# The post-disaster transportation of injured people when hospitals have districts 

Tareq Babaqi<br>Department of Systems Science and Industrial Engineering, State University of New York at Binghamton, Binghamton, New York, USA, and<br>Béla Vizvári<br>Department of Industrial Engineering, Eastern Mediterranean University, Famagusta, Turkey


#### Abstract

Purpose - The total capacity of ambulances in metropolitan cities is often less than the post-disaster demand, especially in the case of disasters such as earthquakes. However, because earthquakes are a rare occurrence in these cities, it is unreasonable to maintain the ambulance capacity at a higher level than usual. Therefore, the effective use of ambulances is critical in saving human lives during such disasters. Thus, this paper aims to provide a method for determining how to transport the maximum number of disaster victims to hospitals on time. Design/methodology/approach - The transportation-related disaster management problem is complex and dynamic. The practical solution needs decomposition and a fast algorithm for determining the next mission of a vehicle. The suggested method is a synthesis of mathematical modeling, scheduling theory, heuristic methods and the Voronoi diagram of geometry. This study presents new elements for the treatment, including new mathematical theorems and algorithms. In the proposed method, each hospital is responsible for a region determined by the Voronoi diagram. The region may change if a hospital becomes full. The ambulance vehicles work for hospitals. For every patient, there is an estimated deadline by which the person must reach the hospital to survive. The second part of the concept is the way of scheduling the vehicles. The objective is to transport the maximum number of patients on time. In terms of scheduling theory, this is a problem whose objective function is to minimize the sum of the unit penalties. Findings - The Voronoi diagram can be effectively used for decomposing the complex problem. The mathematical model of transportation to one hospital is the $P \| \Sigma U_{j}$ problem of scheduling theory. This study provides a new mathematical theorem to describe the structure of an algorithm that provides the optimal solution. This study introduces the notion of the partial oracle. This algorithmic tool helps to elaborate heuristic methods, which provide approximations to the precise method. The realization of the partial oracle with constructive elements and elements proves the nonexistence of any solution. This paper contains case studies of three hospitals in Tehran. The results are close to the best possible results that can be achieved. However, obtaining the optimal solution requires a long CPU time, even in the nondynamic case, because the problem $P \| \Sigma U_{j}$ is $N P$-complete. Research limitations/implications - This research suggests good approximation because of the complexity of the problem. Researchers are encouraged to test the proposed propositions further. In addition, the problem in the dynamic environment needs more attention. Practical implications - If a large-scale earthquake can be expected in a city, the city authorities should have a central control system of ambulances. This study presents a simple and efficient method for the post-disaster transport problem and decision-making. The security of the city can be improved by purchasing ambulances and using the proposed method to boost the effectiveness of post-disaster relief. Social implications - The population will be safer and more secure if the recommended measures are realized. The measures are important for any city situated in a region where the outbreak of a major earthquake is possible at any moment. Originality/value - This paper fulfills an identified need to study the operations related to the transport of seriously injured people using emergency vehicles in the post-disaster period in an efficient way.


Keywords Hospital district, Humanitarian logistics, Partial oracle, Post-disaster period, Scheduling, Voronoi diagram
Paper type Research paper

## Introduction

A disaster is defined as an event that satisfies at least two of the following criteria:
1 it is a sudden event;

The current issue and full text archive of this journal is available on Emerald Insight at: https://www.emerald.com/insight/2042-6747.htm

Journal of Humanitarian Logistics and Supply Chain Management
13/1 (2023) 61-73
Emerald Publishing Limited [ISSN 2042-6747]
[DOI 10.1108/JHLSCM-09-2021-0088]
© Tareq Babaqi and Béla Vizvári. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/4.0/legalcode

Received 11 September 2021
Revised 30 January 2022
26 May 2022
4 October 2022
Accepted 10 November 2022

2 it causes great damage to human lives and the economy; and
3 the local population does not have enough capacity to handle the damage (Papp, 2019).

For example, drought and pandemic satisfies criteria (2) and (3) (Babaqi et al., 2022). This paper deals with disasters that satisfy criteria (1) and (2) regardless of whether they satisfy criterion (3). Many natural disasters (earthquakes, floods and wildfires) that satisfy criteria (1) and (2) cannot be predicted in time. These sudden disasters demand real-time, dynamic, immediate, effective and efficient solutions to protect human beings, animals, buildings and properties in the immediate post-disaster period. Disasters such as earthquakes, floods, hurricanes, wildfires, terrorist attacks and volcano eruptions negatively impact both the city's inhabitants and infrastructure on a large scale. Therefore, rescue operations are challenging, and such a problem is often computationally intractable. Many different approaches and methods have been used in operational research and system management to tackle these problems (Altay and Green, 2006). Every year, many lives are lost around the world because of disasters. According to the United States Geological Survey, there are millions of earthquakes worldwide every year, and in the 15 years leading up to 2015 , more than 800,000 people died as a result of earthquakes. At least 250 million people are affected by 125 disasters worldwide on average each year (United Nations, 2007).

According to Altay and Green (2006), emergency response efforts are composed of four disaster operations management phases: mitigation, preparedness, response and recovery. Mitigation can be defined as a continuing action to preclude or reduce the risk to people and property from a disaster and its effects. Preparedness is the measures taken to prepare for and, when possible, prevent or reduce the effects of disasters (Bullock et al., 2013). Response refers to the use of emergency procedures and resources defined in emergency plans to save lives, the environment, properties, etc. The response, which needs an organization with many functions, must start immediately after the sudden event of the disaster with the available capacities (Vizvári et al., 2019). Recovery is the longterm action taken to restore normalcy and stabilize the community after the immediate impact of the disaster has passed (Whybark, 2015). It is essential to make thorough preparations for a potential disaster situation beforehand. However, once the actual disaster occurs, the managers on the scene have to develop complex response strategies in a compressed time frame without the support of higher-level managers. Implementing plans, establishing command posts and shelters and ensuring the provision of all necessary emergency services are the activities of the disaster response phase.
Emergency medical services (EMS) are the most important services for saving lives during a disaster. However, one of the major related problems that needs to be tackled to increase the number of lives saved is the ambulance routing problem (ARP); it involves coordinating and regulating the flow of ambulances to save those impacted by disasters (Tlili et al., 2017). The transportation of casualties after an earthquake poses a serious problem in a metropolitan city (Pinarbaşi et al.,
2022). Thus, operations research experts have been attracted to EMS decision-making and management problems because of their highly sensitive nature. Over the years, they have explored various problems and solutions in managing the EMS systems (Bozorgi-Amiri et al., 2017; Kadri et al., 2014; Molenbruch et al., 2017).
Recent literature has described the roles and responsibilities of EMS as establishing emergency facilities, transferring injured patients to hospitals, providing medical assistance, rescuing survivors and orchestrating activities across multiple organizations. The various elements of the response to a disaster related to transportation are as follows:

- transportation planning: determining the suppliers' delivery schedule before assigning ambulances;
- ambulance assignment: allocating ambulances to the appropriate emergency points;
- routing: giving routes to the ambulances to transport patients quickly;
- road repair: repairing damaged roadways and restoring lifelines to affected areas; and
- integrated problems: resolving a set of individual problems to achieve one or more common objectives (Wren and Holliday, 1972).

The ARP addresses the routing issue to form the strategy to transport patients from the disaster site to medical facilities.
The primary roles of ambulances in the post-disaster period are to provide first aid to the lightly injured people and transport the more seriously injured to operating hospitals. It is massively complicated to manage the operations of ambulances in the immediate aftermath of a disaster because of the dynamics and uncertainty with which the relevant information changes over time and, thus, leads to changing the plan. The availability of ambulances and the number, location and medical state of people calling for help are the information required to manage ambulance operations and routes.
After a serious disaster in a metropolitan city, there are many patients who need urgent medical care to survive. Every patient has a deadline by which they must get to the operating room, depending on the type and severity of the damage; otherwise, the patient will die. This deadline can be obtained from the expert triage system using the information provided by the person who reports the victim to the emergency center. The location is also acquired in the same way.

When the capacity of the emergency transportation system is high, it is easy for operations management to find a schedule that ensures that everyone arrives at the hospital on time. However, this capacity becomes insufficient in cases of disasters that have a large-scale impact because of the vast difference between the capacities needed in the non-disaster and post-disaster periods. Thus, saving all the injured people in such situations is impossible (Thomson, 1985). Some studies have projected the incorporation of military assets into the emergency task force in the post-disaster period (Kallehauge, 2008). If it is not possible to save everyone, then it is essential to identify the ethical principles to be applied in the transportation process. The simplest and most egalitarian solution to this ethical conundrum is to maximize the number of lives saved (Shavarani and Vizvari, 2018). This solution conforms to many published ethical principles, such as the "Ethical Goals of the

Allocation Framework" section in White et al. (2020). The result of this approach is the thought process that people in isolated places have a slim chance of survival. Nevertheless, this study applies the principle of maximizing the expected number of saved patients after a disaster. This paper does not focus on the investigation of ethical matters.
The summary of the topic of this paper is described in this section. In the event of a sudden disaster, there exists a multifunctional organization that is responsible for the response to the disaster. One of its many functions is the transportation of seriously injured people to hospitals. The main contribution of this paper is the proposal of a method for effectively using the available ambulances. The proposed method can be used in an automated system and the synthesis of several new and known elements. The two most important elements are the problem $P \| \Sigma U_{j}$ of scheduling theory and the Voronoi diagram of geometry. The problem $P \| \Sigma U_{j}$ has been very rarely investigated. Thus, its available theory was not suitable for this application. New mathematical results, including Theorem 1 and Lemma 2 described below, had to be obtained to make the application possible. The notion of a partial oracle, which is the basis of the proposed heuristic method, is also a novel idea used in the proposed method. The Voronoi diagram is applied in its original form.

The structure of the paper reflects the structure of the elements and results. The following section presents the literature overview of the study. The Materials and methods section describes the Voronoi diagram both in a static and dynamic environment. This diagram enables the decomposition of the enormous, intractable problem. A complete mathematical description of the simplified model after decomposition is given in the Mathematical model for a single hospital in the case of a small capacity. This problem is investigated from both theoretical and algorithmic points of view in the Mathematical analysis of the problem $P \| \Sigma U_{j}$ section. The method is applied to three districts of Tehran in the Case study section. The Results and discussion section evaluates the numerical results. The conclusion is presented in the final section of the paper.

## Literature overview

This paper describes a transportation system in the postdisaster period. Therefore, it uses previous results from the following areas related to post-disaster relief management:

- the organization of the relief system;
- triage;
- post-disaster transportation of casualties and evacuation;
- algorithms for calculating the minimal path; and
- mathematical tools of scheduling theory used in modeling the problem.

The main result of this paper is a synthesis of all these areas. Along with the synthesis, several new findings are presented in the mentioned areas:

- After a disaster, a considerable number of victims must be evacuated to temporary resettlement sites. Much information from many different places and sources, such as the police and reports of civilians, is received by Emergency Reply Centers (ERC) through satellite technologies (Delmonteil and Rancourt, 2017). Because of the disaster, many connections may be lost. Thus,
information must be available on the state of the transportation grid (Li and Tsukaguchi, 2003), as the rescue procedure results are significantly influenced by the period between the injury and medical intervention, which requires effective communication (Wang et al., 2009). To maximize the number of lives saved, ERCs use the information to plan the ordering of emergency vehicles and determine which roads are needed.
- Jotshi et al. (2009) divided casualties into four types: slightly injured, moderately injured, severely injured and mortally injured. With this classification, an approximate deadline for patients can be established. Post-disaster decisions must be made quickly to be operative, which dictates that the solution approach must be fast and timely.
- Several publications have focused on the processes of lifesaving and evacuation of patients in the post-disaster period. Aringhieri (2020) suggested a mathematical model and an online heuristic method for determining the movement of ambulances in the post-disaster period. Aringhieri et al. (2021) surveyed several transportation services in the health-care industry, including emergency cases.

Jin et al. (2014) developed a logistical model to transport injured people with various injuries and determine the probabilities of survival.

Liu et al. (2011) proposed a mathematical model for vehicle scheduling to rescue victims from terrorist attacks by choosing the most suitable and available hospital for each victim based on the rescue priority. Amadini et al. (2013) proposed an algorithm and software tool to assign disaster victims to hospitals and schedule emergency vehicles for their transportation, but without addressing any of the ethical concerns.
Ozdamar (2011) described an efficient planning system for helicopter operations in the post-disaster medical care and evacuation of injured people.
Transporting gravely injured people to hospitals is one of the most crucial problems in the post-disaster period. This problem was raised by Shavarani and Vizvari (2018), who presented some models and discussed their mathematical properties.
However, they did not provide a general solution.

- Obtaining the shortest routes between the depots and casualties is the primary phase in calculating the routes for emergency vehicles, and among the casualties themselves, in a situation when the capacity of the vehicle is one or more. However, by evaluating 15 shortest path algorithms, Zhan and Noon (1998) confirmed that applying the shortest path algorithms can be time consuming in large networks, and a proper solution must be carefully chosen.

Chou et al. (1998) projected a hierarchical algorithm to reduce the period required to calculate the shortest route and to discover the distance among the nodes of a large grid.
These approaches have often been used in recent literature (Kula et al., 2012). However, these methods are not fast enough for an extensive grid or an emergency. Therefore, this paper uses the Manhattan-like distance $\left(l_{1}\right)$ as an alternative.

- The main modeling and algorithmic tool used in this paper is task scheduling on identical parallel machines. Scheduling is primarily used to determine the time or order of operations but can also be used to assign tasks to machines. In the case of parallel machines, one machine processes each task. If the machines are identical, then the processing time of each job is the same on each machine. The most famous problem in the area of identical parallel machines is makespan minimization (Pinedo, 1995), that is, the completion time of the last job. This problem is denoted by $P \| C_{\max }$ in the usual notation of scheduling theory. In this paper, the minimization of the sum of the unit penalties, which is denoted by $P \| \sum U_{j}$ in scheduling theory, is applied with the help of its relation to the $P \| C_{\text {max }}$ problem (Pinedo, 1995). A few papers have discussed the minimization of tardy jobs in identical parallel machines, beginning with Bedworth and Bailey (1987), who proposed a heuristic procedure to solve the problem.

Ho and Chang (1995) presented two heuristics with several variations to achieve the same objective. Süer et al. (1993) solved the problem by formulating an integer programming mathematical model and presenting three simple heuristics. They later evaluated their performance. Recently, Najat et al. (2019) discussed the same problem in the case of several maintenance periods needed during the scheduling process; they proposed a mathematical model and heuristic procedure to minimize the number of tardy jobs. All the previous research focuses on manufacturing systems, where reasonable time is available to find answers. To the best of our knowledge, no paper has discussed the problem when an immediate answer is required, as in the situation of a rescue operation, where finding a schedule for saving the maximum number of seriously injured people has to be fast, or people may die.
This paper applies the assumption that all ambulances have the same specifications and their capacities are limited to one, that is, only one seriously injured person can be taken on a route at a time.

## Materials and methods

Post-disaster transportation is a challenging problem, especially in large-scale disasters in metropolitan cities. Overpopulated cities like San Francisco, Istanbul and Tehran can have tens of thousands of disaster casualties. The rescue operations require the participation of many hospitals and the management of a large number of vehicles. Each ambulance vehicle must execute several missions. The post-disaster relief system has a dynamic nature because reports from different locations arrive at different times during the post-disaster period. Thus, postdisaster relief, particularly post-disaster transportation, is a complex management issue that requires quick and efficient solutions.
The method of finding the solution must have two main features. First, the complex problem must be decomposed into smaller subproblems; otherwise, the reaction of the relief system will be too slow and cause unnecessary death. Second, the dynamic nature of the problem implies that there is no single optimal solution as the situation changes continuously. Instead of looking for computationally expensive optimization,
a logical and adequate action plan must be selected and executed every time a decision is needed.

## Voronoi diagram

One way of simplifying the control problem is if the disaster area is divided among the hospitals. This can be done by considering several hospitals that can operate to save injured people after a disaster in an area and then applying a Voronoi diagram to identify which hospital is nearest to each injured person in a normal case.

Georgy Voronoi first defined Voronoi diagrams in 1908 as a two-dimensional representation of the relationship between a set of objects and the plane (Aurenhammer et al., 2013; Voronoi, 1908). They are now used in different fields of study, such as robot navigation (Thrun, 1998) and material physics (Sze and $\mathrm{Ng}, 2006$ ). In the field of robotics, Voronoi diagrams are used to create a protocol for avoiding detected obstacles. They are useful for modeling natural occurrences in zoology, anthropology, archaeology, forestry and ecology and studying the patterns of urban settlements (Drysdale, 1996). The geographical spread of dialects of human languages can also be described by Voronoi diagrams (Goebl, 2010). In this paper, we consider the hospitals in a city and use the Voronoi diagram to find the nearest hospital to each injured person.
A collection of regions that divide up the plane is called a Voronoi diagram of "sites." Each site has its own region, and all the points in the interior of one region are closer to the corresponding site than to any other site. Where there is not one closest point, there is a boundary. In Figure 1, points P1 and P2 are closer to site H 1 than to any other site. Moreover, point P3, which is on the boundary between the sites H3 and H5, is equidistant from both of those sites. The Voronoi diagram depends on the sites and the applied distance as well.

## Voronoi diagram in a dynamic environment

The Voronoi diagram serves only for the imaginary visualization of the decomposed structure of the hospitals. It is not necessary to construct the diagram graphically. Moreover,

Figure 1 Basic example of a Voronoi diagram in a plane

it can be changed during the disaster response time according to the changing conditions.
The notion of the Voronoi diagram may give the impression that the regions of the sites/hospitals are the most important information. This is not true in the application in question. The system demands the converse information, that is, the hospital to which the location where an injured person is found belongs. This can be determined by simply calculating the distances from the location to the working hospitals. The potential patients of each hospital are stored in a list on a computer at the ERC headquarters. If a hospital is reported to be full, its list is deleted, and the remaining patients are distributed to other working hospitals. The lists are created again if the hospital reports that further patients can be received. The calculation of the distances takes negligible time in an automated system. The redistribution method, which is necessary because of the dynamic nature of the system, is based on the properties of the ERC.

The necessary condition for the change of the regions of the Voronoi diagram is a working information system not affected by the disaster. It can be a system that is partially or completely separated from the systems of the general public, including the internet and mobile phones. The hospitals must be able to report the changes in their state if they are full and cannot accept more patients and they become ready to accept patients again. The ERC redistributes the patients based on the reports and shares the changes with the hospitals.

These activities can be formalized as follows:
Notations:
$h, q \quad=$ indices of hospitals;
ACTIVE $=$ the list of hospitals able to accept new patients;
dist $(i, q)=$ the distance between the patient $i$ and the hospital $q$;
$L_{q} \quad=$ the list of patients assigned to a hospital $q$;
$L_{q}-i=$ the deletion of a patient $i$ from the list; and
$L_{q}+i=$ adding a patient $i$ to the list.

```
The redistribution of the patients of hospital \(h\)
            ACTIVE: \(=\) ACTIVE \(-h\)
        for \(i \in L_{h}\) do
        begin
        \(q^{*}=\operatorname{argmin}\{\operatorname{dist}(i, q): q \in A C T I V E\}\)
            \(L_{q^{*}}:=L_{q^{*}}+i\)
        end
Reassignment of patients to hospital h if it can accept patients again
for \(q \in A C T I V E\) do
            begin
                for \(i \in L_{q}\) do
                if
                dist(i, q) >dist(i, h)
                then
                begin
                    \(L_{q}:=L_{q}-i\)
                        \(L_{h}:=L_{h}+i\)
                end
            end
            ACTIVE: = ACTIVE + h
```

The response part of the post-disaster period is a dynamic environment. The reports that arrive at the ERC headquarters continuously change depending on the situation. The following
four types of reports are important from the point of view of the problem in question:

- report on an injured person;
- report on the state of a road;
- report on the mission of an ambulance; and
- report on the capacity of a hospital.

If a hospital reports that it is temporarily full and cannot receive patients, it is not possible to transport people to that hospital. In the Voronoi diagram, that site is deleted from its region. This leads to changes in the other regions and raises the question of whether it is necessary to recalculate the Voronoi diagram from the beginning.

The sites in the diagram are the hospitals in the affected area. A Voronoi diagram is only an informative object in the predisaster period. It is not possible to know beforehand which hospitals will remain undamaged and retain their ability to function post-disaster. Thus, the disaster can change the set of sites in the diagram.

## Mathematical model for a single hospital in case of small vehicle capacity

This section shows that the transportation of injured people and a problem of scheduling theory are closely connected. This scheduling problem is $P \| \Sigma U_{j}$, which is a rarely investigated problem. There is no method that can solve the problem in a short time because the problem is NP-complete (see Appendix).

Assume that the capacity of each ambulance vehicle is one, that is, a vehicle can transport only one seriously injured person at a time. In that case, the vehicles behave like machines in a workshop whose job is to transport a person to a hospital. When one job is completed, the vehicle can immediately go to the next person, that is, it can start the process of the next job (in the language of scheduling theory). The vehicles are assumed to be identical. Each person is transported to the hospital only once, that is, each job is processed by only one machine. In scheduling theory, this situation is called the case of identical parallel machines.

It is also assumed that each person has a deadline. If a person reaches the hospital before the deadline, that person is saved. If the person arrives after the deadline, then that person dies. The declared moral principle is to save as many lives as possible. It is equivalent to minimizing the number of people reaching hospitals after their deadlines. In the language of scheduling theory, this means that every job has a due date. If a job is completed after its due date, then the job is tardy. The objective is to minimize the number of tardy jobs or, equivalently, to minimize the sum of unit penalties in scheduling theory, where the unit penalty is 1 if the job is tardy and 0 if the job is completed on time.
Thus, minimizing the sum of unit penalties on identical parallel machines is the proper mathematical model of the problem. It is denoted by $P \| \Sigma U_{j}$ in scheduling theory, where $P$ stands for identical parallel machines and $U_{j}$ is the unit penalty of job $j$.

More precisely, $P \| \Sigma U_{j}$ is the model of the problem for the static case. The static case means that all the necessary information is available at the beginning, including all jobs, that is, all injured people with their location, transportation times and deadlines. We present the integer linear programming (ILP) model with the notations described below.
$y_{j}=1$ if the patient $j$ is saved, 0 otherwise;
$x_{i j k}=1$ if the patient $j$ is the $k$ th patient of vehicle $i, 0$ otherwise;
$p_{j}=$ the time needed to take patient $j$ to the hospital;
$d_{j}=$ the deadline of patient $j$;
$M=$ a large positive number;
$n$ = the number of patients;
$v=$ the number of vehicles;
$k=$ the position number, where the maximum $k=n+1-v$;
$i=$ index for vehicles, $i=1,2, \ldots$, $v$; and
$j=$ index for patients, $j=1,2, \ldots, n$.
The objective is to maximize the number of saved people in the model, as shown in equation (1).

$$
\begin{equation*}
\operatorname{Max} \sum_{j} y_{j} \tag{1}
\end{equation*}
$$

The constraint of equation (2) ensures that each patient must be assigned to exactly one position.

$$
\begin{equation*}
\sum_{i} \sum_{k} x_{i j k}=1, \forall j \tag{2}
\end{equation*}
$$

Equation (3) guarantees that no more than one patient can be transported in each vehicle trip.

$$
\begin{equation*}
\sum_{j} x_{i j k} \leq 1, \forall i, k \tag{3}
\end{equation*}
$$

Finally, the transportation time required to reach the hospital cannot exceed the deadline of the selected patient, which is ensured by equation (4), where $M$ is a large number. Specifically, when $x_{i j k}=0$, the dominant $M$ ensures that the right-hand side is greater than the left-hand side.

When $x_{i j k}=y_{j}=1$, patient $j$ arrives at the hospital on time in vehicle $i$.

$$
\begin{gather*}
\sum_{u}^{k-1} \sum_{w \neq j} p_{z w} x_{i w u}+p_{j} x_{i j k} \leq d_{j}+M\left(1-x_{i j k}\right) \\
+M\left(1-y_{j}\right), \forall i, j, k \tag{4}
\end{gather*}
$$

The problem $P \| \Sigma U_{j}$ is applied in the more or less chaotic postdisaster period in a dynamic environment, as mentioned during the discussion of the Voronoi diagram. Reports on injured people arrive at different times, and every new report changes the problem. This feature is discussed in the Problem in a dynamic environment section.

## Mathematical analysis of the problem $\boldsymbol{P} \| \Sigma \boldsymbol{U}_{\boldsymbol{j}}$

The algorithmic difficulties of the solution of combinatorial problems are discussed using complexity theory. The problem $P \| \Sigma U j$ is NP-complete according to the classification of complexity theory (see Appendix). It means that even smallsized problems take a long time to be solved. However, this assertion is not immediately clear as the somewhat simpler problem, $1 \| \Sigma U_{j}$, which is also a polynomial, is significantly easier to solve. In general, if a problem is NP-complete, then the existence of a fast algorithm that can solve all instances of
the problem class is unlikely. However, this existence is undecided in mathematics for the time being.
The $1 \| \Sigma U_{j}$ problem is closely related to $P \| \Sigma U_{j}$. The only difference is the number of machines. The former has a single machine, and the latter has more than one machine. The solution method of $1 \| \Sigma U_{j}$ is called either Moore's algorithm, Hodgson's algorithm or the Moore-Hodgson algorithm (Moore, 1968). There is no original publication by Hodgson. However, Moore's paper contains the contribution of Hodgson as well (Hodgson, 2018).

A short description of the $1 \| \Sigma U_{j}$ problem is as follows: The number of jobs is $n$. Every job $j$ has a processing time ( $p_{j}$ ) and a due date $\left(d_{j}\right)$. All jobs are available for processing at time 0 . The machine can process one job at a time. There is no setup or other technical break between two consecutive processes. The objective is to minimize the number of tardy jobs.

## Moore-Hodgson algorithm

The algorithm decides on the schedule of jobs one by one in the earliest due date (EDD) order. The due dates are monotone increasing in this order. If it is not possible to finish a job on time, then a job, either the same one or another one scheduled earlier, becomes tardy. The non-tardy jobs are scheduled after each other in the EDD order. The tardy jobs are scheduled after the non-tardy ones.

Notations. $t_{k}$ is the total processing time of the jobs on time at the end of the $k$ th iteration $\left(t_{0}=0\right)$. The sets of scheduled and tardy jobs are denoted by $S$ and $T$, respectively. In the beginning, $S=T=\varnothing$.

Step 1. Determine the EDD order of the jobs. (It is assumed without loss of generality that this order is the index order of the jobs.) $t_{0}:=0$

```
Step 2. For \(k\) : \(=1\) to \(n\) :
begin
        if \(t_{k-1}+p_{k} \leq d_{k}\)
        then
        begin
                \(S:=S \cup\{k\}\)
                \(t_{k}:=t_{k-1}+p_{k}\)
        end
        else
        begin
            \(i:=\operatorname{argmax}\left\{p_{j}: j \in S \cup\{k\}\right\}\)
            \(S:=(S \cup\{k\}) \backslash\{i\}\)
            \(t_{k}:=t_{k-1}+p_{k}-p_{i}\)
            \(T:=T \cup\{i\}\)
        end
    end
```

The key observation is that if the inequality $t_{k-1}+p_{k}>d_{k}$ is satisfied, then there must be a tardy job in the set $S \cup\{k\}$. It is possible to generalize the structure of the algorithm to the $P \| \Sigma U_{j}$ problem.

In the theory of algorithms, the word "oracle" means a procedure that gives the correct answer for a decision (YES/ NO) problem. The complexity of an oracle-using algorithm is measured many times in the number of uses of the oracle. This time, it helps to explore the structure of the optimal solution. Let $X$ be a subset of jobs. The oracle answers the following question:

## Is it possible to complete all the jobs of $X$ on time?

If the answer is YES, then such a schedule exists. If the answer is NO, then there is at least one tardy job in any schedule of the jobs of $X$. The notation $\operatorname{oracle}(X)$ is used for the answer in the description of the algorithm. Thus, oracle $(X)$ is a Booleanvalued function, that is, its value is either YES or NO. Note that the underlying problem in equation (5) is a subproblem of the complete problem. Thus, it belongs to the same problem class.
The running time of the oracle can be as long as it takes to solve an instance of the NP-complete problem. The concept of this oracle is still algorithmically useful, as discussed in the section Fast heuristics using a partial oracle.

The oracle algorithm has the same basic steps as the MooreHodgson algorithm. It selects the jobs to be scheduled on time in the EDD order. If a tardy job must exist, the algorithm selects the job having the longest processing time (LPT) among the on-time candidates and changes its state from on-time to tardy.

## Structural algorithm of $\boldsymbol{P} \| \Sigma \boldsymbol{U}_{\boldsymbol{j}}$

Step 1. Determine the EDD order of the jobs. (It is assumed without loss of generality that this order is the index order of the jobs.) Set $S=T=\varnothing$.

```
Step 2. For \(k\) : \(=1\) to \(n\) :
begin
    if oracle(S)
    then \(S:=S \cup\{k\}\)
    else
    begin
        \(i:=\operatorname{argmax}\left\{p_{j}: j \epsilon S \cup\{k\}\right\}\)
        \(S:=(S \cup\{k\}) \backslash\{i\}\)
        \(T:=T \cup\{i\}\)
    end
    end
```

Theorem 1 shows that the structural algorithm provides an optimal solution.

Theorem 1. The set $S$ provided by the structural algorithm is a set of on-time schedulable jobs of maximal cardinality.
Proof. Induction is used to prove that the set $S$ has the following two properties in every iteration $k$ :
$1 S$ is a maximal cardinality subset of the first $k$ jobs such that all jobs of $S$ can be scheduled on time.
2 Among the subsets of the first $k$ jobs satisfying property (1), $S$ has the minimal total processing time.

The statement is true for $k=1$. If the first job can be completed without tardiness, then the only maximal cardinality subset is $S=$ $\{1\}$; otherwise, $S$ remains an empty set. In a general iteration $k$, there are three cases. If job $k$ can be added to set $S$ without deleting any other job, then the number of on-time jobs is increased by 1. Assume that a set of jobs denoted by $U$ satisfies properties (1) and (2) in iteration $k$. $U$ must contain job $k$; otherwise, it contains a smaller number of jobs than $S$. Hence, it follows from the inductive assumption that $S$ also satisfies property (2). If job $k$ cannot be added to set $S$ without the deletion of another job and has the maximal processing time, that is, set $S$ is unchanged in iteration $k$, then (1) and (2) are satisfied.
The last case is when job $j$ is changed for job $k$ with $j<k$ and $p_{j}>p_{k}$. Property (1) is satisfied if it is shown that it is not possible to add both job $k$ and another job $l \in T$ to $S$. If $j<l<k$,
then $p_{l}>p_{j}$ as a job $j$ was present in $S$ at the deletion of job $l$. Thus, the repositioning of job $l$ back to $S$ increases the loads of the machines. If $l<j$, then the repositioning of job $l$ back to $S$ creates infeasibility among the jobs contained in $S$ at the end of the iteration when job $l$ was moved to $T$. Thus, another job must be deleted, that is, the cardinality of $S$ is not increased. It follows from the selection of job $j$ that $S$ has the minimal possible total processing time among the sets of jobs that satisfy property (2).

Relationship between the problems $P \| \Sigma U_{j}$ and $P \| C_{\text {max }}$ In the problem $P \| C_{\text {max }}$, the makespan, that is, the greatest completion time, is to be minimized. If it is greater than the maximal due date in the problem $P \| \Sigma U_{j}$, that is, $d_{n}=\max$ $\{d j: 1 \leq j \leq n\}$, then at least one tardy job must exist in $P \| \Sigma U_{j}$. Let $Q$ be an optimization problem. The optimal value of $Q$ is denoted by $\nu(Q)$. More formally, the statement is:

$$
\begin{equation*}
\nu\left(P \| C_{\text {max }}\right)>d_{n} \operatorname{implies} \nu\left(P \| \Sigma U_{j}\right) \geq 1 \tag{6}
\end{equation*}
$$

It is not necessary to solve the $P \| C_{\text {max }}$ problem with a complicated and long algorithm to obtain a conclusion similar to equation (6). The theory of the $P \| C_{\text {max }}$ problem is rich, and there are easy tools to get such a conclusion. The first thing to be mentioned is that there is a well-known lower bound of $\nu\left(P \| C_{\text {max }}\right)$. Let $m$ be the number of machines. The lower bound is:

$$
\begin{equation*}
L B=\max \left\{\frac{\sum_{j=1}^{n} p_{j}}{m}, \max \left\{p_{j}: 1 \leq j \leq n\right\}\right\} \tag{7}
\end{equation*}
$$

The first term in equation (7) is the average load of the machines, and the second term is the longest processing time, where $m$ stands for the number of machines, which is vehicles in this paper, and $p_{j}$ is the processing time of the job $j$, which is patient $j$ in this paper. Both are individual lower bounds. There is an upper bound of $\nu\left(P \| C_{\text {max }}\right)$, which is:

$$
\begin{equation*}
U B=\max \left\{\frac{2 *\left(\sum_{j=1}^{n} p_{j}\right)}{m}, \max \left\{p_{j}: 1 \leq j \leq n\right\}\right\} \tag{8}
\end{equation*}
$$

The first term in equation (8) is double the average load of the machines.

A large part of the literature on $P \| C_{\text {max }}$ started with the celebrated paper by Graham (1969). This paper analyzed a heuristic method of $P \| C_{\max }$ called the LPT list scheduling. The method assigns the jobs to machines in the decreasing order of processing times. Each job is assigned to the temporarily least loaded machine (Graham, 1969). Graham (1969) proved an implicit lemma in the proof of his main theorem.

Lemma 1. If the LPT list scheduling assigns to each machine two jobs at most, then the solution provided is optimal (Graham, 1969).

Lemma 1 implies that the solution to the problem can be obtained easily if the number of jobs is low relative to the number of machines.

The LPT list scheduling and any other heuristic methods can be used in the opposite direction as well. If all jobs are on time
in the solution provided by a heuristic method, then the answer of $\operatorname{oracle}(S)$ is YES. The LPT list scheduling method has some improvements in the performance rate (Dósa and Vizvári, 2004). The optimal positions of the next $k$ jobs are determined in each iteration of this method. However, only the first of the next $k$ is assigned to its optimal position. If $k$ is fixed, then the complexity of the iterations is $O(1)$.
The performance of a heuristic method for a positive minimization problem can be characterized by the so-called performance ratio. This is the largest possible value of the ratio between the value of a solution provided by the heuristic method and the optimal value of the same problem instance. For example, let $C$ (LPT) be the makespan by the LPT list scheduling. Then:

$$
\begin{equation*}
\sup _{P}: \frac{C(L P T)}{\nu\left(P \| C_{\max }\right)} \tag{9}
\end{equation*}
$$

is the performance ratio of the LPT list scheduling.

## Fast heuristics using a partial oracle

As the oracle must solve an NP-complete problem, it is not reasonable to assume that a perfect oracle can be applied in real practice. Therefore, a weaker notion is introduced, which can be a procedure of a computer program. A partial-oracle $(X)$ is a partially defined Boolean-valued function, that is, its value is either YES or NO or UNDEFINED.

The partial oracle contains procedures for proving and disproving feasibility. The construction of a feasible solution is tried by implementing heuristic methods. If a method is successful, then the feasibility is proven. On the other hand, if a lower bound on the makespan is greater than the last due date, a tardy job must exist, that is, the problem is infeasible. It is possible, of course, that none of these methods determines the solution. In that case, the answer of partial-oracle $(X)$ is UNDEFINED, which means "I don't know."

The important question is which algorithmic elements are worth including in the partial oracle. Constructive heuristics should be able to find a feasible solution if one exists. In general, each heuristic method has a "philosophy" about how it is possible to obtain a good solution. Each "philosophy" works for certain types of problem instances. Thus, it is worth having different heuristic methods in the partial oracle. For example, LPT list scheduling and multi-fit heuristics of $P \| C_{\text {max }}$ have different approaches. There can be a third heuristic method superior to these two methods. However, it is difficult to find problem instances such that the performance ratio of the third one is high (Vizvári and Demir, 1992).

## Algorithmic tools for proving the feasibility

The applied heuristic methods have three parts. First, to form an ordered list of patients, an ordering rule is required. The patients are assigned to vehicles in this order. Second, to select an appropriate vehicle for the patient, applying the vehicle selection rule is required. Third, if the selected patient is not saved, then a patient rejection rule is implemented to reject a patient from the scheduled and selected patients, no matter which vehicle the patient is assigned to, and to place the patient in the set of removed patients. That patient is always the one with the longest travel time.

The first and second heuristic approaches are called earliest due-date best fit (EDDBF) and earliest due date worst fit (EDDWF), respectively, BF for the best fit and WF for the worst fit. The EDD dispatching rule is used in both EDDBF and EDDWF. EDDBF assigns the current patient to the least loaded vehicle. EDDWF assigns the patient to the vehicle that can still deliver the patient on time and has the maximum load among such vehicles. The third heuristic method is the LPT list scheduling, which is LPTBF.

If none of the constructive heuristics gives a feasible solution and the disproving tools do not work, then an attempt can be made to improve the solution. This means the exchange of jobs between two machines (M1 and M2). This is shown by an example in Table 1.

Let $C_{j}$ be the completion time of job $j$. The earliness of job $j$ denoted by $E_{j}$ is $E_{j}=d_{j}-C_{j}$. If $E_{j} \geq 0$, then the job is on time. If $E_{j}<0$, then the job is tardy. Assume that the EDDBF algorithm executes five iterations, that is, $k=5$ in the structural algorithm. The current schedule is shown in Table 2.

In the above example, it can be seen that job 5 is tardy. However, job 2 and job 5 can be interchanged such that job 2 becomes the second job of M1 and job 5 moves to the end of the M2 row. This is possible because job 2 does not become tardy in this position and the job which will be processed after job 2 on M1 has an earliness greater than the processing time of job 2 . The new solution is shown in Table 3.
Assume that the schedule of M1 and M2 starts at time 0. The sequence of the jobs on M 1 is $i_{1}, i_{2}, \ldots, i_{k}, \ldots, i_{p}$. Similarly, the jobs on M2 are $j_{1}, j_{2}, \ldots, j_{1}, \ldots, j_{q}$. The completion times of the jobs are as follows:

$$
\begin{equation*}
C_{i_{k}}=\sum_{t=1}^{k} C_{i_{i}} k=1, \ldots, p, C_{j_{l}}=\sum_{u=1}^{l} C_{j_{u}} l=1, \ldots, q . \tag{10}
\end{equation*}
$$

The earliness of the jobs are:

Table 1 Data of the example

|  |  | Job |  |  |  |  |  |
| :--- | :--- | :--- | :--- | ---: | ---: | ---: | :---: |
|  | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ |  |
| $p_{\boldsymbol{j}}$ | 3 | 2 | 5 | 4 | 6 | 5 |  |
| $d_{j}$ | 6 | 8 | 9 | 10 | 11 | 12 |  |

Table 2 Result of the EDDBF algorithm

| M1 | Job | 1 | 4 | 5 | M2 | Job | 2 | 3 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $p_{j}$ | 3 | 4 | 6 |  | $p_{j}$ | 2 | 5 |
|  | $C_{j}$ | 3 | 7 | 13 |  | $C_{j}$ | 2 | 7 |
|  | $E_{j}$ | 3 | 3 | -3 |  | $E_{j}$ | 6 | 2 |

Table 3 Improved schedule after the interchange

| M1 | Job | 1 | 2 | 4 | M2 | Job | 3 | 5 |
| ---: | ---: | ---: | ---: | :--- | :--- | :--- | :--- | ---: |
|  | $p_{\boldsymbol{j}}$ | 3 | 2 | 4 |  | $p_{j}$ | 5 | 6 |
|  | $C_{j}$ | 3 | 5 | 9 |  | $C_{j}$ | 5 | 11 |
|  | $E_{j}$ | 3 | 3 | 1 |  | $E_{j}$ | 4 | 0 |

## Tareq Babaqi and Béla Vizvári

$$
\begin{equation*}
E_{i_{k}}=\max \left\{d_{i_{k}}-C_{i_{k}}, 0\right\} k=1, \ldots, p \tag{11}
\end{equation*}
$$

Assume that the current job is $j_{q}$, which is tardy, that is, $C_{j_{q}}>d_{j_{q}}$. To make the current job on time, a job of M2, say job $j_{l}$, is to be exchanged with a job of M 1 , say job $i_{k}$, where $p_{j l}>p_{i k}$. This change makes sense only if no new tardy job is created on M1. The condition for that is that the earliness of the jobs on M1 behind job $i_{k}$ is not less than the processing time surplus of job $j_{l}$, that is,

$$
\begin{equation*}
p_{j_{l}}-p_{i_{k}} \leq E_{i_{t}} t=k+1, \ldots, p \tag{12}
\end{equation*}
$$

The change is successful if:

$$
\begin{equation*}
\sum_{u=1}^{q} p_{j_{u}}+p_{i_{k}}-p_{j_{l}} \leq d_{j_{q}} \tag{13}
\end{equation*}
$$

## Algorithmic tools for disproving feasibility

The simplest potential way of disproving feasibility is coded by equations (6) and (7). If the inequality:

$$
\begin{equation*}
\max \left\{\frac{\sum_{j=1}^{k} p_{j}}{m}, \max \left\{p_{j}: 1 \leq j \leq k\right\}\right\}>d_{k} \tag{14}
\end{equation*}
$$

holds, there is no feasible solution to the subproblem of iteration $k$.

Shavarani and Vizvari (2018) gave an upper bound on the total working time of the machines. Assume that the index order is the EDD order and all jobs are completed on time. Then, the total working time of the machines is the last $m$ due dates at most (Shavarani and Vizvari, 2018). This implies the following inequality:

$$
\begin{equation*}
\sum_{j=1}^{n} p_{j} \leq \sum_{j=n-m+1}^{n} d_{j} \tag{15}
\end{equation*}
$$

is a necessary condition of the feasibility, that is, the total processing time cannot exceed the last $m$ due dates for each iteration. The opposite inequality disproves the feasibility.

Lemma 2. Assume that the index order is identical to the EDD order. Let $\pi(k, j), j=1, \ldots, k$ be the SPT order of the first $k$ job, that is:

$$
\begin{equation*}
\{1,2, \ldots, k\}=\{\pi(k, 1), \pi(k, 2), \ldots, \pi(k, k)\} \tag{16}
\end{equation*}
$$

and

$$
\begin{equation*}
p_{\pi(k, 1)} \leq p_{\pi(k, 2)} \leq \cdots \leq p_{\pi(k, k)} \tag{17}
\end{equation*}
$$

If $1 \leq t \leq k$ is an integer such that:

$$
\begin{equation*}
\sum_{j=1}^{t} p_{\pi(k, j)} \leq \sum_{j=k-m+1}^{k} d_{j}<\sum_{j=1}^{t+1} p_{\pi(k, j)} \tag{18}
\end{equation*}
$$

then at least $k-t$ jobs are late.
Heuristic methods also give lower bounds on the makespan if the performance ratio is known. The main theorem of Graham (1969) is an exact upper bound on the performance ratio.

Theorem 2. Let $Q$ be an instance of the problem class $P \| C_{\text {max }}$. In the case of $m$ machines, its optimal value and the value
obtained by the LPT list scheduling are $\nu(Q)$ and $L P T(Q)$, respectively (Graham, 1969). Then,

$$
\begin{equation*}
\frac{L P T(Q)}{\nu(Q)} \leq \frac{4}{3}-\frac{1}{3 m} \tag{19}
\end{equation*}
$$

The theorem implies that for a problem instance $Q$,

$$
\begin{equation*}
\nu(Q) \geq \frac{3 m L P T(Q)}{4 m-1} \tag{20}
\end{equation*}
$$

Graham (1969) gave an example where the performance ratio is exactly the upper bound. Dósa (2004) proved that Graham's example is the only one having this performance ratio. He gave some performance ratios for $m=2,3$ that were better than Graham's general performance ratio for the improved version of the LPT list scheduling (Dósa, 2004).

## Problem in a dynamic environment

In the post-disaster period, reports on injured persons are received continuously. Therefore, it would be difficult to solve the exact optimization problem in the case of the dynamic environment because the information needed to set up the model is changing continuously. The exact solution needs a long CPU time, which is not acceptable, as fast decisions must be made on the assignment of patients to the next mission. Thus, the heuristic method is repeated after receiving a new report on an injured person in the case of a real disaster.

## Case study

As a case study, Tehran city, the capital of Iran, is considered. Tehran is in the middle of a triangle of three geographical faults, and each fault can generate an earthquake (Figure 2). Thus, the city is under constant threat of an earthquake. The map and hospitals of the city are acquired from ArcGIS software (ESRI, 2017). For this study, we assume that there are two ambulances at each hospital in the city and consider the average speed of 60 $\mathrm{km} / \mathrm{h}$ for the ambulances. We assume that 115 hospitals operate in the post-disaster period. We generate the deadlines and locations of the seriously injured patients randomly; there are 100 seriously injured patients in each region of the city. The minimum, maximum, mean and standard deviation of the deadlines are $30,300,166.7$ and 83.3 min , respectively.

Because of the high runtime required and the complexity of calculating the distance between patients and hospitals, the distance between every hospital and any patient in its region is determined using the Manhattan-like distance $\left(l_{1}\right)$. The situation is shown in Figure 3, where the city is divided into regions by using ArcGIS to implement a Voronoi diagram between the hospitals of the city on the map. The divided regions have different features, as shown in Figure 3. The area of some of the regions is big, whereas others are smaller. When the area covered is huge, the problem becomes more complicated in the sense of the number of ambulances in the corresponding hospital. For instance, there are many hospitals in the center of the city, which implies that the area covered by each hospital in the region is small. Thus, a lot of people can be saved even with two ambulances.

Figure 2 Hospitals in Tehran city


Figure 3 Voronoi diagram of hospitals and the three studied regions in Tehran city


## Results and discussion

In this section, we report the results of the ILP model obtained by CPLEX 12.8. The proposed heuristics are implemented in $\mathrm{C}++$ programming language and tested in Windows 10 with Intel ${ }^{\circledR}$ Core $^{\text {TM }}$ i5-5200U CPU @ 2.20 GHz with a memory of 6 GB and their solutions are summarized. Ten different scenarios are taken for each size and compared with the EDDBF, EDDWF and LPT heuristics. The results of those scenarios are reported in Table 4. In this table, we indicate the number of patients saved and the computational time (in s) of CPLEX12.8 to find the optimal solution to the ILP problems (1)-(4). Also, let $n$ be the number of patients considered in the simulations. The optimal solutions are performed for two vehicles and ten different scenarios for each $n$.
As shown in Table 4, the computational time increases as $n$ increases because of the high complexity of the problem. To find an optimal solution of size 10 , the CPU time required is approximately 50 min . Thus, problems of size $n=10$ or larger

Table 4 Average solutions for ten different problem instances in case of a small size problem

| Problem size ( $\boldsymbol{n}$ ) | EDDBF | EDDWF | LPT | Opt. CPLEX | Time ( $\mathbf{s}$ ) |
| :--- | :---: | :---: | :---: | :---: | ---: |
| 5 | 3 | 3 | 2 | 4 | 0.06 |
| 6 | 4 | 4 | 2 | 4 | 0.10 |
| 7 | 4 | 4 | 2 | 4 | 0.32 |
| 8 | 4 | 4 | 2 | 5 | 2.42 |
| 9 | 5 | 5 | 2 | 5 | 69.30 |
| 10 | 5 | 5 | 2 | 5 | $3,078.03$ |

cannot be solved precisely in the case of a disaster because of the exponential increase in CPU time (see Figure 4). This phenomenon is well-known in general and is called a combinatorial explosion. The EDDBF and EDDWF heuristics give optimal solutions in most of the instances in each size. However, in a few other instances, the solutions obtained are

Figure 4 CPU times needed to obtain the optimal solution as the function of the number of patients

near-optimal only. The results are significantly less than optimal in all instances solved by the LPT heuristic. The use of fast heuristic methods is justified by the fact that an emergency environment requires immediate answers.
The average percentage and time of the expected saved patients in three different regions of Tehran city are summarized in Table 5. These regions, called 1, 2 and 3, respectively, are shown in Figure 3 from left to right. The number of persons, $n=100$, is fixed, with ten random location distributions in each selected region. Their upper bounds are calculated, and on average, 64, 100 and 58 persons are found in regions 1,2 and 3 , respectively. Thus, it is not possible to save all the patients in regions 1 and 3. This shows that LPT has the smallest mean percentage of expected saved patients of $72.4 \%$, $77.8 \%$ and $75.1 \%$ in regions 1,2 and 3, respectively. EDDBF is very closely followed by EDDWF, with the means of expected saved patients being $91.5 \%$ and $88.7 \%$ in regions 1 and 3, respectively. The results of the simulation also reveal other facts. First, EDDBF performs better than EDDWF in $100 \%$ of the test problems in region 1, performs equally well in region 2 , but performs better than (worse than) EDDWF in $20 \%$ of the test problems in region 3, and is equally good in $60 \%$ of the test problems. Second, EDDWF and EDDBF perform better than LPT in all the test problems.
Table 6 shows the percentage of maximum, minimum and average definite answers (YES + NO) in all problem instances using the tools of a partial oracle. When the three heuristics, EDDBF, EDDWF and LPT, and the lower bounds and upper bounds tools are used in the partial oracle, the maximum definite answer in regions 2 and 3 is $100 \%$. However, this percentage exists in EDDBF and EDDWF in all problem instances of region 2, whereas it is found only once in region 3 in EDDWF. It is interesting to note that region 2 has the

Table 6 Percentage of the definite answers of partial oracle

| Regions | Maximum | Minimum | Average |
| :--- | :---: | :---: | :---: |
| $\mathbf{1}$ | 98 | 46 | 74.73 |
| $\mathbf{2}$ | 100 | 51 | 86.10 |
| $\mathbf{3}$ | 100 | 43 | 66.87 |

maximum number of definite answers compared to the other regions because of the small area of that region.

## Conclusions

After a disaster, a substantial number of injured people must be transported to hospitals, and the time between the injury and the arrival at the hospital plays a significant role in the outcome of the rescue procedures. This study provides the tools for saving the maximum possible number of people. It further provides previously unreported scheduling and obligation algorithms for disaster managers.
The problem of the transportation of injured people to hospitals in large cities like San Francisco and Tehran is largescale, complex and dynamic. Furthermore, it demands many big real-time decisions. Thus, the transportation problem must be decomposed, and fast algorithms giving results of good quality must be elaborated to ensure better post-disaster recovery.
The suggested new method is a synthesis of mathematical modeling, scheduling theory, heuristic methods and the Voronoi diagram of geometry. The results of the paper are as follows:

- The Voronoi diagram can be effectively used for decomposing the complex problem.

Table 5 Average percentage and time of the expected saved patients

|  | Region 1 <br> Expected saved <br> patients (\%) | Completion <br> time $(\mathbf{m i n})$ | Region 2 <br> Expected saved <br> patients $(\%)$ | Completion <br> time $(\mathbf{m i n})$ | Region 3 <br> Expected saved <br> patients (\%) | Completion <br> time ( $\mathbf{m i n}$ ) |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Heuristic | 91.5 | 320.86 | 100 | 253.19 | 88.7 | 313.99 |
| EDDBF | 89.3 | 333.34 | 100 | 311.79 | 87.8 | 328.99 |
| EDDWF | 72.4 | 268.34 | 77.8 | 194.38 | 75.1 | 274.47 |
| LPT |  |  |  |  |  |  |

- The mathematical model of the transportation of casualties to one hospital is the $P \| \sum U_{j}$ problem of scheduling theory. It is a known but very rarely investigated problem. This study provides a new mathematical theorem to describe the structure of an algorithm that provides the optimal solution.
- The study introduces the notion of the partial oracle, an algorithmic tool. It aids in the development of heuristic methods that provide approximations to the precise algorithmic method in (2).
- The realization of the partial oracle with constructive algorithmic elements and elements proves the nonexistence of any solution.
- Lemma 2 belongs to the second kind of elements and is a new mathematical result.
- The paper presents the findings from the example of three hospitals in the metropolitan city of Tehran and examines the heuristics.

It finds that the heuristics based on the EDD order always give better results than the LPT-based method. The results are close to the best possible results that can be achieved. However, obtaining the optimal solution requires a very long CPU time because the problem $P \| \sum U_{j}$ is NP-complete, as proven in the Appendix.
The main aim of this study was to create an insight into the assignment of medical staff to different operating rooms in different hospitals to make these units more successful in postdisaster periods. Some relaxations have been assumed in this study. More detailed models can be created as an outcome of this study in the future. Further research should investigate and address the social circumstances and ethical issues related to the post-disaster period in general and the transportation problem in particular. Moreover, the dynamic environment of the postdisaster period needs further investigation and analysis.

## References

Altay, N. and Green, W.G., III (2006), "OR/MS research in disaster operations management", European Fournal of Operational Research, Vol. 175, pp. 475-493.
Amadini, R., Sefrioui, I., Mauro, J. and Gabbrielli, M. (2013), "Fast post-disaster emergency vehicle scheduling", in Distributed Computing and Artificial Intelligence, Springer, Cham, pp. 219-226.
Aringhieri, R. (2020), "Online optimization in health care delivery: overview and possible applications", Operations Research Proceedings, pp. 357-363.
Aringhieri, R., Bigharaz, S., Duma, D. and Guastalla, A. (2021), "Fairness in ambulance routing for post disaster management", Central European fournal of Operations Research, Vol. 30 No. 1, pp. 189-211.
Aurenhammer, F., Klein, R. and Lee, D.-T. (2013), Voronoi Diagrams and Delaunay Triangulations, World Scientific Publishing Company.
Babaqi, T., Al-Natoor, S.R. and Vizvári, B. (2022), "Design and preparation of mass production system for protective items in the European union for the case of serious pandemics", Disaster Medicine and Public Health Preparedness, pp. 1-5.

Bedworth, D.D. and Bailey, J.E. (1987), Integrated Production Control System, John Wiley and Sons. Inc., New York.
Bozorgi-Amiri, A., Tavakoli, S., Mirzaeipour, H. and Rabbani, M. (2017), "Integrated locating of helicopter stations and helipads for wounded transfer under demand location uncertainty", The American fournal of Emergency Medicine, Vol. 35 No. 3, pp. 410-417.
Bullock, J.A., Haddow, G.D. and Coppola, D.P. (2013), "Mitigation, prevention, and preparedness", Introduction to Homeland Security, Vol. 2013, pp. 435-494.
Chou, Y.-L., Romeijn, H.E. and Smith, R.L. (1998), "Approximating shortest paths in large-scale networks with an application to intelligent transportation systems", INFORMS fournal on Computing, Vol. 10, pp. 163-179.
Delmonteil, F.-X. and Rancourt, M.-E. (2017), "The role of satellite technologies in relief logistics", fournal of Humanitarian Logistics and Supply Chain Management, Vol. 7, pp. 57-78.
Dósa, G. (2004), "Graham's example is the only tight one for P $\|$ Cmax", Annales Universitatis Scientarium Budapestiensis de Rolando Eötös Nominatae, Sectio Mathematica, Vol. 47, pp. 207-210.
Dósa, G. and Vizvári, B. (2004), "The LPT(k) algorithm for the scheduling of identical parallel machines (in Hungarian)", Alkalmazott Matematikai Lapok, pp. 269-289.
Drysdale, S. (1996), "Geometry in action", Donald Bren School of Information and Computer Sciences at University of California, Irvine, available at: www.ics.uci.edu/~eppstein/gina/scot. drysdale.html (accessed 10 November 2018).
ESRI (2017), "ArcGIS Desktop10.7", Environmental Systems Research Institute.
Goebl, H. (2010), "Dialectometry: theoretical pre-requisites, practical problems, and concrete applications (mainly with examples draw from the 'Atlas linguistique de la France', 1902-1910)", Dialectologia: Revista Electrònica, pp. 63-77.
Graham, R.L. (1969), "Bounds on multiprocessing timing anomalies", SIAM journal on Applied Mathematics, Vol. 17, pp. 416-429.
Ho, J.C. and Chang, Y.-L. (1995), "Minimizing the number of tardy jobs for m parallel machines", European Fournal of Operational Research, Vol. 84, pp. 343-355.
Hodgson, T. (2018), "RE: email to the co-author Béla Vizvári. The complete text of the email is as follows: "Mikael J. Moore published the original paper in management science (about 1968). Mike was my office mate at the Univ of Michigan. My modification was included in that paper. Then, since no proof was given for my algorithm, 3 papers were published in different journals proving optimality (which is obvious) and called my version Hodgson's algorithm, then ken baker (Dartmouth) included it in his scheduling book. So, I am famous for my algorithm without a formal reference of my own".
Jin, S., Jeong, S., Kim, J. and Kim, K. (2014), "A logistics model for the transport of disaster victims with various injuries and survival probabilities", Annals of Operations Research, Vol. 230, pp. 17-33.
Jotshi, A., Gong, Q. and Batta, R. (2009), "Dispatching and routing of emergency vehicles in disaster mitigation using data fusion", Socio-Economic Planning Sciences, Vol. 43, pp. 1-24.
Kadri, F., Chaabane, S. and Tahon, C. (2014), "A simulationbased decision support system to prevent and predict strain
situations in emergency department systems", Simulation Modelling Practice and Theory, Vol. 42, pp. 32-52.
Kallehauge, B. (2008), "Formulations and exact algorithms for the vehicle routing problem with time windows", Computers $\mathcal{E}$ Operations Research, Vol. 35, pp. 2307-2330.
Kula, U., Tozanli, O. and Tarakcio, S. (2012), "Emergency vehicle routing in disaster response operations", POMS 23rd Annual Conference, Chicago (April 20-23).
$\mathrm{Li}, \mathrm{Y}$. and Tsukaguchi, H. (2003), "Improving the reliability of street networks in highly densely populated urban areas: proceedings of the 1st international symposium on transportation network reliability (INSTR)", The Network Reliability of Transport, Emerald Group Publishing, Bingley, pp. 261-272.
Liu, Z., Yu, H., Sui, J., Liu, D. and Wang, J. (2011), "A research on vehicle scheduling problem to rescue the victims from chemical and biological terrorist attacks", 2011 IEEE international conference on automation and logistics (ICAL), IEEE, pp. 356-361.
Molenbruch, Y., Braekers, K., Caris, A. and Berghe, G.V. (2017), "Multi-directional local search for a bi-objective dial-a-ride problem in patient transportation", Computers $\mathcal{E}$ Operations Research, Vol. 77, pp. 58-71.
Moore, J.M. (1968), "An $n$ job, one machine sequencing algorithm for minimizing the number of late jobs", Management Science, Vol. 15, pp. 102-109.
Najat, A., Yuan, C., Gursel, S. and Tao, Y.J.P.M. (2019), "Minimizing the number of tardy jobs on identical parallel machines subject to periodic maintenance", Procedia Manufacturing, Vol. 38, pp. 1409-1416.
Ozdamar, L. (2011), "Planning helicopter logistics in disaster relief", OR Spectrum, Vol. 33, pp. 655-672.
Papp, B. (2019), "Disaster risk data and its terminological difficulties: a statistical review", Delta: Vedecko-Odborný Časopis Katedry Protipožiarnej Ochrany, Vol. 13, pp. 5-21.
Pınarbaşı, A., Babaqi, T. and Vizvári, B. (2022), "On the evaluation of the ambulance capacity of the Asian side of Istanbul in the case of a serious earthquake", Disaster medicine and Public Health Preparedness, Vol. 16 No. 2, pp. 510-519.
Pinedo, M. (1995), Scheduling: Theory, Algorithms and Systems, Prentice-Hall, UK.
Shavarani, S.M. and Vizvari, B. (2018), "Post-disaster transportation of seriously injured people to hospitals", Fournal of Humanitarian Logistics and Supply Chain Management, Vol. 8, pp. 227-251.
Süer, G.A., Báez, E. and Czajkiewicz, Z. (1993), "Minimizing the number of tardy jobs in identical machine scheduling", Computers Industrial Engineering, Vol. 25, pp. 243-246.
Sze, S.M. and Ng, K.K. (2006), "pn Junctions", Physics of Semiconductor Devices, John Wiley \& Sons, Vol. 2, pp. 80-89.
Thomson, J.J. (1985), "The trolley problem", The Yale Law fournal, Vol. 94, pp. 1395-1415.
Thrun, S. (1998), "Learning metric-topological maps for indoor mobile robot navigation", Artificial Intelligence, Vol. 99, pp. 21-71.

Tlili, T., Harzi, M. and Krichen, S. (2017), "Swarm-based approach for solving the ambulance routing problem", Procedia Computer Science, Vol. 112, pp. 350-357.
United Nations (2007), "Disaster risk reduction: 2007 global review - UNISDR", available at: www.undrr.org/publication/ disaster-risk-reduction-2007-global-review (accessed 3 March 2021).

Vizvári, B. and Demir, R. (1992), "It is difficult to find a difficult problem for the scheduling of identical parallel machines", Department of Industrial Engineering of Bilkent University, Research Report, IEOR-9212.
Vizvári, B., Golabi, M., Nedjati, A., Gümüşbuğa, F. and Izbirak, G.J. (2019), "Top-down approach to design the relief system in a metropolitan city using UAV technology, part I: the first 48 h", Natural Hazards, Vol. 99, pp. 571-597.
Voronoi, G. (1908), "Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Premier mémoire. Sur quelques propriétés des formes quadratiques positives parfaites", fournal fur die Reine und Angewandte Mathematik, Vol. 134, p. 198.
Wang, F., Cheng, Q., Highland, L., Miyajima, M., Wang, H. and Yan, C. (2009), "Preliminary investigation of some large landslides triggered by the 2008 Wenchuan earthquake, Sichuan province, China", Landslides, Vol. 6, pp. 47-54.
White, D.B., Katz, M., Luce, J., Lo, B., Daugherty-Biddison, L. and Toner, E. (2020), "Allocation of scarce critical care resources during a public health emergency", University of Pittsburgh, Department of Critical Care Medicine, March, 26.
Whybark, D.C. (2015), "Co-creation of improved quality in disaster response and recovery", International fournal of Quality Innovation, Vol. 1, pp. 1-10.
Wren, A. and Holliday, A. (1972), "Computer scheduling of vehicles from one or more depots to a number of delivery points", fournal of the Operational Research Society, Vol. 23, pp. 333-344.
Zhan, F.B. and Noon, C.E. (1998), "Shortest path algorithms: an evaluation using real road networks", Transportation Science, Vol. 32, pp. 65-73.

## Appendix

Theorem. The complexity of the problem $P \| \Sigma U_{j}$ is at least NP-complete.
Proof. Assume that the number of machines is 2 , the sum of the processing times is $2 d$, where $d$ is a positive integer, and all due dates are also $d$. Then, the problem is reduced to finding a subset of the jobs, such that the sum of the processing times of the jobs in the subset is exactly $d$. This problem is a partition problem, which is one of the most well-known NP-complete problems. If a subproblem is NP-complete, then the problem class is at least NP-complete.

## Corresponding author

Tareq Babaqi can be contacted at: ie.tareq@gmail.com

