Chapter 24
Managing a Fleet of Ambulances to Respond to Emergency and Transfer Patient Transportation Demands

Y. Kergosien, M. Gendreau, A. Ruiz, and P. Soriano

Abstract Organizations of pre-hospital emergency medical services have as first mission to provide medical assistance to patients including transport to medical centers if necessary, and a second one that concerns the transfers of patients from one medical facility to another one. Most organizations in Canada and in North-America use two independent fleets to perform these missions. Although operating two separated fleets seems easier to do, it appears to be less efficient than an integrated fleet management approach able to deal with both types of demands. Taking this into consideration, this paper aims to design and evaluate the performance of three management approaches. This is a very challenging problem, since it involves solving simultaneously two difficult vehicle routing problems: an ambulance relocation problem and a dial a ride problem.

Y. Kergosien (✉)
Université François Rabelais Tours, Laboratoire d'Informatique (EA 6300), Equipe Ordonnancement et Conduite (ERL CNRS 6305), 64 avenue Jean Portalis, 37200 Tours, France
e-mail: yannick.kergosien@univ-tours.fr

M. Gendreau
École Polytechnique de Montréal, 2900 boul. Édouard-Montpetit, Montréal,
QC H3T 1J4, Canada
e-mail: michel.gendreau@cirrelt.ca

A. Ruiz
Département opérations et systèmes de décision, Faculté des sciences de l'administration, Université Laval, QC G1K 7P4, Canada
e-mail: angel.ruiz@fsa.ulaval.ca

P. Soriano
Service de l'enseignement des méthodes quantitatives de gestion, HEC Montréal, 3000 chemin de la Côte Sainte-Catherine, Montréal, QC H3T 2A7, Canada
e-mail: patrick.soriano@cirrelt.net

A. Matta et al. (eds.), *Proceedings of the International Conference on Health Care Systems Engineering*, Springer Proceedings in Mathematics & Statistics 61,
DOI 10.1007/978-3-319-01848-5_24,

24.1 Introduction

The primary mission of pre-hospital emergency medical services organizations (EMS) is to provide the first medical assistance and, if required, the transportation of patients to a medical center. Transportation is performed by paramedic teams using ambulances that are generally deployed at strategic places to respond as soon as possible to emergency demands. The territory deserved is divided into areas. An area is recognized as covered if at least one ambulance can reach any demand in this area within a prefixed time. When an urgent demand appears, the nearest available ambulance is sent to the demand's location. It then may happen that idle ambulances need to be *redeployed* to new waiting locations in order to compensate the void created by the departure of the dispatched ambulance. In addition, most of EMS organizations carry out another mission, which is to respond to transportation demands between medical facilities, and sometimes, to or from the patients' homes. These transfer demands arrive in real time, but fairly in advance with respect to the desired patient departure. This slack allows to schedule them to form ambulances' routes.

Both types of demands, "emergency" and "transfer", can be performed by the same teams and ambulances. Nevertheless, Urgences-Santé currently manages these two types of demands independently by dividing their ambulances into two fleets and managing them separately. This solution was chosen to simplify the management of the vehicles, but it appears to be less efficient than an integrated fleet management approach dealing with both types of demands. Therefore, the aim of this paper is to assess whether or not an integrated management approach can lead to better service quality and to a reduction of costs. In this paper, we first present a management approach to manage efficiently two independent fleets. Then, we proposed two management strategies considering a single integrated fleet. The performances of these three models are compared by several simulation experiments.

This paper is structured as follows. The next section presents a short literature review. Section 24.3 introduces the problem considered here: the management of an integrated fleet of ambulances to respond to both emergency and transfer patient transportation demands. Section 24.4 describes the three management strategies proposed. Section 24.5 focuses on the simulation experiments and results' analysis. Finally, conclusions and further research avenues are provided.

24.2 Literature Review

In the literature, there are two well known families of problems which are closely related to the dynamic management of the two transportation requests here described. The first one is named "Dial A Ride Problem" [13]. In this problem, a set of customer's transportation demands has to be performed by a set of vehicles under various constraints like maximum ride time, vehicle capacity, maximum ride time constraints, time windows, etc. The problem consists in finding the routes to be done by each vehicle (sequence of requests to perform) minimizing one or several criteria like the total distance traveled or mean user ride time. This problem has been studied both in a static context by Cordeau and Laporte [12] and Parragh et al. [25] as well as in a dynamic context by Attanasio et al. [3] and Xiang et al. [32]. In our case, however, ambulances can transport only one patient at a time. The dial a ride problem (DARP) has been studied in a medical context in [6]. A DARP with heterogeneous users and vehicles is studied in [24] where different modes (seated, stretcher and wheelchair) and types of vehicles are considered. The problem studied in [21] was inspired by a real case of transportation of patients in a hospital complex in France, where transportations are subject to particular constraints. Among them, its worth to mention priority of urgent demands, disinfection of a vehicle after the transportation of a contagion, respect of the type of vehicle needed and the opportunity to outsource demands to a private company.

The second family of problems concerns the ambulance relocation problem. For each emergency demand, the choice of an ambulance to be dispatched must be made. In some cases, a redeployment of ambulances is applied after the dispatching of an ambulance. Redeployment consists in assigning ambulances to potential sites to provide adequate coverage and in order to respond as quickly as possible to a new emergency demand. Several literature reviews have been published over the past years focusing on both ambulance location and relocation problems [9], and recently [7]. The problem was first studied in its static and determinist versions [31] and [11] with one coverage and with multiple coverage have been developed, like in [14, 16] and [15]. Stochastic approaches have also been proposed in [1, 5, 20] and [8]. In a dynamic version, two approaches can be found in the literature. The first one, multi-period, consists in decomposing the day into several periods and apply a redeployment at the beginning of each period [28] or [4]. Several studies using simulation have been conducted to test different deployment plans for each period, either without taking into account the relocation costs between two successive period [27] or taking them into account [10], or even considering a variation in the size of the ambulance fleet between periods [26]. The second approach consists in applying a redeployment according to the evolution of the system's state, [18] and [2]. In [23], it was proposed a dynamic programming approach combined with a Monte Carlo Simulation to determine where the next available ambulance should be redeployed in order to increase the number of demands reached within a given lapse of time. Recently, in [29], an approximate dynamic programming formulation is proposed to solve a dynamic version of the ambulance dispatching and relocation problem taking into account time-dependent information like variations in the demand volumes and travel times throughout the course of the day.

To the best of our knowledge, these two families of problems have been always studied independently in the literature except in one paper [22], where the authors deal with a problem that has a common feature with our case: the simultaneous management of emergency transportations and of transfer transportations between hospitals. The authors tested different strategies to manage ambulances based on the selection of waiting points, which can only be the hospitals. However they did not solve the associated location problem.

The contributions of this is paper to improve the knowledge on the EMS fleet management problem are twofold. Firstly it proposes three strategies for managing both separated and integrated fleets. Secondly, it proposes efficient solving approaches to tackle the two problems underlying the EMS fleet management: the dial a ride problem and the ambulance relocation problem in a dynamic context.

24.3 Problem Description

A fleet of identical ambulances has to perform two types of transports during a given planning period. A team composed of two technicians having their own work time schedule is associated to each ambulance. At the beginning of their working shift, the team picks an ambulance at a given depot and must return it to the same depot at the end of their shift. Managing the fleet consists basically in assigning transportation requests to ambulances, and locating idle ambulances to standby points in such a way that they will respond as quick as possible to emergency demands. The whole problem can be divided into two subproblems (a dial a ride problem and an ambulance relocation problem) where the pool of ambulances, is shared. To evaluate the performance of a management strategy, we consider three types of objectives. The first one concerns how the fleet is able to perform all the transfer demands respecting their required time windows. This objective is measured by computing the sum of all the transportation delays. A second objective aims at maximize the total urgent demands covered. The last objective seeks at minimizing the operating costs, measured by the sum of empty running of ambulances as well as the technicians' overtime.

24.3.1 Transfer Patient Transportation Problem

At every moment, the set of transfer transportation demands is known. This set can change over time due to the arrival of new demands or canceling of existing demands. Each transfer transportation demand is characterized by:

- An origin point and a destination which are usually a hospital but may also be a patient's home,
- A time window that modelizes the earliest time at which the patient is ready to be transported and the latest time accepted for the beginning of the transportation. After such a latest time, a delay is incurred.
- A time needed to transfer the patient outside of the ambulance (administrative tasks, stretcher transportation, etc.) that has to be taken into account for both the destination and the origin of the transport.

24.3.2 Ambulance Relocation Problem

When the coverage offered by the available ambulances is not acceptable, a new relocation plan has to be computed. The problem consists in finding new standby locations for the ambulances. The potential standby points are known in advance and are usually strategic locations in the region that should be covered. Also, a maximum number of ambulances can be assigned to each standby point. The desired coverage is the same as that defined in [17]: two types of covering constraints in agreement with the United States EMS Act of 1973. These constraints specify that all emergency demands must be satisfied by an ambulance within S' minutes and a proportion α of the total demand is also satisfied within S minutes ($S' > S$). Thereby each zone of the region is characterized by:

- Two sets of potential sites. One represents the potential sites from which an ambulance can reach all points of the zone within S, and within S' for the other set.
- A density of population.
- A probability vector. The probability vector of a zone A gives, for each period, the probability that the next demand occurs in A. Since this probability can fluctuate according to the time of day, a day is decomposed into several periods (2 h long by default).

When an emergency demand occurs, an ambulance should be selected to respond to this demand. An emergency demand is defined by:

- An intervention time at the scene,
- Whether a transport to a hospital is required or not,
- The hospital destination,
- And a time needed at the hospital to transfer the patient to the hospital staff.

Finally, a minimum time has to be respected between two consecutives redeployments of a given ambulance in order to avoid to redeploy an ambulance too many times in a short period.

24.3.3 Performance Evaluation

To evaluate the performance of a given management strategy, we consider the following criteria related to service quality and costs:

- The elapse time between the arrival of an emergency call and the arrival of the ambulance at the scene for emergency demands,
- The delays for transfer demands,
- The workload of each team and overtime if any,
- The number of times that ambulances were redeployed or diverted,
- The number of deployments or redeployments,
- And the total distance traveled by empty ambulances.

24.4 Management Approaches

In order to assess whether or not an integrated management approach may improve service quality and lead to reduction of costs, this section proposes three management strategies. The first one, named Independent management, corresponds to the dominant strategy that consists in dividing the ambulances into separated fleets, which will deal with transfer and urgent requests, respectively. The second and third strategies consider a single fleet to respond to both types of demand. However, the second strategy, named Reactive integrated fleet management, adopts a pure reactive approach (i.e. demands are considered at their execution time) whereas the third one, named Proactive integrated fleet management, uses a proactive scheduling that consists in deciding the execution date of each transfer demand in order to minimize the number of ambulances that will be busy simultaneously. The three strategies as well as the related tools proposed to solve the ambulances relocation and assignment decisions are described in the next subsections.

24.4.1 Independent Management

This management is based on two fleets: the first one responds to emergency demands only and the second one responds to transfer demands. Both fleets are managed independently and the number of ambulances assigned to them is kept constant during the planning horizon. Strategies to manage each fleet are now described.

24.4.1.1 Management of the Emergency Requests Fleet

In order to have an adequate coverage at any time, an ambulance redeployment is computed and applied if and only if not all zones are covered within S and the last redeployment is not too recent (e.g. more than 15 min ago). Therefore, after each event like an ambulance becomes available (e.g. a team starts its shift or an ambulance finishes serving a demand) or becomes unavailable (e.g. a team finishes its shift or an ambulance is assigned to a demand), a redeployments can occur. If an ambulance becomes available and no redeployment is needed, two cases are possible. In the first case at least one zone is not covered within S'; then the ambulance is sent to the site that maximizes the number of zones that are covered. Otherwise: the ambulance is sent to the site that maximizes the sum over all doubly covered zones within S of the probability that the next demand will appear in that zone. When a new demand occurs, the nearest available ambulance is dispatched.

Redeployment decisions are done by solving an integer linear program based on [16] and solved by CPLEX. In [16], the ambulance redeployment problem is modeled with the same types of coverage. The objective function is to maximize the sum of the zones that are covered twice within S minutes weighted by the probability of a new demand occurring in that zone. The covering constraints and the constraints on the minimum time before redeployment for each ambulance are relaxed to avoid getting infeasible solutions. However, their violation is strongly penalized in the objective function. The solution produced indicates the new number of ambulances that should located at each standby point. Then, the specific instructions for each ambulance are decided by minimizing the total travel distance. This is done after solving a min-cost max-flow problem.

24.4.1.2 Management of the Transfer Requests Fleet

To manage the ambulance fleet and the transfer demands, each ambulance is associated to a route (e.g. a sequence of transfer requests to be performed). Routes are recomputed by using a fast and efficient tabu search algorithm each time a new event happens. An event consists in a new request arrival or the cancellation of an existing request. The initial solution for each execution of the tabu search is the best solution found at the previous event, with some updates like the new position of each ambulance, the demands completed since them, etc. When a new demand occurs, it is included in the route of one ambulance before executing the tabu search in such a way that the sum of delays is minimized. The tabu search uses a lexicographic objective function: it first minimizes the sum of transportation delays and crew overtimes, then the sum of traveled distances.

The tabu search algorithm has a structure as in [19]. A solution is the set of routes for each ambulance. The stopping criterion is the maximum number of iterations without improving the best solution found so far. The neighborhood is built by the CROSS exchange operator, which is particularly well suitable for vehicle routing problems with time windows [30]. In CROSS, the neighborhood of a solution is obtained by exchanging all the sub-segments (parts of a route) of all routes. As in [30], the tabu list contains the objective function values. This way of managing the tabu condition helps reducing computation time as well as the risk of cycling between visited solutions, since the likelihood of having two different solutions with the same objective is very low.

24.4.2 Reactive Integrated Fleet Management

This approach considers a single ambulance fleet responding to both types of demands. The main idea of this strategy is to consider all the demands, transfer or urgent, as emergency demands. The "planning" part of the problem (scheduling of the transfer requests) disappears and only the location and relocation plans need to be solved.

Whenever a transfer request arrives to the system, it is transformed into a *dummy* urgent request that will happen at the patient earliest transportation date. Dummy emergency demands are then managed like the real emergency demands but, unlike them, they can be postponed in some cases that will be described later. To take into account dummy demands in the ambulance relocation problem, the probability vector $P_t(a)$ of a zone a, that define the probability that the next demand will arrive in that zone for each period, is changed according to the Eq. 24.1. For a period t, this probability is a weighted average of the probability that a new real emergency demand happens and an average of the probability of a transfer demand occurring in that zone. The former probability is an average between the probability that a transfer demand occurs in one of the hospitals belonging to that zone and the proportion of the transfer demands known in that zone.

$$P_t(a) = (1-\beta)Pe_t(a) + 0.5\beta\left(\sum_{h\in a} Ps_t(h) + \frac{U(a)}{\#dmd}\right) \tag{24.1}$$

- β: Proportion of transfer demands (a parameter computed from historical data).
- $Pe_t(a)$: Probability of a new emergency demand occurring in zone a.
- $Ps_t(h)$: Probability of a transfer demand occurring at hospital h.
- $U(a)$: Number of transfer demands known from hospitals belonging to zone a.
- $\#dmd$: Total number of transfer demands known.

The fleet management strategy corresponds to the one described in Sect. 24.4.1.1 section. However, in order to avoid situations where too many ambulances will be occupied by dummy demands, the number of ambulances devoted to these requests is limited. Beyond this limit, the demands are added to a list of *postponed* demands. Once an ambulance becomes available, and if the number of dummy requests being performing is below the mentioned limit, then an ambulance may assigned to the first request in the postponed list.

24.4.3 Proactive Integrated Fleet Management

The main idea of this management approach is to improve the previous ones by anticipating the best dates to perform the transfer demands. Basically, the proactive management reproduces the strategy in Sect. 24.4.2 without the constraint which limits the number of ambulances that are responding to transfer demands simultaneously. The execution dates for transfer requests are calculated in order to minimize the number of ambulances that will be busy at the same time, and the sum of transfer transportation delays. These execution dates represent the dates at which the dummy emergency demands will occur. The new problem consists then in schedules the transfer demands, and can be modeled as parallel machines tasks scheduling problem where the tasks are the requests and the machines are the ambulances. To solve this problem we propose a method that can be seen as a proactive scheduler for the transfer demands. Processing times of tasks are an estimate of the time required to perform the transfer demands obtained by computing an estimate of the travel time to move to the departure hospital, plus the travel time, plus the patient transfer time to hospitals staff. The machines have some periods of unavailability according to ambulance activities and their work schedules. Once the problem is solved, the solution indicates the execution dates to perform the transfer demands by sending an ambulance, but the specific ambulances/tasks assignments are not used to keep additional flexibility. The ambulance that will be sent will not be the nearest but the one minimizing the deterioration of the coverages.

We used a tabu search algorithm to solve efficiently this problem. An indirect encoding of the solutions, based on a sequence of tasks, was used in order to simplify the method. To build and evaluate an actual solution, each task is placed iteratively in its best place in the order of the encoding sequence. A partial schedule to the problem is evaluated by the sum of assignment costs where the cost of assigning a task i at a date x is given by the Eq. 24.2. The neighborhood is built by a swap operator and the tabu list contains the objective function values.

$$c_i^x = (1-\lambda) \sum_{y=x}^{y=x+p_i} W_y n_y + \lambda \; delay_i(x) \tag{24.2}$$

- p_i: processing time of i.
- W_y: probability that a new emergency demand occurs at time y.
- n_y: number of ambulances used at time y.
- $delay_i(x)$: delay of i.
- λ: weight applied to two criteria (“coverage degradation” and “transport delay”).

24.5 Numerical Experiments and Preliminary Results

To assess the performance of the described strategies, we designed several simulation experiments. To this end, we developed a generic discrete-event simulation model. In order to increase the flexibility of the model, we separated the decision logic from the routines simulating the physical processes (ambulances movements).

To generate realistic instances, we inspired by the real case of the Urgences-Santé, the EMS organisation in Montreal, Canada. Based on previous published works, we proposed parameters and probability distributions of random variables. Some of the main characteristics of this case are now reported. Montreal is divided into 595 zones. Across the city, there are also the 39 potential sites that can hold up to 4 ambulances, 2 depots and the 15 hospitals. A total of 153 paramedical teams were considered. An exponential distribution is used to model the inter-arrival times between two emergency demands. Depending on the period of the day, the mean of the exponential distribution varies between 110 and 278 s. The time at call location and the time needed at the hospital to take care the patient to the end are generated using the gamma distribution Gamma(k;θ) with θ the scale parameter and k the shape parameter. We assumed that, if no transport to hospital is needed, the time at call location in minutes is generated with Gamma(28.5;19.5) with a probability of 0.85 and Gamma(6.5;4) otherwise. If a transport to hospital is needed, the time at call location is generated with Gamma(16.5;7.2) and the time needed at the hospital with Gamma(42;15). The destination hospital is randomly chosen such as the nearest hospitals have a strong probability of selection.

We also assumed that the number of transfer demands and their arrival times follow a combination of two normal distributions, one during the morning and second one during the afternoon. The probability of selection of the morning's normal distribution for the arrival times were considered slightly larger (0.55). If we note $N(\mu;\sigma^2)$ a normal distribution with a mean μ and a variance σ^2, the morning's normal distribution and the afternoon's normal distribution were set respectively to $N(10\,\mathrm{h};2\,\mathrm{h})$ and $N(17\,\mathrm{h};2\,\mathrm{h})$. We also assumed that approximately 5 % of the demands are canceled. The time at which the cancelation is known is randomly generated at a moment between half an hour after the time where the demand is known and the latest date for the demand execution. Each transfer demand can start during a time window. The earliest date is generated between half an hour and 5 h after the time where the demand is known, using a normal distribution $N(2\,\mathrm{h}30;1\,\mathrm{h})$. The size of the time windows is randomly generated between half an hour and 4 h. The probability that the departure or the destination is a hospital is equal to 0.85, but the departure and the destination cannot both be a patient home. The times need at the places of the departure and the destination are generated using a normal distribution $N(20;7)$ (minutes).

Based on these assumptions, we generated 20 instances of 7 days covering 5 months. However, our analyses and results are based only on the 5 middle-days to remove the transitory states of the first and last day. We tested the three management strategies on all instances. To reduce as much as possible the variance in the results, the random events were kept exactly the same for all the three management strategies for a given instance. The results are summarized in Table 24.1. For each management strategy we report the average of the following indicators over the 20 instances: the number of emergency demands without and with transports, the average response times, the coverages according to S and S', the number of transfer demands, the sum of transport delays and the number of late transfer demands. We also noted some criteria related to efficiency of the strategy: the number of diversions, the number of redeployments, the number of ambulances redeployed, the percentage of ambulances empty travels and the overtime of paramedics.

We can conclude that, regarding the quality of services for emergency and transfer demands, the results of Independent fleet management are clearly worse than the ones produced by the other management strategies. However, the fleet efficiency indicators (number of diversions, average ambulances empty travels and the overtime of paramedics) are better when using Independent fleet management. Even if the Reactive fleet management seems a rather simple strategy, it produces

Table 24.1 Results

	Management type		
	Independent	Without planning	Robust
Emergency demands			
Number of demands without transport	589.1	589.1	589.1
Number of demands with transport	1,762.2	1,762.2	1,762.2
Average response time (R.T.) in sec.	469.5	454.6	446.1
Percentage of demands such that R.T.$\leq S$	68.8	71.0	73.4
Percentage of demands such that R.T.$\leq S'$	85.1	86.9	88.3
Percentage of demands such that R.T.$> S'$	14.9	13.1	11.7
Transfer demands			
Number of demands	452.2	452.2	452.2
Sum of transport delays in sec.	2,397.7	473.1	661.0
Number of late demands	46.3	3.2	16.0
Ambulances			
Number of diversions	2,407.9	2,640.3	2,584.4
Number of redeployments	202.6	182.7	182.8
Number of ambulances redeployed	611.2	585.7	594.7
Percentage of average ambulances empty travels	22.8	24.6	24.5
Average overtime of paramedics	1,521.1	1,593.8	1,709.2

a good balance between the service quality of emergency demands and transfer demands: both the delays and the coverages are strongly improved with respect to the Independent fleet management case. Unfortunately, if we decrease the number of ambulances limited to perform the transfer transports, the coverage is slightly improved but the transfer transport delays are considerably increased. Finally, the Proactive fleet management strategy improves the service quality of the emergency demands but it deteriorates the delays of transfer demands.

24.6 Conclusion

This paper proposes new fleet management strategies to deal with two types of demands, transfer and emergency demands, in an integrated manner. We showed that these strategies can improve the quality of the service without increasing the number of ambulance in the fleet. The proposed Proactive fleet management strategy is very promising, and it is also simple to implement in a real setting. One of our future research questions concerns the consideration of the breaks in paramedics schedules. We also intend to formulate a single model for the two problems (DARP and relocation problem) as well as integrated solving approach to improve the results.

References

1. Alsalloum, O.I., Rand, G.K.: Extensions to emergency vehicle location model. Comput. Oper. Res. **33**, 2725–2743 (2006)
2. Andersson, T., Petersson, S., Varband, P.: Decision support for efficient ambulance logistics. ITN research report LiTH-ITN-R-2005-1, Linkopings Universiteit (2007)
3. Attanasio, A., Cordeau, J.F., Ghiani, G., Laporte, G.: Parallel tabu search heuristics for the dynamic multi-vehicle dial-a-ride problem. Parallel Comput. **30**, 377–387 (2004)
4. Başar, A., Catay, B., Unluyurt, T.: A multi-period double coverage approach for locating the emergency medical service stations in Istanbul. J. Oper. Res. Soc. **62**, 627–637 (2011)
5. Batta, R., Dolan, J.M., Krishnamurty, N.N.: The maximal expected covering location problem: revisited. Transp. Sci. **23**, 277–287 (1989)
6. Beaudry, A., Laporte, G., Melo, T., Nickel, S.: Dynamic transportation of patients in hospitals. OR Spectr. **32**, 77–107 (2009)
7. Bélanger, V., Ruiz, A., Soriano, P.: Déploiement et redéploiement des véhicules ambulanciers dans la gestion des services préhospitaliers d'urgence. INFOR. **50**, 1–30 (2012)
8. Beraldi, P., Bruni, M.E., Krishnamurty, N.N.: A probabilistic model applied to emergency service vehicle location. EJOR **196**, 323–331 (2009)
9. Brotcorne, L., Laporte, G., Semet, F.: Ambulance location and relocation models. EJOR **147**, 451–463 (2003)
10. Carpentier, G.: La conception et la gestion d'un réseau de service ambulancier. Mémoire de maîtrise, Université Laval (2006)
11. Church, R.L., ReVelle, C.S.: The maximal covering llocation problem. Pap. Reg. Sci. Assoc. **32**, 101–118 (1974)
12. Cordeau, J.F., Laporte, G.: A tabu search heuristic for the static multi-vehicle dial-a-ride problem. Transp. Res. B Methodol. **37**, 579–594 (2003)
13. Cordeau, J.F., Laporte, G.: The dial-a-ride problem: models and algorithms. Ann. Oper. Res. **153**, 29–46 (2007)
14. Daskin, M.S., ReVelle, C.S.: A hierarchical objective set covering model for emergency medical service vehicle deployment. Transp. Sci. **15**, 137–152 (1981)
15. Doerner, K.F., Gutjahr, W.J., Hartl, R.F., Karall, M., Reimann, M.: Heuristic solution of an extended double-coverage ambulance location problem for austria. EJOR **13**, 325–340 (2005)
16. Gendreau, M., Laporte, G., Semet, F.: Solving an ambulance location model by tabu search. Locat. Sci. **5**, 75–88 (1997)
17. Gendreau, M., Laporte, G., Semet, F.: A dynamic model and parallel tabu search heuristic for real-time ambulance relocation. Locat. Sci. **27**, 1641–1653 (2001)
18. Gendreau, M., Laporte, G., Semet, F.: The maximal expected relocation problem for emergency vehicles. J. Oper. Res. Soc. **57**, 22–28 (2006)
19. Glovert, F.: Future paths for integer programming and links to artificial intelligence. Comput. Oper. Res. **13**, 533–549 (1986)
20. Ingolfsson, A., Budge, S., Erkut, E.: Optimal ambulance location with random delays and travel times. Health Care Manag. Sci. **11**, 262–274 (2008)
21. Kergosien, Y., Lenté, C., Piton, D., Billaut, J.-C.: A tabu search heuristic for the dynamic transportation of patients between care units. EJOR **214**, 442–452 (2011)
22. Kiechle, G., Doerner, K.F., Gendreau, M., Hartl, R.F.: Waiting strategies for regular and emergency patient transportation. In: Operations Research Proceedings 2008, Augsburg, vol. 6, pp. 271–276 (2008)
23. Maxwell, M.S., Restepo, M., Henderson, S.G., Topaloglu, H.: Approximate dynamic programming for ambulance redeployment. INFORMS J. Comput. **22**, 266–281 (2010)
24. Parragh, S.N.: Introducing heterogeneous users and vehicles into models and algorithms for the dial-a-ride problem. Res. C Emerg. Technol. **19**, 912–930 (2011)
25. Parragh, S.N., Doerner, K.F., Hartl, R.F.: Variable neighborhood search for the dial-a-ride problem. Comput. Oper. Res. **37**, 1129–1138 (2010)

26. Rajagopalan, H.K., Saydam, C., Xiao, J.: A multiperiod set covering location model for dynamic redeployment of ambulances. Comput. Oper. Res. **35**, 814–826 (2008)
27. Repede, J.F., Bernardo, J.J.: Developping and validating a decision support system for location emergency medical vehicles in Louisville. EJOR **75**, 567–581 (1994)
28. Schmid, V.: Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming. EJOR **219**, 611–621 (2012)
29. Schmid, V., Doerner, K.F.: Ambulance location and relocation problems with time-dependant travel times. EJOR **207**, 1293–1303 (2010)
30. Taillard, E., Badeau, P., Gendreau, M., Guertin, F., Potvin, J.Y.: A tabu search heuristic for the vehicle routing problem with soft time windows. Transp. Sci. **31**, 170–186 (1997)
31. Toregas, C., Swain, R., ReVelle, C.S., Bergman, L.: The location of emergency service facilities. Oper. Res. **19**, 1363–1373 (1971)
32. Xiang, Z., Chu, C., Chen, H.: The study of a dynamic dial a ride problem under time dependant and stochastic environments. EJOR **185**, 53–551 (2008)