Designing of Large-Scale Public Service Systems by Covering Methods

Marek Kvet*
Department of Transportation Networks
Faculty of Management Science and Informatics
University of Žilina
Univerzitná 8215/1, 010 26 Žilina, Slovakia
Marek.Kvet@fri.uniza.sk

Abstract
This paper deals with the problem of designing the optimal structure of most public service systems, which is often formulated as the p-median problem. The real instances of these problems are characterized by a considerably big number of possible facility locations, which can take the value of several thousands. Current exact approaches must face up to a big demand on computational time, and they often fail when a large instance is being solved. This paper is focused on the approximate approach based on specific model reformulation. It uses the approximation of a common distance by some predetermined distances given by so-called dividing points. The deployment of the dividing points influences the solution accuracy. To improve this approach, we have developed a sequential method of dividing points deployment. Hereby, we study the accuracy of the suggested method, using the upper and lower distance approximations in comparison to the saved computational time.

Categories and Subject Descriptors
G.1.6 [**Optimization**]: Integer Programming

Keywords
Public service system, large p-median problem, covering formulation, approximate approach, sequential method, lower and upper bound, XPRESS

1. Introduction
Large instances of the p-median problem and associated solving approaches form a background of many public service system design problems [13]. The family of public service systems includes medical emergency system [4], [12], fire-brigade deployment, public administration system [10] and many others, overall combinatorial problems, where the quality criterion of the design takes into account network distances. The public service system structure is formed by the deployment of limited number of service centers and the associated objective is to minimize social costs, which are often proportional to distances from serviced customers to the nearest source of provided service. To obtain a good decision on service center location in any serviced area, a mathematical model of the problem can be formulated and some of mathematical programming methods applied to get the optimal solution of the problem. Mathematical models of the public service system design problem are often related to the p-median problem, where the p-median problem is formulated as a task of determination of at most p network nodes as facility locations, so that the sum of distances between each node and the nearest located facility is minimal. With real problems, the number of serviced customers takes the value of several thousands and the number of possible facility locations can take this value as well [1]. For some integer programming algorithms, the number of possible service center locations seriously impacts the computational time [11]. Nevertheless, the location-allocation models, which have been commonly used to describe the location problems with the criterion including distances between serviced customer and the nearest service center, constitute such mathematical programming problems, which resist to any attempt at fast solution.

Another way of the p-median problem representation by means of mathematical programming uses so called radius formulation [3], [5], [6]. This approach avoids assigning individual customers to some of the located facilities, and it deals only with the information whether some facility is or is not located in a given radius from the customer. Mentioned approximate method has appeared in available literature in two different versions. The first approach [5] forms the particular finite system of radii for each customer according to the distances from the customer to possible facility locations. The second approach [8] discussed in this paper forms one common system of radii based on so called dividing points, which are determined in accordance to the estimated distance relevance. The later approach leads to the model similar to the set covering problem, which is easy to solve even for large instances by common optimization software tools. This approach pays for shorter computational time or smaller computer

*Recommended by thesis supervisor: Prof. Jaroslav Janáček. Defended at Faculty of Management Science and Informatics, University of Žilina on August 27, 2013.

Kvet, M. Designing of Large-Scale Public Service Systems by Covering Methods. Information Sciences and Technologies Bulletin of the ACM Slovakia, Vol. 5, No. 4 (2013) 8-14

memory demand by losing its accuracy. The accuracy can be improved by a convenient determination of so called dividing points, which are used in the problem reformulation [8]. Hereby, we present a sequential approach to the dividing points determination and we use them not only for obtaining a good solution of the problem, but also for gaining a good lower bound of the unknown optimal solution. We demonstrate an impact of the sequential approach to the accuracy and computational time in the dissertation thesis. To solve the associated problems, we use common optimization environment XPRESS. To describe the p-median problem on a network we denote J a set of serviced nodes, similarly, we denote I a set of possible service center locations. We use here only formulation of the p-median problem, where it should be determined at most p locations from the set I so that the sum of network distances from each element of J to the nearest located facility is minimal. The network distance between a possible location $i \in I$ and a customer j from J is denoted as d_{ij}. The basic decisions in any solving process of the p-median problem concern location of service centers at network nodes from the set I. To model these decisions at particular nodes, we introduce a zero-one variable $y_i \in \{0,1\}$, which takes the value of 1, if a facility should be located at the place i from I and it takes the value of 0 otherwise. The p-median problem can be formulated as follows:

$$\begin{aligned} &Minimize \quad F(y) = \sum_{j \in J} min\{d_{ij} : i \in I, y_i = 1\} \\ &Subject\ to: \sum_{i \in I} y_i \le p, \ \ y_i \in \{0,1\} \ \ for\ \ i \in I \end{aligned} \tag{1}$$

Reminder of this paper follows the below scheme. Section 2 contains the location-allocation formulation of the p-median problem. Section 3 comprises the approximate approach based on given sequence of dividing points, where a covering model is used to obtain both lower and upper bound of the optimal solution value together with a near-optimal solution. Section 4 presents the dividing points determination, which enables to minimize estimated deviation of the upper and lower bounds from the optimal solution value. Section 5 introduces a dynamic approach which adjusts distances between the dividing points and improves the solution of the original problem step by step. Section 6 reports on various mathematical definitions of distance relevance which plays a very important role in the process of dividing points deployment and it directly influences the solution accuracy. The main goal of this paper is to present different approaches based on a covering formulation which have been studied in the dissertation thesis.

2. Location-Allocation Approach

The above-formulated p-median problem can be modeled using the following notation. Let the decision of service center location at the place $i \in I$ be modeled by a zero-one variable $y_i \in \{0,1\}$, which takes the value of 1, if a center is located at i and it takes the value of 0 otherwise. In addition, the allocation variables $z_{ij} \in \{0,1\}$ for each $i \in I$ and $j \in J$ are introduced to assign a customer j to a possible location i by the value of one. Then the location-allocation model can be formulated as follows:

$$Minimize \quad \sum_{i \in I} \sum_{j \in J} d_{ij} z_{ij} \tag{2}$$

$$Subject\ to: \quad \sum_{i \in I} z_{ij} = 1 \quad for\ j \in J \tag{3}$$

$$z_{ij} \le y_i \quad for\ \ i \in I\ \ and\ \ j \in J \tag{4}$$

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

$$z_{ij} \in \{0,1\} \quad for\ \ i \in I\ \ and\ \ j \in J \tag{6}$$

$$y_i \in \{0,1\} \quad for\ \ i \in I \tag{7}$$

In the above model, the allocation constraints (3) ensure that each customer is assigned to exactly one possible service center location. Link-up constraints (4) enable to assign a customer to the possible location i only if the service center is located at this location and the constraint (5) bounds the number of located service centers.

The problem described by terms (2) - (7) can be rewritten to a form acceptable by a modeler of integrated optimization environment and solved by associated IP-solver. Due to huge number of allocation variables z_{ij}, a commercial software usually fails, when a very large instance of the problem (2) - (7) is solved.

3. Radial Formulation of the p-Median Problem

Let us use the above introduced notation. As above, the variable $y_i \in \{0,1\}$ models the decision of service center location at the place $i \in I$.

The keystone of the approximate approach consists of a relaxation of the assignment of a service center to a customer [8]. The distance between a customer and the nearest facility is approximated unless the facility must be assigned. To obtain an upper approximation of the original objective function value, the range of all possible distances $< 0, max\{d_{ij} : i \in I, j \in J\} >$ is partitioned into $r+1$ zones. The zones are separated by a finite ascending sequence of $m+1$ dividing points $D_0, D_1, \ldots, D_m$ chosen from the sequence, where $D_0 = 0$ and $D_m = max\{d_{ij} : i \in I, j \in J\}$. The zone k corresponds with the interval $(D_k, D_{k+1}]$, the zone one corresponds with the interval $(D_1, D_2]$ and the r-th zone corresponds with the interval $(D_r, D_m]$. The length of the k-th interval is denoted by e_k for $k = 0, \ldots, r$.

In addition, auxiliary zero-one variables x_{jk} for $k = 0, \ldots, r$ are introduced. The variable x_{jk} takes the value of 1, if the distance of the customer $j \in J$ from the nearest located facility is greater than D_k and it takes the value of 0 otherwise. Then the expression $e_0 x_{j0} + e_1 x_{j1} + \cdots + e_r x_{jr}$ constitutes an upper approximation of the distance d_{j*} from the customer j to the nearest located facility. If the distance d_{j*} belongs to the interval $(D_k, D_{k+1}]$, it is estimated by the upper bound D_{k+1}.

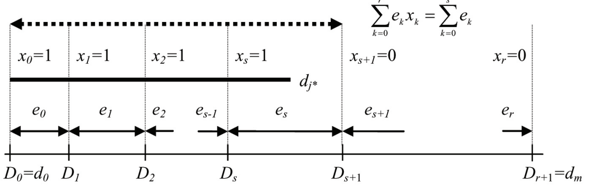

Figure 1: Upper approximation of d_{j*} using zone widths e_k and auxiliary variables x_k. The upper approximation of d_{j*} is denoted by thick dotted line at the top of figure.

Similarly to the covering model, we introduce a zero-one constant a_{ij}^k for each triple $[i, j, k] \in I \times J \times \{0, \dots, r\}$. The constant a_{ij}^k is equal to 1, if the distance between the customer j and the possible location i is less or equal to D_k, otherwise a_{ij}^k is equal to 0. Then a covering-type model can be formulated as follows:

$$Minimize \qquad \sum_{j \in J} \sum_{k=0}^{r} e_k x_{jk} \tag{8}$$

$$Subject\ to: \qquad x_{jk} + \sum_{i \in I} a_{ij}^k y_i \geq 1 \quad for\ j \in J\ and\ k = 0, \dots, r \tag{9}$$

$$\sum_{i \in I} y_i \leq p \tag{10}$$

$$x_{jk} \geq 0 \qquad for\ j \in J\ and\ k = 0, \dots, r \tag{11}$$

$$y_i \in \{0, 1\} \qquad for\ i \in I \tag{12}$$

The objective function (8) gives the upper bound of the sum of the original distances. The constraints (9) ensure that the variables x_{jk} are allowed to take the value of 0 if there is at least one center located in radius D_k from the customer j. The constraint (10) limits the number of located facilities by p.

To obtain a lower bound of the original problem optimal solution, we realize that the interval $(D_k, D_{k+1}]$ given by a pair of succeeding dividing points contains exactly the elements $D_k^1, D_k^2, \dots, D_k^{r(k)}$ of the sequence $d_0, d_1, \dots, d_m$. The elements are strictly greater than D_k and less than D_{k+1}. If a distance d between a customer and a possible service center location belongs to the interval $(D_k, D_{k+1}]$, then the maximum deviation of d from the lower estimation D_{k+1} is $D_{k+1} - D_k^1$. As the variable x_{jk} from the model (8) - (12) takes the value of 1 if the distance of the customer $j \in J$ from the nearest located facility is greater than D_k and this variable takes the value of 0 otherwise, we can redefine the zone coefficients e_k in accordance to $\underline{e}_0 = D_0^1 - D_0$ and $\underline{e}_k = D_k^1 - D_{k-1}^1$ for each $k = 1, \dots, r$. Then the expression $\underline{e}_0 x_{j0} + \underline{e}_1 x_{j1} + \dots + \underline{e}_r x_{jr}$ constitutes a lower approximation of d_{j*}, which corresponds to the distance of the node j from the nearest located facility.

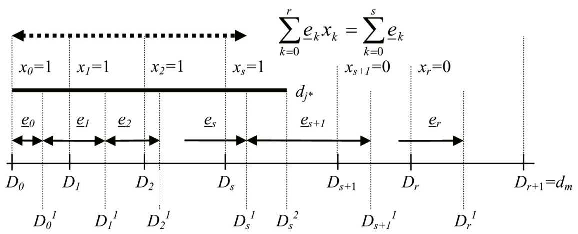

Figure 2: Lower approximation of d_{j*} using zone widths $\underline{e}_k$ and auxiliary variables x_k. The lower approximation of d_{j*} is denoted by thick dotted line at the top of figure.

The optimal objective function value of the following problem gives the lower bound of the objective function value of the original problem [7].

$$Minimize \qquad \sum_{j \in J} \sum_{k=0}^{r} \underline{e}_k x_{jk} \tag{13}$$

$$Subject\ to: \qquad (9) - (12)$$

Having solved both problems, the better of two obtained solutions concerning the original objective function value gives the resulting solution of this approach and the optimal value of (13) gives the lower bound of the unknown optimal solution.Thus, this method allows evaluating the maximal deviation of the approximate solution from theunknown optimal one.

As concern the formulated models, we can easily find that the model (2) - (7) uses $|I| * (|J| + 1)$ decision variables and $(|I| + 1) * |J| + 1$ constraints. The model (8) - (12) contains $|I| + (r + 1) * |J|$ variables and $(r + 1) * |J| + 1$ constraints. It follows, that if we want to keep the approximate model at a moderate size, then the number r of the dividing points must be in order less than the number of possible locations.

4. Optimal Deployment of Dividing Points

The number r of dividing points $D_1, D_2, \dots, D_r$ influences the size of the covering model (8) - (12) as concerns either the number of variables x_{jk} or the number of constraints (9). That is why the number r must be kept in a mediate extent to achieve the resulting solution quickly enough. On the other hand, the smaller the number of dividing points is, the bigger inaccuracy afflicts the approximate solution. Let us focus now on the problem of efficient deployment of given number of dividing points in the set of values $d_0 < d_1 < \dots < d_m$. As before, let us denote $D_0 = d_0$ and $D_m = d_m$. Let the value d_h have a frequency N_h of its occurrence in the matrix $\{d_{ij}\}$. We start from a hypothesis that the distance d_h from the sequence $d_0 < d_1 < \dots < d_m$ occurs in the resulting solution n_h times and that is why the deviation of this distance from its approximation encumbers the total deviation proportionally to n_h.

The distance d between a customer and the nearest located facility can be only estimated taking into account that it belongs to the interval $(D_k, D_{k+1}]$ containing only values $D_k^1, \dots, D_k^{r(k)}$ of the sequence $d_0 < \dots < d_m$. Maximal deviation of the upper estimation D_{k+1} from the exact value d is $D_{k+1} - D_k^1$. If we were able to anticipate a frequency n_h of each d_h in the unknown optimal solution, we could minimize the total deviation of the upper approximation from the unknown optimal solution by deployment of dividing points. The dividing points for the upper approximation follow from the optimal solution of the problem described by (14) - (18).

$$Minimize \qquad \sum_{t=1}^{m} \sum_{h=1}^{t} (d_t - d_h) n_h z_{ht} \tag{14}$$

$$Subject\ to: \quad z_{(h-1)t} \leq z_{ht} \quad for\ t = 2, \dots, m\ and\ h = 2, \dots, t \tag{15}$$

$$\sum_{t=h}^{m} z_{ht} = 1 \qquad for\ h = 1, \dots, m \tag{16}$$

$$\sum_{t=1}^{m-1} z_{tt} = r \tag{17}$$

$$z_{ht} \in \{0, 1\} \qquad for\ t = 1, \dots, m\ and\ h = 1, \dots, t \tag{18}$$

If the distance d_h belongs to the interval ending by a dividing point d_t then the decision variable z_{ht} takes the value of 1. Link-up constraints (15) ensure that the distance d_{h-1} belongs to the interval ending with d_t only if each other distance between d_{h-1} and d_t belongs to this interval. Constraints (16) assure that each distance d_h belongs to some interval and the constraint (17) enables only r dividing points. After the problem (14) - (18) is solved, the nonzero values of z_{tt} indicate the distances d_t which correspond with dividing points.

Similar approach can be used to obtain an efficient deployment of dividing points to determine a good lower bound of the optimal solution of the original problem [7]. Nevertheless, several differences in the way of the approximation must be taken into account. When a lower bound is computed, the expression $\underline{e}_0 x_{j0} + \underline{e}_1 x_{j1} + \cdots + \underline{e}_r x_{jr}$ is used as a lower approximation of d_{j*}. Here $\underline{e}_0 = D_0^1 - D_0$ and $\underline{e}_k = D_{k+1}^1 - D_k^1$ for $k = 1, \ldots, r$. If a relevancy n_h of each d_h is given, we could minimize the total deviation of the lower approximation from the unknown optimal solution in a similar way as in the case of the upper approximation. The dividing points for the lower distance approximation can be obtained from the optimal solution of the problem (19) - (23).

$$Minimize \quad \sum_{t=0}^{m-1} \sum_{h=t}^{m-1} (d_{h+1} - d_{t+1}) n_{h+1} \underline{z}_{th} \tag{19}$$

$$\begin{aligned} &Subject\ to: \quad \underline{z}_{t(h+1)} \leq \underline{z}_{th} \\ &for\ \ t = 0, \ldots, m-1\ \ and\ \ h = t, \ldots, m-1 \end{aligned} \tag{20}$$

$$\sum_{t=0}^{h} \underline{z}_{th} = 1 \qquad for\ \ h = 0, \ldots, m-1 \tag{21}$$

$$\sum_{t=1}^{m-1} \underline{z}_{tt} = r \tag{22}$$

$$\begin{aligned} &\underline{z}_{th} \in \{0, 1\} \\ &for\ \ t = 0, \ldots, m-1\ \ and\ \ h = t, \ldots, m \end{aligned} \tag{23}$$

On the contrary to the upper bound model, it holds here, that if the distance d_h belongs to the interval starting with a possible dividing point d_t then the decision variable z_{th} takes the value of 1. Link-up constraints (20) ensure that the distance d_{h+1} can belong to the interval starting with d_t only if each distance between d_{h+1} and d_t belongs to this interval. Constraints (21) assure that each distance d_h belongs to some interval and the constraint (22) enables that only r dividing points will be chosen. After the problem (19) - (23) is solved, the nonzero values of z_{tt} indicate the distances which correspond with dividing points for the lower bounding process.

5. Sequential Approach

The static approach [8] to the dividing points determination comes out from the estimated frequencies n_h of each d_h in the unknown optimal solution. It is necessary to take into account, that the mentioned sequence N_h of occurrence frequencies does not provide us the information, because it reports only on the elements contained in the matrix $\{d_{ij}\}$. We start here from the hypothesis that the frequency n_h of d_h may be proportional to N_h and to some weight, which decreases with the increasing value of d_h. We formulate this hypothesis in the form of (24).

$$n_h = N_h e^{\frac{-d_h}{T}} \tag{24}$$

In this expression, T is a positive parameter and N_h is the mentioned occurrence frequency, where only $|I| - p + 1$ smallest distances of each matrix column are included. After the frequency determination, models (14) - (18) and also (19) - (23) are used to obtain the series of dividing points for the upper and lower bound respectively. Then the models (8) - (12) and (13), (9) - (12) are used to obtain the upper and lower bounds and the associated resulting solution of the original problem.

To the contrary with the static approach, the presented sequential improvement of the relevancies n_h is based on the idea of making the estimation of the individual distance d_h relevancy more accurate. The distance relevancy here also means a measure of our expectation that this distance value is the distance between a customer and the nearest located service center but this estimation is improved step by step by the following algorithm, which can be used either for the lower or upper bound determination. The input of the algorithm consists of the matrix $\{d_{ij}\}$, the sequence $d_0 < \cdots < d_m$ and the associated sequence of the frequencies N_h for $h = 0, \ldots, m$, and the number p of centers, which are to be located. Further parameters T and r of the algorithm must be given, where T is the shaping parameter and r is the number of the dividing points. The upper bound algorithm can be described by the following steps:

1. Determine the initial values of the relevancies n_h according to (24).
2. Compute the sequence of the dividing points by solving the problem (13)-(17).
3. Using the sequence of the dividing points, determine the constants a_{ij}^k and e_k and solve the covering problem (7)-(11) to the optimality, to obtain the optimal values of the location variables y. Determine the value of the original objective function according to (25), and update the best found solution.
4. If the stopping rule is met, terminate, otherwise go to Step 5.
5. Determine the set I_1 of the active rows according to $I_1 = \{i \in I : y_i = 1\}$. Update the relevancies n_h so that each column of the matrix $\{d_{ij}\}$ is processed and only minimal value over the active rows is included into the set of the relevant distances and their occurrence frequencies. Go to Step 2.

$$\sum_{j \in J} min\{d_{ij} : i \in I, y_i = 1\} \tag{25}$$

The above algorithm can be converted to the lower bound algorithm by replacing the model (13)-(17) in Step 2 by the model (18)-(22) and by replacing the model (8)-(12) in Step 3 with the model (13), (9)-(12).

The effectiveness of the suggested algorithm, and also the time necessary to find the resulting solution both depend on the criteria of terminating the iteration process. The easiest way consists of the basic condition that the computing process is to be stopped whenever no better solution of the original problem is obtained. Since the sequential method may perform too much iteration with very little improvement of the objective function value, we suggest limiting the number of the performed iterations. In the dissertation thesis, we study some additional rules that make the algorithm less time-consuming.

The above algorithm starts with an initial relevancy estimation described in the previous section, and computes relevancies n_h in accordance to the hypothesis formalized by the expression (24). Having obtained the first optimal solution of the covering model following the dividing points deployed according to the initial relevancies, the algorithm updates the relevancies. For this purpose, a set of active matrix rows is defined so that the i-th row of the matrix $\{d_{ij}\}$, is denoted as active, if the location variable y_i of the problem (8)-(12) is equal to one. Then each column j of this matrix is processed, the minimal value over the active rows is included into the set of the relevant distances, and the associated frequency is increased. Thus a new sequence of the distance frequencies n_h is obtained. These new frequencies are used in the next iteration of the algorithm. This process can be repeated as long as better solution of the original problem keeps being obtained, or until the used stopping criterion is met.

6. Distance Relevance

This chapter deals with the fundamental problem of dividing points deployment, which consists in the distance relevance estimation. The distance value relevance expresses the strength of our expectation that the distance belongs to the optimal solution of the p-median problem. There are several ways to express the distance value relevance using mathematical relations. In previous works [6], [7] we based our research on the idea that the distance value relevance dropps exponentially with the distance value. In the dissertation thesis we base our approach to the relevance estimation on so called column ranking, where not only value but also the order of the distance is taken into account. We suggest the formulation of the ranking relevance and study the influence of the associated dividing point deployment on the effectiveness of the approximate approach.

The former approach to the estimation of the relevance follows the hypothesis that the relevance decreases exponentially with the increasing distance value [8]. The hypothesis can be mathematically expressed by (24), where T is a positive parameter and N_h is the above-mentioned d_h occurrence frequency in the matrix $\{d_{ij}\}$. Hereby, we present a new approach to the relevance, where the column ranking evaluation $L_j^{ts}(d_{ij})$ of the distance d_{ij} is used to define the relevance n_h according to

$$n_h = L^{tsh} = \sum_{j \in J} \sum_{\substack{i \in I \\ d_{ij} = d_h}} L_j^{ts}(d_{ij}) \qquad (26)$$

The linear column ranking function $L_j^{ts}(d_{ij})$ is defined as follows: let $P_j(d_{ij})$ be the position of d_{ij} in the ascending sequence of the j-th column items of the distance matrix $\{d_{ij}\}$ and let a denote the cardinality of I. Then $L_j^{ts}(d_{ij}) = a + s * (1 - P_j(d_{ij}))$ for $P_j(d_{ij}) < a + 1 - t$ and $L_j^{ts}(d_{ij}) = 0$ otherwise. The parameters t and s represent a threshold and a step respectively. The threshold influences the number of $t - 1$ largest distances of the j-th column, which are not taken into account, and the step gives the difference between the contributions of the k-th and $k - 1$th item of the ascending sequence of the j-th column items. The parameter t can vary over the range $[p..a - 1]$ of integers, and the step s can take values from the interval $[0, a/(a - t)]$. At the end of this section we introduce the third approach to the relevance, which combines both exponential and ranking approaches. The associated relevance n_h is defined in accordance with (27).

$$n_h = L^{tsh} e^{\frac{-d_h}{T}} \qquad (27)$$

Various approaches to the distance relevance estimation have been studied and compared within the dissertation thesis. This topic presents a very interesting area which will become a subject of future possible research.

The dissertation thesis contains also a big portion of numerical results that prove the usefullness of suggested approaches based on the set covering formulation. The main goal of the numerical comparison was to find appropriate settings of different parameters which influence the solution accuracy of concrete instances. The test problems were taken from commonly used benchmark libraries used in available literature. The aim of this paper is not to present all the numerical results, but to give an overview of suggested principles and to describe the studied research field.

7. Summary

Designing a public service system, including medical emergency system, fire-brigade deployment, public administration system and many others, can often bring along some overall combinatorial problems concerning the system structure. The public service system structure is formed by the deployment of the limited number of service centers, and the associated objective is to minimize social costs, which are proportional to the distances from serviced objects to the nearest source of the provided service. The mathematical models of the public service system design problem are often related to the p-median problem, which is formulated as the task of the determination of most p network nodes as facility locations, so that the sum of the distances between each node and the nearest located facility is minimal. With real problems, the number of serviced customers takes the value of several thousands, and the number of possible facility locations can take this value as well. The number of possible service center locations seriously impacts the computational time. To obtain a good decision on a facility location in any serviced area, the mathematical model of the problem can be formulated, and some of mathematical programming methods can be applied to find an optimal solution. The location-allocation model constitutes a mathematical programming problem which resists any attempt at a fast solution.

In the dissertation thesis, we deal with annother way of the p-median problem representation by means of mathematical programming, which uses so called radius formulation. This approach avoids assigning individual customers to some of the located facilities, and it deals only with the information whether some facility is or is not located in a given radius from the customer. The later approach leads to the model similar to the set covering problem, which is easy to solve even for large instances by common optimization software tools. This approach pays for shorter computational time or smaller computer memory demand by losing its accuracy. The accuracy can be improved by some convenient determination of so called dividing points which are used in the problem reformulation. The dividing points can be determined so that the expected deviation is minimized and the deviation is expressed using so called distance value relevance.

The distance value relevance expresses the strength of expectation that the distance belongs to the optimal solution of the p-median problem. There are several ways to express the distance value relevance. In previous works we based our research on the idea that the distance value relevance dropped exponentially with the distance value. In the dissertation thesis we base our approach to the relevance estimation on so called column ranking, where not only value but also the order of the distance is taken into account. We suggest the formulation of the relevance and study its influence of the associated dividing point deployment on the effectiveness of the approximate approach.

As a complementary algorithm used for making the covering method more accurate we have suggested a sequential approach to the dividing points determination, and we use them not only to obtain a good solution of the problem but also to gain lower bound of the unknown optimal solution. We demonstrate the impact of the sequential approach to the accuracy and computational time. We have performed the numerical comparison of the suggested approximate method to the location-allocation approach. To solve the associated problems, we use the common optimization environment XPRESS.

Suggested methods proved to be very efficient as far as the accuracy is concerned, when solving the instance up to one thousand customers or middle and larger instances originated from real transportation networks. Therefore, we can conclude that the proposed method is a suitable complement to the state-of-the-art methods. The main contribution of the suggested approach lies in its simple implementation without the necessity of programming several algorithms. Just two different models are enough to obtain a good solution in a short time. Furthermore, common software tools can be used instead of special ones.

Further research connected with the suggested dynamical approach will be focused on the initial stage of the approach, where we found a possibility of making the computational process more efficient by a suitable adjustment of the mentioned parameter T.

Acknowledgements. This work was supported by the research grants VEGA 1/0296/12 "Public Service Systems with Fair Access to Service" and APVV-0760-11 "Designing of Fair Service Systems on Transportation Networks". We would also like to thank to "Centre of excellence for systems and services of intelligent transport" (ITMS 26220120028) for built up used infrastructure.

References

[1] Avella, P., Sassano, A., Vassil'ev, I.: *Computational study of large scale p-median problems*, Mathematical Programming 109 pp. 89-114, (2007)

[2] Beasley J. E.: *OR Library: Distributing Test Problems by Electronic Mail*, Journal of the Operational Research Society, 41(11), pp 1069-1072, (1990)

[3] Cornuejols, G., Nemhauser, G.L., Wolsey, L.A.: *A canonical representation of simple plant location problems and its applicaitons*, SIAM Journal on Algebraic and Discrete Methods 1 (3) pp. 261-272, (1980)

[4] Doerner, K.F., Gutjahr, W.J., Hartl, R.F., Karall, M.; Reimann, M.: *Heuristic Solution of an Extended Double-Coverage Ambulance Location Problem for Austria*, Central European Journal of Operations Research, Vol. 13, No 4, pp. 325-340, (2005)

[5] García, S., Labbe, M., Marin, A.: *Solving large p-median problems with a radius formulation*, INFORMS Journal on Computing 23 (4) pp. 546-556, (2011)

[6] Janáček, J.: *Approximate Covering Models of Location Problems*, In: Lecture Notes in Management Science. Proceedings of the 1st International Conference on Applied Operational Research ICAOR ť08, Vol. 1, Sept. 2008, Yerevan, Armenia, pp. 53-61, (2008)

[7] Janáček, J.: *Lower bound on large p-median problem using the sequential zone adjustment method*, In: Mathematical methods in economics 2011: Proceedings of the 29th international conference: September 6-9.2011, Jánska Dolina, Slovakia, Praha: Professional Publishing, 2011, pp. 317-320, (2011)

[8] Janáček, J., Kvet, M.: *Approximate solving of large p-median problems*, In: ORP3=Operational research peripatetic post-graduate programme: Cádiz, Spain, September 13-17, 2011: proceedings, Cádiz: Servicio de Publicaciones de la Universidad de Cádiz, 2011, pp. 221-225, (2011a)

[9] Janáček, J., Kvet, M.: *Sequential zone adjustment for approximate solving of large p-median problems*, In Operations Research Proceedings 2011: Selected Papers of the International Conference on Operations Research (OR 2011), August 30-September 2, 2011, Zurich, Switzerland, Berlin Heidelberg: Springer-Verlag, 2011, pp. 269-274, (2011b)

[10] Janáček, J., Linda, B., Ritschelová, I.: *Optimization of Municipalities with Extended Competence Selection*, In Prager Economic Papers Ű Quarterly Journal of Economic Theory and Policy, Vol. 19, No 1, pp. 21-34, (2010)

[11] Janáčková, M., Szendreyová, A.: *An impact of set of centres of distribution system design problem on computational time of solving algorithm*, In: Mikulski J. (ed): Advances in Transport Systems Telematics, Katowice, pp 109-112, (2006)

[12] Jánošiková, L.: *Emergency Medical Service Planning*, Communications - Scientific Letters of the University of Žilina, Vol. 9, No 2, pp 64-68, (2007)

[13] Marianov, V., Serra, D.: *Location problems in the public sector*, In: Drezner Z (ed.) et al. Facility location. Applications and theory. Berlin: Springer, pp. 119-150, (2002)

Selected Papers by the Author

J. Janáček, M. Kvet. *Sequential zone adjustment for approximate solving of large p-median problems* In Operations Research Proceedings 2011: Selected Papers of the International Conference on Operations Research (OR 2011), August 30-September 2, 2011, Zurich, Switzerland, Berlin Heidelberg: Springer-Verlag, 2011, pp. 269-274

J. Janáček, M. Kvet. *Approximate solving of large p-median problems* In Operational research peripatetic post-graduate programme: Cádiz, Spain. Sept. 13-17, 2011, Cádiz: Servicio de Publicaciones de la Universidad de Cádiz, 2011, pp. 221-225

J. Janáček, M. Kvet. *Dynamic zone adjustment for approximate approach to the p-Median problem* In ORP3 = Operational research peripatetic post-graduate program 2012, July 16-20 2012, Linz, Austria, pp. 555-562

J. Janáček, M. Kvet. *Relevant network distances for approximate approach to the p-median problem* In Operations Research Proceedings 2012: Selected Papers of the International Conference on Operations Research (OR 2012), September 4-7 2012, Leibniz Univesitat Hannover, Germany, Berlin Heidelberg: Springer-Verlag, 2012, to appear

M. Kvet. *Umiestňovanie zdrojov na riešenie krízových javov* In Mladá veda 2010, 9.-10.11.2010, Žilina: University of Žilina, 2010

M. Kvet. *Applications of the set covering problem* In Mathematics for information and communication technologies: 6th winter school of mathematics for ICT: Šachtičky 3.-8.1.2011, Banská Bystrica: Science and Research Institute, MBU, 2011, pp. 67-70

M. Kvet. *Aproximácia matice vzdialeností pre približné riešenie rozsiahlych úloh p-mediánu* In INFOTRANS 2011: 28.4. 2011, Pardubice: Univerzita Pardubice, 2011, pp. 123-128

M. Kvet. *Približné riešenie rozsiahlych úloh mediánového typu* In Úlohy diskrétní optimalizace v dopravní praxi 2011, Pardubice 29.4 2011, Pardubice: Univerzita Pardubice, Dopravní fakulta Jana Pernera, 2011, pp. 18-33

M. Kvet. *Static Zone Adjustment for Approximate Solving of the p-Median Problem* In TRANSCOM 2011: 9-th European conference of young research and scientific workers: Žilina, 27.-29.6.2011, Žilina: University of Žilina, pp. 129-132

M. Kvet. *Adaptive algorithm of parameter adjustment for approximate solving of the p-median problem* In Mathematical methods in economics 2011: proceedings of the 29th international conference: September 6-9.2011, Jánska Dolina, Slovakia, Praha: Professional Publishing, pp. 437-442

M. Kvet. *Využitie IP-solvera na približné riešenie rozsiahlych úloh p-mediánu* In Úlohy diskrétní optimalizace v dopravní praxi 2012, Pardubice: Dopravní fakulta Jana Pernera, Univerzita Pardubice, 18.5.2012, pp. 56-70

M. Kvet. *Public service system design with advanced covering approach* In Technika, technologie, telematika a informatika v dopravních a logistických systémech, Pardubice: Dopravní fakulta Jana Pernera, Univerzita Pardubice, 15.11.2012, to appear

M. Kvet, J. Janáček. *Trade-off the accuracy for computational time in approximate solving technique for the p-median problem* In Quantitative methods in economics 2012: multiple criteria decision making XVI: proceedings of the international scientific conference: 30th May - 1st June 2012, Bratislava, Slovakia, Ekonóm, 2012, pp. 136-140

M. Kvet, J. Janáček. *Generalization of zoning method for approximate approach to large p-median problems* In Methods and Applications of Artificial Intelligence - Proceedings of the International Workshop, 21.-22. September 2012, Bielsko-Biala: Institute of Management and Information Technology, Poland, pp. 5-10