Czech Economic Review 7 (2013) 119–127
Acta Universitatis Carolinae Oeconomica

Comprehensive Approximate Sequential Tool for Public Service System Design

Marek Kvet*, **Jaroslav Janáček****

Received 11 September 2012; Accepted 29 April 2013

Abstract This paper deals with the approximate approach to the p-median problem, which constitutes a background for the public service system design. The real instances are characterized by thousands of possible service center locations. The attempts at exact solving of these problems often fail due to enormous computational time or huge memory demands, when the location-allocation model is used. The presented approach uses an approximation of a common distance by some pre-determined distances given by so-called dividing points. Covering formulation of the problem enables the implementation of the solving technique in the frame of commercial optimization software to obtain near-optimal solution in a short term. As the deployment of the dividing points influences the accuracy of the solution, we have developed the sequential method of dividing point deployment. The main goal of this study is to explore the effectiveness of suggested approximate method measured by the solution accuracy in comparison to the saved computational time.

Keywords p-median problem, approximate covering model, lower bound, sequential method
JEL classification C61

1. Introduction

Designing of almost any public service system includes determination of some centers, from which the associated service is distributed to all users of the system (Current et al. 2002; Janáček et al. 2010; Jánošíková 2007). The service providing must be concentrated to a limited number of the centers due to the economic and technological reasons, which follows from indivisibility of facilities necessary for the service. Regardless of the case, whether the service is delivered to users or the users travel for the service to the nearest center, the communication between the user and the service provider is performed via a transportation network, which covers the serviced area.

Thus the public service system structure is formed by deployment of limited number of service centers and the associated objective is to minimize social costs which are proportional to the distances between serviced objects and the nearest service centers. Mathematical models of the public service system design problem are often related to the p-median problem, where the number of serviced users takes the value of several

* University of Žilina, Faculty of Management Science and Informatics, Univerzitná 8215/1, 01026 Žilina, Slovakia. Phone: +421 41 5134209, E-mail: Marek.Kvet@fri.uniza.sk.
** University of Žilina, Faculty of Management Science and Informatics, Univerzitná 8215/1, 01026 Žilina, Slovakia. Phone: +421 41 5134204, E-mail: Jaroslav.Janacek@fri.uniza.sk.

thousands and the number of possible service center locations can take this value as well. The number of possible service center locations seriously impacts the computational time (Janáčková and Szendreyová 2006). The necessity of solving large instances of the p-median problem has led to the approximate approach, which enables us to solve real-sized instances in admissible time (Janáček and Kvet 2011a,b).

There were developed at least two approximate approaches in the past to avoid introducing vast structure of allocation variables for each pair of possible service center location and service user. Both approaches were based on so called radial model, in which a user is not assigned to a concrete center location, but there is only recognized whether some service center is or is not located in a given radius from the user. The first approach forms a particular finite system of radii for each user according to distances from the user to possible service center locations (see Avella et al. 2007; Cornujelos et al. 1980; García et al. 2011).

The second approach discussed and improved in this paper forms one common system of radii based on a set of so-called *dividing points*, which are employed to estimate the distances between users and the nearest service centers (see Janáček 2008). The dividing points are deployed in the range of the network distance values between possible service centers and user locations. The deployment of dividing points impacts the deviation of the approximate objective function from the original objective function of the p-median problem. The dividing points can be determined so that the expected deviation is minimized and the deviation is expressed using so called distance value relevance (Janáček and Kvet 2012). The distance value relevance expresses the strength of expectation that the distance belongs to searched optimal solution of the p-median problem.

Furthermore, we took into account a usual position of a common public service system designer, who faces the necessity to submit complete design in a short term. This situation excludes the long term development of a special software for this purpose. This is why we concentrate our effort on such approach, which can be implemented in the frame of a commercial optimization software and thus enables a common designer to obtain a near-optimal solution of the designed public service system in a short term.

The presented approach pays for shorter computational time or smaller computer memory demand by a loss of accuracy. The accuracy can be improved by the convenient determination of the dividing points, which are used in the objective function approximation. The suggested sequential approach to the dividing points determination proved to be suitable not only for obtaining a good solution of the problem, but also for gaining a good lower bound of the unknown optimal solution (Janáček 2011). The experiments with solving very large instances by the approximate approach gave enormous precise results, which differed from the exact solution by less than one percent. Unfortunately, the computational time necessary for the lower bound determination was very high. In this contribution, we focus on this phenomenon and suggest a “trade-off” the accuracy for shorter computational time.

2. Sequential approach to the p-median problem

The p-median problem is formulated as the task of determination of at most p network nodes as service center locations so that the sum of distances between each node-user location and the nearest service center is minimal. To describe the p-median problem on a network we denote the set of serviced users by J, similarly, we denote the set of possible center locations by I. The network distance between the possible center location i and the user j from J is denoted as d_{ij}. The basic decisions in any solving process of the p-median problem concern location of centers at possible center location from the set I.

To model these decisions at particular location i, we introduce a zero-one variable $y_i \in \{0,1\}$, which takes the value of 1, if a center should be located at the location i, and which takes the value of 0 otherwise. The discussed approximate approach is based on the relaxation of the assignment of the user to the nearest service center. The distance between the user and the nearest service center is approximated unless the service center must be identified. To obtain an upper or a lower bound on the original objective function, the range $\langle 0, \max\{d_{ij} : i \in I,\, j \in J\}\rangle$ of all possible values $d_0 < d_1 < \ldots < d_m$ of distances from the matrix $\{d_{ij}\}$ is partitioned into $r+1$ zones. The zones are separated by a finite ascending sequence of so called dividing points $D_1, D_2, \ldots, D_r$ chosen from the sequence $d_0 < d_1 < \ldots < d_m$, where $0 = d_0 = D_0 < D_1$ and $D_r < D_m = d_m = \max\{d_{ij} : i \in I,\, j \in J\}$. We introduce a numbering of zones so that the zone k corresponds with the interval $(D_k, D_{k+1}\rangle$. 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 from the user $j \in J$ to the nearest located service center is greater than D_k and it takes the value of 0 otherwise. Then the expression $e_0 x_{j0} + e_1 x_{j1} + \ldots + e_r x_{jr}$ constitutes the upper approximation of the distance d_{j^*} from the user j to the nearest located center. If the distance d_{j^*} belongs to the interval $(D_k, D_{k+1}\rangle$, it is estimated by the upper bound D_{k+1} with the maximal possible deviation e_k.

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

$$\min \quad \sum_{j\in J}\sum_{k=0}^{r} e_k x_{jk} \tag{1}$$

$$\text{s.t.:} \quad x_{jk} + \sum_{i\in I} a_{ij}^k y_i \geq 1 \quad \text{for } j \in J \text{ and } k = 0, \ldots, r \tag{2}$$

$$\sum_{i\in I} y_i \leq p \tag{3}$$

$$x_{jk} \geq 0 \quad \text{for } j \in J \text{ and } k = 0, \ldots, r \tag{4}$$

$$y_i \in \{0,1\} \quad \text{for } i \in I \tag{5}$$

The objective function (1) gives the upper bound of the sum of original distances. The constraints (2) ensure that the variables x_{jk} are enforced to take the value of 1 if there is no center located in radius D_k from the user j. The constraint (3) limits the number of located service centers by p.

To obtain the lower bound of the objective function value of the optimal solution of the original problem, we realize, that the interval $(D_k, D_{k+1}\rangle$ given by a pair of succeeding dividing points D_k and D_{k+1} contains some elements of the sequence $d_0 < d_1 < \ldots < d_m$. Let us denote the elements by $D_k^1, D_k^2, \ldots, D_k^{r(k)}$. The elements are strictly greater than D_k and less or equal to D_{k+1}. If a distance d between a user and a possible center location belongs to the interval $(D_k, D_{k+1}\rangle$, then the maximal possible 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 (1)–(5) takes the value of 1, if the distance from the user $j \in J$ to the nearest located center 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 $k = 1, \ldots, r$. Then the expression $\underline{e}_0 x_{j0} + \underline{e}_1 x_{j1} + \ldots + \underline{e}_r x_{jr}$ is the lower approximation of d_{j^*}, which corresponds to the distance of the serviced node j from the nearest located center i. The redefined objective function value of the optimal solution gives the lower bound of the original problem (Janáček 2011). Having solved both problems, the better of two obtained solutions concerning the original objective function gives the resulting solution.

The number r influences the size of the covering model (1)–(5) and the deployment of the dividing points influences the accuracy of the associated result. The dividing points can be chosen from the set of values $d_0 < d_1 < \ldots < d_m$ of the distance matrix $\{d_{ij}\}$, where $D_0 = d_0$ and $D_m = d_m$. Let the value d_h have a frequency N_h of its occurrence in the matrix $\{d_{ij}\}$. In suggested approximate approach, we start from the hypothesis, that the distance d_h from the sequence $d_0 < d_1 < \ldots < d_m$ 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. If a relevance n_h of each d_h is estimated, we could minimize the deviation using other series of dividing points obtained by solving the following problem:

$$\min \quad \sum_{t=0}^{m-1} \sum_{h=t}^{m-1} (d_{h+1} - d_{t+1}) n_{h+1} \underline{z}_{th} \tag{6}$$

$$\text{s.t.:} \quad \underline{z}_{t(h+1)} \leq \underline{z}_{th} \qquad \text{for } t = 0, \ldots, m-1 \text{ and } h = t, \ldots, m-1 \tag{7}$$

$$\sum_{t=0}^{h} \underline{z}_{th} = 1 \qquad \text{for } h = 0, \ldots, m-1 \tag{8}$$

$$\sum_{t=1}^{m-1} \underline{z}_{tt} = r \tag{9}$$

$$\underline{z}_{th} \in \{0, 1\} \qquad \text{for } t = 0, \ldots, m-1 \text{ and } h = t, \ldots, m \tag{10}$$

If the distance d_h belongs to the interval starting with a possible dividing point d_t, then the decision variable $\underline{z}_{th}$ takes the value of 1. Link-up constraints (7) 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 (8) assure that each distance d_h belongs to some interval and the constraint (9) enables that only r dividing points will be chosen. After the problem (6)–(10) is solved, the nonzero values of $\underline{z}_{tt}$ indicate the distances d_t which correspond with the dividing points for the lower bounding process. The sequential zone adjustment method is based on the idea of making the estimation of individual distance d_h relevance n_h more and more accurate. The distance relevance represents here some measure of our expectation that this distance value is the distance between a user and the nearest located service center, but this estimation is improved step by step (Janáček and Kvet 2011b, 2012).

The initial relevance estimation is related to the frequency N_h of d_h occurrence. To make the relevance estimation more realistic, we suggest making use of the optimal solution of the approximate problem described by (1)–(5) and (6)–(10). Assuming that the optimal solution of the approximate problem is known, especially knowing which y_i is equal to one, only the associated rows of the matrix $\{d_{ij}\}$ are taken as so called *active rows*. Then each column of this matrix is processed and the minimal value over the active rows is included into the set of relevant distances and their occurrence frequencies. Thus a new sequence of the distance frequencies n_k is obtained. The first and last distances d_0 and d_m are taken from the former sequence. The new sequence is used in the problem (6)–(10) and the optimal solution of the problem yields a new set of the dividing points. Based on this new dividing points set, the sequence of zones is readjust and for new zone sequence the approximate problems are resolved. This process can be repeated until the stopping criterion is met (Janáček and Kvet 2012).

3. Stopping rule and lower bound computational time

The above approach was tested on the pool of benchmarks, which size varied from 1 to 17 hundreds of possible service center locations. Solved instances of the problem were obtained from the *OR-Lib* set of the *p*-median problem instances (Beasley 1990). This pool of benchmarks was enlarged by instances obtained from Slovak road network (Janáček and Kvet 2012; Janáčková and Szendreyová 2006). The cardinality of I exceeds the number of 900 in these new instances. The cardinality of J is the same as the cardinality of I and the maximal number p of located service centers was chosen so that the ratio of $|I|$ to p was 2, 3, 4, 5, 10 and 20. The number of dividing points r was set to the value of 20 for each solved instance and also the same stopping criterion was used to terminate the main loop, which performs individual iterations. The stopping criterion combines conditions of reaching maximal number 10 of performed iterations and performing only one iteration without improving the lower bound.

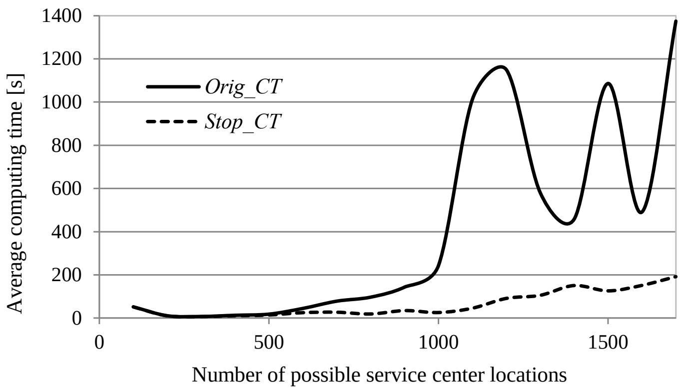
As shown in the Table 1 (see the row *Orig_Gap*) and on the Figure 1 (see the curve *Orig_CT*), this approach proved to be very efficient as concerns the accuracy, but the demanded computational time was unstable and extremely high. This phenomenon evoked to us the idea to make a "trade-off" between the excellent accuracy and the high computational time of the approach (Kvet and Janáček 2012). This compromise was achieved by the adjustment of the stopping criterion, where reaching of a tolerable deviation below two percent of the best found solution is taken as the third condition for the loop termination.

The approach has been tested on the pool of benchmarks and the associated results

Table 1. Average gaps between the best found solutions and the associated lower bounds in percent of the best found solution

$\|I\|$	100	300	500	700	900	1100	1300	1500	1700
Orig_Gap	4.20	2.12	0.75	0.25	0.10	0.07	0.06	0.05	0.13
Stop_Gap	4.20	2.12	0.80	0.36	0.20	0.17	0.24	0.21	0.32

Source: Kvet and Janáček (2012)



Source: Kvet and Janáček (2012)

Figure 1. Time comparison of solving approaches

are plotted in Table 1 and Figure 1. The table 1 contains average gaps between the best found solutions and the associated lower bounds in percent of the best found solution values. Each column corresponds to the size of 6 benchmarks, which were solved by the original approach (*Orig_*) and the second one (*Stop_*). The associated rows are denoted by the corresponding prefixes and this notation is also used for the associated curves in Figure 1, where dependences of average computational times (*CT*) on the benchmark sizes are depicted.

Comparing reported results we can conclude, that the attempt at the trade-off proved to be successful in considerable reduction of computational time under small loss of accuracy.

4. Comprehensive lower and upper approach

In the resulting suggested comprehensive approach based on the mentioned stopping criterion, we make use of the lower bound approach as concerns the dividing point determination. It means that we started with the frequency N_h of d_h occurrence and set the initial relevance at n_h given by the expression (11), where T is a positive parameter.

$$n_h = N_h e^{\frac{-d_h}{T}} \quad \forall h = 0, 1, \dots, m \tag{11}$$

Then the problem (6)–(10) was solved and the initial sequence of dividing points was found. This sequence was used both for the lower bound and upper bound computational processes, where both processes employed the model (1)–(5), each of them did it with the specific zone width coefficients $\underline{e}_k$ and e_k respectively. The best found solution was initialized by the solution, which had lower value of the original objective function. The lower bound solution was used to determine new values of n_h and the process was repeated until the stopping criterion was met. The stopping criterion in the comprehensive approach includes condition that the process is to be stopped whenever the gap between the original objective function value of the best found solution and the highest value of obtained lower bound is less than a given percentage. The best found solution value corresponds to hundred percent. In addition, it combines conditions of reaching maximal number 10 of performed iterations and performing only one iteration without improving the lower bound.

5. Numerical experiments

The original and comprehensive approaches were tested on the pool of instances (Janáček and Kvet 2012), which size varied from 1 to 17 hundreds of possible center locations. The number of dividing points r was fixed to the value of 20 in all solved instances and the parameter T was fixed at the value of 1. The maximal gap from the stopping criterion was set to 2 percent of the best found solution. All experiments were performed using the optimization software FICO Xpress 7.2 (64-bit, release 2011). The associated code was run on a PC equipped with the Intel®Core$^{\text{TM}}$i7 2630QM processor with the parameters: 2.0 GHz and 8 GB RAM.

Both approaches have been tested on the pool of benchmarks and the associated results are plotted in Table 2 and Table 3. Table 2 contains average gaps between the best found solutions and the associated lower bounds in percentage of the best found solution values both for the original approach (*Orig_Gap*) and for the comprehensive

Table 2. Average gaps between the best found solutions and the associated lower bounds in percent of the best found solution

$\lvert I \rvert$	100	300	500	700	900	1100	1300	1500	1700
Orig_Gap	3.23	1.98	0.69	0.34	0.19	0.16	0.16	0.10	0.26
Comp_Gap	3.26	2.12	0.68	0.34	0.20	0.21	0.24	0.21	0.33

Table 3. Average computational times in seconds of the original and comprehensive approaches

$\lvert I \rvert$	100	300	500	700	900	1100	1300	1500	1700
Orig_CT	24.6	11.6	30.3	98.8	21.2	36.7	52.5	62.7	125.9
Comp_CT	24.1	12.9	22.4	94.6	23.2	28.9	49.3	57.3	99.4

one (*Comp_Gap*). Each column corresponds to experiments performed with six benchmarks, where each one has the same size reported in the head of the table. The rows are denoted by the corresponding prefixes and this notation is also used in Table 3, where the average computational times (*CT*) on the benchmark sizes are depicted.

Comparing the results in Table 2 we can conclude that suggested comprehensive approach is able to solve the instances almost at the same accuracy as the original approach. Taking into account that the stopping criterion includes the condition that the computational process terminates whenever the gap is less than two percent, the differences of entries for $|I| \geq 500$ are a consequence of randomness. Some significant differences can be observed in Table 3 between the computational times, especially for bigger instances. From this point of view the comprehensive approach tightly wins.

6. Conclusions

We have introduced the comprehensive approximate approach, which yields a near-optimal solution of the public service system design for large instances in admissible computational time. In addition, this approach provides a lower bound on the optimal solution and thus enables to determine the maximal possible deviation of the unknown optimal objective function value from the objective function value of the found solution. Another contribution from the algorithmic point of view is that the suggested approach is more compact, i.e. one step consists of three optimization processes instead of four optimizations, which are included in one step of the original approach.

Acknowledgment 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 26220120050) for built up the infrastructure, which was used.

References

Avella, P., Sassano, A. and Vassil'ev, I. (2007). Computational Study of Large Scale p-Median Problems. *Mathematical Programming*, 109, 89–114.

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

Cornujelos, G., Nemhauser, G. L. and Wolsey, L. A. (1980). A Canonical Representation of Simple Plant Location Problems and its Applications. *SIAM Journal on Algebraic and Discrete Methods*, 1(3), 261–272.

Current, J., Daskin, M. and Schilling, D. (2002). Discrete Network Location Models. In Drezner, Z. (ed.), *Facility Location. Applications and Theory*. Berlin, Springer, 81–118.

García, S., Labbé, M. and Marín, A. (2011). Solving Large p-Median Problems with a Radius Formulation. *INFORMS Journal on Computing*, 23(4), 546–556.

Janáček, J. (2008). Approximate Covering Models of Location Problems. In Sheibani, K. (ed.), *Lecture Notes in Management Science. Proceedings of the 1st International Conference on Applied Operational Research*, Vol. 1, Sept. 2008. Yerevan, 53–61.

Janáček, J. (2011). Lower Bound on Large *p*-Median Problem Using the Sequential Zone Adjustment Method. In Dlouhý, M. and Skočdopolová, V. (eds.), *Mathematical Methods in Economics 2011: Proceedings of the 29th International Conference*. Praha, Professional Publishing, 317–320.

Janáček, J. and Kvet, M. (2011a). Approximate Solving of Large *p*-Median Problems. In Castano, A., Espejo, I., López, B., Puerto, J., Ramos, C., Rodríguez-Chía, A. M. and Valero, C. (eds.), *Proceedings of the Operational Research Peripatetic Post-Graduate Programme*. Cádiz, Servicio de Publicaciones de la Universidad de Cádiz, 221–225.

Janáček, J. and Kvet, M. (2011b). Sequential Zone Adjustment for Approximate Solving of Large *p*-Median Problems. In Klatte, D., Lüthi, H.-J. and Schmedders, K. (eds.), *Operations Research Proceedings 2011*. Berlin-Heidelberg, Springer-Verlag, 269–274.

Janáček, J. and Kvet, M. (2012). Dynamic Zone Adjustment for Approximate Approach to the *p*-Median Problem. In *Proceedings of the Operational Research Peripatetic Post-Graduate Programme*. Linz, Johannes Kepler University, 555–562.

Janáček, J., Linda, B. and Ritschelová, I. (2010). Optimization of Municipalities with Extended Competence Selection. *Prague Economic Papers*, 19(1), 21–34.

Janáčková, M. and Szendreyová, A. (2006). An Impact of Set of Centers of Distribution System Design Problem on Computational Time of Solving Algorithm. In Mikulski, J. (ed.), *Advances in Transport Systems Telematics*. Katowice, 109–112.

Jánošíková, L. (2007). Emergency Medical Service Planning. *Communications – Scientific Letters of the University of Žilina*, 9(2), 64–68.

Kvet, M. and Janáček, J. (2012). Trade-Off the Accuracy for Computational Time in Approximate Solving Technique for the *p*-Median Problem. In *Quantitative Methods in Economics: Multiple Criteria Decision Making XVI*. Bratislava, Ekonóm, 136–140.