Approximate Covering Models of Location Problems

Jaroslav Janáček

Faculty of Management Science and Informatics, University of Žilina
01026 Žilina, Slovak Republic
jaroslav.janacek@fit.uniza

Abstract. There are two main computer-supported approaches to public service system design. The approaches differ in used quality criteria and the associated models. If the quality criterion does not include relation between a served customer and a servicing facility (e.g., a distance) explicitly, then the problem can be usually modeled as a covering problem and it can be solved quickly by a commercial IP-solver. On the contrary to this favorite case, the location problems, criterion of which takes into account individual distances, are modeled by so-called location-allocation models. Only special algorithms can solve large sized instances of the latter problems and the associated computational times often exceed any admissible limit. A new modeling approach to these bad-solvable problems is presented in this contribution. This approach is based on a specific reformulation of an allocation model to generalized covering model, which consists of less variables than the allocation one and preserves the properties, which enable its fast solving using a standard IP-solver. As a loss of preciseness is paid for this easier solvability, three big instances of the problem were chosen to demonstrate suitability of the approach and to find a relation between size of the instance, preciseness and computational time. In: Sheibani K (ed). Proceedings of the 1st International Conference on Applied Operational Research – ICAOR (2008), pp 53–61. Lecture Notes in Management Science Vol. 1. ISSN 2006-0050.

Keywords: Public service system design, location-allocation problem, covering model, approximate model.

1 Introduction

Most of designers, which try to design an effective public service system like medical emergency system or similar systems, use methods of mathematical programming to get optimal decisions on facility location in a serviced area. The associated models are characterized by considerably big number of possible facility locations, which must be taken in consideration. For some integer programming algorithms, this number represents the substantial size, which impacts computational time necessary for optimal solution. It is well known that
the problems described by models of covering type are easy to solve even if the number of possible locations exceeds the value of thousand. Nevertheless, location-allocation models, which have been commonly used to describe location problems with a criterion including a relation between served customer and the nearest facility, constitute such mathematical programming problems, which resist to any attempt for fast solution. The computational time is usually several times bigger than the computational time necessary for a covering problem solution and, in addition, if the structure of objective function coefficients does not follow a natural structure of transportation network, then the computational time of solving algorithm exceeds every admissible limit.

This observation has evoked to us an idea to reformulate the location-allocation models of public service system to covering models, what enables us to solve real-sized instances in admissible time. It is clear that the used reformulation may cause loss of preciseness and an attempt of its improvement may bring an increase of the problem complexity and of the computational time as well. In this paper we deal with a newly developed modeling technique and demonstrate its exploitation for quick solving of those public service system design problems, which have been described by location-allocation models so far. We give an overview of the relevant location-allocation models and do their classification in accordance to objectives of the associated problems. We prove usefulness of the suggested modeling technique by numerical experiments to find out a relation between computational time and size of solved instance. The same way will be used to explore a relation between preciseness or associated increase of model complexity and computational time.

2 Classical Covering Models

A considerable part of the public service system design problems is formulated such that customers’ equity in access to the service must be respected to some extent. Then the associated problem is described by the following words: The served customers are aggregated in dwelling places, which form a set \( J \). Each dwelling place \( j \) has a demand on service described by a weight \( b_j \), which is often proportional to population of the dwelling place. The servicing facilities (ambulance vehicles, fire brigades, offices etc.) are allowed to be located only in some prescribed finite set \( I \) of possible locations. There is usually given a number \( p \) of facilities, which are to be located. It is necessary to decide on the facility location so that minimal number of customers is uncovered. A customer is denoted as covered, if it lies in a radius \( r_{\text{max}} \) from some located servicing facility. It must be noted that this approach acccents a rule “not to harm the worst located customers”. Let \( d_{ij} \) denote a distance between a possible location \( i \in I \) and a dwelling place \( j \in J \), then the classical approach to model building of this problem follows.
In the designed model, a decision on ambulance location at place \( i \in I \) is modeled by the zero-one variable \( y_i \in \{0,1\} \), which takes the value of 1, if an ambulance should be located at \( i \) and it takes the value of 0 otherwise. In addition, there are introduced zero-one variables \( x_j \), which express by the values of 0 or 1 whether the demand of dwelling place \( j \in J \) is or is not satisfied. To be able to recognize, whether dwelling place \( j \) is or is not accessible from location \( i \), we introduce zero-one constant \( a_{ij} \) for each pair \( i,j \in I \times J \). The constant \( a_{ij} \) is equal to 1 if and only if the dwelling place \( j \) lies in the radius \( D^{max} \) from the possible location \( i \), i.e. \( d_{ij} \leq D^{max} \). Otherwise, the constant \( a_{ij} \) is equal to 0. Then we can formulate a covering model of the problem:

\[
\text{Minimize} \quad \sum_{j \in J} \; b_j \; x_j, \quad (1)
\]

Subject to

\[
x_j + \sum_{i \in I} \; a_{ij} \; y_i \; \geq \; 1 \quad \text{for} \; j \in J. \quad (2)
\]

\[
\sum_{i \in I} \; y_i \; \leq \; p \quad (3)
\]

\[
x_j \in \{0,1\} \quad \text{for} \; j \in J \quad (4)
\]

\[
y_i \in \{0,1\} \quad \text{for} \; i \in I. \quad (5)
\]

The objective function (1) gives the volume of uncovered demands. The constraints (2) ensure that the variables \( x_j \) are allowed to take the value of 0, if there is at least one ambulance vehicle located in the access time \( D^{max} \) from the dwelling place location \( j \). The constraint (3) puts the limit \( p \) on the number of located vehicles [10].

It must be noted that all of the above mentioned problems belong to the family of integer programming problems or, more precisely, to the family of zero-one integer programming problems. As concerns the solving technique for these problems, they can be theoretically solved by any commercial solver, which contains some general integer programming algorithm, e.g. the branch and bound method or the branch and cut method. These general algorithms are able to solve to optimality real-sized covering problems. Some disadvantage of this approach consists in the fact that the objective function is equal to the number of uncovered customers. This is a very imprecise criterion, because it is not able to evaluate any measure of customers’ affliction. The used model enables only binary distinguishing of customer’s coverage, where a customer is said to be covered by the service, if he lies in radius of a given number kilometers from a service center. We have shown in studies [4], [9] that this binary criterion influences almost negligibly such system parameters as the average time, at which the service of a customer starts. Nevertheless, this parameter is substantial for patient survival, if
the public service system provides customers with emergency medical care. Models, which comprehend size of uncovered customer affliction in their objective functions, are discussed in the next section.

3 Location-Allocation Models

On the contrary to the previous formulation of public service system design problem, in which only number of uncovered customers is taken into account, the location-allocation models accent a point of view of an average customer or an average uncovered customer. The preliminaries of the problem formulation are the same as in the preceding problem but quality criterion of searched solution. The objectives of these problems can be formulated as follows: Decide on locations of facilities so that an average distance of customer from the nearest facility is minimal or do it so that average difference between uncovered customer’s distance from the nearest facility and the prescribed distance $d^{\text{min}}$ is minimal.

Minimize \[
\sum_{i \in I} \sum_{j \in J} c_{ij} z_{ij} \tag{6}
\]

Subject to \[
\sum_{i \in I} z_{ij} = 1 \quad \text{for } j \in J \tag{7}
\]

\[
z_{ij} \leq y_i \quad \text{for } i \in I \text{ and } j \in J \tag{8}
\]

\[
\sum_{i \in I} y_i \leq p \tag{9}
\]

\[
z_{ij} \in \{0,1\} \quad \text{for } i \in I \text{ and } j \in J \tag{10}
\]

\[
y_i \in \{0,1\} \quad \text{for } i \in I. \tag{11}
\]

In this model, a decision on facility location at place $i \in I$ is also modeled by the zero-one variable $y_i \in \{0,1\}$, which takes the value of 1 if a facility should be located at $i$ and it takes the value of 0 otherwise. Furthermore, the allocation variables $z_{ij} \in \{0,1\}$ for each $i \in I$ and $j \in J$ are introduced to assign a dwelling place $j$ to a possible location $i$ ($z_{ij}=1$).

Due to huge number of the variables $z_{ij}$, a commercial solver usually fails, when large-sized instance of the problem (6) – (11) is solved.
Nevertheless the problem can be solved by the approach reported in [4] or [8], where the Lagrangean multiplier is introduced for constraint (9), and the constraint is relaxed. Then the problem takes form of the uncapacitated facility location problem. To solve this problem, the procedure BBDual, see [7] was designed and implemented based on principle presented in [3]. This procedure was embedded into the dichotomy algorithm, which was used to find proper value of the Lagrangean multiplier. This way, some large instances of the case of average distance problem were successfully solved in time of about one hour. Unfortunately, this approach proved to be ineffective, if large instances of the case of average difference were solved. This effect might be caused by small differences between lower bounds of inspected branches in the searching process through searching tree. These small differences might follow from zero contribution of covered customers to the objective function.

4 Reformulation of Location-Allocation to Covering Models

The surprisingly excellent performance of general professional IP-solvers on large-sized instances of covering problem inspired us with an idea to exploit this phenomena as a general approach to the public service system design. Accomplishment of this idea requires reformulation of location-allocation problem to a form of the covering problem even if a loss of preciseness must be paid for it. The core of reformulation consists in making loose the relation between a customer and the facility location, which is the nearest one to the customer. This firm relation is modeled by series of binary variables $z_{ij}$ for $i \in I$ in location-allocation models and enables to determine the distance $d_{ij}$ of the customer $j$ from the nearest located facility $i$. Using the variables $z_{ij}$, the distance $d_{ij}$ can be brought into objective function. We try to estimate this distance unless the nearest facility must be determined.

To his purpose, we partition the range $<0, \max \{ d_{ij} \mid i \in I, j \in J \}>$ of all possible distances of the former location-allocation problem into $r+1$ zones. The zones are separated by finite ascending sequence of values $D_0, D_1, \ldots, D_r$, where $0 < D_0$ and $D_r = \max \{ d_{ij} \mid i \in I, j \in J \}$. We introduce a numbering of these zones so that the zero zone corresponds with the interval $<0, D_0>$, the first zone corresponds with the interval $(D_0, D_1]$ and so on, till the $r$-th zone, which corresponds with interval $(D_r, \max \{ d_{ij} \mid i \in I, j \in J \})$. To make this explanation easier, we denote a width of the $k$-th interval by $e_k$ for $k = 0, \ldots, r$. This way, $e_0 = D_0, e_1 = D_0 - D_1, \ldots, e_r = \max \{ d_{ij} \mid i \in I, j \in J \} - D_r$. Hereafter, we define $D_0 = 0$ and $D_r = \max \{ d_{ij} \mid i \in I, j \in J \}$.

In addition to the zero-one variable $y_j \in \{0,1\}$, which takes the value of 1 if a facility should be located at location $i$, and which takes the value of 0 otherwise, we introduce auxiliary zero-one variable $x_{kj}$ for $k = 1, \ldots, r$, which expresses by the value of 1 that the distance of the dwelling place $j \in J$ from the nearest located facility is bigger than $D_k$ and this variable takes the value of 0 otherwise. Then the expression $e_0 + e_1 x_{j1} + e_2 x_{j2} + e_3 x_{j3} + \ldots + e_r x_{jr}$ is an approximation of $d_{ij}$. It means
that if the distance $d_{ij}$ falls to the interval $(D^k_i, D^{k-1}_i]$, it is estimated by its upper bound $D^{k-1}_i$ with a possible deviation $e_{ij}$.

If we want to estimated a value of $d_{ij}$ for $d_{ij} \geq D_{max}$, then it is sufficient to put $D^i_i = D_{max}$ and to use expression $e_{ij}x_{ij} + e_{ij}x_{ij}^2 + e_{ij}x_{ij}^3 + ... + e_{ij}x_{ij}^p$.

Similarly to the covering model, we introduce zero-one constant $a_{ij}^k$ for each triple $<i, j, k \in \{x\}$ and $r$. The constant $a_{ij}^k$ is equal to 1 if and only if the distance between the dwelling place $j$ and the possible location $i$ is less or equal to $D^k_i$, i.e. $d_{ij} \leq D^k_i$. Otherwise, the constant $a_{ij}^k$ is equal to 0. Then the associated covering-type model can be formulated as follows:

\[
\text{Minimize} \quad \sum_{j \in J} \sum_{k=1}^{r} b_j e_k x_{jk} + \sum_{j \in J} b_j e_0 
\]

\[
\text{Subject to} \quad x_{jk} + \sum_{i \in I} a_{ij}^k y_i \geq 1 \quad \text{for } j \in J \text{ and } k=1, ..., r
\]

\[
\sum_{i \in I} y_i \leq p
\]

\[
x_{jk} \geq 0 \quad j \in J \text{ and } k=1, ..., r
\]

\[
y_i \in \{0,1\} \quad \text{for } i \in I.
\]

In this model, the objective function (12) gives an estimation of the volume of uncovered demands. The constraints (13) ensure that the variables $x_{jk}$ are allowed to take the value of 0, if there is at least one facility located in radius $D^k_i$ from the dwelling place location $j$. The constraint (14) puts the limit $p$ on the number of located facilities.

5 Numerical Experiments

We have suggested a sequence of numerical experiments to find out a relation between computational time and size of solved instances and a relation between preciseness and associated increase of computational time. The experiments have been performed for the problem, which was originally described by the model (6) – (11). Objective function of this original problem corresponds width the average customer distance, i.e. coefficients $c_{ij}$ in the expression (6) are defined as $c_{ij} = b_{ij} d_{ij}$. For benchmark construction, we have used the medical emergency system design problem with the data originating at the Slovak road network with 2916 dwelling places. Each dwelling place represents some aggregation of potential patients. The number of all considered inhabitants was equal to 5.4 millions. In accordance to
the real medical emergency system, the number of 223 different ambulance vehicle locations was taken into consideration as the value \( p \) in the derived instances. We derived three instances of this problem by defining of three different sets of possible locations, which include 449, 961 and 2284 possible locations respectively. This way three different matrices of distances between possible locations and dwelling places were computed from the road network description. The corresponding maximal distances, i.e. \( \max \{d_{ij} \mid i \in I, j \in J\} \), were 289, 289 and 432 kilometers respectively.

Six variants of each individual instance were obtained using the same template, accordingly to which the bordering points \( D^1, D^2, \ldots, D^6 \) were chosen to determine the particular zones. The point \( D^r \) was determined in all cases as the value of 64 km. The variants of one instance differ only in the number \( r \) of the bordering points and in the values of the \( r-1 \) bordering points. The number \( r \) takes values 6, 7, 9, 11, 17 and 33 for variants 1, 2, ..., 6 respectively. The points \( D^k \) for \( k=1, \ldots, r-1 \) were determined in accordance to the formula \( D^k = k e \), where the zone width \( e \) takes values of 6, 5, ..., 1 km for variants 1, 2, ..., 6 respectively. This way, the formed zones cover the range of distances from 0 to 32 km with precision from 6 to 1 km.

The experiments were performed using the general optimization software Xpress-IVE. The associated code were run on a personal computer equipped with the Intel Core 2 6700 processor with parameters: 2.66 GHz and 3 GB RAM. The results of numerical experiments are reported in tables 1, 2 and 3, for the 449, 961 and 2284 possible locations respectively.

In these tables, each column corresponds to one variant of the associated instance and is specified by the value of \( r \).

<table>
<thead>
<tr>
<th>( r )</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>11</th>
<th>17</th>
<th>33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>4,11E+07</td>
<td>3,77E+07</td>
<td>3,31E+07</td>
<td>2,91E+07</td>
<td>2,51E+07</td>
<td>2,09E+07</td>
</tr>
<tr>
<td>Time [s]</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Loc</td>
<td>223</td>
<td>223</td>
<td>225</td>
<td>223</td>
<td>223</td>
<td>223</td>
</tr>
</tbody>
</table>

Table 1. Variants of the first instance of the 449 possible locations.

<table>
<thead>
<tr>
<th>( r )</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>11</th>
<th>17</th>
<th>33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>4,03E+07</td>
<td>3,69E+07</td>
<td>3,25E+07</td>
<td>2,86E+07</td>
<td>2,47E+07</td>
<td>2,05E+07</td>
</tr>
<tr>
<td>Time [s]</td>
<td>20</td>
<td>18</td>
<td>25</td>
<td>17</td>
<td>43</td>
<td>24</td>
</tr>
<tr>
<td>Loc</td>
<td>223</td>
<td>223</td>
<td>223</td>
<td>223</td>
<td>223</td>
<td>223</td>
</tr>
</tbody>
</table>

Table 2. Variants of the second instance of the 961 possible locations.
Table 3. Variants of the third instance of the 2284 possible locations.

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>11</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>3.96E+07</td>
<td>3.63E+07</td>
<td>3.20E+07</td>
<td>2.83E+07</td>
<td>2.44E+07</td>
</tr>
<tr>
<td>Time [s]</td>
<td>54</td>
<td>52</td>
<td>70</td>
<td>70</td>
<td>74</td>
</tr>
<tr>
<td>Loc</td>
<td>223</td>
<td>223</td>
<td>223</td>
<td>223</td>
<td>223</td>
</tr>
</tbody>
</table>

The rows contain the objective function value (Objective) of optimal solution, the computation time in seconds (Time [s]) and the number of located ambulances (Loc). We do not represent the results for r=33 in the third case, because, this instance exceeded available memory of the used computer and IP-solver installation.

Taking into account the total number of 5.4 millions inhabitants and the most precise result, which was achieved for r=33 (or r=17 for the last instance), it can be found that the minimal average customer distance is approximately 4 km. This fact explains the similarity of the results obtained for the three instances (compare rows “Objective” in tables 1, 2 and 3). It also explains the almost regular decrease of the value “Objective” by approximately 4 E+06 for each subsequent variant, where zone width decreases by 1 km and thus it changes from the starting 6 km to 1 km. The relevant distances in resulting solutions range near to 4 km and thus the change of the zone width by 1 km causes the proportional change of the objective function value.

6 Conclusions

The presented results of numerical experiments show that the approximate models constitute a promising approach to large-scaled location problems, which have been resisted to recent attempts to their solving. Even when the critical size of solved problems, i.e. the number of possible locations, exceeds considerably two thousands, the computational times are moderate in comparison with the location-allocation approaches. Nevertheless, it can be observed in the presented tables that the objective function values considerably decrease with increasing number of the equidistant zones. It indicates that an addition of next zones or their rearrangement may bring further improvement of the result preciseness. These possible improvements based on use of the distance distribution will be a topic of the future possible research.
Acknowledgment

This work was supported by grant VEGA 1/3775/06.

References