COMMUNICATIONS

Lydia Gabrisova - Jaroslav Janacek *

DESIGN OF CAPACITATED EMERGENCY SERVICE SYSTEM

Optimal design of the emergency service systems mostly consists in the decisions on a deployment of service centers, which provide the nearest system users with the associated service. As the demand for emergency service occurs randomly, the nearest service center may be occupied by servicing some previous demand and thus the current demand must be serviced from some more distant available service center. This phenomenon is dealt with in this paper where the center capacity is considered in the service system design to mitigate frequency of the above-mentioned events. For this purpose, two approaches to the capacitated service center deployment were suggested, implemented and tested on several benchmarks to give a relevant comparison of them.

***Keywords:** Capacitated p-median problem, NP-hard problem, local optimization, integer programming, decomposition heuristic technique, emergency medical service.*

1. Introduction

Solving techniques for large p-median problems represent a technical background of many service system designs [1], [2], [3], [4], [5], [6] and [7]. The p-median problem formulation is used to describe the impact of service center deployment on the average time of the service accessibility. The average time between the moment of demand rising and the start of the service is usually approximated by the sum of distances from the serviced object to the nearest service center. This min-sum objective can be used when the service can be scheduled in advance and no service center capacity is considered, but usage of the p-median problem formulation may fail when an emergency service system is designed or capacitated centers are taken into account [8], [9], [10] and [11]. The emergency service system services the randomly emerging demands as accidents, fires and similar events. In this case, the emergency system works as the queuing system where service center represents a facility with such capacity that the service center is able to service only few demands simultaneously. The limited capacity together with the randomly coming demands may cause that the current demand cannot be served from the nearest service center due to its occupancy by previously occurred demands. The center occupancy incurs that the newly occurred demand must either wait or be serviced from some more distant service center. Both eventualities mean that the service accessibility gets worse. The service accessibility deterioration affects especially those system users that belong to the dense populated districts where the demand frequency is extremely high. To withstand this defect in

emergency service system design, we suggested two approaches which deal with the limited capacity of the deployed service centers. The first approach is based on capacitated p-median problem with partially relaxed service center capacities. The second approach is represented by a sequence of several phases which decompose the capacitated location problem, to obtain such service center deployment and center capacities that the average service time is minimized subject to equal load of the individual service centers.

The remainder of the paper is organized as follows. The section 2 contains the mathematical formulation of the capacitated p-median problem used in the first approach. Section 3 describes the decomposition heuristic. Section 4 contains numerical experiments and comparison of the suggested approaches and Sections 5 gives final conclusions.

2. Mathematical model of problem

When a public service system is designed to service demands of its users spread over some geographical area, then the possible user's demand deployment is usually modeled by a finite set of locations which the user's demands are concentrated in. These locations represent locations of dwelling places and possibly some segments of a road network and so on. It is considered that the size or frequency of the demand at such location is proportional to the number of users concentrated at the locations. It is also considered that a service center can be located only at some location from a finite set of possible locations, which is given in advance.

* **Lydia Gabrisova, Jaroslav Janacek**
  Department of Mathematical Methods and Operations Research, Faculty of Management Science and Informatics, University of Zilina, Slovakia
  E-mail: lydia.gabrisova@fri.uniza.sk

The objective of the weighted capacitated *p*-median problem is to find the locations of at most *p* centers so that the sum of weighted distances from each user location to the nearest located service center is minimal. Furthermore, we assume within the paper that each located center consists of several facilities which are able to provide the users with service. Thus, each decision on center location is accompanied by the decision on the associated number of facilities.

The inputs to the mathematical programming model are as follow:

I   the set of candidate center locations,
J   the set of municipalities (user locations),
p   the total number of facilities to be located,
d_{ij} the distance from a candidate center location $i \in I$ to a user location $j \in J$,
b_j the number of users (user demands) at user location j ∈ J,
a   the capacity limit of one facility.

The decision on placing a facility at the location *i* can be modeled by the nonnegative integer variable y_i introduced for each $i \in I$. The value of y_i gives the number of facilities located at the center location. The assignment of user's demand located at *j* to the candidate center location *i* is modeled by nonnegative variable z_{ij}. Variable z_{ij} takes nonzero value if the municipality *j* will be served by a facility located at *i* on the whole or fractionally.

After these preliminaries, the model of the weighted *p*-median problem with limited center capacity can be written as:

$$Minimize \sum_{i \in I} \sum_{j \in J} d_{ij} \, b_j \, z_{ij} \tag{1}$$

$$Subject\ to \sum_{i \in I} z_{ij} = 1 \ \ for\ j \in J \tag{2}$$

$$z_{ij} \leq y_i \ \ for\ i \in I, j \in J \tag{3}$$

$$\sum_{j \in J} b_j z_{ij} \leq a y_i \ \ for\ i \in I \tag{4}$$

$$\sum_{i \in I} y_i \leq p \tag{5}$$

$$y_i \in Z_0^+ \ \ for\ i \in I \tag{6}$$

$$z_{ij} \geq 0 \ \ for\ i \in I, j \in J \tag{7}$$

The main system criterion (1) is the sum of distances among located centers and users weighted by the volume of demand which are assigned to them. The model constraints (2) ensure that all user demands concentrated at a user location *j* will be assigned to possible service center locations. Constraints (3) ensure the implication that if a non-zero demand from *j* is assigned to a possible center location *i*, then some facility must be located at this place. Constraints (4) limit the total demand served by a center located at *i*. Constraint (5) limits the total number of facilities which can be deployed. The remaining

obligatory constraints (6) and (7) specify the definition domains of the variables.

The capacitated *p*-median problem of linear programming is known to be NP-hard. As a consequence, it cannot be solved to optimality even for moderate sized problem instances [12] and [13].

3. Relaxation and decomposition heuristics using mathematical programming

In this Section we present our contribution to the given problem. First, we try to cope with the bad convergence of the branch-and-bound method for a capacitated location problem solving by a relaxation technique to obtain an approximate solution of the problem. Further, we developed a special decomposition technique to avoid the bad convergence of the branch-and-bound method in the case when the capacity constraints are involved into the solved problem. We made use of our preliminary research presented in [14] and [15].

Our approach is based on the simple idea which can be formulated as follows: if the problem cannot be optimized on the whole, optimize it in parts. This idea can be used for every problem that can be divided into easily solvable sub-problems. The capacitated facility location problem meets this condition. In chapter 3.1 and 3.2 we presented two approaches to the problem solving.

3.1. Relaxation of capacitated constraints

The first approach is based on relaxation of the capacity constraints (4). The relaxation is not complete, but it consists in increasing the capacity *a* by some value a_+. Our preliminary research showed that the convergence of the solving process of the problem (1) – (7) can be considerably improved if the original service center capacity is relaxed. The value a_+ is proportional to the value *a*. New form of the constraints (4) is as follows:

$$\sum_{j \in J} b_j z_{ij} \leq (a + a_+) y_i \ \ for\ i \in I \tag{8}$$

This relaxation enables to obtain a slightly infeasible solution in acceptable computational time. This approach will be denoted by the title *Relax*.

3.2. Decomposition heuristic method

Heuristic method denoted as *Decomp* successively solves the following sub-problems in four phases:
1. In the first phase, the classical uncapacitated *p*-median problem (1) – (3), (5) is solved with the integer obligatory

constraints where $y_i \in \{0,1\}$ for $i \in I$. As can be noticed, the constraints (4) were eliminated from the model ad interim. The solution of this relaxed problem gives locations of p centers which form the set $I_1 = \{i \in I, y_i = 1\}$. For each $i \in I_1$, we define the set $J_i = \{i \in I, z_{ij} > 0\}$ of users' locations assigned to center location i and we define also the sum B_i of demands satisfied from i according to $B_i = \sum_{j \in J_i} b_j z_{ij}$.

2. The value B_i is the volume of demands serviced by the center located at i and it can be larger or lower than the capacity of one facility. In this phase, we solve the problem of facility allocation. We aim to "strengthen capacity" centers which are overloaded (i.e. $B_i >> a$) and close the unutilized centers (i.e. $a >> B_i$) so that the number of located facilities stays the same. We partition set I_1 into two sub-sets $I_1^- = \{i \in I_1, B_i \le a\}$ and $I_1^+ = \{i \in I_1, B_i > a\}$, and denote $B_i^- = B_i$ for a center $i \in I_1^-$ and surplus $B_i^+ = B_i - a$ for a center $i \in I_1^+$.

The strategy of facility reallocation is based on "profitability" of center capacity strengthening subject to closing some of the partially unutilized centers.

In this strategy, the "revenue" at a center j is presented by the part of surplus which is covered by the reallocated capacity of facilities. The total volume of demands, which stays uncovered at the centers closed due to coverage of j, can be denoted as the "cost" paid for the covering. To model the decisions on reallocating the capacities of the closed centers, we introduce the variable $x_{kj} \in [0, 1]$ for each pair k, j where $k \in I_1^-$, $j \in I_1^+$. The variable x_{kj} takes the value of 1 if the center k is to be closed and its capacity a is moved to the center j, otherwise the variable takes the value of 0. An auxiliary nonnegative variables u_j is introduced to model amount of the "revenue" at the center j. Then the model of the problem can be stated as follows:

$$Maximize \sum_{j \in I_1^+} u_j - \sum_{j \in I_1^+} \sum_{k \in I_1^-} B_k^- x_{kj} \qquad (9)$$

$$Subject\ to \sum_{j \in I_1^+} x_{kj} \le 1 \ \ for \ \ k \in I_1^- \qquad (10)$$

$$u_j \le B_j^+ \ \ for \ \ j \in I_1^+ \qquad (11)$$

$$u_j \le a \sum_{k \in I_1^-} x_{kj} \ \ for \ \ j \in I_1^+ \qquad (12)$$

$$x_{kj} \in \{0, 1\} \ \ for \ \ k \in I_1^-, j \in I_1^+ \qquad (13)$$

$$u_j \ge 0 \ \ for \ \ j \in I_1^+ \qquad (14)$$

The expression (9) represents the maximized "profitability" subject to the system of constraints where constraints (10) ensure that unutilized center capacity can be reallocated at most to one overloaded center. The constraints (11) and (12) assure that the "revenue" at overloaded center j cannot

exceed the associated surplus and also the total capacity reallocated to the center j cannot be surpassed. Even if the solution of the problem (9) – (14) provides us with a solution of the capacitated p-median problem, we use only information about the total number of closed centers.

3. The reduced number $p_1 = p - \sum_{j \in I_1^+} \sum_{j \in I_1^-} x_{kj}$ of located centers is taken as output from the second phase. The new number p_1 is used in the uncapacitated p- median problem (1) – (3), (5) where constraint (5) is replaced by constraint (15).

$$\sum_{i \in I} y_i \le p_1 \qquad (15)$$

The result of the problem determines the final deployment of the p_1 centers, whereas the numbers of facilities located at these centers will be determined by the phase 4.

4. The set of new p_1 locations is denoted by I^*. The sum B_i of demands satisfied from i can be computed for all $i \in I^*$ in the same way as it was done in the phase 1. We assume that each center $i \in I^*$ is equipped by one facility with capacity a. Remaining free $p - p_1$ facilities are allocated according to the further procedure. We introduce an integer variable x_i for each $i \in I^*$ to model the number of additional facilities assigned to the center i. Additionally, we introduce an auxiliary variable h to express the lower bound of all ratios of allocated capacity to surplus at a given center. Then, we solve the max-min problem described by the following model.

$$Maximize\ h \qquad (16)$$

$$Subject\ to\ x_i + 1 \ge B_i \cdot h \ for \ i \in I^* \qquad (17)$$

$$\sum_{i \in I^*} x_i \le p - p_1 \qquad (18)$$

$$x_i \in Z_0^+ \ for \ i \in I^* \qquad (19)$$

The resulting solution assigns the remaining facilities to the centers, and thus $y_i = x_i + 1$ facilities will be located at center $i \in I^*$.

4. Numerical experiments

To compare the two approaches mentioned in the Section 3, several experiments were performed. The benchmarks were derived from real emergency health care system which was originally designed for eight self-governing regions of the Slovak Republic. The original designs will be referred to as "*Original*". The instances are organized so that they correspond to the administrative organization of Slovakia (Bratislava - BA, Banska Bystrica - BB, Kosice - KE, Nitra - NR, Presov - PO, Trencin -

TN, Trnava - TT and Zilina - ZA). The number of inhabitants in municipality j is rounded to hundreds and denoted as b_j. The facility mentioned in the previous Sections corresponds to one ambulance vehicle. The capacity limit of the facility was set at the number of inhabitants which falls upon one ambulance, i.e. $a = \sum_{j \in J} b_j / p$.

The number p of facilities corresponds to the real number of ambulances deployed in the given self-governing region. To be able to evaluate the results of the both approaches and the current design, we take into account that the result of each of the approaches can be described by a vector y which consists of integer components y_i for $i \in I$. The value of y_i gives the number of facilities which are located at the location i. For each solution y, $I_1(y)$ denotes the set of located centers, i.e. $I_1(y)=\{i \in I, y_i \geq 1\}$. Furthermore, $J_i(y)$ denotes the set of user locations which are assigned to the center location i. The cluster $J_i(y)$ of the located center $i \in I_1(y)$ can be defined by the equality $J_i(y)=\{ j \in J, d_{ij} = min[d_{kj}, k \in I_1(y)]\}$. Then the volume of demands served by center i is $B_i(y) = \sum_{j \in J_i(y)} b_j$ and the part which must be served by one facility - ambulance located at this center is denoted by $\underline{B}_i = B_i / y_i$. The basic characteristics of used benchmarks are described in Table 1 where the column denoted as $|I|$ contains numbers of possible service center locations. This number is equal to the number of the user's locations. The column denoted by p contains maximal number of ambulance vehicles which are to be deployed. The column "Inhabitants" gives number of inhabitants of the self-governing region in hundreds. This number is considered as the total demand of the region. The column "avg" contains the portion of demand which must be served by one ambulance on average.

The original design (*Original*) and the two designs obtained by the mentioned approaches of the emergency service system design (*Relax* and *Decomp*) were evaluated and obtained characteristics of the designs are plotted in the following Tables 2, 3 and 4. In those tables, column $|I_1|$ denotes the number of located centres. The double column "Number of people per ambulance" gives minimal and maximal portion $\underline{B}_i$ of the served demands per ambulance in hundreds. The column "max" contains the maximum distance from user location to the nearest service center in kilometers. The column "OF" contains the value of objective function value computed in accordance to (20).

$$\sum_{i \in I_1(y)} \sum_{j \in J_i(y)} d_{ij} b_i \qquad (20)$$

The double column "Workload per ambulance" gives average and maximal workload per one ambulance vehicle computed according to (21) per each service center $i \in I_1(y)$.

$$\sum_{j \in J_i(y)} d_{ij} b_i / y_i \qquad (21)$$

The description of the benchmarks — Table 1

Region	$|I|$	p	inhabitants	avg
BA	87	25	6063	243
TT	249	22	5552	253
TN	276	26	5942	229
NR	350	36	6896	192
ZA	315	36	6896	192
BB	515	46	6601	144
PO	664	44	8158	186
KE	460	38	7930	209

Description of the *Original* design — Table 2

Region	$|I_1|$	Number of people per ambulance		max [km]	OF	Workload per ambulance	
		min	max			avg	max
BA	14	81	415	19	21497	860	2746
TT	18	73	560	38	30288	1377	3548
TN	21	65	537	30	24763	952	3338
NR	27	42	300	21	36285	1008	1783
ZA	29	37	368	24	29774	827	1683
BB	36	21	386	23	30178	656	1772
PO	32	25	395	34	42634	969	2692
KE	32	14	769	25	34693	913	2255

The description of the *Relax* design obtained for $a_+=0.085*a$ Table 3

Region	$|I_1|$	Number of people per ambulance		max [km]	OF	Workload per ambulance	
		min	max			avg	max
BA	19	67	365	16	15666	627	1363
TT	20	93	416	30	26060	1185	2977
TN	22	102	339	30	23565	906	1639
NR	28	96	263	24	31490	875	1371
ZA	29	61	286	26	28364	788	1865
BB	35	43	217	26	28822	627	1435
PO	34	69	328	42	38948	885	1573
KE	33	85	338	28	33151	872	1933

The description of the *Decomp* design   Table 4

Region	$|I_1|$	Number of people per ambulance		max [km]	OF	Workload per ambulance	
		min	max			avg	max
BA	17	89	365	16	16148	646	1363
TT	18	116	443	24	27079	1231	2977
TN	18	104	335	30	25316	974	1695
NR	26	92	310	24	32172	894	1861
ZA	24	61	272	26	30256	840	1586
BB	33	50	203	26	29240	636	1396
PO	29	88	260	42	40980	931	1537
KE	25	70	292	25	35567	936	1910

To be able to compare regularity of demand or workload distribution over the user clusters of the individual facilities, we drew the following three bar graphs. In Fig. 1 we present maximal number of people served by one ambulance for a comparison of all systems.

Fig. 2 The maximal workload per ambulance served by one ambulance for the three system designs
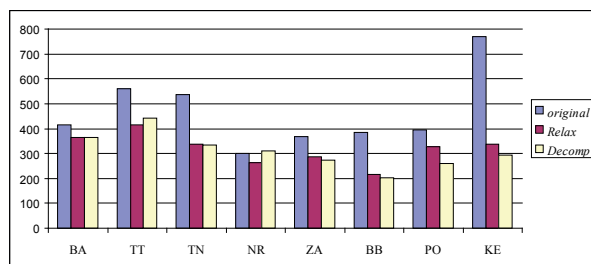
Fig. 1 The maximal number of people served by one ambulance for the three system designs

In Fig. 2 we present maximal workload per ambulance and in Fig. 3 we present average workload per ambulance served by one ambulance for a comparison of all systems.
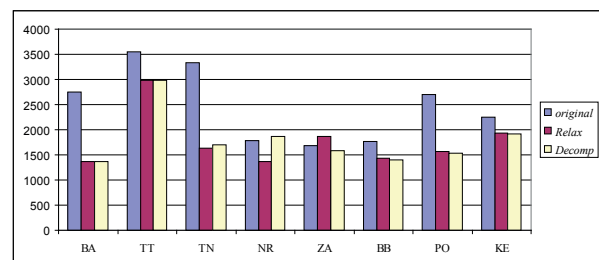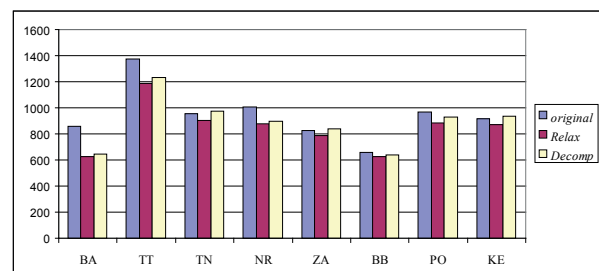
Fig. 3 The average workload per ambulance served by one ambulance for the three system designs

The suggested approaches were implemented in the visual development environment Xpress-IVE using solver Xpress-Optimizer v2.2.3. The experiments were performed on a personal computer equipped with the Intel Core i7 processor with 1.60 GHz and 8 GB of RAM. The sum of computation times did not exceed 27 minutes for *Decomp* and 40 hours for the *Relax* approaches respectively. The computational times necessary for obtaining the resulting solution are given in Table 5.

Computational times of the two suggested methods   Table 5

Region	*Relax* [min]	*Decomp* [min]
BA	0.04	0.05
TT	19.4	0.95
TN	32.0	0.85
NR	0.97	1.7
ZA	28.9	1.4
BB	7.1	5.3
PO	1203.1	12.5
KE	1102.6	3.6

5. Conclusions

We have suggested two approaches to the capacitated emergency service system design to mitigate frequency of the events where a user demand must be served from a service center which is more distant than the nearest one. The approaches have been suggested so that they are implementable on a common personal computer equipped with a commercial IP-solver which enables to compress the long terms of a software tool development when an emergency public service system is designed. The suggested approach *Relax* outperforms significantly the original system design and it also outperforms the second suggested approach *Decomp* even if the difference is not as considerable as concerns the objective function value. If regularity of demand or workload distribution over the user clusters of the individual facilities is taken into account, it can be found that the both suggested approaches provide much better results in comparison with the original design. From the point of applicability the decomposition approach *Decomp* can be recommended due to its very low computational time comparing to the relaxation approach. As the decomposition approach solves only uncapacitated facility location problems, the approach has big potential to be accelerated by usage of a faster method for *p*-median problem solving. This idea will be a topic of our further research in this field.

References

[1] MARIANOV, V., SERRA, D.: *Location Problems in the Public Sector,* Facility location. Applications and theory (by Drezner Z (ed.) et al.). Berlin, Springer, pp.119-150, 2004.

[2] CHANTA, S., MAYORGA, M. E., MCLAY, L. A.: *Improving Emergency Service in Rural Areas: A Bi-objective Covering Location Model for EMS Systems*, Annals of Operations Research. [online] DOI 10.1007, 2011.

[3] BROTCORNE, L., LAPORTE, G., SEMET, F.: Ambulance Location and Relocation Models, *European J. of Operational Research,* vol. 147, pp. 451-463, 2003.

[4] JANACEK, J., LINDA, B., RITSCHELOVA, I.: Optimization of Municipalities with Extended Competence Selection, *Prager Economic Papers - J. of Economic Theory and Policy,* vol. 19, No. 1, pp. 21-34, 2010.

[5] JANOSIKOVA, L.: Emergency Medical Service Planning, in *Communications - Scientific Letters of the University of Zilina,* vol. 9, No. 2, pp. 64-68, 2007.

[6] KOHANI, M.: Designing of Zone Tariff in Integrated Transport Systems, *Communications - Scientific Letters of the University of Zilina,* vol. 15, No. 1, pp. 29-33, 2013.

[7] PESKO, S.: Minimal Total Area Convex Set Partitioning Problem, *Communications - Scientific Letters of the University of Zilina,* vol. 11, No, 3, pp. 39-42, 2009.

[8] HOLMBERG, K., RONNQVIST, M., YUAN, D.: *An Exact Algorithm for the Capacitated Facility Location Problems with Single Sourcing, European J. of Operational Research,* vol. 113, pp. 544-559, 1999.

[9] JANACEK, J., JANOSIKOVA, L.: Computability of the Emergency Service System Design Problem, *Communications - Scientific Letters of the University of Zilina,* vol. 10, No. 2, pp. 5-9, 2008.

[10] JANOSIKOVA, L., ZARNAY, M.: *Location of Emergency Stations as the Capacitated p-median Problem*, Proc. of the Intern. Scientific Conference Quantitative Methods in Economics-Multiple Criteria Decision Making XVII. Virt, pp. 116-122, 2014.

[11] PIRKUL, H., SCHILLING, D.: The Capacitated Maximal Covering Location Problem with Backup Service, *Annals of Operations Research,* vol. 18, pp. 141-154, 1989.

[12] JANACEK, J.: *The Medical Eservice System Design*, Advances in Transport Systems Telematics. Katowice, Jacek Skalmierski Computer Studio, pp. 443-449, 2006.

[13] JANACEK, J. et al.: *Designing territorially extensive service systems (in Slovak),* EDIS: University of Zilina, 404 p. 2010.

[14] JANACEK, J., GABRISOVA, L.: Lagrangean Relaxation Based Approximate Approach to the Capacitated Location Problem, *Communications - Scientific Letters of the University of Zilina,* vol. 8, No. 3, pp. 19-24, 2006.

[15] JANACEK, J., GABRISOVA, L.: A Two-phase Method for the Capacitated Facility Problem of Compact Customer Sub-sets, *Transport: J. of Vilnius Gediminas Technical University and Lithuanian Academy of Sciences,* vol. 24, No. 4, pp. 274-282, 2009.