FAIR FACILITY ALLOCATION IN EMERGENCY SERVICE SYSTEM

Jaroslav JANÁČEK[1], Lýdia GÁBRIŠOVÁ[2], Miroslav PLEVNÝ [ID][3*]

[1, 2]Department of Mathematical Methods and Operations Research,
Faculty of Management Science and Informatics, University of Žilina, Žilina, Slovakia
[3]Department of Economics and Quantitative Methods, Faculty of Economics,
University of West Bohemia in Pilsen, Univerzitni 22, 306 14 Plzeň, Czech Republic

Abstract. The request of equal accessibility must be respected to some extent when dealing with problems of designing or rebuilding of emergency service systems. Not only the disutility of the average user but also the disutility of the worst situated user must be taken into consideration. Respecting this principle is called fairness of system design. Unfairness can be mitigated to a certain extent by an appropriate fair allocation of additional facilities among the centers. In this article, two criteria of collective fairness are defined in the connection with the facility allocation problem. To solve the problems, a series of fast algorithms for solving of the allocation problem was suggested. This article extends the family of the original solving techniques based on branch-and-bound principle by newly suggested techniques, which exploit either dynamic programming principle or convexity and monotony of decreasing nonlinearities in objective functions. The resulting algorithms were tested and compared performing numerical experiments with real-sized problem instances. The new proposed algorithms outperform the original approach. The suggested methods are able to solve general min-sum and min-max problems, in which a limited number of facilities should be assigned to individual members from a finite set of providers.

Keywords: emergency system design, collective fairness, equal accessibility, allocation problem, dynamic programming, polynomial approach.

JEL Classification: C610, R530.

Introduction

An important task of state administration and self-government is to ensure public safety and health. Organization called emergency services and rescue services are established for this purpose to address different emergencies. The problems that have to be solved are multi-faceted, complex, and seemingly insoluble. That's a way the policymakers are increasingly looking for new options (Marsh et al., 2017; Koval et al., 2019). The three primary emergency services that can be summoned directly by the public are emergency medical services, fire-

*Corresponding author. E-mail: *plevny@kem.zcu.cz*

fighters, and police. These service systems are designed and operated to provide the population of a geographical region with the emergency service. The availability of emergency services depends very heavily on location.

Probably, the most important efficacy measurement of emergency services is the response time, i.e. the amount of time that it takes for emergency responders to arrive at the scene of an incident after activating the emergency system. Fast response times are often a crucial component of the emergency service system (Davis, 2005).

As said above, to manage such a system is a complex problem involving, from the public's point of view, very sensitive issues such as security, saving of lives or property protection on the one hand, and the technical and especially economic capabilities of the emergency service provider on the other (Staňková et al., 2018; Cyganska, 2017). One of the possible ways to solve these problems is to use the tools of mathematical optimization. An essential factor in using this approach is the precise formulation of the model and, in particular, the criteria for assessing suitable solutions and the question of the solvability of the constructed models in a reasonable time.

The most common organizational model for providing an emergency service system is possible to describe as follows. A finite set of service centers located at nodes of the road network of the region represents a structure of the emergency system. Demands for service originate mostly at dwelling places of the region. The set of dwelling places is denoted as the set of system users, where each user is specified by his location and frequency of demands. When a service system structure is to be designed, the number of service centers is usually given. Disutility perceived by an individual user is proportional to the network distance between the user and the nearest service center.

In this paper, a problem of fair allocation of additional facilities in the given set of emergency service centers is dealt. The objective of the associated problem is to deploy the additional facilities in the fairest way. To solve the problem, a series of fast algorithms is suggested that were tested and compared performing numerical experiments with real-sized problem instances.

The paper is organized as follows. Section 1 presents literature review on designing the emergency system, Section 2 introduces the linear model approach to modelling collective fairness in emergency systems followed by dynamic programming approach in Section 3. An alternative polynomial approaches to the fair facility deployment is described in Section 4. Section 5 presents results of numerical experiments comparing the methods developed in previous sections, and the final section summarizes presented findings and concludes.

1. Literature review

A large number of papers can be found in published literature dealing with various aspects relating to the design and optimization of emergency service systems. Several aspects affecting the availability and response time of these services were examined in these contributions.

The issues concerning the patient flow within the health system were studied by Lau et al. (2018). Mitigating the patient flow bottlenecks can enhance workflow efficiency and reduce patient wait-time. Similarly, Carmen, Van Nieuwenhuyse, and Van Houdt (2018) studied

the impact of inpatient boarding on emergency department congestion and capacity using a semi-open queueing network model with limited resources and discontinuous patient service.

Another aspect of designing the emergency service systems studied in several papers is the problem of emergency service bases location. Nogueira, Pinto, and Silva (2018) deal with this problem on a case of ambulance base locations in Brazil. Similarly, Stanimirovic et al. (2017) studied the problem to choose locations for establishing emergency units on real-life instances obtained from two networks of police units in Montenegro and Serbia. A two-stage stochastic programming location-allocation model is proposed by Boujemaa et al. (2018) to simultaneously determine the location of ambulance stations, the number and the type of ambulances to be deployed, and the demand areas served by each station. Research carried on the ambulance location and management in the Milano area (Italy) is reported by Aringhieri, Carello, and Morale (2016).

The limited personal and/or technical resources can be considered probably the most common aspect in practice concerning performance and perceived quality of an emergency service system. This topic was studied in a number of papers. For instance, Vermuyten et al. (2018) studied the staff scheduling problem encountered in practice at an Emergency Medical Services system. Analogically, Garner and van den Berg (2018) estimated the optimal helicopter emergency medical services base locations within New South Wales using advanced mathematical modeling techniques. A new double standard model, along with a genetic algorithm, is introduced by Liu et al. (2016) for solving the emergency medical service vehicle allocation problem that ensures acceptable service reliability with limited vehicle resources.

In the following text, the above-mentioned aspect is considered with limited resources where a service system structure is to be designed and the number of service centers is given. Depending on the applied objective, the service system structure design is formulated as a kind of either the weighted p-median or p-center problem (Brotcorne et al., 2003; Marianov & Serra 2002; Ingolfsson, et al., 2008; Chanta et al., 2014). A host of publications has been devoted to the problem formulation and to the associated methods of the problem solving (Jánošíková, 2007; Janáček & Gábrišová, 2009). Once an emergency system structure is determined, clusters of users are also uniquely designated according to the rule that a user's demand is allocated to the nearest service center. As cluster demands are primarily satisfied by facilities located at the cluster service center, a disutility volume perceived by users involved in the cluster is inversely proportional to the number of facilities allocated at the service center. This proposition follows from the simple fact that if the facilities are currently occupied by earlier demands, the newly emerged demands must wait. This way the average response time in the cluster increases, when there is located an insufficient number of facilities regarding the given frequency of demand occurrence and the average time-distance from the center to a user. Due to the uneven distribution of the population in the region, the populations of the individual clusters are exposed to various degrees of disutility.

Perceived disutility is influenced by the total cluster workload per one facility located at the cluster center. The amount of a cluster workload is quantified by the so-called transportation performance necessary for the satisfaction of the cluster demands. The cluster workload is defined as a sum of time distances from the cluster users to the nearest service center mul-

tiplied by the demand frequencies. Unfairness following from the uneven workload distribution can be mitigated to a certain extent by an appropriate distribution of additional facilities among the centers. This phenomenon called "collective fairness" has been addressed by only a few authors so far. Two cases of the problem formulation were studied by Janáček and Gábrišová (2017). Similarly, Toro-Diaz et al. (2015) included some fairness considerations when analyzing the planning of large-scale emergency medical service systems.

The impact of the facility deployment on an individual cluster is characterized by the ratio of the cluster workload and the number of facilities assigned to the cluster. The cases differ in the definition of the minimized objective function. The first case objective involves the minimization of the squared deviation of the cluster workload ratio from a fixed value, such as the average workload per one facility in the whole serviced region. The second case objective consists of minimizing the maximum of the cluster workload ratios. In this case, the ratio for the worst situated cluster is minimized not regarding the ratio values of the clusters, which are situated a bit better. In contrary to the first case, the second case prefers to improve the worse situated cluster instead of improving the situation in each one. An integer linear programming model can be formulated and solved by an arbitrary integer-programming solver (IP-solver) equipped with a branch-and-bound procedure (Janáček & Gábrišová, 2017) for both cases mentioned above.

In this paper, the problem formulation is generalized and formulated both the generalized min-sum and min-max problems in the terms of mathematical programming. Furthermore, dynamic programming procedures are suggested not only for the min-sum problem but also for the min-max one. In addition, another approach based on convexity and monotony of decreasing nonlinearities in objective functions is presented.

2. Models of collective fairness and linear integer programming approach

Let m denote the number of clusters and the associated service centers of a considered emergency service system and let p denote the number of additional facilities, which are to be deployed among the service centers. Both cases of the fair facility deployment mentioned in Section 1 can be formulated using a series of strictly convex decreasing functions $f_i(y)$ defined on the interval $[0, p]$ for $i = 1, \ldots, m$. Examples which demonstrate the two mentioned cases were given in Janáček and Gábrišová (2017), where the function was defined by (1) for the case of minimization of the average squared deviation of the cluster ratio from the average workload per one facility in the whole system. The constant P_i in (1) and (2) denotes a workload connected with the cluster (or service center) i:

$$f_i(y) = \frac{(P_i)^2}{y+1}. \tag{1}$$

The second case corresponding to a minimization of the maximum of cluster ratios defines function $f_i(y)$ according to (2):

$$f_i(y) = \frac{P_i}{y+1}. \tag{2}$$

To be able to describe facility deployment, the integer variable y_i for each service center $i = 1, …, m$ is introduced. The variable y_i models a decision on the number of additional facilities located at the service center i.

The first case of the fair facility deployment problem can be formulated as follows for $f_i(y)$ defined by (1):

$$\min\{\sum_{i=1}^{m} f_i(y_i) : \sum_{i=1}^{m} y_i = p, \, y_i \in Z^+ \text{ for } i = 1,..., m\} . \tag{3}$$

Model (4) describes the second case of the fair facility deployment problem regarding $f_i(y)$ defined by (2):

$$\min\{\max\{f_i(y_i) : i = 1,..., m\} : \sum_{i=1}^{m} y_i = p, \, y_i \in Z^+ \text{ for } i = 1,..., m\} . \tag{4}$$

Both models (3) and (4) are nonlinear. Originally by Janáček and Gábrišová (2017), they were reformulated to linear integer programming models as follows.

Let the constants e_{ik} for $i = 1, …, m$ and $k = 0, …, p$ are defined by (5), and then a series of auxiliary variables $z_{ik} \in \{0, 1\}$ for $i = 1, …, m$ and $k = 0, …, p$ is introduced, where the variable z_{ik} will take the value of one, if k additional facilities are to be located at the center i:

$$e_{ik} = f_i(k) \text{ for } i = 1, ..., m, \, k = 0, ..., p . \tag{5}$$

After these preliminaries, the non-linear problem (3) for the first case can be rewritten as the following linear form:

$$\text{Minimize } \sum_{i=1}^{m} \sum_{k=0}^{p} e_{ik} z_{ik} ; \tag{6}$$

$$\text{Subject to } \sum_{i=1}^{m} \sum_{k=1}^{p} k z_{ik} = p ; \tag{7}$$

$$\sum_{k=0}^{p} z_{ik} = 1 \text{ for } i = 1, ..., m ; \tag{8}$$

$$z_{ik} \in \{0, 1\} \text{ for } i = 1, ..., m, k = 0, ..., p . \tag{9}$$

The constraint (7) allows distributing exactly p additional facilities in the set of m centers and constraints (8) assure that only one of possible facility numbers is assigned to the center i.

Using the above-defined constants e_{ik} and zero-one variables z_{ik}, the non-linear problem (4) for the second case can be reformulated to the following linear one. To complete the min-max model, a variable h is introduced to model upper bound for the values $f_i(y_i)$:

$$\text{Minimize } h \tag{10}$$

Subject to (7), (8), (9) and

$$\sum_{k=0}^{p} e_{ik} z_{ik} \le h \text{ for } i = 1, ..., m ; \tag{11}$$

$$h \ge 0 . \tag{12}$$

Another approach to linearization of the problems (3) and (4) is presented below. Let the constants c_{ik} for $i = 1, …, m$ and $k = 1, …, p$ are defined by (13):

$$c_{ik} = f_i(k-1) - f_i(k) \quad \text{for} \quad i = 1, …, m, \ k = 1, …, p . \tag{13}$$

As the functions $f_i(y)$ are decreasing, the constants c_{ik} are positive. Due to assumed strict convexity of the functions, also $c_{ik} > c_{ik+1}$ holds for each $k = 1, …, p\text{-}1$.

To be able to substitute the value of variable y_i in the models (3) and (4), a series of auxiliary zero-one variables $x_{ik} \in \{0, 1\}$ for $i = 1, …, m$ and $k = 1, …, p$ is introduced and the relation between variables y_i and variables x_{ik} for $k = 1, …, p$ is defined by (14):

$$\sum_{k=1}^{p} x_{ik} = y_i \quad \text{for} \quad i = 1, …, m . \tag{14}$$

The value of $f_i(y_i)$ can be expressed by (15) using the auxiliary variables x_{ik}, which meet conditions $x_{ik} \geq x_{ik+1}$ for each $k = 1, …, p\text{-}1$:

$$f_i(y_i) = f_i(0) - \sum_{k=1}^{p} c_{ik} x_{ik} \quad \text{for} \quad i = 1, …, m . \tag{15}$$

After these preliminaries, the non-linear problem (3) can be rewritten into the following linear form:

$$\text{Minimize} \ \sum_{i=1}^{m} \left(f_i(0) - \sum_{k=1}^{p} c_{ik} x_{ik} \right) \tag{16}$$

$$\text{Subject to} \ \sum_{i=1}^{m} \sum_{k=1}^{p} x_{ik} = p ; \tag{17}$$

$$x_{ik} \geq x_{ik+1} \quad \text{for} \quad i = 1, …, m, \ k = 1, …, p-1 ; \tag{18}$$

$$x_{ik} \in \{0, 1\} \quad \text{for} \quad i = 1, …, m, \ k = 1, …, p . \tag{19}$$

The model has several special properties, which might make the problem solving easier. First, the constraints (18) can be relaxed due to the strict convexity of the functions $f_i(y)$. Furthermore, the model (16)–(19) obviously has integral property, which means that the demand for integrality in constraints (19) can be also relaxed and then, constraints (20) can be applied instead of (19):

$$x_{ik} \in [0, 1] \quad \text{for} \quad i = 1, …, m, \ k = 1, …, p . \tag{20}$$

This way, the problem (15)–(19) is reduced to a simple linear programming problem (16), (17), and (20).

Using the above-defined constants c_{ik} and zero-one variables x_{ik} and substitution (15), the non-linear problem (4) can be reformulated to the following linear one. To complete the min-max model, a variable h is introduced to model upper bound of the values $f_i(y_i)$.

$$\text{Minimize} \ h \tag{21}$$

Subject to (17), (18), (19) and

$$f_i(0) - \sum_{k=1}^{p} c_{ik} x_{ik} \leq h \quad \text{for} \ \ i = 1, ..., m \, ; \tag{22}$$

$$h \geq 0 \, . \tag{23}$$

This model does not have an integral property at all, but the constraints (18) can be also released due to convexity. Thus the model (21), (17)–(19), (22) and (23) can be reduced to the problem of minimizing (21) subject to (17), (19), (22) and (23).

Using the above described linear programming models, both cases of the fair facility deployment problem can be solved by any linear programming (LP) or IP solvers.

3. Dynamic programming approach to the fair facility deployment

As the problem (4) belongs to the family of min-sum problems, whose objective functions have obviously additive property, the problem can be easily formulated and solved as a dynamic programming problem (Janáček & Gábrišová, 2017). The timeline of the computational process can be modeled by a discrete set of instants $i = 1, …, m$, which correspond to subscripts of clusters and the associated functions $f_i(y)$. Let variable y_i model the basic decision on the number of the additional facilities assigned to the service center i. In addition, state s_i of the computational process at the instant i is introduced. The value of s_i corresponds to the total number of facilities allocated by the decisions $y_1, …, y_{i-1}$ forgoing the decision y_i. The relation between s_i and s_{i+1} is given by the transition equation $s_{i+1} = s_i + y_i$ for $i = 1, …, m$-1. In the computational process, two values of further introduced functions will be iteratively computed. The formula (24) shows the Bellman's function $B_i(s)$ that gives the optimal objective function value of the problem (24):

$$B_i(s) = \min\{\sum_{k=i}^{m} f_k(y_k) : \sum_{k=i}^{m} y_k = p - s, \ y_k \in Z^+ \ \text{for} \ \ k = i, ..., m\} \, . \tag{24}$$

The control function $Z_i(s)$ gives the value y_i^* of variable y_i in the optimal solution of the problem (24). The computational process starts with the initialization of $B_m(s)$ and $Z_m(s)$ for $s = 1, …, p$. The values are defined by (25) and (26):

$$B_m(s) = f_m(s) \quad \text{for} \ \ s = 0, ..., p \, ; \tag{25}$$

$$Z_m(s) = p - s \quad \text{for} \ \ s = 0, ..., p \, . \tag{26}$$

Then, a step of the iterative process can be defined by (27) and (28):

$$B_i(s) = \min\{f_i(y) + B_{i+1}(s + y) : y = 0, ..., p - s\} \quad \text{for} \ \ s = 0, ..., p \, ; \tag{27}$$

$$Z_i(s) = \arg\min\{f_i(y) + B_{i+1}(s + y) : y = 0, ..., p - s\} \quad \text{for} \ \ s = 0, ..., p \, . \tag{28}$$

Having performed the iterative steps subsequently for $i = m$-1, m-2, …, 1, the value $B_1(0)$ gives the optimal objective function of (3) and the associated optimal values of y_i

can be obtained by a backtracking recursive process. In the process, s_1 is defined by $s_1 = 0$ and y_1 is determined by $y_1 = Z_1(s_1)$, then the following sequence of steps $s_{i+1} = s_i + y_i$ and $y_{i+1} = Z_{i+1}(s_{i+1})$ are performed for $i = 2, \ldots, m-1$.

As concerns the min-max problem (4), the dynamic programming approach as a solving process can be used also. In this case, the Bellman's function $B_i(s)$ for a given subscript i and a state s are defined by (29):

$$B_i(s) = \min\{\max\{f_k(y_k): k=i,\ldots m\}: \sum_{k=i}^{m} y_k = p - s, \ y_k \in Z^+ \ \text{for} \ k=i, \ldots, m\}. \quad (29)$$

The control function $Z_i(s)$ gives the value y_i^* of variable y_i of the optimal solution of the problem (29). The values of $B_m(s)$ and $Z_m(s)$ at the initial stage m will be defined once more by (25) and (26), respectively. To formulate the iterative formulae, the associativity of the operation max is made use of, and the iterative step by (30) and (31) is defined:

$$B_i(s) = \min\{\max\{f_i(y), B_{i+1}(s+y)\}: y=0, \ldots, p-s\} \quad \text{for} \ s=0,\ldots, p; \quad (30)$$

$$Z_i(s) = \arg\min\{\max\{f_i(y), B_{i+1}(s+y)\}: y=0, \ldots, p-s\} \quad \text{for} \ s=0,\ldots, p. \quad (31)$$

After the recursive process has been completed, the value $B_1(0)$ equals to the optimal objective function of (4) and the associated optimal values of $y_1^*, \ldots, y_m^*$ can be obtained by the backtracking recursive process mentioned above. The computational complexity in both cases is obviously $O(mp^2)$.

4. Alternative polynomial approaches to the fair facility deployment

The properties of convexity and integrality concerning the model (16), (17) and (20) enable us to develop a smart exact algorithm for the problem (3). The algorithm makes use of constants c_{ik} and zero-one auxiliary variables x_{ik} for $i = 1, \ldots, m$ and $k = 1, \ldots, p$, and it takes into account the substitution (14) as the relation between variables y_i and x_{ik}. The suggested algorithm performs according to the following steps.

Step 0.   Order all constants c_{ik} for $i = 1, \ldots, m$, $k = 1, \ldots, p$ in descending order and get the associated sequence $c_{i(r)k(r)}$ for $r = 1, \ldots, mp$. Set $y_i = 0$ for $i = 1, \ldots, m$.

Step 1.   For $r = 1, \ldots, p$ set $y_{i(r)} = y_{i(r)} + 1$.

The complexity of the algorithm is given by the complexity of Step 0, which is $O(mp\log_2(mp))$. Realizing that the series of c_{ik} are pre-ordered and exactly p biggest members of the ordered sequence are used, the complexity could be considerably reduced.

Concerning the problem (4), the model (21), (17), (19), (22) and (23) is made use of. The following proposition can be easily derived.

Let us denote $p_i(h)$ the minimum integer t, for which inequality (32) holds:

$$f_i(0) - \sum_{k=1}^{t} c_{ik} < h. \quad (32)$$

If an objective function value of an arbitrary feasible solution of (4) is less than h, then $y_i \geq p_i(h)$ holds for $i = 1, \ldots, m$.

Based on the proposition, the following algorithm can be suggested:

Step 0. Set $y_i = 0$ for $i = 1, \ldots, m$ and order the values of $f_i(y_i)$ in descending order so that the sequence $f_{i(r)}(y_{i(r)})$ for $r = 1, \ldots, m$ is obtained. Set $s = p$.

Step 1. If $s = 0$, terminate. Otherwise, go to Step 2.

Step 2. Set $h = f_{i(2)}(y_{i(2)})$ and determine the value of $p_{i(1)}(h)$.
If $p_{i(1)}(h) - y_{i(1)} < s$, then update $s = s - (p_{i(1)}(h) - y_{i(1)})$ and $y_{i(1)} = p_{i(1)}(h)$ and reorder the sequence $f_{i(r)}(y_{i(r)})$. Else, update $y_{i(1)} = y_{i(1)} + s$ and set $s = 0$.
Go to Step 1.

Ordering at Step 0 has complexity $O(m log_2(m))$. Reordering in the Step 2 can be repeated at most p-times. It follows that the overall complexity of the algorithm is maximum of $O(m log_2(m))$ and $O(mp)$.

5. Numerical experiments

In the previous sections, several approaches to the problem of fair facility distribution among service centers were suggested. The problem was formulated in both min-sum and min-max variants, where the variance of workloads per one facility over the set of service centers is minimized in the min-sum variant. Maximal workload per one facility is minimized in the min-max variant. Three new exact methods were developed for each considered variant in this paper and another one was taken from Janáček and Gábrišová (2017). The goal of the performed numerical experiments was to compare the quadruple of methods developed for each variant and decide about the best method for solving a variant problem instance.

The methods developed for the first (min-sum) variant are successively denoted as **FD_Mpo**, **FD_MP**, **FD_DP**, and **FD_GE**, where the first label designates the original method (Janáček & Gábrišová, 2017) based on the mathematical programming model (6) – (9), which was solved by an IP solver employing branch-and-bound method. The optimization software FICO Xpress 7.9 (64-bit, release 2015) was used to complete this study. The second label **FD_MP** is used for the method based on the mathematical programming model (16)–(19), where the associated problem was also solved by the above-mentioned IP solver. The label **FD_DP** denotes the dynamic programming approach, which follows the recursive computational process described by (24)–(28). The process was implemented in the programming language Mosel in the environment of FICO Xpress, unless employing any auxiliary optimization procedures embedded in the associated IP-solver. The label **FD_GE** denotes the polynomial approach described in the first part of Section 4.

The methods for the min-max variant are successively denoted by **FM_Mpo**, **FM_MP**, **FM_DP** and **FM_GE**, where the first label corresponds with the original method (Janáček & Gábrišová, 2017) based on the mathematical programming model (7)–(12) of the problem, which was solved by the above-mentioned IP-solver. The method denoted by **FM_MP** is also based on usage of the IP-solver, but, contrary to **FM_Mpo**, the model (17)–(19), (21)–(23) is used for the problem presentation. The method denoted by **FM_DP** is based on the dynamic programming principle, where the associated recursive process is determined by equations (29)–(31). The quadruple of methods for the min-max variant ends with the polynomial approach denoted by **FM_GE**. An algorithm of the approach is described at the end of Section 4.

The experiments were run on a PC equipped with the Intel® Core™ i7 4510U processor with the parameters: 2.00 GHz and 8 GB RAM. The procedures programmed in language Mosel were also run under the FICO Xpress.

The benchmarks used for comparison were derived from the real emergency health care system of the Slovak Republic (Jánošíková et al., 2017). Users of the systems are represented by cities, villages, and hamlets with the corresponding population, which was rounded to hundreds when the workload of individual clusters was computed. The current state of the above-mentioned real emergency system is characterized by 273 facilities (ambulance vehicles), which are located at 208 service centers in the Slovak Republic. The maximal value of facility workload of this original system equals to 2,526 and the variance of these facility workloads is 176.337. Twelve problem instances were derived in the following way. The first instance denoted as **OptOrg208** was created so that the original service center locations were accepted but only one facility to each of the centers was assigned, and $p=65$ facilities remained to be deployed by the tested methods. To create the next eleven instances, the number m of service centers was varied in the range {188, 192, 196, 200, 204, 208, 212, 216, 220, 224, 228} in such a way that some centers were abolished or newly created, and the associated weighted m-median problem was solved for the set of possible center locations equal to the set of all dwelling places of the Slovak Republic. Various sets of center locations were performed by this optimization provided according to (Kvet, 2014, 2015). This way, eleven instances denoted by the label **OptDerm** were obtained, where m corresponds to the number of service centers. The number p of additional facilities was determined according to $p = 273 - m$.

The results of min-sum variant methods are displayed in Table 1, which is organized so that each row corresponds to one of all solved instances. The first three columns describe the instance by its label, number m of service centers and number p of additional facilities, respectively. The two-column section **FD_Mpo** contains the variance of facility workloads of the optimal facility deployment obtained by each of the tested methods in the column "variance" and computational time of the method in seconds in the column "time[s]". The other one-column sections **FD_MP**, **FD_DP**, and **FD_GE** contain only the columns "time[s]" referring computational time of the method in seconds.

Table 1. Optimal objective function values and associated computational times of the min-sum variant methods (source: own calculations)

			FD_Mpo		FD_MP	FD_DP	FD_GE
Instance	m	p	variance	*time*[s]	*time*[s]	*time*[s]	*time*[s]
OptOrg208	208	65	96,268	0.219	0.125	0.0997	0.0030
OptDer188	188	85	31,987	0.250	0.141	0.1625	0.0031
OptDer192	192	81	33,174	0.219	0.141	0.1438	0.0047
OptDer196	196	77	35,880	0.219	0.141	0.1359	0.0031
OptDer200	200	73	39,111	0.203	0.125	0.1234	0.0047
OptDer204	204	69	39,757	0.204	0.125	0.1125	0.0032

			FD_Mpo		FD_MP	FD_DP	FD_GE
Instance	m	p	variance	*time*[s]	*time*[s]	*time*[s]	*time*[s]
OptDer208	208	65	42,366	0.203	0.125	0.1016	0.0032
OptDer212	212	61	45,758	0.187	0.125	0.0968	0.0031
OptDer216	216	57	48,502	0.188	0.125	0.0844	0.0016
OptDer220	220	53	51,411	0.172	0.125	0.0734	0.0016
OptDer224	224	49	55,532	0.157	0.110	0.0641	0.0016
OptDer228	228	45	58,353	0.156	0.094	0.0563	0.0015

The results of the min-max variant methods are presented in Table 2, which is organized in the same way as Table 1, with the exception that maximal facility workload is reported in the column "*max h*". The entries 600* denotes the case, when the branch-and-bound process was prematurely stopped unless it proved optimality of the best found solution. The process solving instances **OptOrg208** and **OptDer196** was stopped at the gaps of 1.60643% and 0.736842%, respectively.

Table 2. Optimal objective function values and associated computational times of the min-max variant methods (source: own calculations)

			FM_Mpo		FM_MP	FM_DP	FM_GE
Instance	m	p	max h	*time*[s]	*time*[s]	*time*[s]	*time*[s]
OptOrg208	208	65	1245	0.766	600*	0.1091	0.0019
OptDer188	188	85	941	0.500	0.609	0.1688	0.0016
OptDer192	192	81	941	0.390	0.765	0.1562	0.0031
OptDer196	196	77	950	0.532	600*	0.1453	0.0016
OptDer200	200	73	950	0.453	0.563	0.1375	0.0031
OptDer204	204	69	950	0.469	0.672	0.1203	0.0016
OptDer208	208	65	965	0.360	0.562	0.1110	0.0032
OptDer212	212	61	988	0.375	0.500	0.1015	0.0016
OptDer216	216	57	992	0.359	0.421	0.0907	0.0031
OptDer220	220	53	993	0.344	0.312	0.0812	0.0016
OptDer224	224	49	1043	0.266	0.422	0.0719	0.0031
OptDer228	228	45	1059	0.266	0.312	0.0610	0.0016

Comparing the resulting computational times of the min-sum methods reported in Table 1, monotonously decreasing values in most instances were observed. It is obvious that the new mathematical model solved by IP-solver outperforms the original approach. The dynamic programming approach proved to be a bit quicker than the mathematical program-

ming approach, but the polynomial approach **FD_GE** outperformed the other approaches in orders. A different situation emerges when the performance of min-max variant methods is studied. The new mathematical programming approach showed the worst performance of all compared methods and, in addition, it showed some kind of instability of computational process convergence, when the best found solution optimality is being proved. As concerns the dynamic programming approach and the polynomial approach **FM_GE**, these methods outperformed the mathematical programming ones in the same order as in the first variant comparison.

Conclusions

In this study, the problem of fair allocation of additional facilities in the given set of emergency service centers was investigated. The objective is to deploy the additional facilities in the fairest way. The problem formulation was generalized and formulated both the generalized min-sum and min-max problems in terms of mathematical programming. In addition, an approach based on convexity and monotony of decreasing nonlinearities in objective functions was presented. Several fast algorithms and dynamic programming procedures were suggested to solve not only the min-sum problem but also the min-max one. These approaches were tested and compared performing numerical experiments with real-sized problem instances. The new mathematical model solved by IP-solver outperforms the original approach. The dynamic programming approach proved to be a bit quicker than the mathematical programming approach, but the polynomial approach using the properties of convexity and integrality concerning the constructed model shows the best performance.

Although the contribution was originally devoted to the fair deployment of additional facilities in an emergency service system, the suggested methods are able to solve general min-sum and min-max problems, in which a limited number of facilities should be assigned to individual members from a finite set of providers. The only assumption is that the functions expressing the impact of the number of assigned facilities are monotonically decreasing and convex.

The suggested methods, especially those based on dynamic programming or preordering, may be a significant part of more complex tools for public service system designing. The relatively low time-consuming nature of the proposed procedures may be advantageous when dealing with problems of designing or rebuilding of emergency service systems in geographical larger regions.

The presented research comes from the assumption that each demand for service is satisfied from the nearest located service centre. This limitation prevents us from taking into account a restricted capacity of a service centre. The limited centre capacity may cause that a current demand may be satisfied from the second or third nearest service centre instead of from the nearest one. Possible generalization of the approach can be included in the further research, which will be aimed at further exploitation of the suggested methods for optimal dimensioning of service center capacities in service systems.

Funding

Disclosure statement

Authors declare they do not have any competing financial, professional, or personal interests from other parties.

References

Aringhieri, R., Carello, G., & Morale, D. (2016). Supporting decision making to improve the performance of an Italian Emergency Medical Service. *Annals of Operations Research*, *236*, 131–148. https://doi.org/10.1007/s10479-013-1487-0

Boujemaa, R., Jebali, A., Hammami, S., Ruiz, A., & Bouchriha, H. (2018). A stochastic approach for designing two-tiered emergency medical service systems. *Flexible Services and Manufacturing Journal*, *30*, 123–152. https://doi.org/10.1007/s10696-017-9286-6

Brotcorne, L., Laporte, G., & Semet, F. (2003). Ambulance location and relocation models. *European Journal of Operational Research*, *147*, 451–463. https://doi.org/10.1016/S0377-2217(02)00364-8

Carmen, R., Van Nieuwenhuyse, I., & Van Houdt, B. (2018). Inpatient boarding in emergency departments: Impact on patient delays and system capacity. *European Journal of Operational Research*, *271*(3), 953–967. https://doi.org/10.1016/j.ejor.2018.06.018

Chanta S., Mayorga, M. E., & McLay, L. A. (2014). Improving emergency service in rural areas: a bi-objective covering location model for EMS systems. *Annals of Operations Research*, *221*(1), 133–159. https://doi.org/10.1007/s10479-011-0972-6

Cyganska, M. (2017). Analysis of high cost outliers in a Polish reference hospital. *E&M Ekonomie a Management*, *20*(4), 59–69. https://doi.org/10.15240/tul/001/2017-4-005

Davis, R. (2005, May 20). *The price of just a few seconds lost: People die. USA Today.* Retrieved February 12, 2019, from https://usatoday30.usatoday.com/news/nation/ems-day2-cover.htm

Garner, A. A., & van den Berg, P. L. (2017). Locating helicopter emergency medical service bases to optimise population coverage versus average response time. *BMC Emergency Medicine*, *17*, 31. https://doi.org/10.1186/s12873-017-0142-5

Ingolfsson, A., Budge, S., & Erkut, E. (2008). Optimal ambulance location with random delays and travel times. *Health Care Management Science, 11*(3), 262–274. https://doi.org/10.1007/s10729-007-9048-1

Janáček, J., & Gábrišová, L. (2009). A two-phase method for the capacitated facility problem of compact customer sub-sets. *Transport, 24*(4), 274–282. https://doi.org/10.3846/1648-4142.2009.24.274-282

Janáček, J., & Gábrišová, L. (2017). Collective fairness in emergency system designing. In *SOR´17: Proceedings of the 14th International Symposium on Operational Research* (pp. 135–140). Bled, Slovenia.

Jánošíková, Ľ. (2007). Emergency medical service planning. *Komunikácie*, *9*(2), 64–68.

Jánošíková, Ľ., Jankovič, P., & Kvet, M. (2017). Improving emergency system using simulation and optimization. In *SOR´17: Proceedings of the 14th International Symposium on Operational Research* (pp. 269–274). Bled, Slovenia.

Koval, O., Nabareseh, S., & Chromjaková, F. (2019). Standardization in services: Assessing the impact on customer satisfaction. *E&M Ekonomie a Management*, *22*(3), 186–203. https://doi.org/10.15240/tul/001/2019-3-012

Kvet, M. (2014). Computational study of radial approach to public service system design with generalized utility. In *Digital Technologies 2014: the 10th International IEEE Conference*. Zilina, Slovakia. IEEE. https://doi.org/10.1109/DT.2014.6868713

Kvet, M. (2015). Exact and heuristic radial approach to fair public service system design. In *2015 International Conference on Information and Digital Technologies*. Zilina, Slovakia. IEEE. https://doi.org/10.1109/DT.2015.7222971

Lau, H., Dadich, A., Nakandala, D. Evans, H., & Zhao, L. (2018). Development of a cost-optimization model to reduce bottlenecks: A health service case study. *Expert Systems, 35*(6), e12294. https://doi.org/10.1111/exsy.12294

Liu, Y., Li, Z. Z., Liu, J. X., & Patel, H. (2016). A double standard model for allocating limited emergency medical service vehicle resources ensuring service reliability. *Transportation Research Part C: Emerging Technologies, 69*, 120–133. https://doi.org/10.1016/j.trc.2016.05.023

Marianov, V., & Serra, D. (2002). Location problems in the public sector. In Z. Drezner, & H. Hambacher (Eds.), *Facility location: Applications and theory* (pp. 119–150). Springer. https://doi.org/10.1007/978-3-642-56082-8_4

Marsh, I., Crowley, K., Grube, D., & Eccleston, R. (2017). Delivering public services: locality, learning and reciprocity in place based practice. *Australian Journal of Public Administration, 76*(4), 443–456. https://doi.org/10.1111/1467-8500.12230

Nogueira, L. C., Pinto, L. R., & Silva, P. M. S. (2016). Reducing Emergency Medical Service response time via the reallocation of ambulance bases. *Health Care Management Science, 19*(1), 31–42. https://doi.org/10.1007/s10729-014-9280-4

Stanimirovic, Z., Miskovic, S., Trifunovic, D., & Veljovic, V. (2017). A two-phase optimization method for solving the multi-type maximal covering location problem in emergency service networks. *Information Technology and Control, 46*(1), 100–117. https://doi.org/10.5755/j01.itc.46.1.13853

Staňková, P., Papadaki, Š., & Dvorský, J. (2018). Comparative analysis of the perception of the advantages and disadvantages of hospital horizontal integration. *E&M Ekonomie a Management*, *21*(1), 101–115. https://doi.org/10.15240/tul/001/2018-1-007

Toro-Diaz, H., Mayorga, M. E., McLay, L. A., Rajagopalan, H. K., & Saydam, C. (2015). Reducing disparities in large-scale emergency medical service systems. *Journal of the Operational Research Society, 66*(7), 1169–1181. https://doi.org/10.1057/jors.2014.83

Vermuyten, H., Rosa, J. N., Marques, I., Belien, J., & Barbosa-Povoa, A. (2018). Integrated staff scheduling at a medical emergency service: An optimisation approach. *Expert Systems with Applications*, *112*, 62–76. https://doi.org/10.1016/j.eswa.2018.06.017