Ambulance Assignment for Medical Emergencies

João Bernardo Correia da Silva

Thesis to obtain the Master of Science Degree in

Information Systems and Computer Engineering

Supervisors: Professor Vasco Miguel Gomes Nunes Manquinho
Professor Maria Inês Camarate de Campos Lynce de Faria

Examination Committee

Chairperson: Professor José Luís Brinquete Borbinha
Supervisor: Professor Vasco Miguel Gomes Nunes Manquinho
Member of the Committee: Professor Luís Manuel Silveira Russo

November 2021

Acknowledgments

Firstly, I would like to thank everyone that directly contributed to the elaboration of this document, including the Professors in charge of supervising this Thesis to other Professors or students that provided assistance throughout the development process.

I would also like to thank my family for supporting me and helping whenever they could, making this hard assignment just a little more tolerable. This goes especially for the closest family, my mother, father and sister.

Lastly, I would also like to thank everyone who in one way or another contributed something to make it possible for me to have elaborated this Thesis and reach the end (for now) of this academic journey.

This work was partially supported by Portuguese national funds through FCT, Fundação para a Ciência e a Tecnologia, under projects UIDB/50021/2020, DSAIPA/AI/0044/2018 and PTDC/CCI-COM/31198/2017.

Resumo

No mundo médico, um dos propósitos das ambulâncias é providenciar auxílio a situações de emergência e transporta pessoas para instalações onde estas possam ter acesso a melhores cuidados médicos. No entanto, existe a possibilidade de haverem demasiadas emergências para as ambulâncias disponíveis numa certa região, ou que demasiadas ambulâncias estejam disponíveis, o que resultaria num desperdício de recursos. Encontrar um balanço entre a disponibilidade de recursos, providenciar auxílio a todas as emergências médicas e a minimização dos tempos de resposta tem sido demonstrado ser uma tarefa árdua.

Neste documento, analisamos a literatura sobre o problema de alocação de ambulâncias dos Sistemas de Emergência Médica, falando de diversos modelos apreentados em trabalhos de investigação anteriores que focam no assunto dos Sistemas de Emergência Médica. Alguns definem e formulam modelos com maior detalhe enquanto outros trabalhos de investigação apresentam técnicas inovadoras.

Adicionalmente, definimos um modelo matemático que consegue representar qualquer situação num período de tempo definido com multiplas emergências com vários veículos que podem prestar ajuda a essas emergências. Desenvolvemos este modelo através da utilização de Multi-Objective Combinatorial Optimization (MOCO) focado na minimização do numero total de recursos utlizados, bem como na minimização dos tempos de resposta. Aplicamos este modelo a dados que recolhemos de três distritos em Portugal num total de três anos, quatro dias por ano. Finalmente resolvemos as instâncias criados por este modelo nestes cenários e inferimos conclusões e possiveís melhorias à operação por parte do Sistema de Emergência Médica nessas situações, assim como em outras possiveís situações.

Palavras-chave: Optimização Combinatória Multi-Objectivo, Problema de Alocação de Ambulâncias, Sistemas de Emergência Médica

Abstract

In the medical world, one of the purposes of ambulances is that of providing aid to emergency situations and transporting people to a facility where they can get further medical attention. However, there is a possibility that there are too many emergencies for the ambulances available in a certain region or that too many ambulances are made available resulting in resource wastage. Finding a balance between the availability of resources, coverage of all medical emergencies and minimizing response times has been shown to be a difficult task.

In this document, we analyze the literature on the Emergency Medical Services (EMS) ambulance location problem, addressing several models presented in previous research work addressing the EMS subject. Some define and formulate models with greater insight while other research works present innovative techniques

Additionally, we define a mathematical model that is able to represent any situation on a defined period of time with multiple emergency occurrences with several vehicles that can provide aid to said occurence. We do this using Multi-Objective Combinatorial Optimization (MOCO) to tackle the ambulance location problem while focusing on minimizing the overall number of resources used, as well as minimizing the response times. We apply this model to real data retrieved from three different districts in Portugal, in various time periods. We then solve the instances created by our model in these scenarios and inferred conclusions and possible improvements to the Emergency Medical System in those situations.

Keywords: Multi-Objective Combinatorial Optimization, Ambulance Location Problem, Emergency Medical System

Contents

List of Figures

List of Tables

Chapter 1

Introduction

An Emergency Medical System (EMS) can be defined as a system that aims to provide urgent treatment or stabilization in medical emergencies. Structurally, a control facility serves as an integrating part of this system, serving the purpose of receiving incoming emergency calls, which can be placed by any person, and then assigning emergency vehicles to the required location, depending on the number of people in need of immediate medical assistance and/or on the seriousness of the emergency at hand. It is important that this service is provided in the least amount of time possible in order to provide medical care to the person in need in the fastest way possible. Once the patient's situation is stabilized and he is transferred to the established health facility, the ambulance that was dispatched becomes available once more as soon as it returns to its base, where it can be given a new task, and all the procedures have been completed for the ambulance to be available again.

In the context of EMS, the vehicle location problem consists of locating the vehicles in some potential service sites in order to reduce the delay of covering emergency service demands [24]. However, in real-world scenarios, there is the need to keep the solutions feasible, while still granting a satisfiable level of optimality. This work focuses on analysing proposed solutions to the vehicle location problem, as the basis to then formulate a model to apply to existing data from the Instituto Nacional de Emergência Médica (INEM) and try to see if there is a more efficient way of providing aid to emergency situations.

If we imagine a one dimensional space which has two emergency sites, represented by the crosses, and two vehicle locations, represented by the dots, arranged as seen in Figure 1.1, and we imagine a scenario where emergency 1 occurs first and only after does emergency 2 occur, it is easy to see why assigning the closest vehicle, the vehicle in location A in this case, to emergency 1 might seem like a decent choice, since the distance to that emergency, which is four kilometers, is much smaller when opposed to the twelve kilometers from location B to emergency 1. However, since vehicle 1 will then be unavailable when emergency 2 occurs, we can only assign vehicle 2, which will now have to travel $12 + 4 + 10 = 26$ kilometers, which means that overall, to provide aid to both emergencies, both vehicle travelled a total of $26+4 = 30$ kilometers. Since our goal is to minimize the total distance travelled by both our vehicles, we can choose to assign vehicle A to emergency 1 and vehicle B to emergency 2, which will lead to a total distance travelled of $12+10 = 22$ kilometers, which is a much better scenario than the one previously described. This is the simplest example in which we can demonstrate what we are looking for when we talk about optimality in the context of the Emergency Medical System performance. In bigger cases, we will have to test every possible assignment of vehicles to every emergency we analyse, which will make the problem grow exponentially.

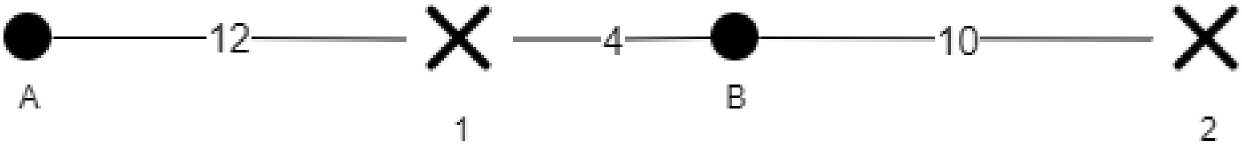

Figure 1.1: Illustrative Example

Additionally we also perform an aggregate demand analysis where we divide a vehicle allocation problem from a large time period, for example a full day, into smaller time periods which are easier to solve with the objective of finding a solution to the bigger problem by combining information from the solutions to the smaller instances of the problem.

Although this situation is based on the assumption that we know where future emergencies will happen, our goal is to create a model that will allow this optimization to be calculated for past solutions where we have complete knowledge of what happened and want to access the performance of the Emergency Medical System in use, or used in conjunction with a predictive model as a way to predict and allocate vehicles in a way that creates a better expected performance in future situations.

This document is organized as follows. In Chapter 3, the fundamental concepts of the optimization types are introduced. In Chapter 2, a set of contextual models and solving methods, as well as their properties, are introduced, providing references from authors that explored the same subject. Chapter 4 contains the modelling and mathematical process that went into the creation of the final model and Chapter 5 provides insight and further investigation of the application of the created model. Chapter 6 concludes the document and Chapter 7 presents possible enhancements or ideas to continue the study of this subject.

Chapter 2

Related Work

One of the ways to improve the procedure that allows an hospital or medical care center to answer medical emergency calls is to assess the ambulance workflow. This workflow can be tackled in a number of different ways, some of them described in this section.

2.1 Overview and early works

The EMS system has an optimization problem regarding the allocation of the emergency vehicles (ambulances) which has been tackled in a number of different ways throughout the years, all of them ultimately aiming at increasing the efficiency of resource usage. The motivation for these works come from the fact that EMS's exist with the purpose of not only assigning ambulances to emergencies when these happen, but also conveying a distribution of these ambulances in a way that allows them to maximize the area covered, as it was first described, by Church & ReVelle [10], in their work that considered a fixed size fleet of ambulances. After this work, there have been an enormous number of authors dwelling upon the intricacies of the ambulance location, relocation and assignment problem, some of them suggest the usage of a dynamic approach, in which each ambulance is able to communicate with other ambulances (multi-agent approach) or with an ambulance coordinator (centralized approach) in order to decide where to go after it has been dispatched to an emergency and therefore maximize the coverage of each zone after each emergency has been dealt with [25, 22, 9]. This dynamic approach is presented with many computational or scalability problems when applied to either large fleets of ambulances or extensive land coverage. This means that dynamic approaches may not be achievable in the expected amount of time, effectively harming the end goal of reducing the overall response time to emergencies.

Since the allocation of emergency vehicles is a very complex problem that involves a great deal of variables and different scenarios, it is important to specify the context for each proposed solution. The details of several of relevant works are present in an very complete and interesting set of tables present in the work of Bélanger [6] in which the author divides and classifies the different works based on the specific details they cover. These tables cover variables such as the number of different types of ambulances available, a list of covering and standby site constraints and the objectives of the work developed in the literature. Different types of ambulances define an interesting alteration to the standard scenario, for example, a certain type of ambulance is more effective in heart disease related emergencies, and the system should, therefore, prioritize assigning these ambulances to emergencies of that nature.

However, other approaches have been taken in order to tackle this problem, namely some deterministic models, probabilistic models [7], and some more recent approaches that use the heuristics, for example, the tabu search heuristic[17]. Discussing all the papers in depth would make this work too extensive and difficult to follow, hence, we will only be discussing a few examples and we will use the tables in the next pages in order provide an overview on which papers cover what topics and which techniques are used.

2.2 Static Ambulance Location Models

The ambulance location model can be defined as a graph $G = (V, E)$ where $V = N \cup M$, $N = (v_1, ..., v_n)$ and $M = v_{n+1}, ..., v_{n+m}$ being two vertex sets representing, respectfully, demand zones and potential standby sites. whereas E is an edge set where each edge $\{(v_i, v_j) : v_i, v_j \in V, i < j\}$ is associated with a travel time or distance t_{ij}. Since most of the models use the notion of coverage, the sets M_i and M'_i correspond to the sets of standby sites that can cover a demand zone v_i respectively within a defined time limit S and S' , $S' > S$. The set N_j will correspond to the set of demand zones that can be reached by a vehicle located in v_j within S . The number of ambulances, when given, is set as p. This model definition to several models throughout this paper, therefore, the nomenclature will be reused further ahead.

2.2.1 Deterministic single coverage Models

The first emergency vehicle location problem explicitly formulated using the notion of coverage is in the work of Toregas, Swain, ReVelle, and Bergman[33], and it is stated in that work that a demand zone is covered if and only if it can be reached by at least one vehicle within a previously established time or distance frame. The objective this work aims to achieve is that of minimizing the number of vehicles in a way that all zones are adequately covered. For this purpose, it uses Boolean variables x_j that take the value 1 when an ambulance is located at $v_j \in M$, the same vertex representing the potential standby sites than in the Static Ambulance Location Model. This is called the location set covering problem (LSCP) and it is formulated as seen in Fig.2.1:

However, the amount of vehicles needed to achieve this level of coverage might not be realistic in practice. Moreover, real-world problems often have a given vehicle fleet, therefore we should take this number of vehicles into account. Having this in mind, Church and ReVelle[10] formulated the maximal covering location problem (MCLP) which seeks to maximize the demand covered by an estabilished number of vehicles p in a fleet. Another Boolean variable y_i which takes the value of 1 when the demand zone v_i is covered by at least one vehicle within the time limit S and a_i which is the demand associated with zone v_i, was added in order to formulate the MCLP, as shown in Fig.2.2.

$$\min \sum_{j=1}^{M} x_j \tag{2.1}$$

$$\text{subject to: } \sum_{j \in M_i} x_j \geq 1, \qquad j = 1, ..., n, \tag{2.2}$$

$$x_j \in \{0, 1\}, \qquad j = 1, ..., m. \tag{2.3}$$

Figure 2.1: Definition of the Location Set Covering Problem

$$\max \sum_{i=1}^{n} a_i y_i \tag{2.4}$$

$$\tag{2.5}$$

$$\text{subject to: } \sum_{j \in M_i} x_j \geq y_i, \qquad i = 1, ..., n, \tag{2.6}$$

$$\sum_{j=1}^{m} x_j = p, \tag{2.7}$$

$$x_j \in \{0, 1\}, \qquad j = 1, ..., m, \tag{2.8}$$

$$y_i \in \{0, 1\}, \qquad i = 1, ..., n. \tag{2.9}$$

Figure 2.2: Definition of the Maximal Covering Location Problem

The MCLP has since then been studied and applied to a number of different scenarios in the ambulance location problem, some[15] only applied it to a concrete case. However, for the purposes of this work, we will only reference those who contributed with different approaches to solve the model, as did Galvão and ReVelle[16] when they proposed a LaGrangean heuristic approach to solve the MCLP or the addition of a second type of vehicle to the problem by Schilling, Elzinga, Cohon, Church, and ReVelle[30]. Both these works had the MCLP as a base model.

2.2.2 Deterministic multiple coverage Models

The core difference in relation to Deterministic single coverage models is that in Deterministic multiple coverage models there is the possibility that no vehicles are available to answer a call, i.e., all the vehicles are alocated to emergency calls and a new call is received. Single coverage models might not be able to deal with these cases which often occur in real-life situations. In order to solve this problem, multiple coverage problems seek to have a smaller likelihood that a zone is not being covered by any vehicles by increasing the number of vehicles available to cover said zone. This is an improvement over single coverage problems as it accounts for the natural randomness of emergency demands by changing vehicle availability.

In 1981, Daskin and Stern[12] proposed the hierarchical objective set covering problem (HOSC), a model that minimizes the amount of vehicles needed to provide complete coverage and maximizes the amount of vehicles that can cover a zone. However, this can cause unwanted effects since each additional vehicle has the same effect on the objective function. Additionally the HOSC is prone to leaving harder to cover zones covered by only one vehicle. This happens due to the fact that the HOSC does not take into account the demand in each zone. Later, the regrouping problem stated previously was tackled by Eaton, Sanchez, Lantigua, and Morgan[14] by considering each zone's demands. Further drawbacks of the HOSC were assessed by Hogan and ReVelle[21] who created two additional models (BACOP1 and BACOP2) that seek to maximize the demand covered twice, instead of just once, given a number of vehicles to locate.

As a consequence of the integration of the double coverage model and different coverage radii, the double standard model (DSM) is proposed by Gendreau, Laporte, and Semet[17]. This new model aims to set the location of a given number of vehicles in an attempt to maximize the demand covered twice within a time frame S. This model was then extended by a variety of other authors that integrated different concepts to the main DSM algorithm or used the very same algorithm for different purposes, most notably when Storbeck[31] proposed the maximal-multiple location covering problem (MMLCP), a goal programming formulation that aims to locate a given number of vehicles with the objectives of minimizing the demand that is left uncovered and maximizing the number of demand zones covered by more than one vehicle.

2.2.3 Probabilistic and Stochastic Models

As previously described, deterministic multiple coverage models came as a significant improvement over deterministic single coverage models. However, even though they brought improvement and solved some of the problems of single coverage models, it was still hard to ensure a satisfying level of service. Furthermore, the double coverage model excels at congested environments, which refers to very active city centers, in this context, meaning that when applied to non-congested examples, it did not provide such a big edge over the single coverage models. It is at this point that some authors considered introducing sources of uncertainty, hence leading to several probabilistic and stochastic models being proposed.One of the most important models, the maximum expected covering location problem (MEXCLP), aims to locate a given number of vehicles in a way that maximizes the expected coverage, while taking into account that there is a probability that a vehicle might be unavailable to respond to an emergency call. This probability is called the busy fraction. The MEXCLP and its variants[8, 13] have three main assumptions at their core: the busy fraction is given and equal for all vehicles, the busy fraction is independent of the location of the vehicle and each vehicle operates independently. A drawback in MEXCLP is that these conditions are not generally met in real-world situations. In order to tackle this issue, Batta[4] proposed two variants on the original MEXCLP which allowed for some relaxation on the assumptions stated above.

So far, the models previously described assume deterministic or static travel times. However, in real situations this is very often not the case due to a number of factors that can modify the travel time, like traffic, for example. Daskin[11] proposed the first model in which not only the location of the vehicles and respective demand zones were set to achieve the maximum expected coverage, but also the route each vehicle should take was taken into account, considering these are not deterministic.

2.3 Dynamic Relocation Models

The set of models discussed so far all assume that after completing an assignment, each vehicle returns to the base that dispatched it. However, there might be cases where it is beneficial that the vehicles go to a different location than the one they were stationed at. From this possibility, the relocation problem arose, and it consists in relocating vehicles that are not responding to any emergencies to more a location where it has a better impact in the operation of the EMS. In this particular section, only the relocation models that consider the state of the system at the exact moment the decisions must be made. This problem can be analysed in several ways, in the next few sections, namely: real-time relocation models, compliance table policies and the use of dynamic programming as an algorithmic solving techinque.

It is important to keep in mind that even though relocation may improve performance, it also leads to additional ambulance movements that can be costly. Therefore, from a practical standpoint, dynamic relocations are generally only acceptable within certain limits.

2.3.1 Real-time relocation models

Real-time relocation models (also called online relocation models) occur when a dynamic relocation takes place at the exact moment a decision must be taken. A solution or an approximate solution must be calculated every time a decision needs to be taken. Gendreau, Laport and Semet[18] proposed the first relocation model that explicitly takes into account the dynamic nature of EMS, and this model is based on the DSM proposed by the same authors, which was referred earlier in this paper. It aims to maximize the demand covered by at least two vehicles within a time frame and minimize the relocation costs. As a consequence, the objective functions contains a penalty term that tracks the relocation history of the vehicles. This serves the purpose of avoiding round-trip or excessively long relocations as well as moving the same vehicles repeatedly. The ambulance relocation problem (RP^t) formulation uses three Boolean variable, namely: u_i that takes the value 1 if the demand zone i is covered at least twice, y_i that takes the value 1 if the demand zone i is covered at least once and x_{jk} equal to 1 if the vehicle k is in location j. It also contains a penalty term M^t_{jk} related to the action of relocating a vehicle k to a new location j at time t. Hence, the (RP^t) is formulated as shown in Fig.2.3:

As stated before, the RP^t is intended to be solved every time a decision needs to be made, which corresponds to the times when a vehicle is dispatched to an emergency call. These calculations, however, yield great computational time needed to solve, and this might be too long for the system to rely on every time a vehicle needs to be dispatched. In order to solve this problem the same tabu search heuristic used by Gendreau[17] was proposed in combination with the use of parallel computing in order to reduce the amount of time it takes to solve the different relocation problems. This methodology was successfully applied to real data from Montreal, Canada.

The models described so far used the coverage measure as the way to assess the performance of the system, Andersson and Värbrand[3] decided to diverge from this measure and introduced a real-time ambulance relocation model that uses the notion of preparedness instead. This measure is defined as

$$\max \sum_{i=1}^{n} a_i u_i - \sum_{j=1}^{m}\sum_{k=1}^{p} M_{jk}^t x_{jk}, \tag{2.10}$$

$$\text{subject to: } \sum_{j \in M_i'} \sum_{k=1}^{p} x_{jk} \geq 1, \qquad i = 1, ..., n, \tag{2.11}$$

$$\sum_{i=1}^{n} a_i y_i \geq \alpha \sum_{i=1}^{n} d_i, \tag{2.12}$$

$$\sum_{j \in M_i} \sum_{k=1}^{p} x_{jk} \geq y_i + u_i, \qquad i = 1, ..., n, \tag{2.13}$$

$$y_i \geq u_i, \qquad i = 1, ..., n, \tag{2.14}$$

$$\sum_{j=1}^{m} x_{jk} = 1, \qquad k = 1, ..., p, \tag{2.15}$$

$$\sum_{k=1}^{p} x_{jk} \leq p_j, \qquad j = 1, ..., m, \tag{2.16}$$

$$y_i, u_i \in \{0, 1\}, \qquad i = 1, ..., n, \tag{2.17}$$

$$x_{jk} \in \{0, 1\}, \qquad j = 1, ..., m, k = 1, ..., p. \tag{2.18}$$

Figure 2.3: Definition of the RP^t

$$\varrho_i = \frac{1}{a_i} \sum_{k=1}^{K_i} \frac{\gamma^k}{t_i^k} \tag{2.19}$$

Figure 2.4: Mathematical definition of preparedness

the capacity of a system to answer future demands. The preparedness of a zone i has a weight, a_i, that mirrors the demand for ambulances in the zone, a given number of vehicles, K_i, that will be used in the computation of the preparedness of the zone, the travel time of a vehicle k to a zone i, t_i^k, and γ^k, a contribution factor of each of the vehicles considered. Considering all of these parameters, the preparedness of a zone, ϱ_i, is defined as seen in Fig.2.4:

When applied, the level of preparedness for each demand zone is verified regularly, and a relocation takes place whenever the level of preparedness drops below a predetermined value. To determine what the relocation plans are, the authors proposed the DYNAROC model. This model aims to minimize the maximal travel time required to perform a relocation. Similar to the RP^t, DYNAROC has a set of constraints to limit the travel times and the number of relocated vehicles. Additionally, it aims to achieve a minimum level of preparedness for each demand zone, ϱ_{min}. The DYNAROC formulation contains a Boolean variable, x_i^k, that takes the value of 1 if and only if the vehicle k is relocated to a standby site in zone i, a set of zones N_k that can be reached by a vehicle k within a predetermined time frame S, and finally, a parameter, P_{max}, that denotes the maximum number of relocated vehicles allowed. Having said this, the DYNAROC is formulated as seen in Fig.2.5.

In order to solve the DYNAROC model, the authors proposed a tree-search heuristic and tested this solution using data from Stockholm, Sweden.

$$
\begin{aligned}
\text{min} \quad & z && && (2.20)\\
\text{subject to:} \quad & z \geq \sum_{i \in N_k} t_i^k x_i^k, && k = 1, ..., P, && (2.21)\\
& \sum_{i \in N_k} x_i^k \leq 1, && k = 1, ..., P, && (2.22)\\
& \sum_{k=1}^{P} \sum_{i \in N_k} x_i^k \leq P_{max}, && && (2.23)\\
& \frac{1}{a_i} \sum_{l=1}^{K_i} \frac{\gamma^k}{t_i^l(x_1^1, ..., x_N^P)} \geq \varrho_{max}, && i = 1, ..., n, && (2.24)\\
& x_j^k \in \{0, 1\}, && i = 1, ..., n, k = 1, ..., P. && (2.25)
\end{aligned}
$$

Figure 2.5: Definition of the DYNAROC model

Mason[26] follows the work of Gendreau[18] to determine the location of available vehicles in a way that maximizes the service quality and minimized relocation costs. Therefore, they propose the real-time multi-view generalized cover repositioning model (RtMvGcRM) as a way to address the dynamic ambulance relocation problem. This model was implemented using Optima Live, an EMS management software. The objective function in this model is based on a general concave piecewise linear function that specifies the reward attributed to each demand zone with relation to the number of vehicles covering it. The objective funtion also handles several types of vehicles with varying performances, effectively expanding the performance measures that Gendreau[18] introduced, even if still using the concept of coverage.

Naoum-Sawaya and Elhedhli[29] proposed a two stage stochastic programming approach to address the dynamic relocation problem. This approach aims to minimize the cost related to vehicles' relocation and the cost associated to demands that cannot be served in the prescribed delay. The first stage of this process concerns decisions on the initial location of the vehicles and the goal here is to minimize the number of future relocations. A set of scenarios, estabilished based on data from Waterloo, Canada, represent the uncertainty in this first stage. The second stage concerns the assignment of vehicles to emergency demands once the emergency calls arrive and, additionally, it identifies the emergencies that cannot be reached within a known time frame. This problem, applied to the same data referred earlier, was solved using the CPLEX, with relatively short computation times.

Focusing on a different aspect within the dynamic ambulance relocation problem, Jagtenberg, Bhulai, and van der Mei[23] proposed a dynamic version of the MEXCLP with the goal of minimizing the expected fraction of late arrivals, meaning, the emergency calls for which the maximum allowed response time is exceeded. Additionally, this work featured an important change relative to other works in this area, the vehicles can only be relocated at the end of a mission. The results obtained from the application of this model to data from Utrecht, Netherlands, showed that the fraction of late arrivals was, in fact, lower, when compared to a static policy featuring no relocations.

$$\max \sum_{k=1}^{P} \sum_{i=1}^{n} a_i q_k y_{ik}, \tag{2.26}$$

$$\text{subject to: } \sum_{j \in M_i} x_{jk} \geq y_{ik}, \qquad i = 1, ..., n, k = 0, ..., P, \tag{2.27}$$

$$\sum_{j=1}^{m} x_{jk} = k, \qquad k = 1, ..., P, \tag{2.28}$$

$$x_{jk} - x_{j,k+1} \leq u_{jk}, \tag{2.29}$$

$$\sum_{j=1}^{m} u_{jk} \leq \alpha_k, \qquad k = 1, ..., P-1, \tag{2.30}$$

$$x_{jk} \in \{0, 1\}, u_{jk} \in \{0, 1\}, \qquad j = 1, ..., m, k = 1, ..., P, \tag{2.31}$$

$$y_{ik} \in \{0, 1\}, \qquad i = 1, ..., n. \tag{2.32}$$

Figure 2.6: Definition of the Maximal Expected Relocation Problem

2.3.2 Compliance table policies

In the previous section the relocation models needed to be solved in real-time multiple times during the day. This poses a dimensionality problem in which the larger the problem at hand is, the stronger the computational power needs in order to solve it in a timely manner. This is where the need for relocation plans or compliance tables comes in. These tables represent sets of scenarios applied to the system that have been solved a priori, in other words, they represent relocation plans for the vehicles in each situation. Taking this offline approach proves much easier to implement in real-life situation as it comes closer to the actual EMS practices. Nonetheless, the number of potential states that need to be solved can also be significantly large. Gendreau, Laporte, and Semet[20] proposed the maximal expected relocation problem (MECRP), which is one of the first dynamic offline relocation model used to create a set of relocation plans. This model aims to determine the appropriate location plan for each possible state of the system, which depends on the number of available vehicles. The authors also introduced a constraint on the number of vehicles that can be relocated between states. The model has a Boolean variable x_{jk} that takes the value 1 if, and only if a vehicle is located at j when the system is in state k, a Boolean variable y_{ik} that takes the value 1 if, and only if a demand zone i is covered by at least one vehicle when the system is in a state k, a Boolean variable u_{jk} that takes the value 1 if, and only if a location j is no longer used when the system goes from state k to state $k+1$ and finally q_k, a probability of reaching state $k, k = 0, ..., P$ where P is the total number of vehicles. The MECRP is formulated as seen in Fig.2.6. In this model, constraint 2.27 mean that a demand zone is covered only if a least one vehicle is located in particular site, and constraint 2.28 control the number of vehicles used in the solution. Constraints 2.29 and 2.30 control the number of waiting site changes when the system changes states.

These constraints are similar to those formulated in the MCLP. Nevertheless, other constraints need to be added to control the number of vehicles relocated between the states. This model is solved once, a priori, and the compliance table referring to each state is applied when needed. In order to solve this model the authors used CPLEX and validated the solution with data from Montreal, Canada.

van Barneveld[34] suggested a model that combines both the MECRP and the MEXCLP in order to account for vehicle unnavailability. This model is referred to as the minimum expected penalty relocation problem (MEXPREP). It features a non-decreasing penalty function that depends on the response time that in the objective function, allowing the consideration of several performance measures. After the initial step of calculating the compliance tables for this model, an online assignment model is solved in order to find the best ambulance movements to reach compliance. This application outperformed both the static policy with no relocation as well as the solution provided by the MECRP, on most performance measures. The same authors later proposed an extension of this model that accounted for two different types of vehicles and a bound on the time needed to perform relocation between states.

Nair and Miller-Hooks[28] proposed a relocation model similar to the one proposed by Gendreau[19] to position the vehicles. This model considers the evolution of the system's state overtime. However, unlike Gendreau[19], the states are defined by the incoming call probability distributions, the number of available vehicles and the travel time between locations, as opposed to being defined only by the number of vehicles available. This model aims to maximize the double coverage and also to minimize relocation costs. The impact that relocation has is therefore incorporated in the objective function instead of being a set of constraints. The model is solved a priori to establish a compliance table for each possible state, just like in the MECRP. Application of the model to data from Montreal, Canada, achieved improvements depending on the number of vehicles.

With the same intent to maximize the expected coverage but also to limit the number of relocation between states, Sudtachat, Mayorga, and McLay [32] proposed a dynamic relocation model that used a set of nested compliance tables where only one ambulance is moved whenever the system goes from one state to another. This model has foundations on other works, like Batta[4], that proposed an extended version of the MEXCLP where servers are not independent and might have different busy probabilities and Alanis, Ingolfsson, and Kolfal[2] which used Markov Chains to model a EMS system. Inspired by these two works to compute and analyze the performance of a fixed compliance table policy, the model uses an adapted version of the Markov Chains to approximate the steady-state probabilities of the system for each state and then incorporates these values in the proposed nested compliance table model. This model is formulated as an integer programming model that maximizes the expected coverage for each state, defined as the number of busy vehicles and the state of the system with respect to compliance. It also ensures that, for each possible state, available vehicles are located and that the coverage is accounted for. When applied to data from Virginia, USA, it showed improvements when compared to the static case where no relocation is allowed.

2.3.3 Approximate Dynamic Programming approaches

Using dynamic programming enables a proper understanding of the random evolution of the system through time, which, in this EMS context, is extremely relevant. One downside is that dynamic programming is usually limited to small problems, which is rarely the case in the EMS context. This is where

$$J(s) = \min_{x \in X(s)} \{E[c(s, x, f(s, x, w(s, x))) + \alpha^{\tau(f(s,x,w(s,x))) - \tau(s)} J(f(s, x, w(s, x)))]\} \quad (2.33)$$

Figure 2.7: Optimality equation

approximate dynamic programming (ADP), which is a method that requires a preparatory tuning process that can be computationally expensive to then be able to operate in real-time situations, comes in handy.

Maxwell, Restepo, Henderson, and Topaloglu[27] were the first to apply ADP to the dynamic ambulance relocation problem. In their definition of the problem, relocations can only occur in vehicles that just completed their mission. This was an attempt to reduce the inconveniences on a practical, human resources' standpoint, at the cost of a reduction of the number of possible decisions, from a mathematical standpoint. Upon mission completion, a vehicle can be relocated to another site, the problem lies in determining what this site will be such that the number of high-priority calls that can be reached within a given time frame is maximized. The model assumes a level methodology for call queuing where the calls with the most priority are answered first and within the same level, a first-in, first-out policy is followed. Calls that cannot be served are queued and the nearest ambulance is always dispatched to a call. The optimality equation that defines this model aims to minimize the discounted total expected cost given an initial state, where the cost is defined as the number of high-priority calls that cannot be reached in a timely manner. The model also accounts for the systems' state, s, which is defined according to the current time and event. The state of each vehicle is defined in a vector A, the call state is described in a vector C. Additionally, the model uses $X(s)$, the set of all feasible decisions given a state s, a transition cost, $c(s_k, x_k, s_{k+1})$, from state s_k to state s_{k+1} given a decision x_k, the transfer function, $f(s, x, w(s, x))$ that depends on the system state, the decision made and random elements w(s,x), a fixed discount factor, $\alpha \in [0, 1[$, and the time at which the system visits state s, $\tau(s)$. The optimality equation expressed in Fig.2.7 can be used to compute the value function to determine the policy that minimized the discounted total expected cost given and initial state s.

In this case, the set $X(s)$ is relatively small, since only one vehicle is eligible for relocation. The number of emergency locations that will not be reached within the time frame are calculated based on the transition costs, which take the value 1 if the next event is of the form "ambulance i arrives at scene of call j" and the corresponding call is urgent and the time frame is exceeded, and 0 otherwise. This value function evaluation is a challenging task nonetheless, and the high dimensionality of the state variable makes it so that classic dynamic programming cannot be applied directly.

The challenge of using approximate dynamic programming consists of selecting the right values for the parameters needed to determine an adequate approximation of the value function. When such an approximation is found, it is possible to identify the optimal policy by enumerating each possible decision and evaluating the corresponding expected value, in this work, the Monte Carlo simulation was used. The results obtained from this study showed an improvement over the static policy, where no relocation is allowed. It was also shown that the performance can be enhanced by considering more frequent locations and involving more vehicles in the relocation process, but this would significantly increase computational times.

2.4 Analysis Works

The works mentioned thus far describe a summarized state of the art of the EMS location problem. There have been various ways of formulating and solving the problem throughout the years. However, for simplicity, we selected and analyzed those which we deemed more important in order the understand how the state of the art has evolved. Table 2.1 contains a collection of papers described in this work classified by a set of relevant aspects. This table serves as both a summary of Chapter 2 and a quick and comprehensive mapping of what papers cover which aspects within the EMS location problem context.

There have also been studies and reviews on the effectiveness, advantages and disadvantages of several approaches and techniques used to try to solve the EMS ambulance location or relocation problem. These studies are particularly helpful as an introduction to the matter at hand because they are often very generic in their descriptions in order to fit most models that are referenced in it and also because they allow the reader to become aware of which papers address what specific topics before actually reading them, effectively serving as a filtering method for whoever needs to search for works in this area of study[5, 1].

Paper	Objective	Solution Strategy	Model Type
Toregas et al.(1971)	Min. Number of Vehicles	Branch and Cut	Deterministic
Church and ReVelle (1974)	Max. demand covered once	Branch and Bound & Greedy Heuristic	Deterministic
Schilling et al. (1979)	Max. demand covered once	Branch and Cut	Deterministic
Daskin and Stern (1981)	Min. number of vehicles & Max. demand covered more than once	Branch and Cut	Deterministic
Storbeck (1982)	Max. demand covered once & Max. demand covered more than once	Not presented	Deterministic
Eaton et al. (1985)	Max. demand covered once	NaN	Deterministic
Hogan and ReVelle (1986)	Max. demand covered once & Max. demand covered twice	Branch and Bound	Deterministic
Galvão and ReVelle (1996)	Max. demand covered once	Lagrangean Heuristic	Deterministic
Gendreau et al. (1997)	Max. demand covered twice	Tabu Search	Deterministic
Bianchi and Church (1988)	Min. uncovered demand	Branch and Bound & Heuristic Method	Probabilistic, Expected Coverage
Batta et al. (1989)	Max. expected covered demand	Heuristic Method & Descent Method	Probabilistic, Expected Coverage
Galvão et al. (2005)	Max. expected covered demand & Max. demand covered with reliability α	Simulated Annealing	Probabilistic, Expected Coverage, Chance-Constrained
Gendreau et al. (2006)	Max. expected covered demand	General-purpose Solver	Dynamic
Naoum-Sawaya and Elhedhli (2013)	Min. number of relocated vehicles & Min. number of calls that cannot be reached in time	General-purpose Solver	Dynamic
Jagtenberg et al. (2015)	Min. number of calls that cannot be reached in time	Heuristic Method & Discrete Event Simulation	Dynamic
van Barneveld (2017)	Max. expected covered demand	General-purpose Solver	Dynamic
Sudtachat et al. (2016)	Max. expected covered demand	General-purpose solver & Discrete Event Simulation	Dynamic
Schmid (2012)	Min. average response time	Approximate dynamic programming	Dynamic

Table 2.1: Classification of several papers.

Chapter 3

Preliminaries

All of the data used in this paper has been retrieved from a database containing detailed information about the medical emergencies in Portugal. This database was provided to us by the INEM (Instituto Nacional de Emergência Médica). In this chapter, a brief description is made about the work that has been done after retrieving the data about emergencies and vehicles from the INEM database as well as insights on the process of deducting further information from the data we were able to retrieve.

3.1 Data Analysis

After cleaning, organising and getting the information we wanted from the INEM database, we needed to draw conclusions that could be used to model our solution for the Ambulance Location problem. Objectively, we were aiming at getting the values of the minimum dispatch time for each individual vehicle and also a measure for the radius of the action zone for each emergency station, which would come from the analysis of the average distance to an emergency.

Dispatch Time Measure

For the purposes of defining an average dispatch time that we could use to model our solution we needed to see the disparities between highly populated districts and districts with less population density, therefore we divided all the eighteen districts in two groups, the first one containing Lisboa and Porto and the other group containing the remaining 16 districts and we noticed that the distribution of dispatch times for the same vehicles was much more elongated in the less populated districts, for example Beja, Évora and Viana do Castelo 3.2 as opposed to a much more steep descent in the distribution in Lisboa and Porto 3.1. In both of these groups of graphs we can see these two similar distributions, with the previously mentioned more elongated descent in the less populated districts of Beja, Évora and Viana do Castelo. It is also important to note that Lisboa and Porto also have significantly more occurences depicted in these graphs as opposed the the other three districts.

We noticed that for the highly populated districts, the most common time between dispatches was between one and two hours, with a steep descent in occurrences of higher values between each dispatch.

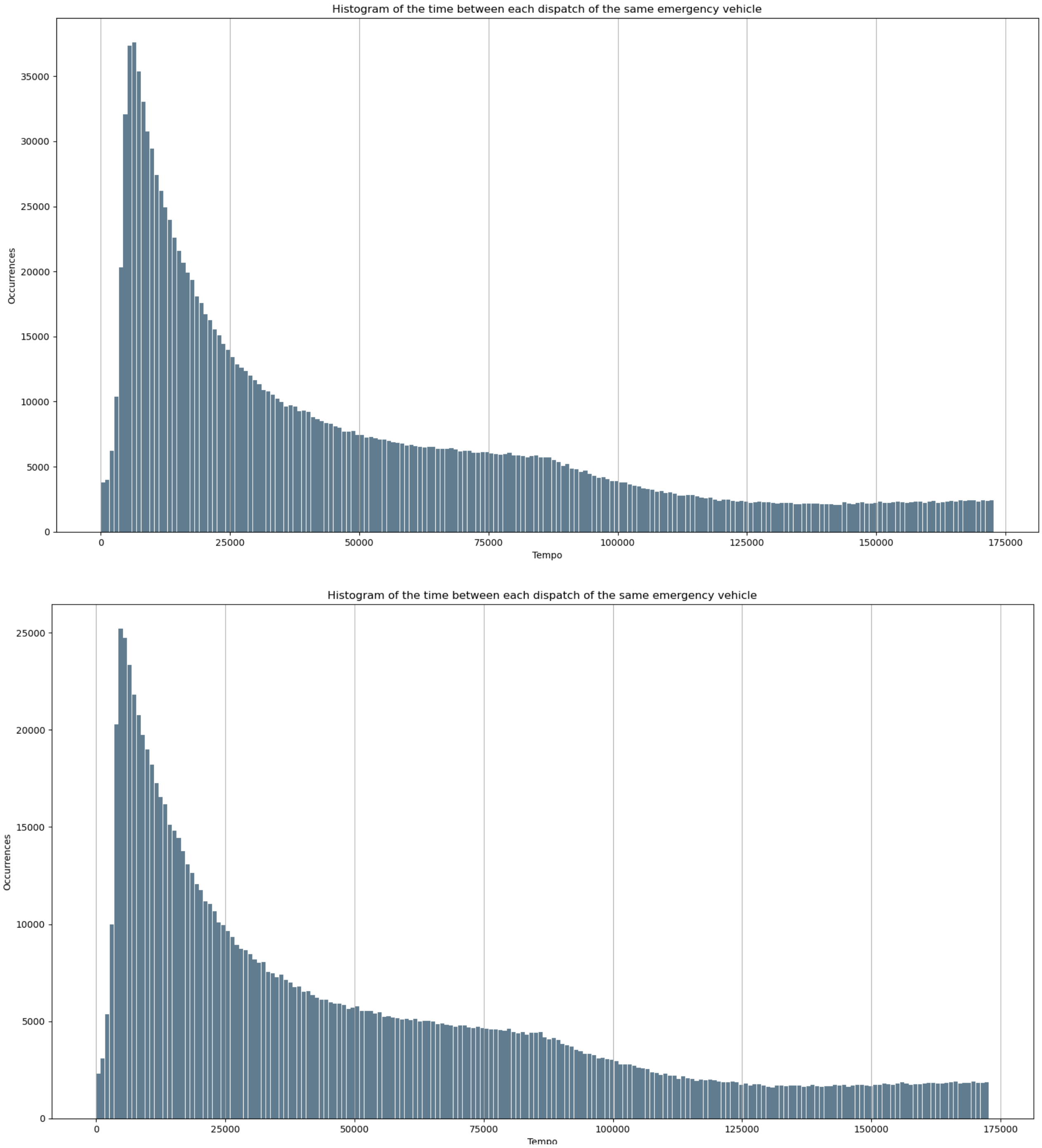

Figure 3.1: Histograms of the time between each dispatch of the same vehicle in Lisboa and Porto (More populated districts

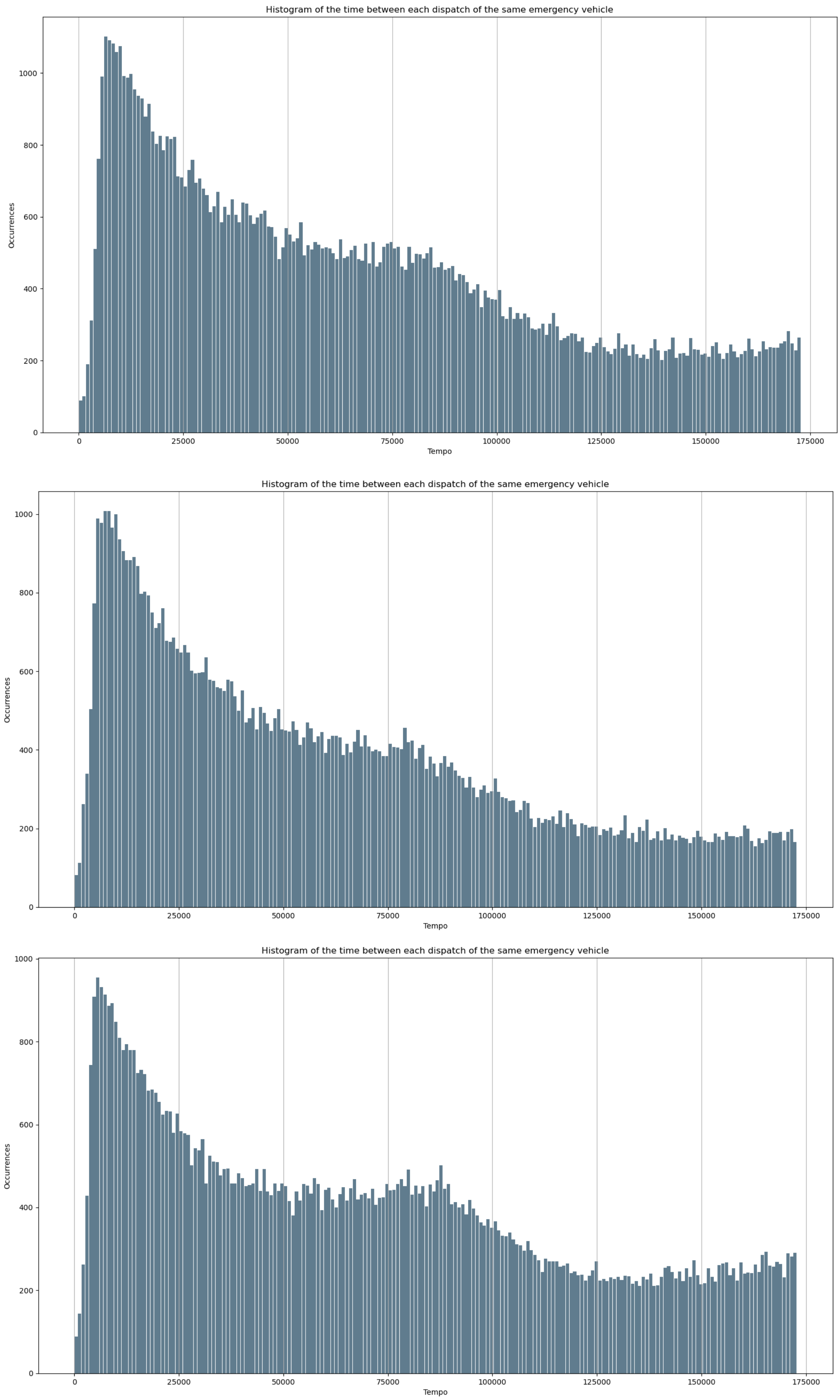

Figure 3.2: Histograms of the time between each dispatch of the same vehicle in Beja, Évora and Viana do Castelo, all three lesser populated districts than Lisboa and Porto

In the less populated districts the most common values had the same distribution with a much less steep descent, meaning that there are more vehicles that take more than two hours to be dispatched again. However, the most common values still oscillate around one to two hours in most districts. This difference between results might be a consequence of a higher density of emergencies in the more populated districts due to a sheer difference in population number. However, it can also be the case that in the less populated districts there is a higher rate of emergency vehicles to number of emergencies, indicating a less than optimal distribution of resources.

3.2 Data Preparation

The database used, courtesy of the Instituto Nacional de Estatistica (INEM) has a variety of different fields that track a number of different variables regarding each emergency, vehicle, station and possibly other fields that were not used for the purposes of this study.

From an analytical point of view, we wanted to retrieve two important measures: the average distance that each vehicle traveled to get to an emergency site and the average time between each dispatch of the same vehicle. The first problem we came across had to do with missing values and incorrectly formatted records, and this problem occurred in the process of reviewing both measures. Since there is no way to infer or calculate the missing values, we discarded any records that were either empty or had out of context information.

Distance Measure

This initial data cleaning allowed us to perform the first proper analysis of the data we were working with and start to draw conclusions from the results of that same analysis. However, in the average distance measure, we found values that were inconsistent with what we were expecting. More specifically, and given the fact that we were analysing data from continental Portugal, in which the longest straight line that connects two parts of the country is just below six hundred (600) kilometers long, we found records that read distances over four thousand (4000) kilometers. After analysing these records, we discovered that they all the same emergency coordinates corresponding to the point at zero degrees latitude and zero degrees longitude (0°N 0°E), as seen in Fig.3.3. Because of this we decided to not take into account the records that used this point as their emergency site.

Even though we had dealt with these erroneous records, we still had a number of emergencies occurring at more than fifteen hundred kilometers (1500) which were still out of context. After further investigation, we found out that these records referred to marine emergencies in the Atlantic Ocean. These records still appear on this database because the Portuguese territorial sea area is about fifty one thousand squared kilometers (50957), sixteen thousand (16460) of those belong to the continental portion of Portugal and the rest belonging to the Açores and Madeira archipelagos, as seen in Fig.3.4. Since we are only dealing with terrestrial emergencies in this study, we imposed a coordinates limit that ensured that the emergency sites were inside continental Portugal.

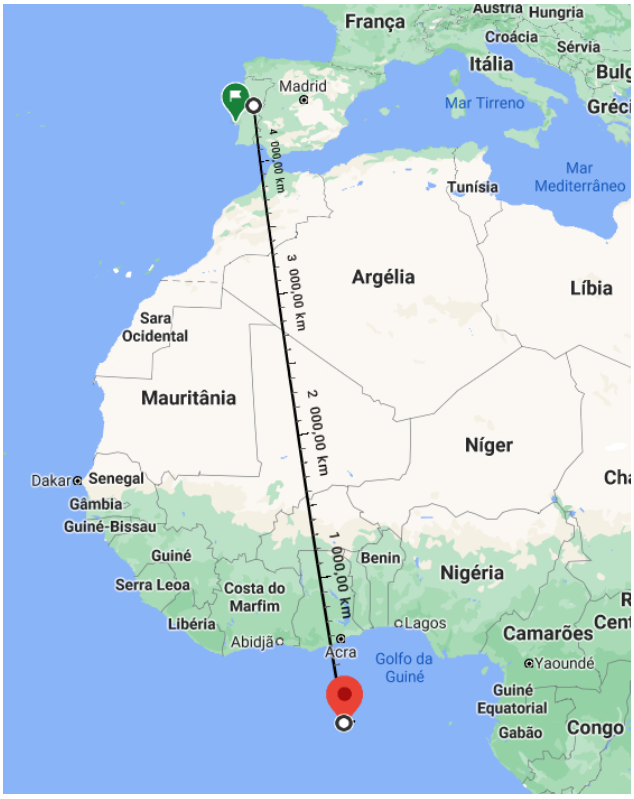

Figure 3.3: Example of emergency being recorded at position (0°N 0°E)

After cleaning all the data, we ended up with the complete graph of all the distances covered by each vehicle when providing aid to an emergency, as seen in Fig.3.5

Dispatch Time Measure

The preparations for calculating the average dispatch time measure were not as extensive as the ones for the average distance. In order to get the records for this measure we first separated and organised the records according to the vehicle identifier and we arranged these same records in a chronological order, which made it trivial to calculate the consequent dispatch times and calculate the average dispatch time for all the vehicles. After getting this initial value, we tried to get more detailed data, namely isolating each one of the eighteen districts in Portugal and after that getting data by vehicle type and emergency priority levels.

Figure 3.4: Example of emergency in the Madeira Archipelago

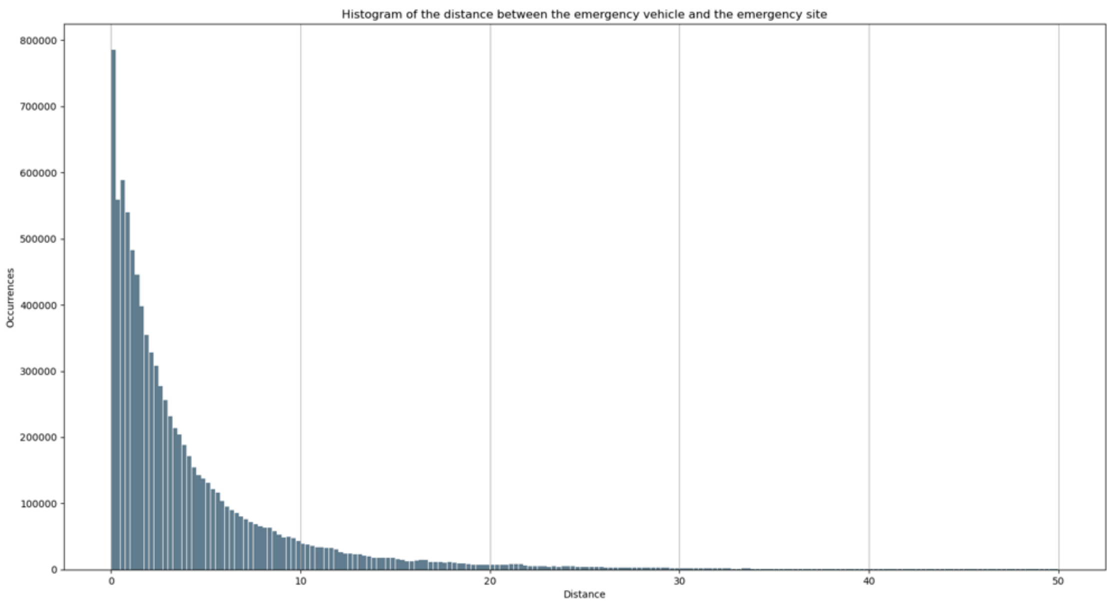

Figure 3.5: Distance Covered by each vehicle when providing aid to emergencies

Chapter 4

Ambulance Assignment

In this section we will go over the several models we developed at different stages of the elaboration of this document, explaining our thought process behind each change we made along the way, as well as providing insight as to what implications those changes had in the way we treated data and had to adapt our process.

After the initial analysis and preparation of the data retrieved from the INEM database, the first step was to create a very simple, traceable model with which we could check for implementation mistakes as well as start testing very small examples in the solver. These first examples were handmade and not derived from data. Our main goals with this approach were testing specific cases and how the solver would react to them while also optimizing the generation of an instance that represents each problem mathematically. As a result, we had to make several changes to both the way the data was being inserted in the programmatic generation file, as well as the generation process itself, since it was not accounting for a number of cases we had not initially considered, for example, the case where two vehicles were being used in subsequent time periods and therefore were unavailable when they were needed in these limited models where we did not have more vehicles than emergencies, causing the solver to deem this example as unsatisfiable.

4.1 Initial Model

The ambulance location model can be defined as a graph $G = (V, E)$ where $V = N \cup M$, $N = (v_1, ..., v_n)$ and $M = v_{n+1}, ..., v_{n+m}$ being two vertex sets representing, respectfully, emergency sites and standby vehicles with and associated position. Additionally, E is an edge set where each edge $\{(v_i, v_j) : v_i, v_j \in V, i < j\}$ is associated with a travel time or distance t_{ij}. The variable x_{ij} will take the value 1 if and only if, the the vehicle j provides aid to the emergency site i, and the value 0 otherwise.

For simplicity purposes, we designed a simple example with a given set of emergency sites, $N = A, B, C, D$, as well as a set of standby vehicles $M = 1, 2, 3$ spread randomly on an example map, represented in Fig. 4.1. This map does not represent any specific zone, it merely serves as an example.

In this map, we implemented a baseline model. It is meant to be a blueprint for Static Models, where

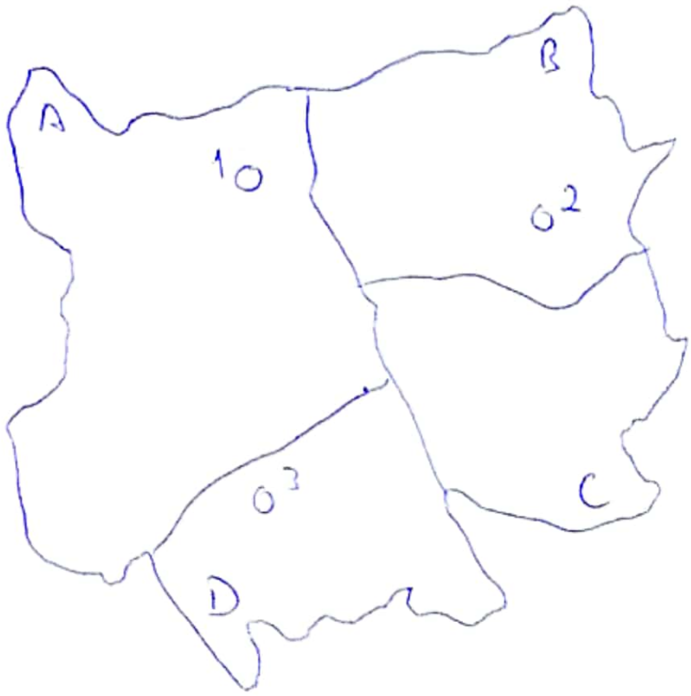

Figure 4.1: Example Map

we only considered edges that represent connections between emergency sites and standby vehicles. The simplified graph for this implementation of a static model is represented in Fig. 4.2.

The objective this type of models aims to achieve is that of minimizing the number of vehicles and the overall travel time in a way that all emergencies are accounted for. For simplicity, we will consider that a vehicle is unavailable for a period of time correspondent to $2t_{ij}$ in these static models. These time periods represent the minimum amount of time that a vehicle takes to answer an emergency call and go back to a standby site.

At this stage we can already say that we will have to impose thresholds on both the values we want to minimize due to the fact that they represent conflicting objectives, meaning that we have to find a balance between the two.

In order to better understand each goal, we are first going to formulate two separate model definitions that represent single objective models. In the first model, represented in Fig. 4.3, the goal is to minimize the travel time throughout all operations within a time frame. For simplicity, we consider only the parcel t_{ij} associated with each variable x_{ij}.

After this, we defined a model in which the goal was to minimize the number of vehicles being used. In order to do that we will define a variable y_j which will take the value of 1 if and only if a vehicle j is used to provide aid to an emergency. A reduction in the number of vehicles used can mean that that area in particular has a more vehicles than it should, and these can be relocated to areas that are struggling with more emergencies. Using this information, we can also cut our vehicle fleet if we realise that there are vehicles that are never used in the long term. This model definition is represented in Fig.4.4.

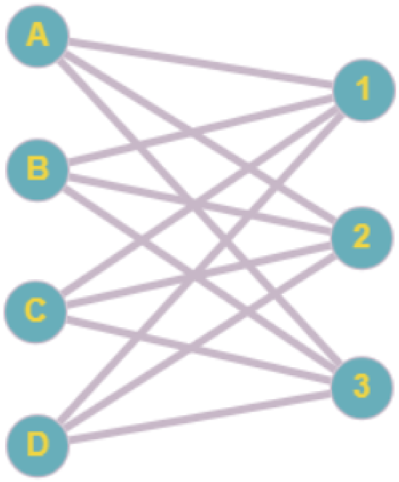

Figure 4.2: Static Model Simplified Graph

$$\min \sum_{j=1}^{M} x_{ij} t_{ij} \tag{4.1}$$

$$\text{subject to: } \sum x_{ij} \geq 1, \qquad i = 1, ..., n, \tag{4.2}$$

$$x_{ij} \in \{0, 1\}, \qquad i = 1, ..., n, j = 1, ..., m. \tag{4.3}$$

Figure 4.3: Distance covered minimizing Model definition

$$\min \sum_{j=1}^{M} y_j \tag{4.4}$$

$$\text{subject to: } \sum y_j \geq n, \qquad j = 1, ..., m, \tag{4.5}$$

$$y_j \in \{0, 1\}, \qquad j = 1, ..., m. \tag{4.6}$$

Figure 4.4: Vehicle usage minimizing Model definition

$$\min \sum_i \sum_j x_{ij} t_{ij} \tag{4.7}$$

$$\min \sum_{j=1}^{M} y_j \tag{4.8}$$

$$\text{subject to:} \sum x_{ij} \geq n, \qquad i = 1, ..., n, j = 1, ..., m, \tag{4.9}$$

$$\sum y_j \leq 1, \qquad j = 1, ..., m, \tag{4.10}$$

$$x_{ij} \in \{0, 1\}, \qquad i = 1, ..., n, j = 1, ..., m. \tag{4.11}$$

$$y_j \in \{0, 1\}, \qquad j = 1, ..., m. \tag{4.12}$$

Figure 4.5: Double Objective Model definition

Since these two objectives are conflicting, meaning that a bigger number of vehicles used will result in a smaller value for the distance covered and vice-versa. Our goal is then to find a Pareto front, or a set of optimal solutions that minimizes both the number of vehicles used and the distance covered by those vehicles.

After both these models are defined, the Multi-Objective problem definition consists of joining both the single objective formulations into a single one, represented in Fig. 4.5.

4.2 Different Time Periods and Vehicles Unavailability

With the initial model we were able to test some examples for small time periods where each vehicle could not be called to two distinct emergencies. However, in a real scenario, it is important that the vehicle can be dispatched to another emergency as soon as it is available again. In order to apply this concept to our model we divided a shift sized time frame into smaller fractions. This alteration meant that in our model we had to alter our set of constraints so that there was a variable for each vehicle for every one of the smaller time frames, represented by u_{ij}. Additionally, we also added a constraint that made sure that a vehicle could not be used in two consecutive time frames, effectively giving each vehicle at least one small time frame where it is unavailable between each dispatch. These time frames might differ in size depending on the situation and the value has to come from an analysis to the different time between dispatches in several similar real case scenarios, as well as the data analysis described in Chapter 3.

In order to mathematically model this addition, we divided our large time period into a group of smaller time periods $E = E_1, ..., E_k$ where k represents the index of the period of time within a certain group E, which contains e time periods. This index k also serves the purpose of identifying a vehicle assignment u_{ijk}, which is a new variable that takes the value 1 if and only if the vehicle j provides aid to the emergency i in the time period k. As for the model itself, we adapted the previous version to account for the addition of time periods and we added a constraint that makes sure each vehicle can only be assign in a certain time period if it has not been assigned in the previous time period as seen in Fig.4.6.

Additionally, because of this alteration, we have also had to make changes to the expression we are

$$\sum u_{ijk} \geq 1, \qquad i = 1, ..., n, j = 1, ...m, k = 1, ..., e \tag{4.13}$$

$$u_{ijk} + u_{ij(k+1)} \leq 1, \qquad i = 1, ...n, j = 1, ...m, k = 1, ..., e \tag{4.14}$$

Figure 4.6: Time Period Constraints

$$\sum u_{ijk} - y_j \geq 0, \qquad i = 1, ...n, j = 1, ...m, k = 1, ..., e \tag{4.15}$$

Figure 4.7: New Time Minimization Expression

trying to minimize that represents the usage of each vehicle. With these changes, if a vehicle is used in any period of time, the corresponding variable y_j should be equal to one. Hence we have added a constraint as seen in Fig.4.7

4.3 Emergency priority and vehicle type

Emergencies are classified by a priority group that is associated with severity. INEM has an eight point priority rating system which can be converted into three major priority groups, due to a very negligible amount of emergencies in some of the categories in the eight point system. Depending on the priority value assigned to each emergency, we wanted to have the more commonly assigned vehicles for each of the three major priority groups be preferred when selecting the vehicle that is going to be dispatched. Therefore, we retrieved the type of each vehicle and the priority of each emergency, and imposed constraints that prevent certain sets of vehicle types to respond to certain emergencies depending on their priority. Mathematically, this means that we will have three sets V_1, V_2, V_3 of vehicle types corresponding to the allowed vehicles to our three major priority groups $p1, p2, p3$, which will contain the priority level of each emergency p_{ij}. When building the model, we will only consider vehicles for a certain emergency if their vehicle type is contained within the set for the specific priority level of that emergency.

As for the model itself, we added a constraint that represents the vehicle exclusivity described above, where only a certain group of vehicles can be assigned to a particular type of emergencies as seen in Fig.4.8.

$$u_{ijk} = 0, \qquad i \notin V_1, j \in E_1, k = 1, ..., e \tag{4.16}$$

$$u_{ijk} = 0, \qquad i \notin V_2, j \in E_2, k = 1, ..., e \tag{4.17}$$

$$u_{ijk} = 0, \qquad i \notin V_3, j \in E_3, k = 1, ..., e \tag{4.18}$$

Figure 4.8: Vehicle Priority Constraint

$$\min \sum_{i} \sum_{j} u_{ijk} t_{ij} \tag{4.19}$$

$$\min \sum_{j=1}^{M} y_j \tag{4.20}$$

$$\text{subject to:} \sum u_{ijk} \geq n, \qquad i = 1, ..., n, j = 1, ..., m, k = 1, ..., e \tag{4.21}$$

$$u_{ijk} + u_{ij'(k+1)} \leq 1, \qquad i = 1, ...n, j = 1, ..., m, j' = 1, ..., m, k = 1, ..., e \tag{4.22}$$

$$\sum u_{ijk} - y_j \geq 0, \qquad i = 1, ...n, j = 1, ..., m, k = 1, ..., e \tag{4.23}$$

$$y_j - u_{ijk} \geq 0, \qquad i = 1, ...n, j = 1, ..., m, k = 1, ..., e \tag{4.24}$$

$$u_{ijk} = 0, \qquad i \notin V_1, j \in E_1, k = 1, ..., e \tag{4.25}$$

$$u_{ijk} = 0, \qquad i \notin V_2, j \in E_2, k = 1, ..., e \tag{4.26}$$

$$u_{ijk} = 0, \qquad i \notin V_3, j \in E_3, k = 1, ..., e \tag{4.27}$$

$$u_{ijk} \in \{0, 1\}, \qquad i = 1, ..., n, j = 1, ..., m. \tag{4.28}$$

$$y_j \in \{0, 1\}, \qquad j = 1, ..., m. \tag{4.29}$$

Figure 4.9: Double Objective Model definition

4.4 Final Model

All these different additions led us to have a model that aims to provide the best assignment for the available vehicles to the emergencies that occur in a determined amount of time, taking into account several relevant details about both the vehicles, namely their vehicle type and their availability as to not overuse a small number of vehicles, as well as the emergencies, namely their level of priority, the distance to each vehicle station, the number of vehicles needed and even taking into account district borders as to not have vehicles from different cities attend to emergencies in other cities if it is not expected.

Even though this model covers a variety of conditions, it does not scale well, meaning that the solver gets exponentially slower at coming up with optimum or even satisfiable results. Because of this, we decided to apply the model to 8 hour shifts, thus dividing the day in 3 equal parts, as a way of allowing the solver to be able to come up with solutions that can be applied in real situations in a reasonable amount of time.

Chapter 5

Experimental Results

Upon having a reliable model built and tested, we started comparing the results of real case scenarios to applications of a solver algorithm to the instances we created of those same scenarios. However, after we analysed a small amount of these scenarios, and consistently getting a Pareto front of solutions that was better than the real case scenario, we wanted to try to infer some additional conclusions as well as come up with some possible solutions to help improve the efficiency of the ambulance assignment process for future situations.

5.1 Benchmarking

Firstly, our main goal was to compare real case scenarios, analyse how many vehicles were used and how much distance they covered, and then compare these results with the Pareto front solutions from the application a solver to the instance we created under the same conditions. From this comparison, we expected to verify if the real case scenarios were sub-optimal, and try to quantify how much the solutions on our Pareto front were performing better than these scenarios. In order to do this we analysed cases with different periods of vehicle inactivity, namely thirthy minutes and one hour and different acceptable distances for a vehicle to provide aid to an emergency, namely ten and fifteen kilometers. We selected four separate days across three different years, namely the first day of February, May, August and December, from 2017, 2018 and 2019. Each of the days was divided in intervals of one, two, four and six hours, giving us a total of 46 instances created per day analysed. Cumulatively, we generated a total of 6624 instances, given that we generated all of the previously mentioned conditions in three different districts, namely Faro, Guarda and Lisboa.

The districts we chose to conduct our analysis are different in population size where Lisboa is the most populated district in Portugal with 2275591 people residing there as of 2021. Faro and Guarda, in the same census from INE, have 467495 and 143019 habitants, respectively.

We also wanted to choose at least one region that had a bigger seasonality than the rest, in this case that region is the district of Faro, which is a common destination for tourists in the summer time. In table 5.1 we see that the month of August is the month that consistently has more emergencies in the

		Faro	Guarda	Lisboa	
2017	01-02	211	74	860	
	01-05	180	50	675	
	01-08	258	63	720	
	01-12	201	45	785	
2018	01-02	216	47	814	
	01-05	176	59	699	
	01-08	232	55	739	
	01-12	184	53	850	
2019	01-02	191	60	907	
	01-05	200	56	724	
	01-08	266	61	724	
	01-12	157	44	794	
		2472	667	9291	12430

Table 5.1: Number of Emergencies

district of Faro, which we speculate happens because of the seasonality present at these times of year. As far as the other two districts, only Lisboa has a similar pattern in the month of February, but not as accentuated as the one in Faro.

Upon running our instances in a generalization of openWBO, a solver developed by João Cortes, that specializes in solving Multi-Objective Combinatorial Optimization instances, we came across some undesirable situations, the first one being the complexity of the problem, which made our solver unable to provide acceptable solutions to a large number of instances within the one hour time limit we set for each instance. In these cases, the solution would be to split the instance problem into smaller problems and then solve them individually. This happened mostly with the larger examples from Lisboa, namely the instances that covered periods of four and six hours in their entirety. However, even among the smaller one and two hour period instances there were several solutions that were not acceptable as these were not better than the original scenario. In the examples of periods of one hour for Lisboa, we stopped getting acceptable solutions when we had about fifty vehicles available for periods with close to one hundred emergencies. Although it is hard infer when exactly the solver is no longer able to provide acceptable solutions in a one hour time period, when we moved to larger periods of time, there were very instances that had conclusive results for the district of Lisboa.

On the other hand there were also several instances that represented periods where there were no emergencies, which do not provide any information, as well as instances that represented single emergencies where the solution is trivial and is often the real case scenario.

In order to illustrate the analysis we will be describing in the following subsections, we added Figure 5.1 as an example, in which the red crosses represent the Pareto Front solutions found for this example of the instance created for Faro, in the first of August, 2017, from 18h to 24h. This particular example we used a thirty minute vehicle inactivity time and a fifteen kilometer range for vehicle action. The black dot represents the real case scenario.

In this particular scenario, we can see that in the real situation eleven vehicles were used and they covered a total of 74806 meters to get to all the emergencies. We can also see that there is only one of our Pareto Front solutions that is better than the real case scenario in both the number of vehicles used

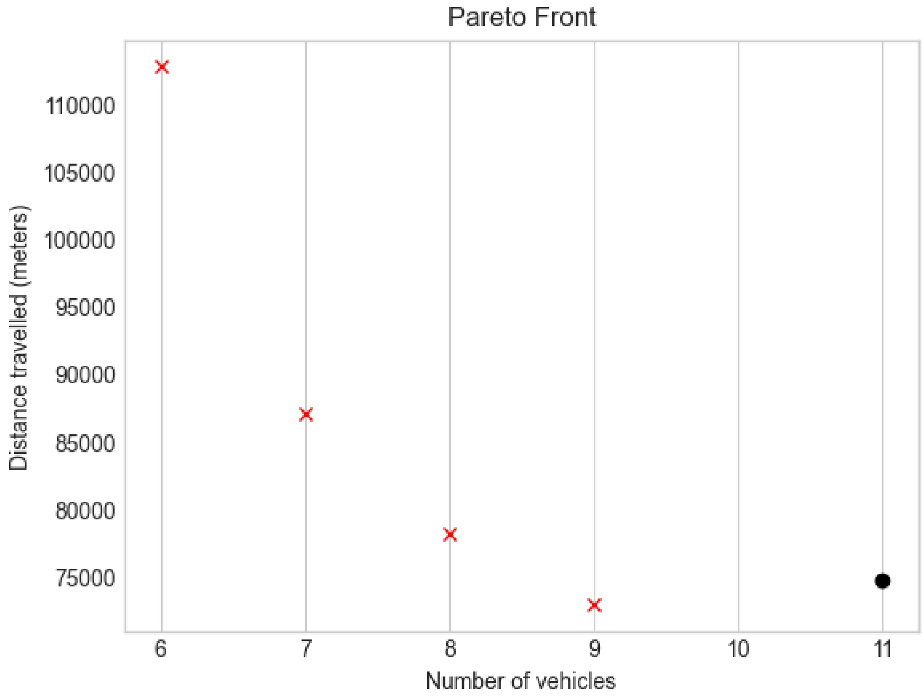

Figure 5.1: Pareto Front example for Faro, 01/08/2017

and in the distance travelled between them, which let's us know that even though the real case scenario was not optimal, it was still rather good when compared to larger examples where it is progressively harder to approach a near optimal situation.

Our goal with this analysis is to compare solutions between districts and seasons and try to understand whether or not we can allocate the vehicles in a different way in order to try to improve the overall performance in all of the places we are analysing.

5.2 Vehicle Inactivity

We wanted to see how big the impact was for the overall solutions if we changed the time a vehicle becomes inactive after it provided aid to an emergency, therefore we tested using a thirty minute inactivity period and using a one hour inactivity period. We used examples with only a ten kilometer radius of activity for the vehicles of each emergency. These two values, thirty minutes and one hour, both came from the analysis of data from previous emergencies, described in Chapter 3.

As expected, the execution of the solver on the instances with the thirty minutes inactivity time took more time than the instances with the one hour inactivity time due to the amount of vehicles available for each emergency. This is especially true for larger examples. In Figure 5.2, we see an example from a two hour period in the district of Lisboa where in the first one we use a thirty minute vehicle inactivity period, represented by the green crosses, and a one hour vehicle inactivity period, represented by the red crosses. Since this is a larger example, it is the only one where the solutions differ when we change the inactivity time. If we look at Figures 5.3 and 5.4 on the other hand, the solutions found are exactly

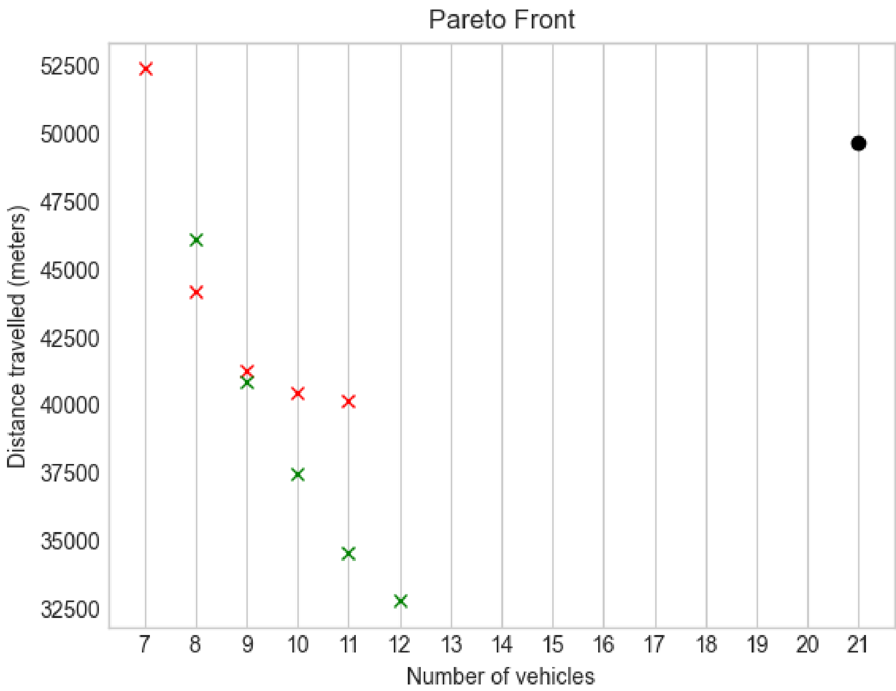

Figure 5.2: Comparison between inactivity times in Lisbon. The green crosses represent the thirty minute vehicle inactivity period and the red crosses represent the one hour vehicle inactivity period.

the same for periods of four hours, hence the green crosses and the red crosses are in the same positions. Although these are larger time periods, because of the difference in number of emergencies, the examples in Lisboa are far larger, and give a perspective of the impact the vehicle inactivity period has on the search of a solution.

Overall, the results in Faro and Guarda did not change when we changed the vehicle inactivity period, so in these examples, which are considered small compared to Lisboa, we can argue that using the one hour inactivity time period is better because it reduces the time execution on the solver. However, when we look at the examples in Lisboa, the general case is that we find better solutions when we use the thirty minute vehicle inactivity period, and therefore, there needs to be a balance between the optimality of our solutions and the time it takes for the solver to execute on each instance in particular. Generally, since we ran our instances with a one hour time limit, it would be better to run the instances with the thirty minute vehicle inactivity time to achieve the solutions closer to optimality.

In real case scenarios, there may be times where thirty minutes are not enough for a vehicle to be able to provide aid to a new emergency after having just provided aid to another emergency. If this is the case, the vehicle inactivity period should be adapted to a value that is compatible with the situation at hand. One hour and thirty minutes were plausible values in these three situations in particular, hence the usage of these values.

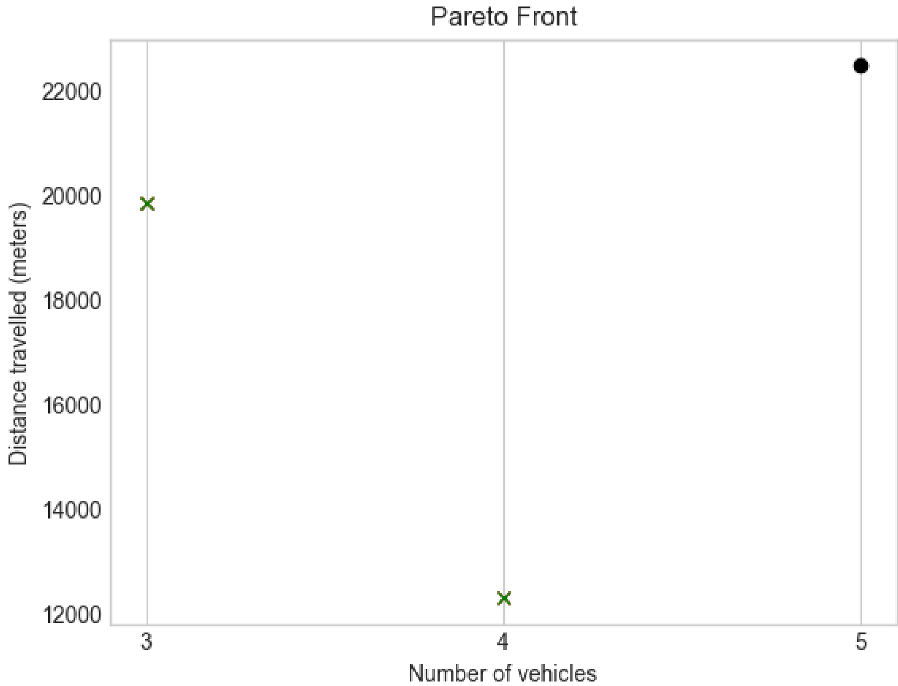

Figure 5.3: Comparison between inactivity times in Guarda. The green crosses represent the thirty minute vehicle inactivity period and the red crosses represent the one hour vehicle inactivity period.

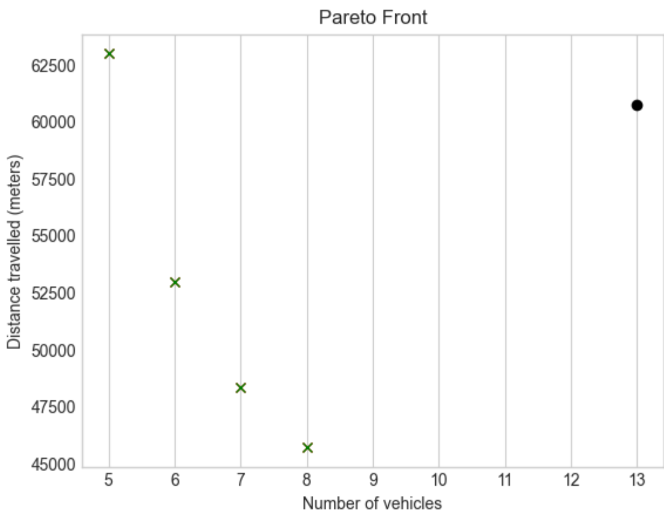

Figure 5.4: Comparison between inactivity times in Faro. The green crosses represent the thirty minute vehicle inactivity period and the red crosses represent the one hour vehicle inactivity period.

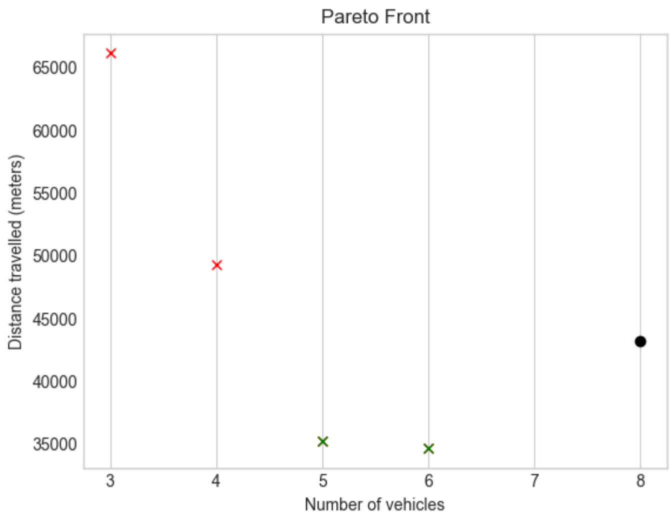

Figure 5.5: Comparison between distance times in Guarda. The green crosses refer to a ten kilometer radius and the red crosses refer to a fifteen kilometer radius

5.3 Distance from the emergency

Analogously, we wanted to test the impact of the radius of activity for the vehicles of each emergency, or in other words, how far away from the emergency a vehicle can be to be taken into account as a possible vehicle to provide aid to an emergency. When we first started this test we had three values for this distance, however, upon several tests, we decided to rule out the distance of five kilometers because a lot of emergencies were getting no available vehicles to provide aid to them. Therefore we ended up with the values of ten and fifteen kilometers and running the tests on these two values.

In some cases, making the radius of search bigger allows the solver to find more solutions that normally use less vehicles but increase the distance travelled by these vehicles overall. These additional solutions are usually cases in which the distance traveled is larger than the real case scenario, and therefore are not interesting when we have other solutions that can both reduce the number of vehicles used and the distance traveled. In Figure 5.5 we can see that the usage of fifteen kilometers adds two solutions, respectfully using three and four vehicles and combined distances of 49348 and 66176 kilometers, but these represent cases that we do not want to consider exactly because the two solutions that we already had from the ten kilometer radius case, respectfully using five and six vehicles and combined distances of 35239 and 34673 kilometers, reduce both the distance travelled and the number of vehicles used when compared to the real case scenario which used eight vehicle and with a combined distance of 43197 kilometers.

On top of this, using a fifteen kilometer radius also largely increases the execution time of the solver for the same scenarios, especially for larger examples. In most cases for Lisboa, the solutions found by the solver were not even close to being optimal as they were worse than the real case scenario. This happened because adding five kilometers means that there are a lot more vehicles to consider for each emergency, and since Lisboa has a larger number of emergencies than the Faro and Guarda and the complexity of this problem makes it grow exponentially fast, the solver could not provide acceptable solutions in the established amount of time of one hour.

Practically, this means that the ten kilometer range is more appropriate in most scenarios, since it grants enough vehicles to reach optimal solution and also because it reduces the universe of vehicles we would have to consider using, making the decision of which vehicle to assign less complex. We could not infer how the Emergency Medical System decided which vehicles were considered for each emergency from the data we had, however, if like us, they use a radius around the emergency, we would suggest the usage of a ten kilometer range as the standard from the three hypothesis we tested. Eventually, if the population density drops heavily there might be a need to use the fifteen kilometer range since there should be less vehicles in that area. Among our three districts, Guarda is the one with the least population density and we did not find any example time period in which there was a need for the fifteen kilometer range, so we infer this is a very unlikely scenario.

Ultimately, it is better to use a radius of ten kilometers as opposed to fifteen, because although we got more solutions from the fifteen kilometer radius examples, these were not solutions that bring any interest in the search for minimization of both our criteria.

5.4 Seasonal Variations

Given the results obtained from the previous sections, all the results presented in this and further sections will be using a vehicle inactivity period of thirty minutes and a radius of ten kilometers.

Provided that Portugal has a big affluence of tourists towards the South in the summer, we wanted to see if this had an influence on the number of occurrences when compared to other times of the year, especially between Faro and Lisboa.

The first thing to note about the examples shown in Figures 5.6 and 5.7, which represent periods of one hour and four hours from the same time period of the day in the year 2019, in Lisboa and Faro respectively, is that the seasonality is present in both districts. This means that both districts account for a consistently different number of emergencies during the month of August as opposed to the month of January. However, in Lisboa January is the month with more emergencies as opposed to Faro, where August is the month with more emergencies. We present only the results for the year 2019, because in the remaining two years of 2017 and 2018, the results were analogous.

We wanted to see if there was a possibility that these seasonal changes were not being addressed properly. In order to infer this, we looked ate the difference between the solutions in our Pareto front and the real case scenarios in both districts and we can see that in both districts, our real case scenario is closer to the optimal solutions in the month of August as opposed to the month of January. However, this difference is more evident in the district of Faro, which can mean that there are fewer vehicles available than there should be for this district at this time of year. This difference could be attenuated if some of the vehicles that are allocated to Lisboa in the month of August were reallocated temporarily to Faro, in an attempt to allow for a better response to a month where there are more emergencies at the expense of a slightly less optimal performance in that month in the district of Lisboa. In the months where the Faro has less emergencies, these vehicles would then be reallocated to Lisboa again as there will be a

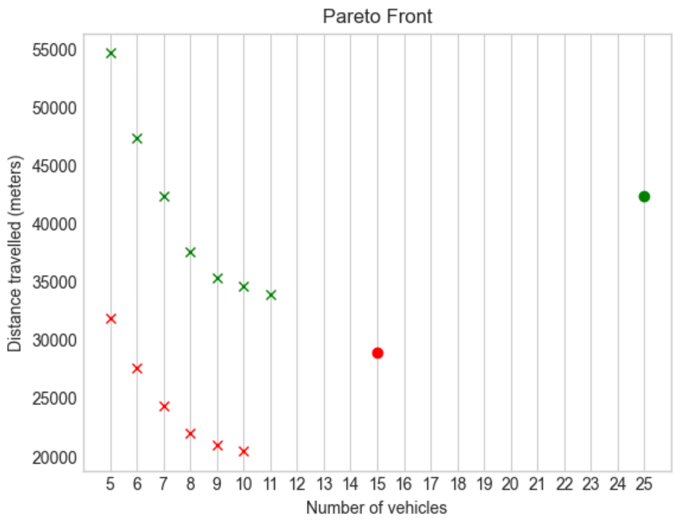

Figure 5.6: Seasonality comparison in Lisboa. The red figures refer to the month of August and the green figures refer to the month of January

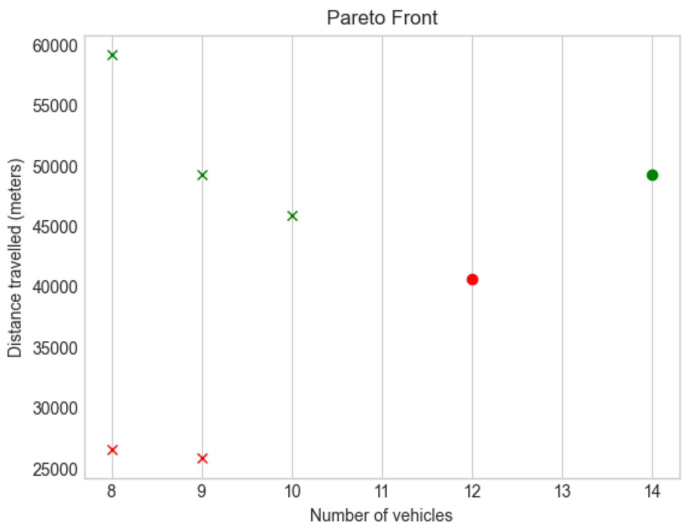

Figure 5.7: Seasonality comparison in Faro. The red figures refer to the month of August and the green figures refer to the month of January

higher need for them there.

Chapter 6

Conclusions

Firstly, when we first set out to do this project, the main goal was to be able to create a model that could accurately represent any situation in the context of the ambulance assignment problem and use it to determine what the optimal situations would be and compare them to the data we had from INEM. This objective was achieved successfully as we created a working model that can in fact represent every situation we have idealized.

Apart from the main goal, we discussed several possible smaller possibilities for further research using the model we had created. Ultimately, we ended up conducting an analysis on three different districts of Portugal and inferring information and possible adjustments to the current way the ambulances are being allocated by the Portuguese Emergency Medical System. This analysis required an analysis of the records we had access to in a variety of different measures like number of emergencies, number of vehicles available for each emergency, emergency priority levels, specific vehicle information, as well as an extensive stage of creation of instances followed by the solving of these same instances. This allowed us to contextualize each situation and be able to critically analyse and compare the real case scenarios and the solutions we obtained from running our examples through the solver in order to come up with possible solutions for less optimal situations we encountered.

Additionally we also perform an aggregate demand analysis where we divide a vehicle allocation problem from a large time period, for example a full day, into smaller time periods which are easier to solve with the objective of finding a solution to the bigger problem by combining information from the solutions to the smaller instances of the problem.

We were hoping to find large inconsistencies in the way the Emergency Medical System works in Portugal, especially when comparing zones which we expected to have a larger seasonality. However, we came to the conclusion that the resources are relatively well distributed to account for these seasonal variances.

Unfortunately, we did not get the chance to pair the application of our model to the results of a predictive model in order to suggest specific resource allocations in order to improve performance. This would have also allowed us to further elaborate on possible improvements that we could not have seen without the use of a predictive model.

Chapter 7

Future Work

In the development process of our algorithm we always considered a given situation in which the total number of vehicles available as well as total number of emergencies were known variables. This means that we have privileged insight when looking for a solution that the operators who choose which vehicles to assign to each emergency. With this in mind, it would be interesting to develop a real-time vehicle assignment mechanism that would better emulate the situation in which the operators have to make the assignment decisions. Afterwards, it would be possible to compare both performances to the solution given by the model developed in this paper. Furthermore, that real-time vehicle assignment mechanism could then benefit from information retrieved from various solutions given by our model like identifying periods of time more prone to a large number of emergencies in a certain area, and trying to preserve more vehicles in that area, even if at the cost of a more lengthy assignment on some other emergency beforehand.

Furthermore, we have talked about using a predictive model to predict possible future emergencies and then apply our model to allocate a higher amount of vehicles in the zones we predict are going to have more emergencies in a certain time period, and leaving zone we deem to be less likely to have emergencies with a smaller amount of vehicles. This would serve as an attempt at optimizing the performance of the Emergency Medical System even further, focusing on probabilities of what will happen in the future, and not only on data from what has happened in the past.

Bibliography

[1] L. Aboueljinane, E. Sahin, and Z. Jemai. A review on simulation models applied to emergency medical service operations. *Computers & Industrial Engineering*, 66(4):734 – 750, 2013. ISSN 0360-8352. doi: https://doi.org/10.1016/j.cie.2013.09.017. URL http://www.sciencedirect.com/science/article/pii/S0360835213003100.

[2] R. Alanis, A. Ingolfsson, and B. Kolfal. A markov chain model for an ems system with repositioning. *Production and Operations Management*, 22(1):216–231, 2013. doi: https://doi.org/10.1111/j.1937-5956.2012.01362.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1937-5956.2012.01362.x.

[3] T. Andersson and P. Värbrand. Decision support tools for ambulance dispatch and relocation. *Journal of the Operational Research Society*, 58(2):195–201, 2007. doi: 10.1057/palgrave.jors.2602174. URL https://doi.org/10.1057/palgrave.jors.2602174.

[4] R. Batta, J. M. Dolan, and N. N. Krishnamurthy. The maximal expected covering location problem: Revisited. *Transportation Science*, 23(4):277–287, 1989. doi: 10.1287/trsc.23.4.277. URL https://doi.org/10.1287/trsc.23.4.277.

[5] V. Bélanger, Y. Kergosien, A. B. Ruiz, and P. Soriano. An empirical comparison of relocation strategies in real-time ambulance fleet management. *Comput. Ind. Eng.*, 94:216–229, 2016. doi: 10.1016/j.cie.2016.01.023. URL https://doi.org/10.1016/j.cie.2016.01.023.

[6] V. Bélanger, A. B. Ruiz, and P. Soriano. Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles. *Eur. J. Oper. Res.*, 272(1):1–23, 2019. doi: 10.1016/j.ejor.2018.02.055. URL https://doi.org/10.1016/j.ejor.2018.02.055.

[7] P. Beraldi and M. E. Bruni. A probabilistic model applied to emergency service vehicle location. *Eur. J. Oper. Res.*, 196(1):323–331, 2009. doi: 10.1016/j.ejor.2008.02.027. URL https://doi.org/10.1016/j.ejor.2008.02.027.

[8] G. Bianchi and R. L. Church. A hybrid fleet model for emergency medical service system design. *Social Science & Medicine*, 26(1):163 – 171, 1988. ISSN 0277-9536. doi: https://doi.org/10.1016/0277-9536(88)90055-X. URL http://www.sciencedirect.com/science/article/pii/027795368890055X.

[9] H. Billhardt, M. Lujak, V. Sánchez-Brunete, A. Fernández, and S. Ossowski. Dynamic coordination of ambulances for emergency medical assistance services. *Knowl. Based Syst.*, 70:268–280, 2014. doi: 10.1016/j.knosys.2014.07.006. URL https://doi.org/10.1016/j.knosys.2014.07.006.

[10] R. Church and C. ReVelle. The maximal covering location problem. *Papers of the Regional Science Association*, 32(1):101–118, 1974. doi: 10.1007/BF01942293. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-0345427032&doi=10.1007%2fBF01942293&partnerID=40&md5=f126038002c87da43b00a050acf47080. cited By 1537.

[11] M. Daskin. Location, dispatching, and routing models for emergency services with stochastic travel times. spatial analysis and location-allocation models, a. ghosh and g. rushton, eds, 1987.

[12] M. S. Daskin and E. H. Stern. A hierarchical objective set covering model for emergency medical service vehicle deployment. *Transportation Science*, 15(2):137–152, 1981. doi: 10.1287/trsc.15.2.137. URL https://doi.org/10.1287/trsc.15.2.137.

[13] M. S. Daskin, K. Hogan, and C. ReVelle. Integration of multiple, excess, backup, and expected covering models. *Environment and Planning B: Planning and Design*, 15(1):15–35, 1988. doi: 10.1068/b150015. URL https://doi.org/10.1068/b150015.

[14] D. Eaton, S. Héctor Ml, R. Lantigua, and J. Morgan. Determining ambulance deployment in santo domingo, dominican republic. *Journal of the Operational Research Society*, 37(2):113–126, 1986. doi: 10.1057/jors.1986.21. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-0022662649&doi=10.1057%2fjors.1986.21&partnerID=40&md5=83aad006c8e20018ef7208a4e6957365. cited By 73.

[15] D. J. Eaton, M. S. Daskin, D. Simmons, B. Bulloch, and G. Jansma. Determining emergency medical service vehicle deployment in austin, texas. *INFORMS Journal on Applied Analytics*, 15(1):96–108, 1985. doi: 10.1287/inte.15.1.96. URL https://doi.org/10.1287/inte.15.1.96.

[16] R. D. Galvão and C. ReVelle. A lagrangean heuristic for the maximal covering location problem. *European Journal of Operational Research*, 88(1):114 – 123, 1996. ISSN 0377-2217. doi: https://doi.org/10.1016/0377-2217(94)00159-6. URL http://www.sciencedirect.com/science/article/pii/0377221794001596.

[17] M. Gendreau, G. Laporte, and F. Semet. Solving an ambulance location model by tabu search. *Location Science*, 5(2):75 – 88, 1997. ISSN 0966-8349. doi: https://doi.org/10.1016/S0966-8349(97)00015-6. URL http://www.sciencedirect.com/science/article/pii/S0966834997000156.

[18] M. Gendreau, G. Laporte, and F. Semet. A dynamic model and parallel tabu search heuristic for real-time ambulance relocation. *Parallel Computing*, 27(12):1641 – 1653, 2001. ISSN 0167-8191. doi: https://doi.org/10.1016/S0167-8191(01)00103-X. URL http://www.sciencedirect.com/science/article/pii/S016781910100103X. Applications of parallel computing in transportation.

[19] M. Gendreau, G. Laporte, and F. Semet. The maximal expected coverage relocation problem for emergency vehicles. *J. Oper. Res. Soc.*, 57(1):22–28, 2006. doi: 10.1057/palgrave.jors.2601991. URL https://doi.org/10.1057/palgrave.jors.2601991.

[20] M. Gendreau, G. Laporte, and F. Semet. The maximal expected coverage relocation problem for emergency vehicles. *Journal of the Operational Research Society*, 57(1):22–28, 2006. doi: 10.1057/palgrave.jors.2601991. URL https://doi.org/10.1057/palgrave.jors.2601991.

[21] K. Hogan and C. ReVelle. Concepts and applications of backup coverage. *Management Science*, 32(11):1434–1444, 1986. doi: 10.1287/mnsc.32.11.1434. URL https://doi.org/10.1287/mnsc.32.11.1434.

[22] S. Ibri, M. Nourelfath, and H. Drias. A multi-agent approach for integrated emergency vehicle dispatching and covering problem. *Eng. Appl. Artif. Intell.*, 25(3):554–565, 2012. doi: 10.1016/j.engappai.2011.10.003. URL https://doi.org/10.1016/j.engappai.2011.10.003.

[23] C. Jagtenberg, S. Bhulai, and R. van der Mei. An efficient heuristic for real-time ambulance redeployment. *Operations Research for Health Care*, 4:27 – 35, 2015. ISSN 2211-6923. doi: https://doi.org/10.1016/j.orhc.2015.01.001. URL http://www.sciencedirect.com/science/article/pii/S2211692314200075.

[24] G. Laporte, F. V. Louveaux, F. Semet, and A. Thirion. Application of the double standard model for ambulance location. In J. A. Nunen, M. G. Speranza, and L. Bertazzi, editors, *Innovations in Distribution Logistics*, pages 235–249, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. ISBN 978-3-540-92944-4.

[25] B. López, B. Innocenti, and D. Busquets. A multiagent system for coordinating ambulances for emergency medical services. *IEEE Intelligent Systems*, 23(5):50–57, 2008. doi: 10.1109/MIS.2008.76.

[26] A. J. Mason. Simulation and real-time optimised relocation for improving ambulance operations. In *Handbook of Healthcare Operations Management*, pages 289–317. Springer, 2013.

[27] M. S. Maxwell, M. Restrepo, S. G. Henderson, and H. Topaloglu. Approximate dynamic programming for ambulance redeployment. *INFORMS Journal on Computing*, 22(2):266–281, 2010. doi: 10.1287/ijoc.1090.0345. URL https://doi.org/10.1287/ijoc.1090.0345.

[28] R. Nair and E. Miller-Hooks. Evaluation of relocation strategies for emergency medical service vehicles. *Transportation research record*, 2137(1):63–73, 2009.

[29] J. Naoum-Sawaya and S. Elhedhli. A stochastic optimization model for real-time ambulance redeployment. *Computers & Operations Research*, 40(8):1972 – 1978, 2013. ISSN 0305-0548. doi: https://doi.org/10.1016/j.cor.2013.02.006. URL http://www.sciencedirect.com/science/article/pii/S0305054813000385.

[30] D. Schilling, D. J. Elzinga, J. Cohon, R. Church, and C. ReVelle. The team/fleet models for simultaneous facility and equipment siting. *Transportation Science*, 13(2):163–175, 1979. doi: 10.1287/trsc.13.2.163. URL https://doi.org/10.1287/trsc.13.2.163.

[31] J. E. Storbeck. Slack, natural slack, and location covering. *Socio-Economic Planning Sciences*, 16(3):99 – 105, 1982. ISSN 0038-0121. doi: https://doi.org/10.1016/0038-0121(82)90020-9. URL http://www.sciencedirect.com/science/article/pii/0038012182900209.

[32] K. Sudtachat, M. E. Mayorga, and L. A. Mclay. A nested-compliance table policy for emergency medical service systems under relocation. *Omega*, 58:154 – 168, 2016. ISSN 0305-0483. doi: https://doi.org/10.1016/j.omega.2015.06.001. URL http://www.sciencedirect.com/science/article/pii/S030504831500122X.

[33] C. Toregas, R. Swain, C. S. ReVelle, and L. Bergman. The location of emergency service facilities. *Oper. Res.*, 19(6):1363–1373, 1971. doi: 10.1287/opre.19.6.1363. URL https://doi.org/10.1287/opre.19.6.1363.

[34] T. Van Barneveld, S. Bhulai, and R. D. van der Mei. A dynamic ambulance management model for rural areas. *Health care management science*, 20(2):165–186, 2017.