Journal of Industrial and Systems Engineering
Vol. 12, No. 1, pp. 95-128
Winter (January) 2019

Ambulance routing in disaster response scenario considering different types of ambulances and semi soft time windows

Hamid Tikani[1], Mostafa Setak[1*]

[1]*Department of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran*

hamid.tikani@yahoo.com, setak@kntu.ac.ir

Abstract

This paper studies the ambulance routing problem (ARP) in disaster situations when a large number of injured people from various locations require receiving treatments and medical aids. In such circumstances, many people summoning the ambulances but the capacity and number of emergency vehicles are not sufficient to visit all the patients at the same time. Therefore, a pivotal issue is to manage the fleet of ambulances to meet all the requests promptly and consequently mitigate human suffering. We considered three different categories of patients with various requirements. Moreover, the support ambulances are segmented into various classes based on their capabilities. A mathematical formulation is presented to obtain route plans with the aim of minimizing the latest service completion time among the patients. Since the patient's condition gets worse and becomes life threatening over the time, semi-soft time window constraint is incorporated to reflect the penalties on late arrivals using survival function. Since the presented model belongs to the class of NP-hard problems, two efficient meta-heuristic algorithms based on genetic algorithm and tabu search are proposed to cope with real size problems. The experiments show that the proposed model could present proper routes and adopt the types of ambulances with the patients' needs to increase the service quality. Moreover, the proposed metaheuristics are capable to find acceptable solutions for the problem in reasonable computational times.

Keywords: Ambulance routing problem, disaster response phase, survival function, vehicle classification, metaheuristic algorithms

1-Introduction

A disaster can be considered as any unprecedented incidents, which causes major damages, destruction, ecological disruption, loss of human life, human suffering, deterioration of health and health service on a scale adequate to warrant an extraordinary response from outside the affected area (Barbarosoğlu and Yasemin 2004). Some examples of disasters are earthquakes, hurricanes, tornadoes, fire, floods, blizzard, drought, terrorism, volcanic eruptions and generally all occurrences that have substantial devastating effects in terms of human lives and damage to a society's buildings and infrastructures. It has been widely investigated that the severity of a misadventure can be mainly influenced by the efficacy of the relief efforts during the response phase (Berkoune et al. 2012).

*Corresponding author
ISSN: 1735-8272,

Unquestionably, the response and recovery phase must be processed under acutely challenging conditions: Treatment capacity limitation (transportation vehicles, manpower and hospital capacity), damaged roadways or transportation infrastructure, together with uncertain information about number and locations of injured people (see e.g., Najafi et al. 2013 and Edrissi et al. 2015). Thereupon, it is necessary to initiate the logistics relief operations quickly and managing and controlling the efficient flows of reliefs, and services to meet the urgent requirements of the affected people. Consequently, there is an indispensable need for decision support tools that provide solutions to the underlying optimization problems instantly (Talarico et al. 2015).

Several studies address issues of locating, dispatching, and the fleet of ambulances as emergency medical services (EMS). The main concern of EMS is immediate patient care before arriving at the hospital. By growing the demands for EMS, it becomes one of the active research areas in the transportation and health care management. An efficient EMS system seeks to provide prompt medical cares to seriously sick or wounded people (Gholami-Zanjani et al. 2018). Among variant issues in the context of EMS, ambulance management problem can substantially improve healthcare systems. Herein, the short response time and assigning a suitable ambulance can be utilized to evaluate the EMS. In detail, any delay in treatment could affect the patients' conditions (Lam et al. 2015). Beside these complicated factors, ambulance planning in disaster events is more complex than normal conditions due to the casualties and lack of appropriate vehicles. Despite the importance of this issue, a few works study vehicle fleet routing problem in post-disaster (see Luis et al. (2012), Pedraza-Martinez and van Wassenhove (2012) and Talarico et al. (2015)). Research gaps in the existing literature are further discussed in the following sections.

In this regard, this study considers ambulance allocation-routing problem which strives to design a proper ambulance routing scheme in order to minimize the times of all rescue operations and the related penalties for late arrivals and inefficacies in post disaster. In detail, we provide a mathematical modeling for suitable allocation of ambulances and determining the routes of available ambulances. We consider different types of ambulances with various features. Moreover, the income requests are classified into three groups based on urgency and their severity. Proposed model strives to assess an optimal solution for routing procedure and presents the optimum service start and completion times for each patient based on their requirements. We also incorporate semi-soft time windows (SSTW) constraints to reflects the urgency of actions and a apply penalty for possible delays. The thresholds of the time window are determined using survival function for each group of patients. Since the proposed ARP should be solved repeatedly in disaster situations but the commercial exact solvers are not capable to solve the problem in an acceptable time, we developed two efficient meta-heuristic methods to find a near optimal solution for the problem in a reasonable computational time.

The remainder of the paper is as follows. In section 2 relevant literature is reviewed. In section 3, we define the problem in detail and notations are explained. Moreover, an illustrating example is presented to clarify the problem. Proposed mathematical model is presented and explained in section 4. The principles of the genetic algorithm (GA) and tabu search (TS) algorithm are discussed in section 5. Section 6 provides various numerical examples to test the proposed mathematical model and evaluate the performance of proposed algorithms. In section 7, a comprehensive sensitivity analysis is performed to determine the role and importance of different factors in the proposed model. Finally, in section 8, paper conclusions are summarized and some further remarks are presented.

2-Literature review

Brotcorne et al. (2003) and Farahani et al. (2012) introduced survey of models and algorithms in locating ambulances where the aim is to deployment or establishment some sites for the vehicles within an urban area. Under these circumstances, we can be sure that it is possible to reach the potential emergency sites in a certain response time. Gong and Rajan (2007) considered ambulance allocation and reallocation models in post-disaster. They proposed two iterative algorithms to optimize the makespan and the weighted total flow time, respectively.

Andersson and Värbrand (2007) studied ambulance dispatching based on the urgency of calls and the location of an ambulance to the site of an incident. Some various models are developed with the purpose of capturing realistic planning situations like traffic-dependent traveling times and congestion phenomena. For example, Schmid and Doerner (2010) studied travel times, which vary in the course of a day for an ambulance location problem. By considering such variations, the coverage fulfilled throughout a day by a certain deployment of ambulances changes dynamically that calls for relocations.

Knight et al. (2012) formulated a model to locate ambulances in order to maximize the overall expected survival probability of multi classes of patients. The patients are differentiated based on targeted response time and their medical states. Schmid (2012) considered emergency service providers that locate ambulances such that emergency patients can be reached in a time-efficient manner. In the proposed model after emerging a request, a vehicle needs to be dispatched instantly to the requests' site. After having served a request, the vehicle needs to be relocated to its next waiting location. In dispatching process, incoming emergency requests should assign to the ambulances. However, it can be solved in combination with the ambulance location problem. Toro-DíAz et al. (2013) developed a mathematical model that integrates location and dispatching decisions. They incorporated the queuing elements and congestion phenomena to the dispatching decisions by considering a fixed priority list for each customer.

Zhang et al (2015) proposed a patient transportation problem based on a real-world application. They formulated the problem as a multi-trip dial-a-ride problem and provided a modified memetic algorithm for solving the problem. Tlili et al. (2017) formulated the ambulance routing problem as an open VRP and a VRP with pickup and delivery. They proposed a cluster-first route-second method based on the petal algorithm and the particle swarm optimization to improve the emergency response-time of medical service providers. In addition to the mentioned researches, various studies in the literature address the transportation of patients and planning of health care services in not urgent situations. For example, transporting patients among hospitals, transporting a patient from home to a hospital or transportation of elderly people to their destination. Mentioned problem is known as Dial-A-Ride Problem (DARP) in the literature. For example in a recent study, Detti et al. (2017) formulated a multi-depot dial-a-ride problem in the healthcare application. They studied the problem under non-emergency situations considering different features such as heterogeneous vehicles, vehicle-patient compatibility, etc. See e.g., Parragh (2011), Parragh et al. (2012), Coppi et al. (2013) and Marcon et al. (2017). We also refer to study Nable et al. (2016) for recent researches and trends in emergency medicine systems.

In addition to the ground ambulance, some recent studies work on air medical transport industry. For example, Carnes et al. (2013) developed a planning tool to assign the requests to aircraft. Proposed model derives an optimized plan by with the purpose of minimizing the costs of system. Bozorgi-Amiri et al. (2017) presented an integrated model for locating the helicopter stations and helipads by considering uncertainty of demands. Shahriari et al. (2017) proposed a multi-objective model for air emergency medical systems. They utilized helicopters for delivering patients to hospitals in some cases because it is much faster than ground ambulances. The first objective maximizes the service level while the second one minimizes the maximum demand weighted transportation time.

In numerous studies, vehicle routing problem is applied for dispatching relief commodities in disaster situations. The general purpose of such studies is dispatching emergency packages like food, water, medications, tents, and other survival equipment from some distribution centers to affected areas considering limited capacities for transportation. Luis et al. (2012) presented a recent literature survey on this subject. Özdamar et al. (2004) formulated a mathematical model for emergency logistics planning in disaster conditions. Their proposed model is a hybrid mathematical model that integrates the vehicle routing problem with multi-commodity network flow problem. It determines the dispatch orders for located vehicles in different areas.

Yi and Özdamar (2007) integrated logistics support and evacuation operations in post disaster situations. They considered hospitals and some temporal emergency centers that wounded people have to be brought to them. Campbell et al. (2008) provided various models based on traveling salesman (TSP) and vehicle routing (VRP) for disaster response actions, which aims at minimizing the maximum arrival time and the average arrival time. He applied insertion and local search techniques to solve the problem. Özdamar and Demir (2012) coordinated vehicle routing and evacuation actions. They applied a hierarchical cluster and route procedure to solving the aggregated problem. Proposed model concerns a fast delivery of supplies and strives to allocate hospitals and warehouses to the demand cluster centers. Huang et al. (2012) focused on vehicle routing problem and allocation decisions in humanitarian relief operations. They described metrics performance for in relief efforts and concentrate on efficacy and equity to achieve a quick and sufficient distribution for all recipients. Rath and Gutjahr (2014) formulated a multi-objective location–routing model in relief operations. Presented model aims at minimizing strategic costs, operative costs, and uncovered demand. Finally, We the refer the reader to Caunhye et al (2012), Abounacer et al (2014), Camacho et al (2015), Özdamar and Ertem (2015) and Zheng et al (2015) for review and some recent works on the related subject.

It can be concluded from the literature that managing post-disaster actions is a substantial active scope of research. However, many researches are mainly focused on the distribution of supplies in response the phase. Nonetheless, researches in the field of ambulance routing for disaster response operations are taken into consideration recently. To the best of authors' knowledge, a few papers exist in the literature, studying ambulance fleet management in disaster response operations; see Pedraza-Martinez and van Wassenhove (2012) and Talarico et al. (2015).

According to the reviewed literature, this study provides three main contributions as follows:

1. This paper studies an ARP in disaster response scenario, where the patients are categorized into different groups based on the severity of injury and service requirements. We consider various shape of survival function for each group of patients. The functions are utilized to determine some thresholds values that reflect the corresponded emergencies in relief operations using semi-soft time window.

2. Ambulance is a transportation vehicle, which is used to provide medical care to patients outside of the health care centers and to transmit the patients to a hospital for further treatments. In the current study, we incorporate different types of ambulances with different features. The ambulances should be operated for the patients based on their capabilities.

3. Since the problem studied in the current paper is NP-hard. We present two meta-heuristic methods based on tabu search and genetic algorithm to solve the real-size problems in an efficient way. We investigate the performance of algorithms from various aspects.

In the subsequent section, the structure of the problem is explained in more details.

3-Problem statement

Proposed model is an extended formulation based on Talarico et al. (2015) to provide a decision support approach for the routing of ambulances in response to a disaster. It incorporates various types of ambulances with different capabilities and considers constraint on delays in serving patients using semi soft time windows. The major task of managing ambulances in a disaster response phase is to appoint first aid to slightly injured people and to transfer severely and moderately severe injured people to the hospitals with the appropriate vehicle.

Managing the plans of ambulances promptly aftermath of a disaster is abundantly complicated and confusing owing to the dynamics and uncertainty in nature of the relevant information that frequently changes over the period of time. Add to the complexity of managing this issue that the managers should consider the availability of ambulances, types of ambulances, the capacity of the nearby health centers and adopting the ambulance type with patient's needs. ARP in such situation can be considered as a static or a dynamic problem.

In the static version of the problem, at first, some property of emergencies are collected and afterward the vehicle routing problem is solved for this set of requests. While in the dynamic version, the routes of ambulances are updated by arriving new requests, which can decrease the response time. Whereas in dynamic approach communication with the ambulances is necessitous at all time, which might not be possible in a disaster condition and the rescue teams may feel under pressure to quickly take action during the operations.

Therefore, in this study, we consider a four-step response process to solve a static ARP. As can be seen from figure 1 the process is performed by a dispatching center that collects requests and manages ambulance operations. This process performs repeatedly to cover all emergency requests. The dispatcher collects several requests until a certain time limit has elapsed or a certain number of requests has been collected and then in a next step the requests should be categorized to different groups according to their severity. The availability of ambulances, the locations and types of them should be recognized. Finally, the priority of patients is taken into account to manage the ambulances. Nevertheless, the time spent for the first three steps is too short, because requests arrive in a short time in the case of a disaster event and the classification of requests can be carried out directly meanwhile answering an emergency call or automatically from the collected data.

The four-step response process is illustrated in figure 1. As can be seen from figure 1, the process is performed by a dispatching center that collects requests and manages ambulance operations. As mentioned earlier, the process performs repeatedly to serve all the demands. This paper focuses on the ambulance fleet management step of the response process and solves routing problem that occurs in the response phase. The routing scheme should consider the priority of patients together with assigning appropriate ambulances to them. Each delay in serving the patient imposes a penalty cost to the solution. These penalties are considered in the model using semi-soft time window, which is described in section 3.2 with more details.

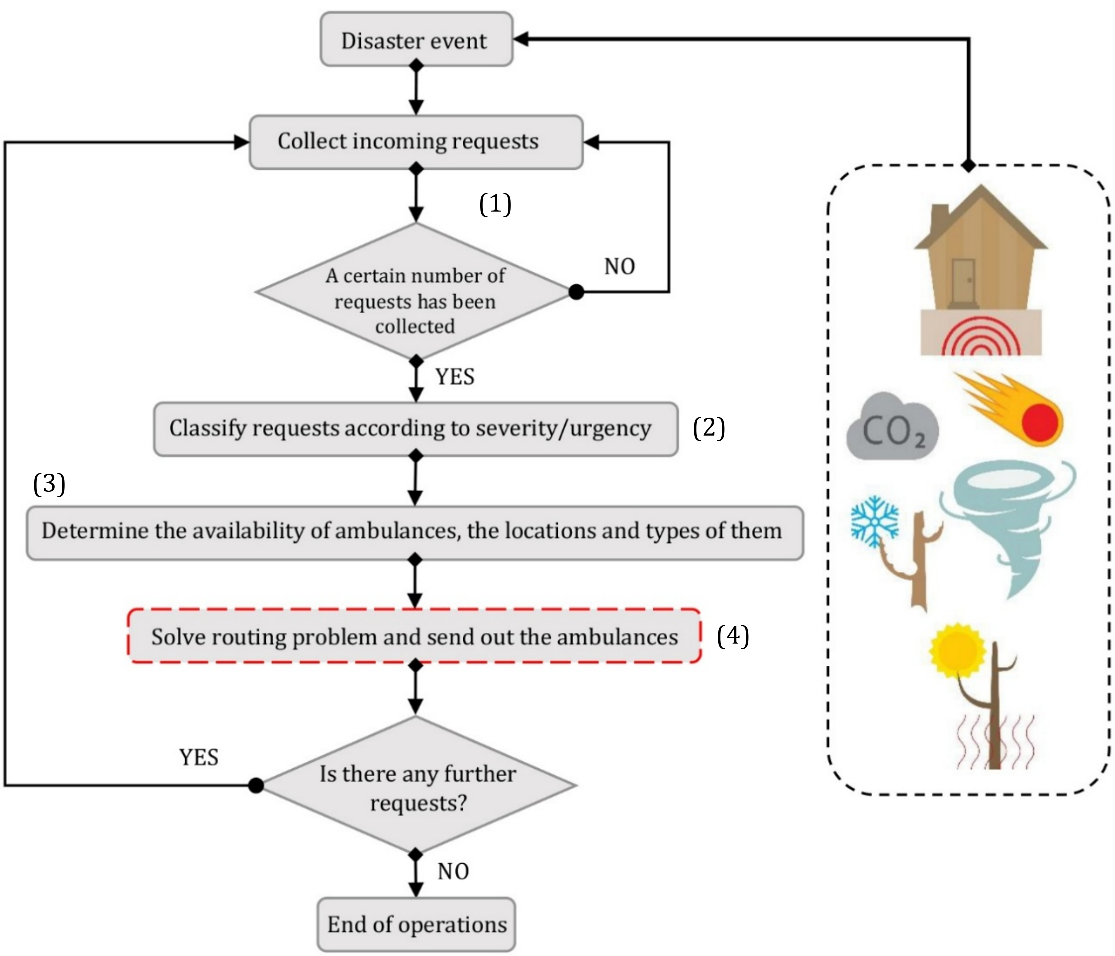


Fig 1. Disaster response flowchart

3-1-Patients grouping and ambulances classification

As described before, here we assume that each ambulance provides medical first aid so as to assist slightly injured people in the location of the patient. Meanwhile, emergency patients should accompany by the medical staff on their way to the hospital to visit skilled doctors in the hospital. Accordingly, we define three types of patients as follows:

- *Green code patient*: An individual with green code classification is slightly injured patient, which can be treated in the place.
- *Yellow code patient*: A person with yellow code classification is moderately severe injured individual, which should be transferred to the hospital for further treatments. Dedicated ambulance to this transportation does not need any special features.
- *Red code patient*: A person with red code classification is seriously injured and needs to be brought to a hospital by an ambulance, which equipped with the medical equipment at least for basic life support.

Figure 2 shows the process of patient segmentation.

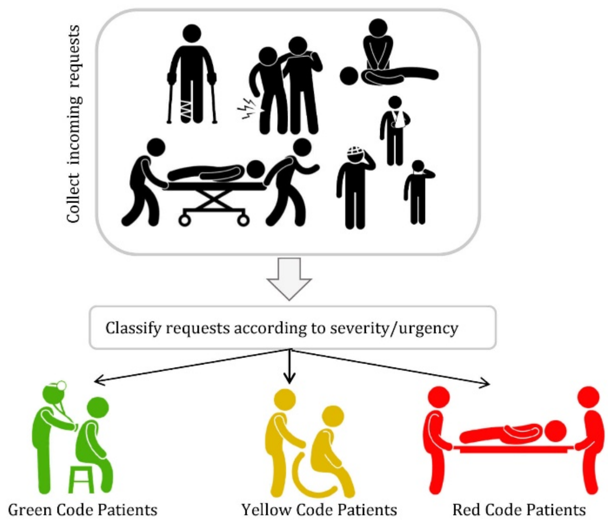


Fig 2. Patient segmentation process

There are various schemes for classifying patients and several systems for categorizing and prioritizing patients in emergency situations with limited time and scarce medical resources. We refer the reader to study Thomas and Elaine (2008) and Andersson and Peter (2007) for further information on the mentioned subject. Due to the different needs of patients, ambulances are customized to use in different situations. In a general definition, an ambulance is a transportation vehicle for carrying sick or injured people. In most cases, there are three general ambulance types with several subtypes. In the following, we presented four types of ambulances that the first three of them are more common based on various standards:

- Type I- Patient transport ambulance: these types of ambulances should have a basic specialized equipment for the first aid and nursing cares.
- Type II- Basic life support ambulance (BLS): these types of ambulances have to be equipped with the medical equipment. It is designed for patients who require medical transportation and continuous medical supervision.
- Type III- Advance Life Support Ambulance (ALS): These ambulances are operated by highly skilled and trained personnel and carry the essential medical equipment required to stabilize, treat and transport patients to a hospital.

- Type IV- Specialty care transport (SCT): These ambulances have been equipped to employ in a hospital to hospital transportation. SCT is urgent when a beneficiary's condition requires ongoing care that must be equipped by one or more health professionals.

When a disaster occurs, we face the lack of professional and skilled personals, in addition, ambulances Type III and Type IV are expensive and available scarcely and using them for transportation is costly. Therefore, we assumed that the manager confines the system to employs first two types of ambulances. In routing of ambulances, both ambulances, types one and two can be applied to visit green and yellow code patients. While red code patients must be visited by ambulance type II. Both of ambulances are able to carry one patient at a time and each patient is directly brought to a hospital after visiting by ambulance. The decisions about caring which patient to which hospital is part of the ARP and depends on the capacities and locations of hospitals. Due to the structure of the problem, an ambulance's route can be interpreted as a tour that begins at one hospital or medical center, serves one or more patients in a pre-specified sequence, and ends in a hospital that could be either the starting one or another hospital. The aim of the problem is to minimize the sum of the latest service completion time among the patients with different codes and related penalties for delays. Moreover, we applied weights for the latest completion times of the patient groups that utilized to determine which code patients should be served sooner. Presented weighted objective function supports such a tradeoff of the priority of a help request and the medical resources required for it. In the next subsection, we described that how the semi-soft time window is incorporated into the problem based on survival function to reflect the penalties for delays.

3-2-Semi-soft time windows constraint

Ambulance services play a vital role in the critical conditions and help many people with serious or life-threatening conditions. In general, an ambulance is supposed to reach the site within 10 minutes. Nevertheless, in disaster situations, due to the lack of vehicles and blocked roads, ambulances have some delays to reach the patients. To this end, we utilized survival function to determine the thresholds of semi-soft time windows constraints. Figure 3 shows a general form of the survival function. As can be seen from the figure, the treatment operations and presence of ambulances after a special time can become useless. Consider τ be the time taken for the emergency team to reach the patient. Let $\xi(\tau)$ represents the survival functions. As depicted in figure 3, any increase in (τ) yields to decrease the efficiency of emergency response. The effects of delays on patients' condition become more dramatic by passing the $\hat{\tau}$. The reverse S-shape in the figure implies that the initial units of time in post-disaster plays an important role in saving people. The value of $\xi(\tau)$ converges to *d* and follows the following function (Edrissi et al. 2013):

$$\xi(\tau) = ae^{b\tau^c} + d \quad a, c, d > 0, \quad b < 0, \quad a + d < 1, \quad d \ll a \tag{1}$$

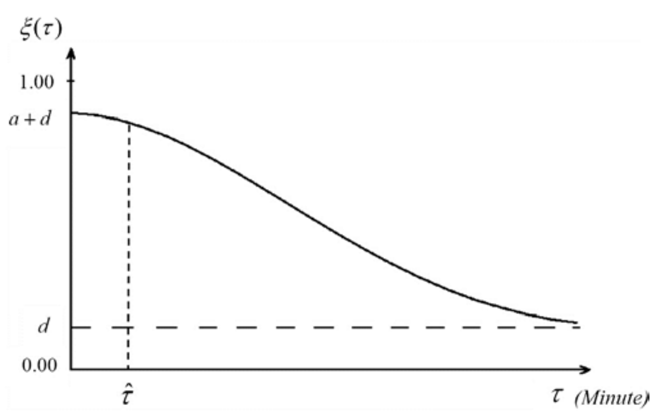


Fig 3. A general shape for survival function

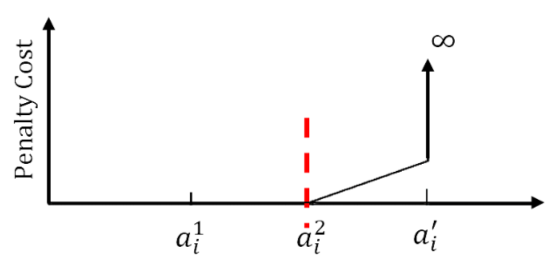


Fig 4. Penalty cost function for SSTW

Figure 4 reflects the procedure of imposing the penalties in semi-soft time windows constraint. In this approach, the latest possible service start time is by extended from a_i^2 to a_i' as shown in the Figure. In particular, we consider that from a_i^1 to a_i^2 patient's condition does not change significantly and between a_i^2 to a_i' the patient's condition gets worse and become life-threatening until a_i'. After $a_i'$the victims usually tend to perish because of a lack of medical aid at the right time aftermath of a disaster. The unit late arrival penalty cost PC is taken to impose the penalty of delay in serving a patient. Consider that the visiting time of patient i is b_i. Also, C represents unit late arrival penalty cost. The penalty cost of delay for patient i is defined as a function of b_i according to equation 2 (Qureshi et al. 2010).

$$\mathrm{PC} = \begin{cases} 0, & if\ b_i \le a_i^2 \\ C(b_i - a_i^2) & if\ a_i^2 \le b_i \le a_i' \\ \infty & if\ b_i > a_i' \end{cases} \tag{2}$$

In order to determine the soft and hard thresholds values for each code patients, we applied three various survival function for the patients, which represents the corresponded emergencies in operations. According to the figure 5, the slope of patient's status for a red patient is greater than other patients groups. For example, we suppose that if the patient's survival rate becomes less than 0.7 then a penalty should be imposed to the quality of operations (see figure 5). Moreover, fleet operations should be managed in the way that for each patient, patient's survival rate does not become less than 0.2. As can be seen from the figure the status of green code patient does not change considerably over the time. Therefore, there is no need to consider time window constraints for them. Nevertheless, for red and yellow code patients, we can formulate the interpretation as semi-soft time windows, which are shown in figure 6.

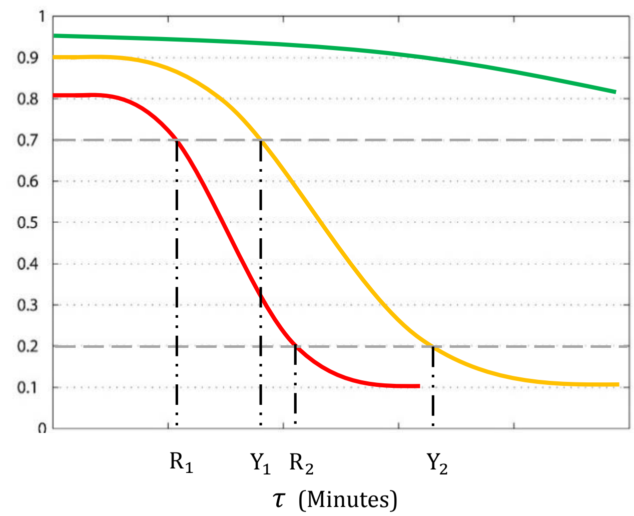


Fig 5. Survival function for various code patients

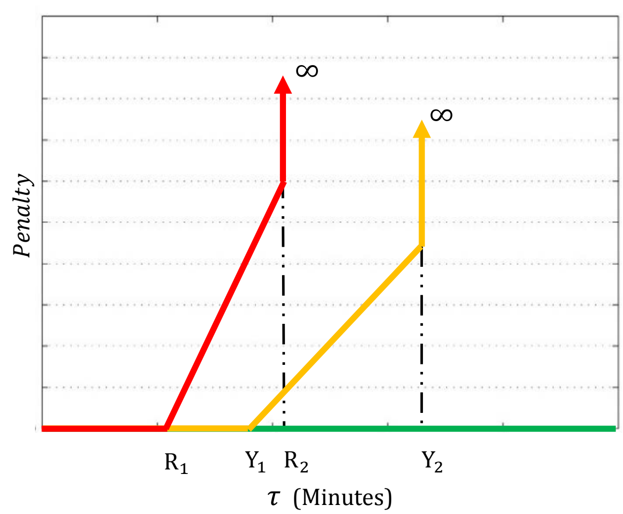


Fig 6. Semi soft time windows corresponded to figure 5

3-3- Illustrative example

In this section, we clarify the model by an illustrative example on a small artificial case. First, the notations and parameters of the problem are provided in table1.

Table 1. Notations and parameters used in the problem

Sets:	
$\mathcal{R}$	set of red code patients
$\mathcal{G}$	set of green code patients
Y	set of yellow code patients
$\mathcal{P}$	set of all patients, $\mathcal{P} = \mathcal{R} \cup \mathcal{G} \cup Y$
$\mathcal{H}$	set of hospitals
$\mathcal{K}_h$	set of ambulances that are initially located at hospital h
$\mathcal{K}$	set of all ambulances, $\mathcal{K} = \cup_{h \in \mathcal{H}} \mathcal{K}_h$, $k \in (Type\ I \cup Type\ II)$
$\mathcal{A}$	set of arcs in a problem, $\mathcal{A} = \{\mathcal{P} \times \mathcal{P}\} \cup \{\mathcal{H} \times \mathcal{P}\}$
Parameters:	
f_h^k	binary parameters, 1 if ambulance k is initially located at hospital h (i.e. $k \in \mathcal{K}_h$)
t_{ij}	travel time from i to j $(i,j) \in \mathcal{A}$
d_i	the service time of patient $i \in \mathcal{P}$
d_h^i	transfer time to drop off a red or yellow code patient i at hospital $h \in \mathcal{H}$, $i \in \mathcal{R} \cup Y$
c_h	capacity of hospital $h \in \mathcal{H}$
w_R	priority given to red code patients
w_Y	priority given to yellow code patients
w_G	priority given to green code patients
C_R	unit late arrival penalty for red code patient $i \in \mathcal{R}$
C_Y	unit late arrival penalty for yellow code patient $i \in$ Y
R_1	the latest possible service start time for red code patient without any penalty
R_2	the latest allowable service start time for red code patient
Y_1	the latest possible service start time for yellow code patient without any penalty
Y_2	the latest allowable service start time for yellow code patient
Decision variables:	
x_{ij}^k	binary, 1 if ambulance k serves patient i directly before patient j
b_i	visiting time of patient $i \in \mathcal{P}$
PR_i	the amount of late arrival of red code patient $i \in \mathcal{R}$
PY_i	the amount of late arrival of yellow code patient $i \in Y$
e_R	latest service completion time among all red code patients
e_Y	latest service completion time among all red code patients
e_G	latest service completion time among all green code patients

In the current instance, we considered three red code patients $R = \{r_1, r_2, r_3\}$ and two yellow code patients $Y = \{\mathcal{Y}_1, \mathcal{Y}_2\}$ and five green code patients$G = \{\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_4, \mathcal{G}_5\}$. In addition, two hospitals $\mathcal{H} = \{h_1, h_2\}$ are ready to service the red and yellow code patients, each has 3 capacities. At the beginning of the horizon one ambulances with type II a_3 is initially located at hospital h_2 and two ambulance with type I and type II, a_1 and a_2 are located at hospitalh_1. The locations of all patients and hospitals are depicted in figure 7. We suppose that the travel time t_{ij} between to node i and j equals to the Euclidean distance between them. The service times d_{I} are assumed to be 30 time units for red code patients $i \in \mathcal{R}$ and 10 time units for remain code patients$i \in G \cup Y$. Dropping off a red cod patient at a hospitals is set to 10 time units for red code patients and 5 time units for yellow code patients. The complete information of illustrative example is presented in table 2. Figure 7 indicates the potential routes for the ambulances.

Table 2. Input parameters for the illustrative example

	Input Parameters			
Patients	***Code patient***	***R***	***Y***	***G***
	Service time	30	10	10
	Latest possible service start time (without penalty)	90	120	∞
	Latest allowable service start time	200	300	∞
	Transfer time to drop off	10	5	0
Hospitals	***Hospital***	$\boldsymbol{H_1}$		$\boldsymbol{H_2}$
	Capacity of hospital	3		2
Ambulances	***Ambulance***	$\mathcal{K}(1)$	$\mathcal{K}(2)$	$\mathcal{K}(3)$
	Initial location	H_1	H_2	H_1
	Ambulance type	*Type II*	*Type I*	*Type II*

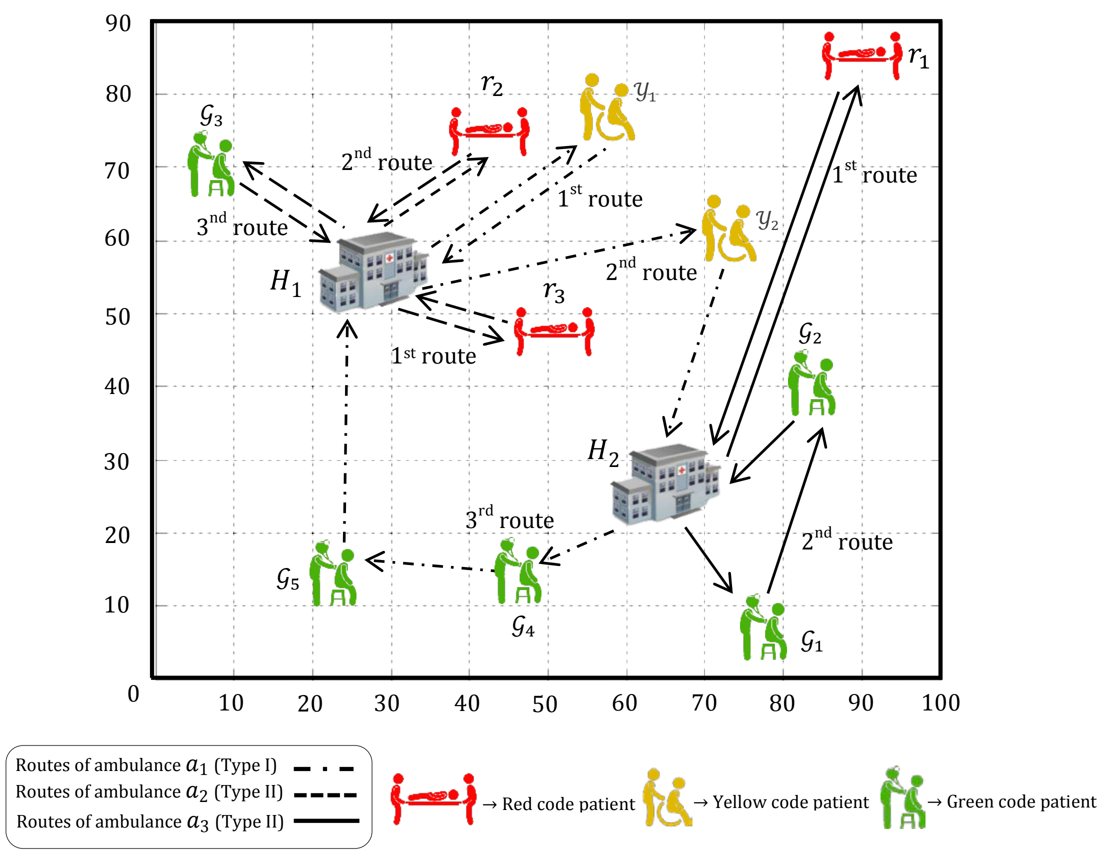


Fig 7. Illustration of all patients' and hospitals' locations and ambulances routes.

In the current solution, ambulances a_1 and a_2 perform three routes while ambulance a_3 performs two routes. Ambulance a_1 (type-I) begins its first route at hospital h_1, and then picks up yellow code patient

$\mathcal{Y}_1$ and brings this patient to hospital h_1. Afterward, it picks up yellow code patient $\mathcal{Y}_2$ and brings this patient to hospital h_2. In the following on its third route, it serves two green code patients $\mathcal{G}_4$ and $\mathcal{G}_5$. Ambulance a_2 (type-II) in the first rout depart from h_1 to picks up red code patient r_2 and brings it to hospital h_1. Subsequently, it transfers patient r_2 to hospital h_1. Finally, it serves patient $\mathcal{G}_3$ and returns to hospital h_1. Ambulance a_3 (type-II) begins from hospital h_2 brings patient r_2 to the hospital h_2. In the next route, it serves two green code patients $\mathcal{G}_1$ and $\mathcal{G}_2$ and goes back to the hospital h_2. In order to assess the quality of the obtained solution, the time-space diagram is shown in figure 8. Provided diagram exerts the positions of all ambulances over the time of disaster response. Blue rectangles in the figure represent dropping off a patient at a hospital. As can be seen from the figure 8 the latest drop off of red and yellow code patients at hospitals take place at time e_R = 182.21 and e_Y = 191.96, respectively. The latest completion time of serving a green code patient is e_G = 254.64. As mentioned earlier, a route can start and end at different hospitals, thereupon it would be accessible to achieve high quality solutions and serving patients as rapidly as possible. In the presented example, it happens for the routes of ambulance a_1.

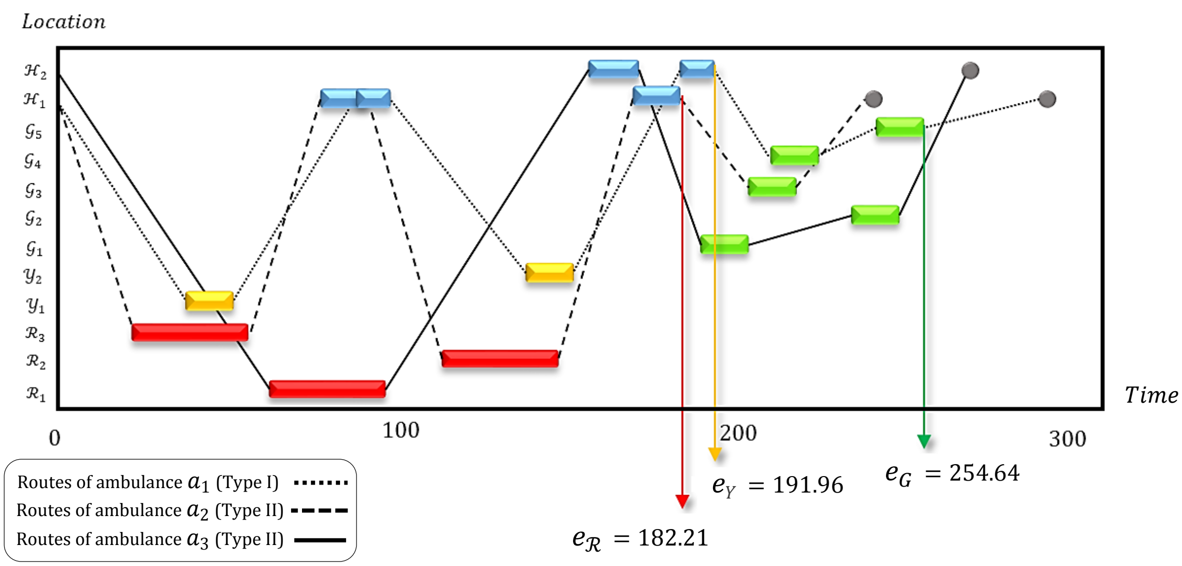


Fig 8. Time-space demonstration of the given instance

4-Model formulation

In this section, we propose the mathematical formulation for the ARP. As described earlier, the model is formulated using binary decision variable x_{ij}^k , which implies the sequence of visiting patients and takes value 1 if the ambulance k serves patient i directly before patient j and 0 otherwise. A positive continuous variable b_i is defined to represent arrival time at the patient i. Other notations and parameters are introduced in table 1. By these variables and notations, we can obtain the problem formulation as follows:

$$Min \;\; w_R \, .e_R \; + w_Y \, . \, e_Y \; + w_G \, . \, e_G \tag{3}$$

$$Min \;\; C_{\mathcal{R}} . \sum_{\forall i \in \mathcal{R}} PR_i \tag{4}$$

$$Min \; C_Y \sum_{\forall i \in Y} PY_i \tag{5}$$

$S.t:$

$$\sum_{j\in\mathcal{P}\cup\mathcal{H}} x_{hj}^k = f_h^k \qquad \forall h \in \mathcal{H}; k \in \mathcal{K} \tag{6}$$

$$\sum_{k\in\mathcal{K}}\sum_{j\in\mathcal{P}\cup\mathcal{H}} x_{ji}^k = 1 \qquad \forall i \in \mathcal{P} \tag{7}$$

$$\sum_{j\in\mathcal{P}\cup\mathcal{H}} x_{ji}^k = \sum_{j\in\mathcal{P}\cup\mathcal{H}} x_{ij}^k \qquad \forall i \in \mathcal{P};\ k \in \mathcal{K} \tag{8}$$

$$\sum_{h\in\mathcal{H}}\sum_{k\in\mathcal{K}} x_{ih}^k = 1 \qquad i \in R\cup Y \tag{9}$$

$$\sum_{i\in R\cup Y}\sum_{k\in\mathcal{K}} x_{ih}^k \le c_h \qquad \forall h \in \mathcal{H} \tag{10}$$

$$\sum_{k\in Type\ II}\sum_{i\in\mathcal{P}} x_{ij}^k = 1 \qquad j \in \mathcal{R} \tag{11}$$

$$b_i + d_i + t_{ij} \le b_j + \left(1 - \textstyle\sum_{K\in\mathcal{K}} x_{ij}^k\right).M \qquad \forall i \in \mathcal{G}\cup\mathcal{H};\ j \in \mathcal{P} \tag{12}$$

$$b_i + d_i + d_i^h + t_{hj} \le b_j\left(2 - \textstyle\sum_{k\in\mathcal{K}} x_{ij}^k - \sum_{k\in\mathcal{K}} x_{ih}^k\right).M \qquad \forall i \in R\cup Y;\ j \in \mathcal{P};\ h \in \mathcal{H} \tag{13}$$

$$0 \le b_i \le \mathrm{R}_2 \qquad \forall i \in \mathcal{R} \tag{14}$$

$$0 \le b_i \le Y_2 \qquad \forall i \in Y \tag{15}$$

$$b_i - R_1 \le PR_i \qquad \forall i \in \mathcal{R} \tag{16}$$

$$b_i - Y_1 \le PY_i \qquad \forall i \in Y \tag{17}$$

$$e_G \ge b_i + d_i \qquad \forall i \in G \tag{18}$$

$$e_R \ge b_i + d_i + \textstyle\sum_{k\in\mathcal{K}} x_{ih}^k \,.\, \left(t_{ih} + d_h^i\right) \qquad \forall i \in \mathcal{R}; h \in \mathcal{H} \tag{19}$$

$$e_Y \ge b_i + d_i + \textstyle\sum_{k\in\mathcal{K}} x_{ih}^k \,.\, \left(t_{ih} + d_h^i\right) \qquad \forall i \in Y; h \in \mathcal{H} \tag{20}$$

$$b_i \ge 0 \qquad \forall i \in \mathcal{P}\cup\mathcal{H} \tag{21}$$

$$x_{ij}^k \in \{0,1\} \qquad \forall (i,j) \in \mathcal{A}; k \in \mathcal{K} \tag{22}$$

$$PR_i, PY_i \ge 0 \qquad \forall i \in \mathcal{P} \tag{23}$$

$$e_R, e_Y, e_G \ge 0 \tag{24}$$

The objective function (3) strives to minimize the weighted sum of the latest service completion time among all patients. In the second and third objective functions (4) and (5) the total penalties related to semi-soft time windows of red and yellow code patients are minimized, respectively. In fact, we impose some penalties for inefficiency. Constraints (6) guarantee that each ambulance starts from the hospital where it is initially located. Constraints (7) state that each patient is visited exactly once by one of the ambulances. Constraints (8) ensure that an ambulance visiting a patient also has to leave that patient's location. Constraints (9) enforce that each red or yellow code patient is assigned to exactly one hospital. Constraints (10) specify that the number of patients, which assigned to a hospital should be less than the capacity of that hospital. Constraints (11) ensure that red code patients, need to be brought to a hospital by a type II ambulance. Constraints (12-13) express the arrival times of ambulances at the locations of patients. Constraints (14-15) state that the late arrival should not exceed the predetermined upper bound for red and yellow code patients, respectively. Constraints (16-17) determine the delays in serving the patients for red and yellow code patients, respectively. Constraints (18) state the latest service completion time e_G between all green code patients. Constraints (19-20) state the latest service completion time e_R and e_Y of all red and yellow code patients, respectively. It is notable that the completion service time of a red and yellow code patient equals to the time when the patient is dropped off at the dedicated hospital. Constraints (21-24) define the domains of the decision variables.

All variables e_G, e_R, e_R, PR_i and PY_i are based on time, while $C_{\mathcal{R}}.\sum_{\forall i\in\mathcal{R}} PR_i$ and $C_Y.\sum_{\forall i\in Y} PR_i$ express the related penalties. Thus, in order to solve the problem with single objective, we can pluralize the objective functions (4) and (5) with some special weights to objective function (3). To this end, we consider the following priorities:

- Reaching the red code patients before any delays is the first priority ($\sum_{\forall i\in\mathcal{R}} P(i)$).
- Reaching the yellow code patients before any delays is the second priority ($\sum_{\forall i\in Y} P(i)$).
- Finally, the model strives to find the best routing scheme that minimizes the total weighted completion times with regard to the first and second priorities.

According to the mentioned description, the objective function can be rewritten by using special weights E_R and E_Y as follows:

$$\text{Min } E_R(C_{\mathcal{R}}.\textstyle\sum_{\forall i\in\mathcal{R}} PR_i) + E_Y(C_Y \textstyle\sum_{\forall i\in Y} \ PY_i) + (w_R \,.\, e_R + w_Y \,.\, e_Y + w_G \,.\, e_G) \tag{25}$$

The objective function (25) can be considered as a hierarchical function where minimizing total penalties are prior to total weighted completion time. In particular, we assume that $E_{\mathcal{R}} \gg E_Y \gg 1$ (for example see Setak et al 2015).

The proposed ARP model contains the parameter M in the equations (12-13) that compute the arrival times at the patient locations. Here, we describe how to estimate a sufficient value for the parameter M. The general idea for this purpose is to calculate the maximum time T_i^{max}, which is required to reach and serve each patient. Herein, the worst case for service start time of red and yellow code patients are R_2 and Y_2, respectively. For $i \in \mathcal{R}$, T_i^{max} can be computed by $T_i^{max} = R_2 + d_i + \max_{h\in\mathcal{H}}\{t_{ih} + d_h^i\}$, including the latest allowable service start time for red code patient, the service time of patient i at the patient's location, and longest possible time to drop off the patient at any hospital. Similarly, for the yellow code patient the T_i^{max} can be computed by $T_i^{max} = Y_2 + d_i + \max_{h\in\mathcal{H}}\{t_{ih} + d_h^i\}$. For green code patients, $T_i^{max} = \max_{k\in\mathcal{G}\cup\mathcal{H}}\{t_{ki} + d_i\}$. In fact, the corresponding ambulance can be at any hospital or be located at another green code patient. Consequently, a sufficient choice for the parameter M can be $M = \sum_{i\in\mathcal{P}} T_i^{max}$.

5-Proposed method solutions

Vehicle routing problem is categorized to the NP-hard in the strong sense (Toth and Vigo 2002). Thereupon, it is not possible to solve the large or even medium size problems with exact methods in a reasonable computational time. In details, by enlargement of problem scale, the solution time increases non-polynomially. Consequently, we proposed two efficient meta-heuristic algorithms include GA and TS to tackle the problem in large scales. These two methods are explained in the next sub-sections in more details. We also applied a heuristic method to evaluate the performance of proposed algorithms.

5-1-Genetic Algorithm

Genetic algorithm is an efficient meta-heuristic search algorithm, which has emerged as a powerful robust optimization in finding the proper solutions in many real-world applications. The initial concept was proposed by Holland (1975). The algorithm inspired by the mechanisms of evolution and the survival of the fittest concept of Darwin. It imitates natural genetics to generate an initial population. Each population is consist of various individuals (chromosomes), which are encoded solutions of the original problem. The fitness of each individual is computed according to the objective function. Afterward, the individuals of a given population go through some evolution process contains selection, cross over and mutation operators.

The cross over and mutation operators are applied to create a new population with better features. The process of reproduction and generation replacement is continued until a well-defined stopping criterion is satisfied. The following subsections are dedicated to describe the solution encoding and genetic operators.

5-1-1-Chromosome representation

A chromosome (individual) in genetic algorithm presents an encoded solution for the problem. The representation of a solution for proposed vehicle routing is an integer string with the length of N that contains the sequence of patients in which N denotes the number of patients plus some delimitations. An example for encoding chromosome is depicted in figure 9. The zeros in the chromosome delimit the routes. Each route relates to one vehicle. The start point of each route is specified based on the initial location of the ambulance.

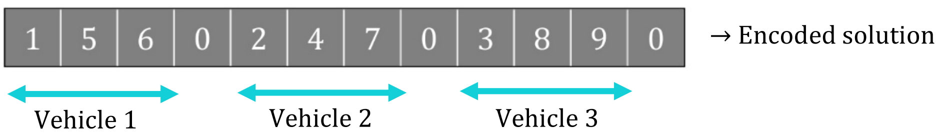


Fig 9. Sample structure of a chromosome

As can be seen from the figure 9, the presented structure only determine the sequence of patients for each vehicle. In the next step, this solution should be modified to specify that each red and yellow code patient should be assigned to which one of the hospitals based on their capacities. Here, we provide a process to change the chromosome structure to a solution, which satisfies the capacity of hospitals and assigned a proper ambulance according to the patient's condition. We named this algorithm with FS-CH-AT. The flowchart of proposed FS-CH-AT is shown in figure 10.

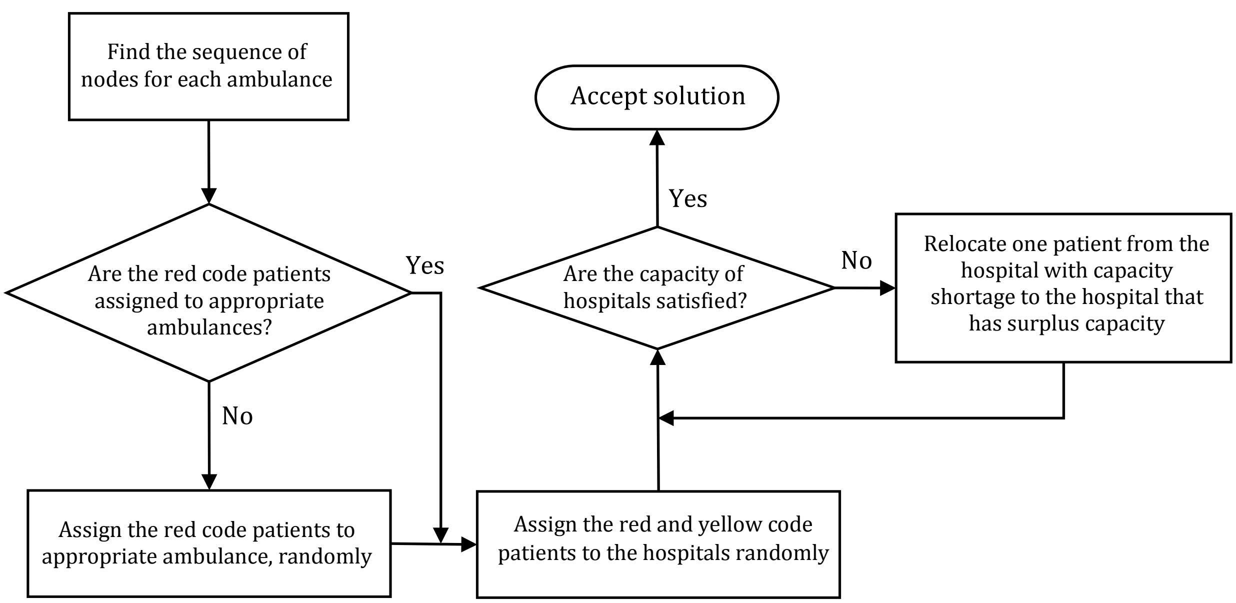


Fig 10. The process of producing a feasible solution with regard to the capacity constraints and ambulance types (FS-CH-AT)

In this approach, the algorithm first checks that if the red code patients are assigned to the proper ambulances. Otherwise, it relocates the red code patients to reach a solution that considers proper vehicles for each patient. Afterward, the algorithm specifies that each patient should bring to which hospital based on the capacities.

In order to clarify the process, consider an example with 7 patients in which the patient number 2 is red code, patients 3, 5 are yellow code, and patients 1,4,6,7 are green cod patients. Other parameters are like as table 2. A solution for the sequence of nodes for this special problem and corresponded decoded solution is provided in figure 11.

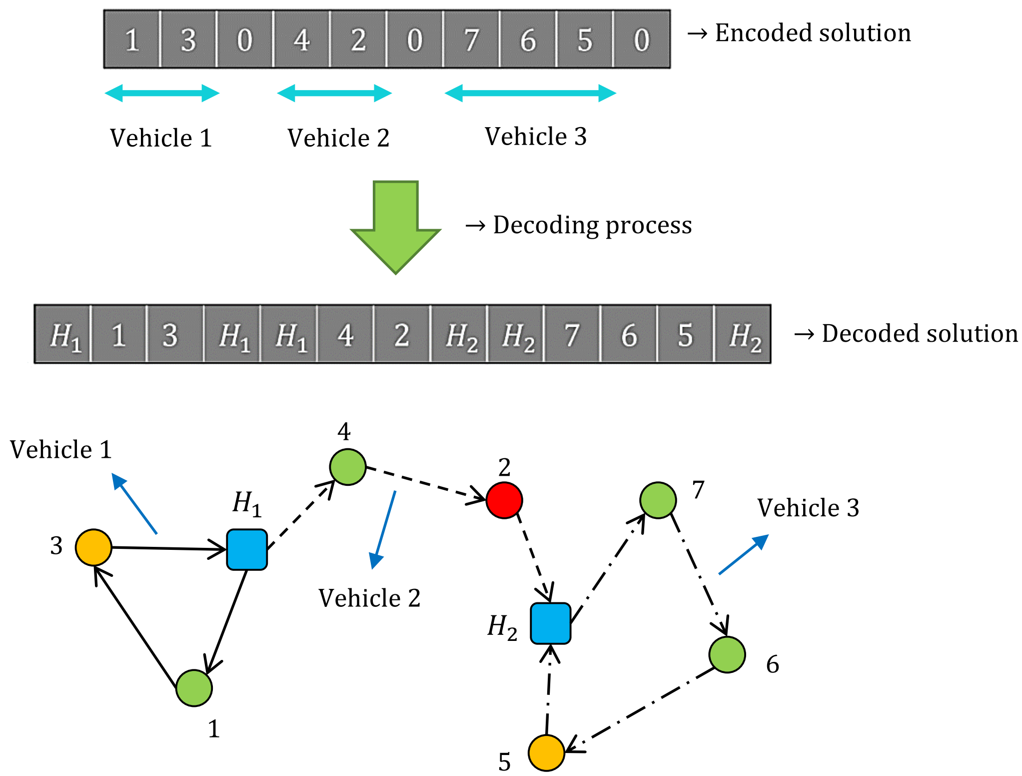


Fig 11. An illustrative example for decoding process

5-1-2-Crossover operator

In crossover operator, two chromosomes are randomly chosen among the population to act as parents. Then these individuals mate to create offspring. There are various methods to combine values of given parents in vehicle routing problem. Here, we applied classical order crossover to generate two offspring at each time. In this method, first, the genes with zero values in the chromosomes are deleted and then in the final step, these genes are located at their initial locations based on parent's chromosome. The procedure of proposed order crossover is as follows:

- Step 1: Delete the zero values from the chromosomes of parents.
- Step 2: Select a subset of genes from the first parent randomly.
- Step 3: Produce an offspring by replacing the subset of genes into the related positions in the offspring.
- Step 4: Delete selected genes of Step 2 from the second parent. The resulting substring determine the sequence of other nodes.

- Step 5: arrange the genes into the unfilled positions of the offspring from left to right based on sequence of genes, which is determined in Step 4. Put the zeros as delimitations in the offspring same as the selected parent in Step2.
- Step 6: Use the second parent and repeat Steps 1–5 to generate another offspring.

5-1-3-Mutation operator

Mutation operator avoids premature convergence in the population and diversifies the searching process. It applies some random exchanges to the chromosome to generate new individual. To this end, we present three mutation operators. The operators are explained as follows:

- Two-point swap: This operator selects two genes in the chromosome and swaps their locations to produce a new individual.
- Insertion: The operator selects one gene in the chromosome randomly and then relocate the position of the gene in the chromosome.
- Reversion: This operator selects two-point cuts in the chromosome randomly, and then reverse the cutting part.

The operations of above mutations are illustrated in figure. 12**.**

5-1-4-Penalized objective function using constraint relaxation

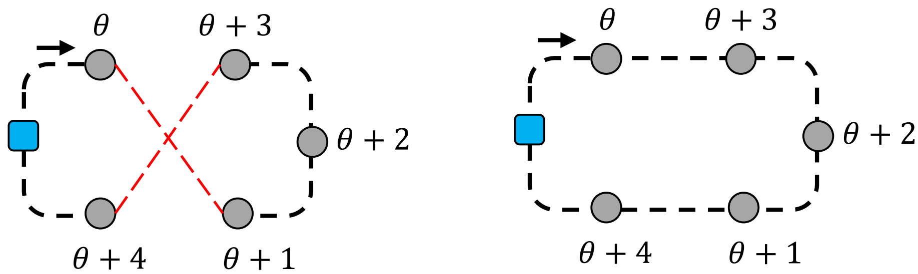
One of the applicable methods for overcoming the constraints in meta-heuristic algorithms is to allow the existence of infeasible solutions in the solution space by relaxing some of the constraints (Lai et al. 2016). In detail, we impose a penalty cost for infeasible solutions in the objective function. The achieved objective function named as penalized objective function. Here we relaxed the constraints, which related to the hard time windows and then, incorporate them into the objective function using specific coefficients. For example, let $\hat{X}$ represents an infeasible solution for the proposed ARP. Consider $\alpha(\hat{X})$ and $\beta(\hat{X})$ be the total violations of hard time window constraints for red and yellow code patients, respectively. Corresponded violations can be obtained by Equations (26-27) as follows:

$$\alpha(\hat{X}) = \sum_{i \in \mathcal{R}} \max\{(b_i - \mathrm{R}_2), 0\} \tag{26}$$

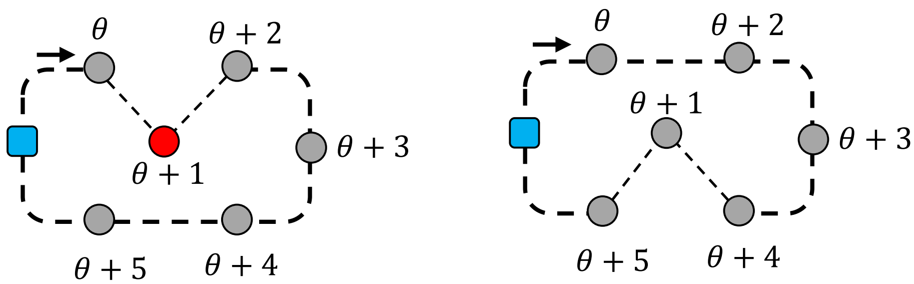
$$\beta(\hat{X}) = \sum_{i \in Y} \max\{(b_i - \mathrm{Y}_2), 0\} \tag{27}$$

Assume $C(\hat{X})$ be the value of objective function without consideration of penalties. Accordingly, the penalized objective function with constraint relaxation can be achieved as equations (28). In the equations (28), θ and φ are two non-negative coefficients, which are initialized by the user.

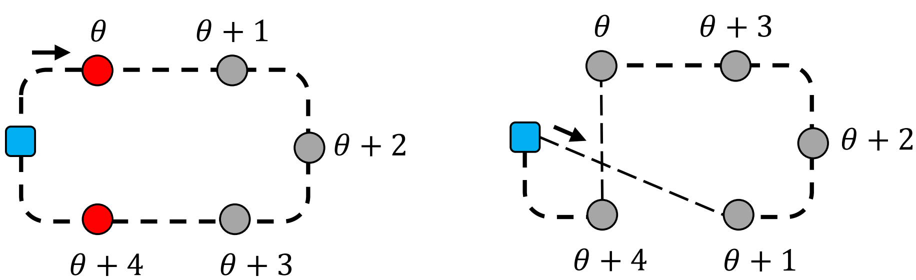
$$Z(\hat{X}) = C(\hat{X}) + \theta.\alpha(\hat{X}) + \varphi.\beta(\hat{X}) \tag{28}$$



(a). Two-point swap operator (swap between node $\theta + 1$ and $\theta + 3$)



(b). Insertion operator (insert node $\theta + 1$ to the fifth position in the sequence)



(c). Reversion operator (two-point cuts are θ to and $\theta + 4$)

Fig 12. The route structures on the left side present the examples for a primary solution and the route structures on the right side present the modified solution after mutation. The demand nodes are shown by circles and hospitals are shown by blue squares. Red elements are selected points in the chromosome.

The procedure of proposed GA is described in algorithm 1 in figure 13.

Algorithm 1. Pseudo code for the proposed genetic algorithm

```
Input: Parameter Popsize, mutation percentage, crossover percentage, non-negative coefficients θ and φ
Generate an initial random population
Decode the individuals of the initial population using FS-CH-AT
Evaluate each individual's fitness value using penalized objective function
Set Iter ← 1;
While (termination criteria is met) do
        Generate a random number p~ U(0,1)
        If  p ≤ 0.5
                Select parents from the population using Roulette wheel Selection
        Else
                Select parents from the population using Tournament Selection
        End If
        Apply crossover operator to generate the offsprings
        Generate a random number μ ~ U(0,1)
        If  μ ≤ 0.33
                Perform mutation using Two-point swap operator
        Else If  0.33 < μ ≤ 0.66
                Perform mutation using Reversion operator
        Else
                Perform mutation using Insertion operator
        End If
        Decode the generated individuals using FS-CH-AT
        Evaluate each individual's fitness value using penalized objective function
        Create new population using elitism
        Set Iter ← Iter + 1
End While
Return The best individual in all generations
```

Fig 13. Pseudo code for the proposed genetic algorithm

5-2-Tabu search

The tabu search method is a meta-heuristic technique, which is presented by Glover (1989, 1990). It employs an adaptive memory called tabu list to evade from the earlier local optimum. The algorithm seeks the solution space by iteratively evaluating some neighbor solutions. These solutions are generated with simple local modifications to the existed solution. The flexible memory collects the employed operations and applies stored information during the searching process to forbid recently visited solutions for a specified tenure (Bräysy and Gendreau 2002).

Basic elements of TS contain neighborhood-generating mechanism, choice of initial solution and tabu list. The length of tabu list, the number of iteration and the number of searched solutions at each iteration is specified using Taguchi method in section 6.1. The mechanism of generating neighborhoods and creating an initial solution is expressed as follows:

Neighborhood-generating mechanism: On each iteration of TS, the algorithm should employ the current solution to generate and evaluate some new neighbor solutions. These neighbors usually obtained by a single-movement through the current solution (Brandão 2016). Proposed algorithm employs three strategies includes swap, insertion, and reversion. These operators are described earlier in section 5.1.3.

The choice of an initial solution: In order to create an initial solution for the problem in TS, we applied nearest-neighbor heuristic (NNH) method. NNH begins to construct a route with the closest node to the depot. In the subsequent iterations, the algorithm seeks for the nearest neighbor demand to the last demand node and adds it to the route. In this problem, we applied the NNH method that starts from one of the hospitals randomly. In fact, we only create a large tour and then some zeros (according to the number of vehicles) are located in the sequence randomly to delimit the routes. Then the obtained solution go through the algorithm FS-CH-AT (see figure 10) and finally send to TS.

The decoding process of solutions and the designated objective function in this algorithm is like as proposed GA. The Pseudo code of proposed TS is presented in algorithm 2 in figure 14.

Algorithm 2. Pseudo code for the proposed tabu search algorithm

Input: *number of iterations, number of neighbors, size of the tabu list, non-negative coefficients θ and φ*
Set $Iter \leftarrow 1$;
Generate an initial solution using NNH algorithm (a sequence of nodes → S)
Run FS-CH-AT to obtain a solution with regard to the capacity constraint and ambulance type
Put $S_{best} \leftarrow S$ and calculate the penalized objective function $Z(S_{best}) \leftarrow C(S_{best}) + \theta.\alpha(S_{best}) + \varphi.\beta(S_{best})$
  While (stopping criteria is not satisfied) **do**
      Create neighborhoods (N) by randomly using swap or reverse or insertion strategies
      Choose the best solution among generated neighborhoods $\bar{S} \in N$ such that:
        → the sequence $\bar{S}$ that minimizes the penalized objective function
        → the sequence $\bar{S}$ should be non-tabu
      Calculate the violations $\alpha(\bar{S})$ and $\beta(\bar{S})$ and estimate the $Z(\bar{S})$
      Update the tabu list (let the used move to be tabu for τ iterations)
      If the solution $\bar{S}$ is better than S_{best} **then**
          Put $S_{best} \leftarrow \bar{S}$
          Put $Z(S_{best}) \leftarrow Z(\bar{S})$
      End If
      Set $Iter \leftarrow Iter + 1$
  End While
Return S_{best} and $Z(S_{best})$

5-3- Greedy heuristic method

In order to evaluate the performance of proposed method solutions and provide a comprehensive comparison among the presented approaches, in this section, we provided a greedy heuristic approach to solve the proposed ARP. The presented greedy heuristic algorithm contains three phases includes route construction, obtaining a feasible solution with regard to capacity and ambulance type and improvement phase. There exist various construction heuristic method in the literature to generate initial solutions for VRPs. Here, we applied the nearest-neighbor heuristic (Solomon, 1987) to find an initial sequence of nodes. This method is explained earlier in section 5.The three main phases of the greedy algorithm can be described as follows:

Step 1. Generate an initial solution for the problem using the nearest-neighbor heuristic.

Step 2. Run the algorithm FS-CH-AT to obtain a feasible solution with regard to the capacity of hospitals and ambulance type

Step 3. Improve the achieved solution of **Step 2** using 2-opt optimization method (by reallocating demand nodes).

In step 3, 2-opt exchange algorithm is applied to improve the solution. The 2-opt is a local search method which takes a solution and strives to modify it in order to reach an alternative solution with lower cost. In detail, it relocates the positions among all possible pairwise of patients served by individual ambulance and checks if an overall improvement in the objective function is possible.

6-Computational results

In this section, we validate the performance of the presented model and report the results of various computational experiments. Moreover, we assess the efficiency of proposed metaheuristics. Since no benchmark instances are presented around the ARP in the literature, we inevitably generate various instances to investigate the problem. The computational experiments contain two parts. At first, in section 6.1 a Taguchi design of experiments is performed to determine more effective parameters for the metaheuristics. In section 6.2, we assess the performance of the proposed algorithms. And a computational comparison between different methods is also provided in this section. In addition, in section 7 some sensitivity tests are performed to investigate the relationship between the problem structures and achieved solutions.

In order to evaluate the proposed mathematical model and the meta-heuristic methods in this paper, various instances with different planning situations are generated. In particular, the examples have different numbers of red, green and yellow code patients, ambulances with various types, and different numbers of hospitals with variant capacities. For each example, the locations of hospitals and patients are randomly placed in an area of size 100×200. Moreover, the travel times correspond to the Euclidean distance. Table 3 represents the features of sample problems. The service duration relates to red and yellow code patients are set to 30 and 10, respectively. The needed time to dropping off red and yellow code patients are set to 10 and 5, respectively. The capacity of hospitals and the availability of ambulances with various types are shared randomly among the hospitals. Each ambulance is able to carry one red or yellow code patient at a time. These patients should pick up with a proper vehicle and must be brought to a hospital, directly. The decision about assigning a patient to which hospital is a part of the ARP and should be performed based on related capacities.

Table 3. Features of the sample problems.

Instance NO	#of nodes	#of Hospitals	#of ambulances	# of patients			Weights		
1	6	1	2	1	2	3	0.60	0.30	0.10
2	6	1	2	1	2	3	0.45	0.35	0.20
3	8	2	3	2	2	4	0.60	0.30	0.10
4	8	2	3	2	2	4	0.45	0.35	0.20
5	10	3	3	2	3	5	0.60	0.30	0.10
6	10	3	3	2	3	5	0.45	0.35	0.20
7	12	4	4	3	4	5	0.60	0.30	0.10
8	12	4	4	3	4	5	0.45	0.35	0.20
9	15	4	5	4	5	6	0.60	0.30	0.10
10	15	4	5	4	5	6	0.45	0.35	0.20
11	20	5	5	5	7	8	0.60	0.30	0.10
12	20	5	5	5	7	8	0.45	0.35	0.20
13	25	5	6	6	8	11	0.60	0.30	0.10
14	25	5	6	6	8	11	0.45	0.35	0.20
15	30	6	7	7	10	13	0.60	0.30	0.10
16	30	6	7	7	10	13	0.45	0.35	0.20
17	35	6	7	9	12	12	0.60	0.30	0.10
18	35	6	7	9	12	12	0.45	0.35	0.20
19	40	6	8	10	13	17	0.60	0.30	0.10
20	40	6	8	10	13	17	0.45	0.35	0.20
21	50	6	8	12	16	22	0.6	0.30	0.10
22	50	6	8	12	16	22	0.45	0.35	0.20
23	60	7	9	15	20	25	0.6	0.30	0.10
24	60	7	9	15	20	25	0.45	0.35	0.20
25	70	7	10	17	24	29	0.60	0.30	0.10
26	70	7	10	17	24	29	0.45	0.35	0.20
27	80	8	10	20	26	34	0.60	0.30	0.10
28	80	8	10	20	26	34	0.45	0.35	0.20
29	90	9	10	22	30	38	0.60	0.30	0.10
30	90	9	10	22	30	38	0.45	0.35	0.20
31	100	10	12	25	33	42	0.60	0.30	0.10
32	100	10	12	25	33	42	0.45	0.35	0.20

All provided instances are performed on an Intel core i5-3337U (1.8 GHz) with 6GB of RAM. In order to solve the problem with exact method, we applied GAMS software with ILOG CPLEX 12.5 64-Bit optimization routines. Meanwhile, due to the NP-hardness of the problem, only the small-sized instances can be solved using MIP solver CPLEX. Thereupon, the heuristic and metaheuristics algorithms are coded in MATLAB 2014 software to solve the problem in a reasonable computational time. Before solving the problem, in the next subsection, we calibrate the parameters of algorithms to ensure the best performance.

6-1- Parameter tuning

Calibration of parameters on each algorithm has a great impact on the quality of solutions. In this paper, we employed the Taguchi experimental design method to determine more effective factors. Taguchi introduced the mentioned method in the early 1960s in which the orthogonal arrays are utilized for evaluation of a large number of factors with a few experiments. In particular, it seeks to minimize variances of quality characteristics acquired from S/N ratio (Roy 2010). Quality characteristic in the current paper is evaluated by relative percentage deviation (RPD), which changes the obtained objective values to non-scale data. Hence, we choose "the smaller-the better" type (Tikani et al. 2018). RPD is computed by the following equation:

$$RPD = \frac{|F_i - F_{best}|}{|F_{best}|} \times 100 \tag{29}$$

Where F_{best} and F_i are the best achieved objective values for one special instance and the objective value achieved for the ith replication, respectively. Moreover, the following formula can be used for “the smaller-the better” characteristic in Taguchi method:

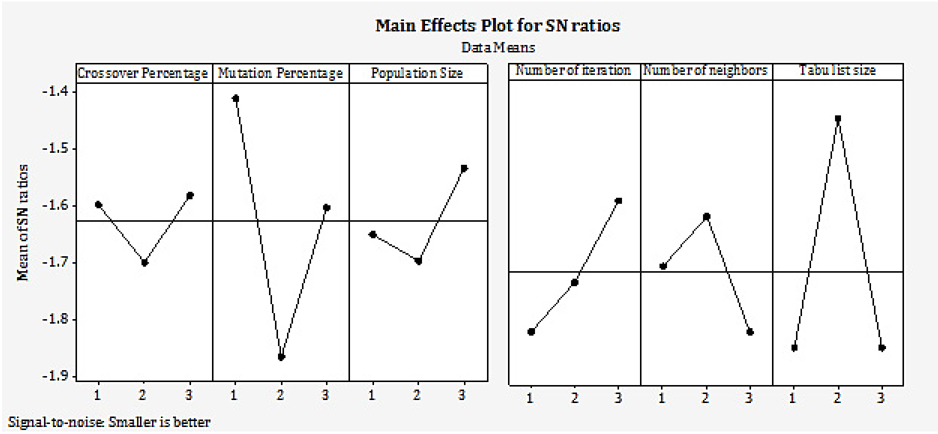
$$S/N = -10 \log(\frac{1}{n}\sum_{i=1}^{n} y_i^2) \tag{30}$$

Where y_i denotes the response value in the ith trial and n represents the total number of experiments. In our problem, we consider three factors that have salient effects on the algorithms with three levels. These levels are listed in table 4. Accordingly, the L9 design of the Taguchi method is performed by using the Minitab 16.2 software.

Table 4. Considered Levels of parameters for meta-heuristics

	Parameter	Level 1	Level 2	Level 3
GA	Crossover percentage	0.60	0.70	0.80
	Mutation percentage	0.10	0.15	0.20
	Population size	50	75	100
TS	Maximum number of iterations	10N	15N	20N
	Number of neighbors	3N	4N	5N
	Tabu list size	N/2	N/3	N/6

The achieved mean S/N ratio for each level of control factors is depicted in figure 15. Larger values of the S/N ratio in each graph indicate the desired levels.



(a) (b)

Fig 15. Mean S/N ratios for the proposed GA (a) and TS (b)

6-2-Experimentation

In the current section, the results of solving the problem with different method solutions are reported. The studied examples are based on table 5. In addition to the parameters, which are specified by Taguchi method; the stopping criteria of the genetic algorithm is designated proportioned to the sample size. In particular, GA terminates if the best-obtained solution does not change after $MAX = (2 \times Sample\ size)$ iterations. Furthermore, we consider maximum running time of 1 hour for CPLEX method and meta-heuristics. While the stopping criteria for the proposed heuristic method is running time of 2 hours. In table 5, the first column shows the instance's number. The optimum solutions obtained by CPLEX are listed in the second column if existed otherwise it is marked with a dash. The related computational times are also presented in the next column. We run the algorithms 10 times then the best-obtained value of the proposed algorithm and average total time needed to terminate the algorithm is given in column *Time* (based on second). The mark × in the last three columns indicates that the heuristic method was not able to find a feasible solution due to the time window constraints. The relative percentage gap is also utilized to evaluate the quality of solutions using the following equation:

$$GAP = \frac{Sol_{Alg} - Sol_{Best}}{Sol_{Best}} \tag{31}$$

Where Sol_{Alg} indicates the best solution achieved by the algorithm during the executions and Sol_{Best} is the best-found solution among all methods.

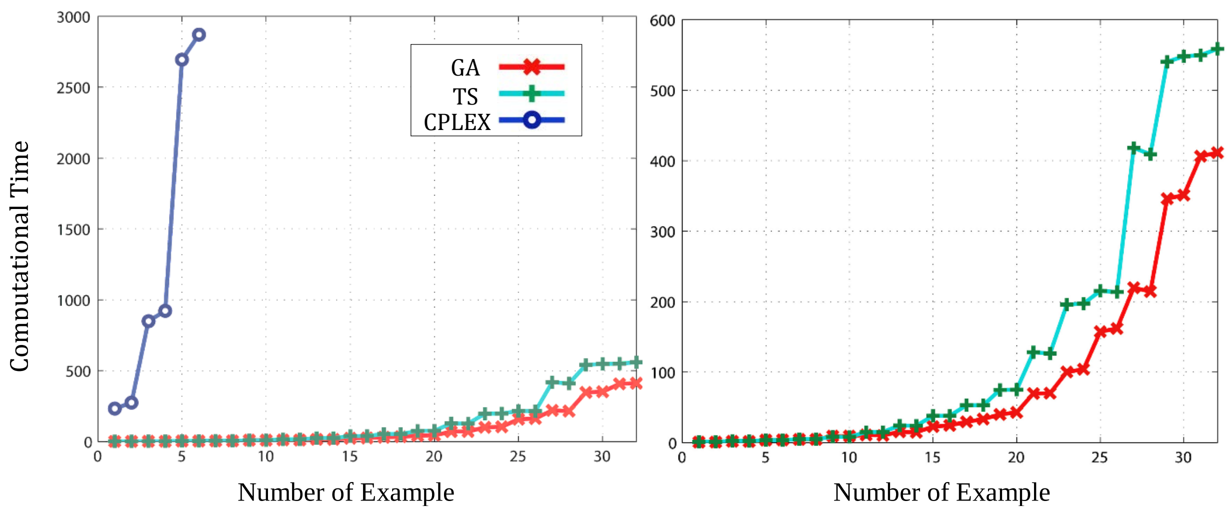
Table 5. Results of CPLEX and the proposed algorithms

Instance NO	CPLEX	Time	GA	Time	GAP	TS	Time	GAP	Heuristic	Time	GAP
1	215.346	230.80	215.346	0.673	0	215.346	1.189	0	226.939	0.0081	0.0538
2	190.318	274.19	190.318	0.719	0	190.318	1.203	0	239.476	0.2072	0.2583
3	205.080	849.17	205.080	1.758	0	205.080	2.502	0	221.054	2.0179	0.0779
4	210.64	920.34	210.64	1.694	0	210.64	2.176	0	230.526	4.0164	0.0944
5	220.863	2695.13	220.863	3.677	0	220.863	3.579	0	298.716	30.028	0.352
6	236.941	2870.24	236.941	3.779	0	236.941	3.378	0	275.244	22.701	0.1617
7	_	_	160.028	4.490	0.0035	159.473	4.908	–	258.136	66.030	0.6187
8	_	_	161.200	4.616	0	161.200	4.9571	–	203.153	78.046	0.2603
9	_	_	164.89	9.263	0.0314	159.875	8.3045	–	200.101	140.51	0.2516
10	_	_	173.024	9.177	0.0159	170.312	8.0693	–	208.862	125.16	0.2263
11	_	_	225.18	10.7501	0.0757	209.328	14.870	–	229.207	302.19	0.0950
12	_	_	236.639	10.156	0.0189	232.255	15.027	–	350.742	194.39	0.5102
13	_	_	221.109	15.340	0.0352	213.590	24.211	–	316.954	296.17	0.4839
14	_	_	242.642	14.660	0.0661	227.59	24.017	–	419.926	477.19	0.8451
15	_	_	206.172	22.757	0.1044	186.687	38.137	–	360.679	916.40	0.9320
16	_	_	212.584	24.461	0.0566	201.197	38.153	–	375.976	1304.9	0.8687
17	_	_	241.129	29.493	0.0354	232.886	53.081	–	425.063	4016.5	0.8252
18	_	_	278.356	33.411	0.0475	265.721	53.119	–	496.306	5640.1	0.8678
19	_	_	341.128	40.025	0.0587	322.23	74.971	–	506.781	7200	0.5728

Table 5. Continued

Instance NO	CPLEX	Time	GA	Time	GAP	TS	Time	GAP	Heuristic	Time	GAP
20	–	–	340.419	43.027	0.0605	321.012	75.017	–	527.232	7200	0.6424
21	–	–	411.584	69.588	0.0540	390.512	128.276	–	642.921	7200	0.6464
22	–	–	415.429	70.110	0.0195	407.501	126.419	–	647.519	7200	0.5890
23	–	–	339.853	100.188	0.0892	312.032	195.885	–	×	×	×
24	–	–	344.130	104.150	0.0083	341.30	197.257	–	×	×	×
25	–	–	388.425	157.485	0.0412	373.073	215.190	–	×	×	×
26	–	–	340.733	161.910	0.0521	323.854	213.814	–	×	×	×
27	–	–	360.304	219.749	0.0359	347.813	418.298	–	×	×	×
28	–	–	390.271	214.390	0.0141	384.837	409.090	–	×	×	×
29	–	–	394.330	346.180	0.0343	381.265	540.293	–	×	×	×
30	–	–	460.171	351.203	0.0149	453.423	548.0458	–	×	×	×
31	–	–	400.191	406.391	0.0570	378.593	549.645	–	×	×	×
32	–	–	434.950	411.431	0.0070	431.894	558.629	–	×	×	×
Average				90.5219	0.0324		142.678	0		1.9280e+03	0.4651
Maximum				411.4310	0.1044		558.629	0		7200	0.932

It is clear from the table 5 that TS is superior to other methods in all instances with regard to the quality of solutions. The behaviors of these algorithms are more analyzed in figures 16–18. Refer to table 5, variation in weights of patients does not make a meaningful effect on the total computational time. As can be seen from the table, CPLEX method was only able to solve the problem with only 10 number of nodes in less than one hour. The exponential growth in computation time of exact solution using CPLEX solver in compare with the total required time of GA and TS is depicted in figure 16(a). Moreover, figure 16 (b) demonstrates a more detailed comparison between proposed meta-heuristics.



(a). Comparing computational time of meta-heuristics

(b). Comparing computational time of exact solution method with meta-heuristics

Fig 16. Comparison the computational time of various method solutions

Furthermore, the relative gaps of solutions obtained by GA and the heuristic method according to proposed TS are shown in figure 17. It shows that the heuristic method was not capable to find proper solutions in compare with meta-heuristic methods.

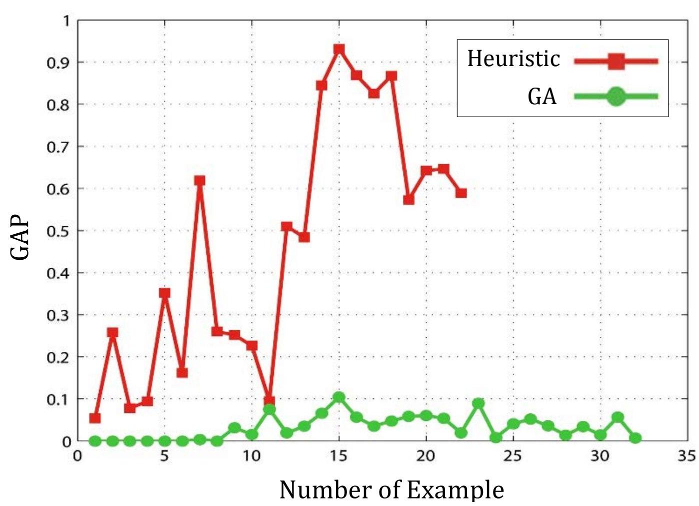


Fig 17. The gap proposed GA and heuristic for various instances

The convergence curves corresponded to the best optimization execution of each meta-heuristic algorithm for the examples 16 and 26 are presented in figure 18.

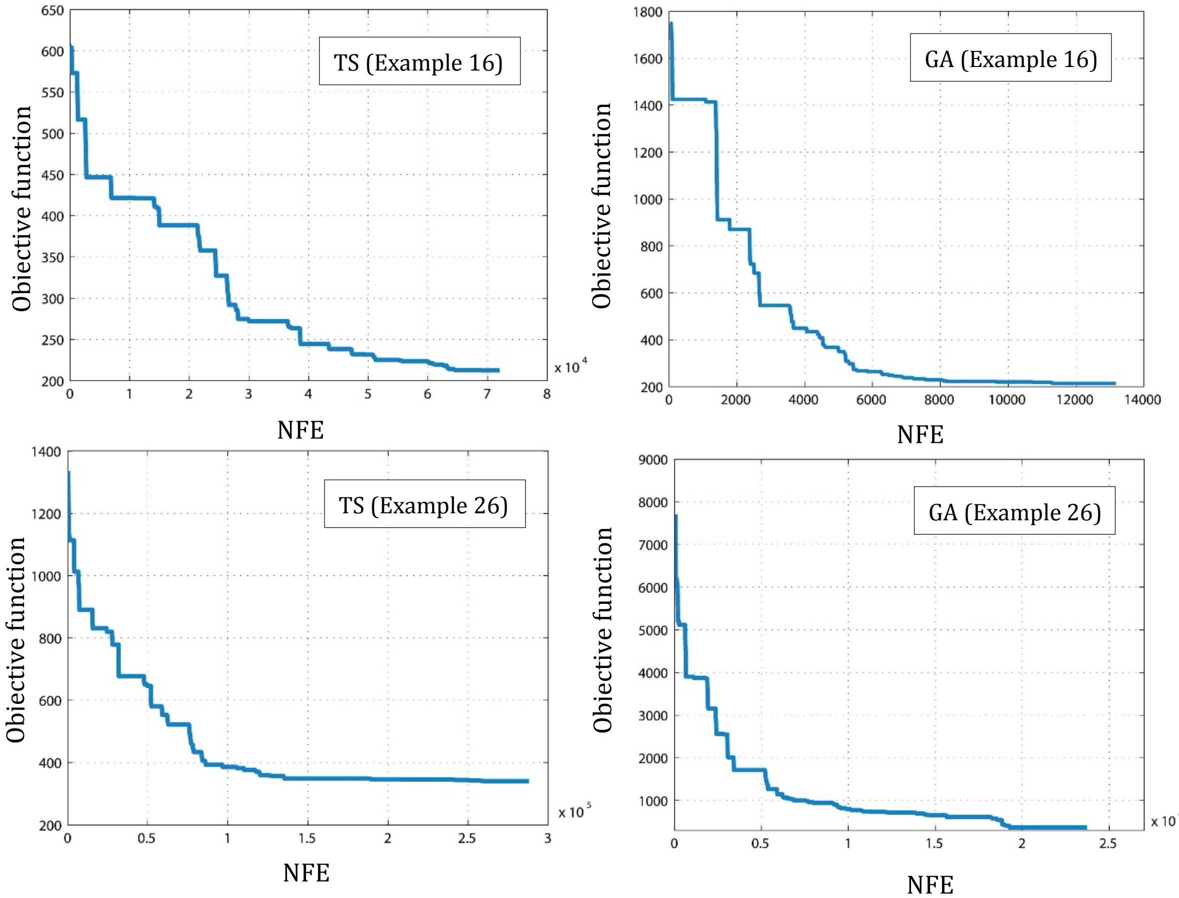


Fig 18. Comparison of convergence curves of metaheuristics (NFE: number of function evaluation)

7-Sensitivity analysis

In the current section, we investigate the relationships between the structure of the problem and the achieved solutions. To this end, we provide five scenarios for each parameter and evaluate the associated objective function. All of the examples are generated based on example 19 in table 3 with 40 numbers of patients. For this purpose, we distinguish some subsets of instances with different number of hospitals, number of ambulances (both type I and II), number of red code patients, number of yellow code patients and the weights of patients. The parameters and their variations are tabulated in table 6.

Table 6. Different parameters and their considered scenarios

Parameter		Scenario 1 (very low)	Scenario 2 (low)	Scenario 3 (Medium)	Scenario 4 (High)	Scenario 5 (very High)
Number of Hospitals		6	7	8	9	10
Number of Ambulances		6	8	10	12	14
Number of Red Code Patients		5	10	15	20	25
Number of Yellow code Patients		8	13	18	23	28
Weights (R/Y/G)		0.7/0.2/0.1	0.6/0.3/0.1	0.5/0.3/0.2	0.45/0.35/0.2	0.4/0.35/0.25

In these examples, the total number of patients is constant and we only change green code patients to red or yellow code patients. The results of solving the problem are presented in Table 7. In Table 7 CT indicates the completion time for red, yellow and green code patients and the row WSCT presents the weighted sum of completion times.

Table 7. The results of solving various instances for sensitivity analysis

Parameter	CT	Scenario 1 (very low)	Scenario 2 (low)	Scenario 3 (Medium)	Scenario 4 (High)	Scenario 5 (very High)
Number of Red Code Patients	e_R	192.3286	270.4244	318.2399	386.0776	454.3673
	e_Y	299.4503	374.7168	550.6829	644.2197	689.7410
	e_G	501.9688	475.6162	378.5731	118.2529	40.2485
	WSCT	255.4291	322.2413	394.0061	436.7377	483.5675
Number of Yellow Code Patients	e_R	227.3237	270.4244	279.5216	302.7025	320.1191
	e_Y	357.4435	374.7168	483.8750	590.0324	598.7323
	e_G	507.6920	475.6162	223.5936	171.9448	97.3613
	WSCT	294.3964	322.2313	335.2348	375.8257	381.4272
Number of Ambulances	e_R	345.4445	270.4244	203.7193	192.3286	192.3286
	e_Y	582.1858	374.7168	285.1371	259.3984	259.3984
	e_G	652.1411	475.6162	317.6596	298.5329	275.9758
	WSCT	447.1365	322.2313	239.5386	223.0699	220.8142
Number of Hospitals	e_R	270.4244	225.3106	171.1262	154.7536	154.4331
	e_Y	474.7168	410.1773	386.6848	375.7344	356.4442
	e_G	575.6162	556.7972	463.7769	455.5417	401.9703
	WSCT	362.2313	313.9192	265.0588	252.9263	237.3901
Weights	e_R	230.4148	270.4244	284.5114	284.5114	297.9954
	e_Y	410.1455	390.7168	402.0462	394.8036	402.9069
	e_G	428.9500	450.6162	415.3607	428.6599	409.7663
	WSCT	314.2875	324.7413	332.8567	332.0109	340.6459

To investigate the behavior of proposed model with regard to the different parameters, the results of table 7 are depicted in figure 19 with more details. The following conclusions can be derived from the figure:

- According to the figure 19(a-b), by increasing the number of red or yellow code patients, the latest service completion time increases for both red and yellow code patients because more patients have to be brought to hospitals. As a result, the objective function is also increasing. However, as can be seen from the figures the slope of WSCT in figure 19 (b) is less than figure 19(a). Moreover, in both instances, the latest service completion time of green code patients decreases from scenario 1 to 5. The reason is that there exist fewer patients with green code, which need assistance.

- As expected, more available ambulances eventuate a better completion service time, (it means a lower objective function). As can be seen from the figure 19 (c), the latest completion times of both red and yellow code patients decrease significantly from scenario 1 to 3, after that the slope of diagrams are declined. In particular, the availability of very high number ambulances does not affect the completion service times, notably.

- Figure 19 (d) shows the impact of the number of hospitals on the completion times of patients and the objective function. The results indicate that a larger number of hospitals yields to serve red and yellow code patients faster. In detail, the duration of trips for transporting patients to hospitals is reduced. Obviously, the objective function time decreases with a larger number of available hospitals. Nevertheless, according to the slope of the objective function from scenario 3-5, existence of very high number of hospitals causes a minor effect on the achieved solutions.

- In order to analyze the impact of patients' weights on their completion times, we provide five scenarios with various combinations of weights. Figure 19 (e) depicts the achieved solutions. As seen in the figure by increasing the relative importance of each code patient the corresponded latest service completion time decreases.

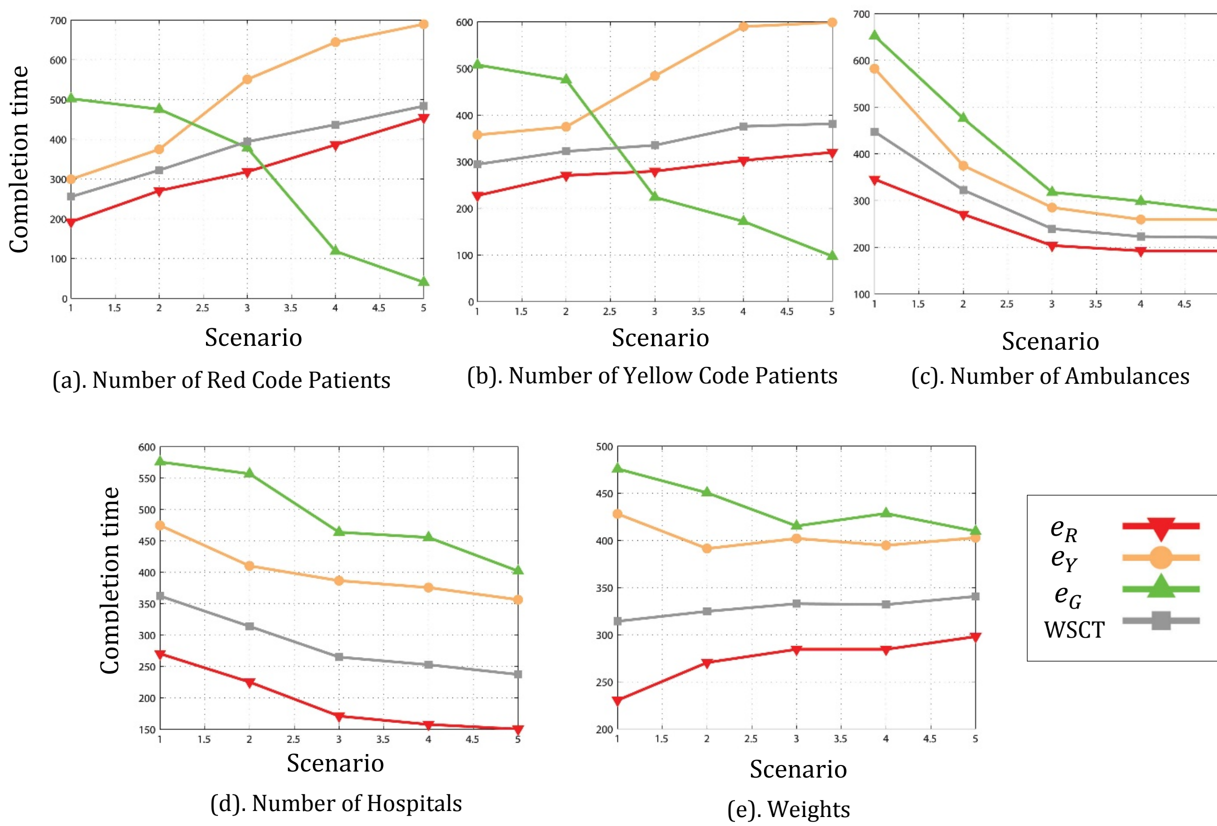


Fig 19. Sensitivity analysis for various parameters

7-1-Effect of patient distribution on the completion times

In this subsection, we examine the impact of different distribution of patients on the completion times. To this end, we generate different instances that randomly dispread in areas with various sizes. Table 8 addresses the generated instances and the related regions. The results are also reported in table 8. According to the table, the objective function increases by enlarging the areas.

Table 8. The examples and their results

	Instance 1	Instance 2	Instance 3	Instance 4	Instance 5
Space	50*100	100*200	200*300	300*400	400*500
e_R	126.6889	270.4244	320.9505	491.1886	693.3655
e_Y	206.3104	374.7168	616.1070	765.1493	1252.801
e_G	241.3917	475.6162	610.6359	822.2283	841.9758
WSCT	162.0457	322.2313	409.266	606.4808	874.0564

To clarify the generated instances, the examples are illustrated in figure 18. These examples are generated base on example 19 in table 3. In all instances, the number of hospitals is equal to 6 and there exist 40 numbers of patients. The blue squares in the figures represent the location of hospitals, while patients are shown with cycles with different colors. Moreover, figure 18 (f) depicts the completion times of red, yellow and green code patients and the objective function. As expected by enlarging the areas we observe a significant increase in the latest service completion time of each code patient and the obtained objective function.

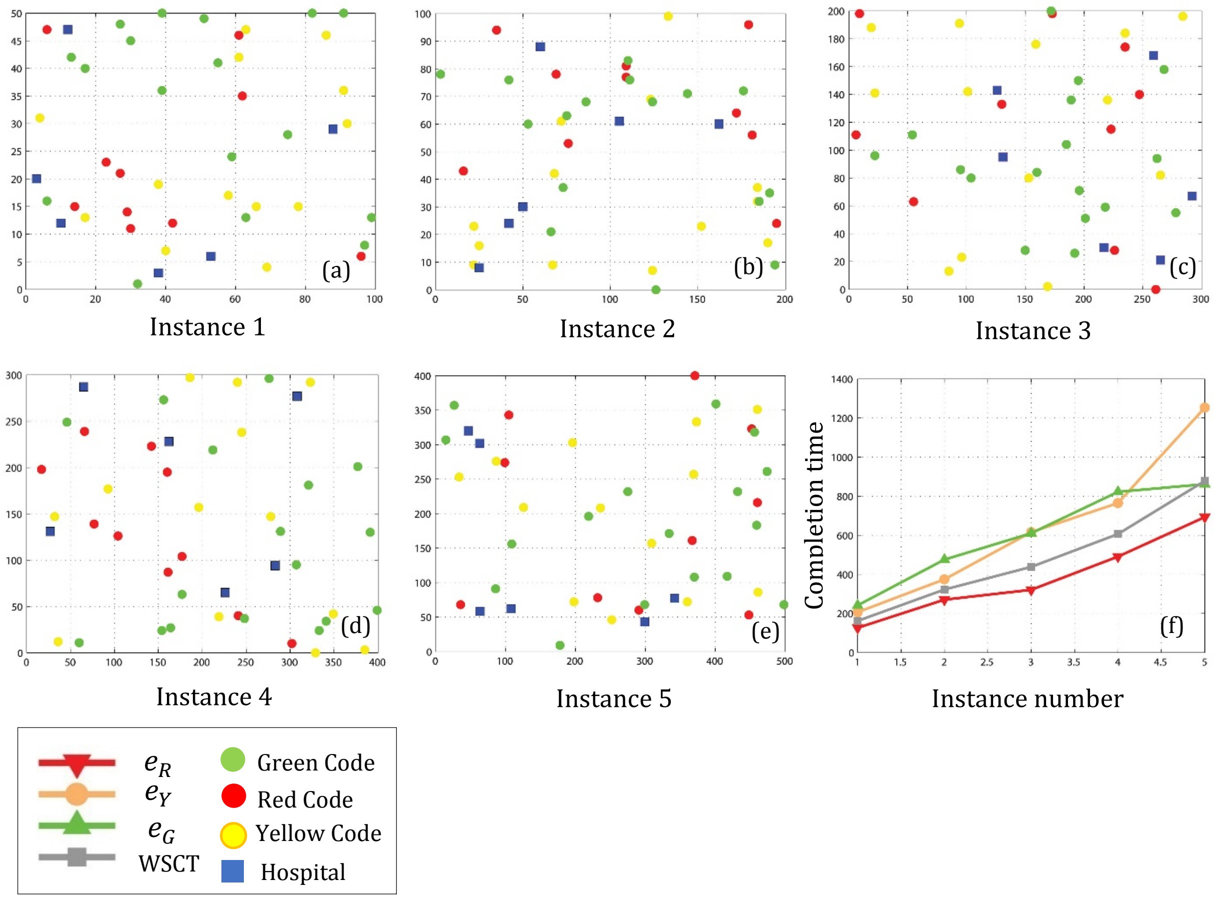


Fig18. Examples with different region sizes and the achieved solutions

8-Conclusion and future remarks

In this paper, we have developed and solved an ambulance routing problem for disaster response phase where the patients and vehicles are classified into different groups based on their requirements and features, respectively. We consider three groups of patients: slightly injured people who can be assisted directly in the field, moderately severe and severely injured individuals, which should be transferred to the hospital for further treatments. Ambulances are recognized as a scarce resource in disaster situations that utilized to transport medical personnel and patients. Since their efficient usage is utmost vital, we distinguish various types of ambulances with different capabilities. The proposed model adopts the ambulance type with patients' needs to increase the total service quality. Moreover, due to the lack of vehicles and blocked roads in disaster situations, ambulances have some delays to reach the patients. Therefore, we incorporate the semi-soft time window in the model to reflect the penalties arising from delays. To this end, the survival function is utilized to determine the thresholds of semi-soft time windows constraints for each code patient. The proposed model aims at minimizing the latest service completion time among the people waiting for help and penalties in serving the patients. Presented ARP should be solved at high frequency in post-disaster with acceptable computation time. However, due to the NP-hardness of the problem, the computational times increase non-polynomially with enlargement of the problem sizes. Thus, in order to solve the problem more efficiently, two meta-heuristics include genetic algorithm and tabu search are proposed to solve the larger instances in a reasonable time. The achieved solutions are also compared with CPLEX method and a well-known heuristic approach. The results confirm the performance and effectiveness of the proposed TS in compare with other method solutions. In the following, we provide some further experiments to analyze the effects of various structural parameters on the obtained solutions.

In future researches, other important actual factors such as the travel time reliability, demand uncertainty and reliability of critical links can be incorporated into the proposed model.

References

Abounacer, R., Rekik, M., & Renaud, J. (2014). An exact solution approach for multi-objective location–transportation problem for disaster response. *Computers & Operations Research*, *41*, 83-93.

Andersson, T., & Värbrand, P. (2007). Decision support tools for ambulance dispatch and relocation. *Journal of the Operational Research Society*, *58*(2), 195-201.

Barbarosoğlu, G., & Arda, Y. (2004). A two-stage stochastic programming framework for transportation planning in disaster response. *Journal of the operational research society*, *55*(1), 43-53.

Berkoune, D., Renaud, J., Rekik, M., & Ruiz, A. (2012). Transportation in disaster response operations. *Socio-Economic Planning Sciences*, *46*(1), 23-32.

Bozorgi-Amiri, A., Tavakoli, S., Mirzaeipour, H., & Rabbani, M. (2017). Integrated locating of helicopter stations and helipads for wounded transfer under demand location uncertainty. *The American journal of emergency medicine*, *35*(3), 410-417.

Brandão, J. (2016). A deterministic iterated local search algorithm for the vehicle routing problem with backhauls. *TOP*, *24*(2), 445-465.

Bräysy, O., & Gendreau, M. (2002). Tabu search heuristics for the vehicle routing problem with time windows. *Top*, *10*(2), 211-237.

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

Camacho-Vallejo, J. F., González-Rodríguez, E., Almaguer, F. J., & González-Ramírez, R. G. (2015). A bi-level optimization model for aid distribution after the occurrence of a disaster. *Journal of Cleaner Production*, *105*, 134-145.

Campbell, A. M., Vandenbussche, D., & Hermann, W. (2008). Routing for relief efforts. *Transportation Science*, *42*(2), 127-145.

Carnes, T. A., Henderson, S. G., Shmoys, D. B., Ahghari, M., & MacDonald, R. D. (2013). Mathematical programming guides air-ambulance routing at ornge. *Interfaces*, *43*(3), 232-239.

Caunhye, A. M., Nie, X., & Pokharel, S. (2012). Optimization models in emergency logistics: A literature review. *Socio-economic planning sciences*, *46*(1), 4-13.

Coppi, A., Detti, P., & Raffaelli, J. (2013). A planning and routing model for patient transportation in health care. *Electronic notes in discrete mathematics*, *41*, 125-132.

Detti, P., Papalini, F., & de Lara, G. Z. M. (2017). A multi-depot dial-a-ride problem with heterogeneous vehicles and compatibility constraints in healthcare. *Omega*, *70*, 1-14.

Edrissi, A., Poorzahedy, H., Nassiri, H., & Nourinejad, M. (2013). A multi-agent optimization formulation of earthquake disaster prevention and management. *European Journal of Operational Research*, *229*(1), 261-275.

Edrissi, A., Nourinejad, M., & Roorda, M. J. (2015). Transportation network reliability in emergency response. *Transportation research part E: logistics and transportation review*, *80*, 56-73.

Farahani, R. Z., Asgari, N., Heidari, N., Hosseininia, M., & Goh, M. (2012). Covering problems in facility location: A review. *Computers & Industrial Engineering*, *62*(1), 368-407.

Gholami-Zanjani, S. M., Pishvaee, M. S., & Torabi, S. A. (2018). OR Models for Emergency Medical Service (EMS) Management. In *Operations Research Applications in Health Care Management* (pp. 395-421). Springer, Cham.

Gennarelli, T. A., & Wodzin, E. (2008). *Abbreviated injury scale 2005: update 2008*. Russ Reeder.

Gong, Q., & Batta, R. (2007). Allocation and reallocation of ambulances to casualty clusters in a disaster relief operation. *IIE Transactions*, *39*(1), 27-39.

Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press.

Huang, M., Smilowitz, K., & Balcik, B. (2012). Models for relief routing: Equity, efficiency and efficacy. *Transportation research part E: logistics and transportation review*, *48*(1), 2-18.

Knight, V. A., Harper, P. R., & Smith, L. (2012). Ambulance allocation for maximal survival with heterogeneous outcome measures. *Omega*, *40*(6), 918-926.

Lai, D. S., Demirag, O. C., & Leung, J. M. (2016). A tabu search heuristic for the heterogeneous vehicle routing problem on a multigraph. *Transportation Research Part E: Logistics and Transportation Review*, *86*, 32-52.

Lam, S. S. W., Nguyen, F. N. H. L., Ng, Y. Y., Lee, V. P. X., Wong, T. H., Fook-Chong, S. M. C., & Ong, M. E. H. (2015). Factors affecting the ambulance response times of trauma incidents in Singapore. *Accident Analysis & Prevention*, *82*, 27-35.

Luis, E., Dolinskaya, I. S., & Smilowitz, K. R. (2012). Disaster relief routing: Integrating research and practice. *Socio-economic planning sciences*, *46*(1), 88-97.

Marcon, E., Chaabane, S., Sallez, Y., Bonte, T., & Trentesaux, D. (2017). A multi-agent system based on reactive decision rules for solving the caregiver routing problem in home health care. *Simulation Modelling Practice and Theory*, *74*, 134-151.

Nable, J. V., Lawner, B. J., & Brady, W. J. (2016). 2016: emergency medical services annotated literature in review. *The American journal of emergency medicine*, *34*(11), 2193-2199.

Najafi, M., Eshghi, K., & Dullaert, W. (2013). A multi-objective robust optimization model for logistics planning in the earthquake response phase. *Transportation Research Part E: Logistics and Transportation Review*, *49*(1), 217-249.

Özdamar, L., & Ertem, M. A. (2015). Models, solutions and enabling technologies in humanitarian logistics. *European Journal of Operational Research*, *244*(1), 55-65.

Özdamar, L., & Demir, O. (2012). A hierarchical clustering and routing procedure for large scale disaster relief logistics planning. *Transportation Research Part E: Logistics and Transportation Review*, *48*(3), 591-602.

Özdamar, L., Ekinci, E., & Küçükyazici, B. (2004). Emergency logistics planning in natural disasters. *Annals of operations research*, *129*(1-4), 217-245.

Parragh, S. N. (2011). Introducing heterogeneous users and vehicles into models and algorithms for the dial-a-ride problem. *Transportation Research Part C: Emerging Technologies*, *19*(5), 912-930.

Parragh, S. N., Cordeau, J. F., Doerner, K. F., & Hartl, R. F. (2012). Models and algorithms for the heterogeneous dial-a-ride problem with driver-related constraints. *OR spectrum*, *34*(3), 593-633.

Pedraza-Martinez, A. J., & Van Wassenhove, L. N. (2012). Transportation and vehicle fleet management in humanitarian logistics: challenges for future research. *EURO Journal on Transportation and Logistics*, *1*(1-2), 185-196.

Qureshi, A. G., Taniguchi, E., & Yamada, T. (2010). Exact solution for the vehicle routing problem with semi soft time windows and its application. *Procedia-Social and Behavioral Sciences*, *2*(3), 5931-5943.

Rath, S., & Gutjahr, W. J. (2014). A math-heuristic for the warehouse location–routing problem in disaster relief. *Computers & Operations Research*, *42*, 25-39.

Roy, R. K. (2010). *A primer on the Taguchi method*. Society of Manufacturing Engineers.

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

Schmid, V. (2012). Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming. *European journal of operational research*, *219*(3), 611-621.

Setak, M., Habibi, M., Karimi, H., & Abedzadeh, M. (2015). A time-dependent vehicle routing problem in multigraph with FIFO property. *Journal of Manufacturing Systems*, *35*, 37-45.

Shahriari, M., Bozorgi-Amiri, A., Tavakoli, S., & Yousefi-Babadi, A. (2017). Bi-objective approach for placing ground and air ambulance base and helipad locations in order to optimize EMS response. *The American journal of emergency medicine*, *35*(12), 1873-1881.

Solomon, M. M. (1987). Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations research*, *35*(2), 254-265.

Talarico, L., Meisel, F., & Sörensen, K. (2015). Ambulance routing for disaster response with patient groups. *Computers & Operations Research*, *56*, 120-133.

Tikani, H., Honarvar, M., & Mehrjerdi, Y. Z. (2018). Developing an integrated hub location and revenue management model considering multi-classes of customers in the airline industry. *Computational and Applied Mathematics*, *37*(3), 3334-3364.

Tlili, T., Harzi, M., & Krichen, S. (2017). Swarm-based approach for solving the ambulance routing problem. *Procedia Computer Science*, *112*, 350-357.

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

Toth, P., & Vigo, D. (2002). The vehicle routing problem, Philadelphia: society for industrial and applied mathematics.

Yi, W., & Özdamar, L. (2007). A dynamic logistics coordination model for evacuation and support in disaster response activities. *European Journal of Operational Research*, *179*(3), 1177-1193.

Zhang, Z., Liu, M., & Lim, A. (2015). A memetic algorithm for the patient transportation problem. *Omega*, *54*, 60-71.

Zheng, Y. J., Chen, S. Y., & Ling, H. F. (2015). Evolutionary optimization for disaster relief operations: A survey. *Applied Soft Computing*, *27*, 553-566.