SUMMARY

In life-threatening situations where every second counts, the ability of ambulance service providers to arrive at the emergency scene within a few minutes to provide medical aid can make the difference between survival or death. In the Netherlands the response-time target is 15 minutes for incidents of the highest urgency. To realize short response times at affordable costs, adequate planning of ambulance services is crucial. An essential element herein is an efficient distribution of ambulances over the region. After all, the location of emergency vehicles at the time an incident is reported at the emergency control center determines to a large extent whether the response-time target is achieved. A complicating factor in addressing these ambulance location problems is the omnipresence of uncertainty in the ambulance service-provisioning process. Especially the uncertainty regarding the availability of emergency vehicles is crucial. Ambulances are not always dispatchable to an incident during their duty time, for instance, due to the treatment of a patient at the emergency scene of an earlier incident or due to the transportation of such a patient. This unavailability can result in a temporarily inefficient distribution of ambulances over the region.

A way to resolve the problem mentioned above is a temporary redeployment of one or more ambulances. This is the core of Dynamic Ambulance Management (DAM). An important advantage of DAM is the ability to anticipate future incidents in a more flexible way. Response times, and hence, mortality and morbidity, can be reduced through the repositioning of ambulances in real time. This dissertation is concerned with several models and algorithms for the optimization of the coverage by means of performing so-called *proactive relocations*.

The optimization methods in this thesis can be roughly classified into two main categories: (1) *online* and (2) *offline* algorithms. The key difference between both is the moment at which the (majority of the) computational work is done. For online methods this happens at the moment at which a relocation decision needs to be taken. This class of methods can handle a very detailed state description, because just for one specific state of the EMS system at the time the relocation decision is computed. In the offline approach the majority of the computations is done a priori and for each possible state the corresponding relocation decision is stored. When a certain situation (i.e., state) occurs, the relocation decision is retrieved and applied. The computation time highly depends on the number of possible states. Therefore, the state description is typically less sophisticated than in the online approach in order to keep the computation time manageable.

The first part of this dissertation (Chapters 2-4) is concerned with the online approach to the ambulance relocation problem. Chapter 2 focuses on rural regions. These usually differ from their urban counterparts due to a smaller number of incidents, a smaller number of ambulances on duty, and the geographical spread of incidents over the region. The relocation problem is modeled as a Markov decision problem, in which the response-time dependent performance objective can be chosen arbitrarily through the selection of a suitable *penalty function*. This function assigns to each possible response time a certain penalty. Besides, this chapter describes a heuristic for the calculation of good relocation decisions, based on the presented model. Chapter 2 concludes with the illustration of this heuristic for several penalty functions using a realistic EMS system corresponding to a Dutch region as test bed.

A frequently mentioned drawback of DAM is the increase in the crew's workload, due to the additional number of relocations. Chapter 3 addresses this issue: the trade-off between the number of relocations and the response-time performance is studied. To this end the penalty heuristic is proposed in this chapter. This heuristic uses the concept of penalty function. Moreover, this chapter introduces the use of so-called chain relocations: a otherwise long trip can be split into multiple shorter ones to attain the desired ambulance configuration (as computed by the penalty heuristic) faster. The chapter is concluded with an extensive numerical study regarding the mentioned trade-off. The consequences of restrictions on the number of ambulance relocations on the performance are studied, based on a large number of realistic situations, using a specific penalty function provided by ambulance practitioners. The response-time performance increases significantly if only a few ambulance relocations are carried out. However, frequent repositioning can possibly result in a performance loss.

In Chapter 4 several insights concerning the implementation of relocation strategies in practice are presented. To this end, the proposed penalty heuristic of Chapter 3 is combined with the Dynamic MEXCLP algorithm of Jagtenberg et al. (2015). The following five aspects are discussed: (1) the frequency of redeployment decision moments, (2) the inclusion of busy ambulances in the state description of the system, (3) the performance criterion on the quality of the relocation strategy, (4) the use of chain relocations, and (5) time bounds on the relocation time. The chapter continues with an extensive simulation study regarding the practical implementation of these facets.

The offline approach for solving the ambulance relocation problem is the topic of the second part of this dissertation (Chapters 5 and 6). Chapter 5 describes a integer linear programming formulation for the computation of *ambulance compliance tables*, called MEXPREP. This model is an extension of the MECRP model by Gendreau et al. (2006) in two different directions: (1) it accounts for the fact that ambulances might become busy during the execution of the compliance table policy, and (2) a generic performance objective can be chosen through the definition of the corresponding penalty function. This chapter also introduces an adjusted version (called AMEXPREP) that relaxes the assumptions made on the busy fraction. A section with numerical results, based on the simulation of the compliance tables computed by MEXPREP, concludes the chapter.

Chapter 6 is devoted to the computation of optimal compliance tables as well. Two types of medical response units are considered in this chapter: vehicles with and without transport capability. The last class usually consists of motor cycles, so this type of unit is typically present faster at the emergency scene. This complicates the calculation of compliance tables, as an extra dimension is added to the state space of the EMS system. This chapter presents a integer linear programming formulation for the computation of so-called *two-dimensional compliance tables*. In this model the number of relocations required to be carried out at once is bounded. Moreover, there are also restrictions on the time a specific relocation may last. Subsequently, several regimes are tested for different fleet mixes in a rural region by simulation. This study shows that imposing a time bound that is equal to the expected interarrival times of incidents seems a good choice.

The last chapter of this dissertation, Chapter 7, presents a unified view on the online and offline approaches for solving the ambulance relocation problem. To this end, representants proposed in the previous chapters are chosen. These are the combination of the penalty heuristic and DMEXCLP (Chapter 4) and the compliance tables obtained through solving AMEXPREP (Chapter 5) for the online and offline approach, respectively. Both methods are simulated for different fleet sizes and performance objectives. In this study, the chosen representants show similar performance, both from a patient and a crew perspective.