-36.

[3] Barneveld, van, T. C., Bhulai, S., & van der Mei, R. D. (2016). The effect of ambulance relocations on the performance of ambulance service providers. *European Journal of Operational Research*, *252*(1), 257-269.

[4] Barneveld, van, T. (2016). The minimum expected penalty relocation problem for the computation of compliance tables for firetruck vehicles. *INFORMS Journal on Computing*, *28*(2), 370-384.

[5] Batta, R., Dolan, J. M., & Krishnamurthy, N. N. (1989). The maximal expected covering location problem: Revisited. *Transportation Science*, *23*(4), 277-287.

[6] Berg, van den, P., Legemaate, G., & van der Mei, R. (2016). Boosting the Responsiveness of firefighter Services with Mathematical Programming. *ERCIM NEWS*, (105), 20-21.

[7] Brotcorne, L., Laporte, G., & Semet, F. (2003). Firetruck location and relocation models. *European journal of operational research*, *147*(3), 451-463.

[8] Budge, S., Ingolfsson, A., & Zerom, D. (2010). Empirical analysis of firetruck travel times: the case of Calgary emergency medical services. *Management Science*, *56*(4), 716-723.

[9] Church, R. L., & Davis, R. R. (1992). The fixed charge maximal covering location problem. *Papers in Regional Science*, *71*(3), 199-215.

[10] Church, R., & ReVelle, C. (1974, December). The maximal covering location problem. In *Papers of the Regional Science Association* (Vol. 32, No. 1, pp. 101-118). Springer-Verlag.

[11] Daskin, M. S. (1983). A maximum expected covering location model: formulation, properties and heuristic solution. *Transportation science*, *17*(1), 48-70.

[12] De Maio, V. J., Stiell, I. G., Wells, G. A., Spaite, D. W., & Ontario Prehospital Advanced Life Support Study Group. (2003). Optimal defibrillation response intervals for maximum out-of-hospital cardiac arrest survival rates. *Annals of emergency medicine*, *42*(2), 242-250.

[13] Erkut, E., Ingolfsson, A., & Erdoğan, G. (2008). Firetruck location for maximum survival. *Naval Research Logistics (NRL)*, *55*(1), 42-58.

[14] Fitch, J. (2005). Response times: myths, measurement & management. *JEMS: a journal of emergency medical services*, *30*(9), 47-56.

[15] Gendreau, M., Guertin, F., Potvin, J. Y., & Taillard, E. (1999). Parallel tabu search for real-time vehicle routing and dispatching. *Transportation science*, *33*(4), 381-390.

[16] Gendreau, M., Laporte, G., & Semet, F. (2001). A dynamic model and parallel tabu search heuristic for real-time firetruck relocation. *Parallel computing*, *27*(12), 1641-1653.

[17] Gendreau, M., Laporte, G., & Semet, F. (2006). The maximal expected coverage relocation problem for emergency vehicles. *Journal of the Operational Research Society*, *57*(1), 22-28.

[18] Ingolfsson, A., Budge, S., & Erkut, E. (2008). Optimal firetruck location with random delays and travel times. *Health Care Management Science*, *11*(3), 262-274.

[19] Ingolfsson, A., Erkut, E., & Budge, S. (2003). Simulation of single start station for Edmonton EMS. *Journal of the operational research society*, *54*(7), 736-746.

[20] Jagtenberg, C. J., Bhulai, S., & van der Mei, R. D. (2015). An efficient heuristic for real-time firetruck redeployment. *Operations Research for Health Care*, *4*, 27-35.

[21] Jenelius, E., & Koutsopoulos, H. N. (2013). Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research Part B: Methodological*, *53*, 64-81.

[22] Kolesar, P., & Walker, W. E. (1974). An algorithm for the dynamic relocation of fire companies. *Operations Research*, *22*(2), 249-274.

[23] Koole, G. (2010). Optimization of business processes: An introduction to applied stochastic modeling. *Amsterdam, The Netherlands: VU University Amsterdam*.

[24] Larsen, M. P., Eisenberg, M. S., Cummins, R. O., & Hallstrom, A. P. (1993). Predicting survival from out-of-hospital cardiac arrest: a graphic model. *Annals of emergency medicine*, *22*(11), 1652-1658.

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

[26] Larson, R. C. (1975). Approximating the performance of urban emergency service systems. *Operations Research*, *23*(5), 845-868.

[27] Larson, R. C., & Odoni, A. R. (1981). *Urban operations research* (No. Monograph).

[28] McLay, L. A., & Mayorga, M. E. (2010). Evaluating emergency medical service performance measures. *Health Care Management Science*, *13*(2), 124-136.

[29] Marianov, V., & Revelle, C. (1994). The queuing probabilistic location set covering problem and some extensions. *Socio-Economic Planning Sciences*, *28*(3), 167-178.

[30] Marianov, V., & ReVelle, C. (1995). Siting emergency services. *Facility Location: a survey of applications and methods*, *1*, 199-223.

[31] MathWorks, Inc. (2005). *MATLAB: the language of technical computing. Desktop tools and development environment, version 7* (Vol. 9).

[32] Maxwell, M. S., Henderson, S. G., & Topaloglu, H. (2009, December). Firetruck redeployment: An approximate dynamic programming approach. In *Winter Simulation Conference* (pp. 1850-1860). Winter Simulation Conference.

[33] Repede, J. F., & Bernardo, J. J. (1994). Developing and validating a decision support system for locating emergency medical vehicles in Louisville, Kentucky. *European journal of operational research*, *75*(3), 567-581.

[34] ReVelle, C., & Hogan, K. (1989). The maximum availability location problem. *Transportation Science*, *23*(3), 192-200

[35] Robbins, T. R., Medeiros, D. J., & Harrison, T. P. Evaluating the Erlang C and Erlang A Models for Call Center Modeling.

[36] Schmid, V., & Doerner, K. F. (2010). Firetruck location and relocation problems with time-dependent travel times. *European journal of operational research*, *207*(3), 1293-1303.

[37] Shanesy, C. E. (1975). An On-Line Program for Relocating Fire-Fighting Resources.

[38] Stiell, I. G., Wells, G. A., Field III, B. J., Spaite, D. W., De Maio, V. J., Ward, R., ... & Maloney, J. (1999). Improved out-of-hospital cardiac arrest survival through the inexpensive optimization of an existing defibrillation program: OPALS Study Phase II. *Jama*, *281*(13), 1175-1181.

[39] Sudtachat, K., Mayorga, M. E., & Mclay, L. A. (2016). A nested-compliance table policy for emergency medical service systems under relocation. *Omega*, *58*, 154-168.

[40] Team, R. C. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013.

[41] Toro-DíAz, H., Mayorga, M. E., Chanta, S., & Mclay, L. A. (2013). Joint location and dispatching decisions for emergency medical services. *Computers & Industrial Engineering*, *64*(4), 917-928.

[42] Valenzuela, T. D., Roe, D. J., Cretin, S., Spaite, D. W., & Larsen, M. P. (1997). Estimating effectiveness of cardiac arrest interventions. *Circulation*, *96*(10), 3308-3313.

[43] Waalewijn, R. A., de Vos, R., Tijssen, J. G., & Koster, R. W. (2001). Survival models for out-of-hospital cardiopulmonary resuscitation from the perspectives of the bystander, the first responder, and the paramedic. *Resuscitation*, *51*(2), 113-122.

[44] Westgate, B. S., Woodard, D. B., Matteson, D. S., & Henderson, S. G. (2013). Travel time estimation for firetrucks using Bayesian data augmentation. *The Annals of Applied Statistics*, *7*(2), 1139-1161.

[45] J. Naoum_Sawaya & S. Elhedhli. A stochastic optimization model for real-time ambulance redeployment. Computers & Operations Research, 40:1972-1978, 201.

[46] D.S. Matteson, M.W. McLean, D.B. Woodard, and S.G. Henderson. Forecasting emergency medical service call arrival rates. The Annals of Applied Statistics, 5(2B):1379–1406, 2011.

[47] M.S. Maxwell, M. Restrepo, S.G. Henderson, and H. Topaloglu. Approx- imate dynamic programming for ambulance redeployment. INFORMS Journal on Computing, 22(2):266–281, 2010.