How does the complexity of a road network affect optimal facility locations?

Author 1: Xiaoyun Zhao
Author 2: Pascal Rebreyend
Author 3: Johan Håkansson
Editor: Hasan Fleyeh
How does the complexity of a road network affect optimal facility locations?

Xiaoyun Zhao*, Pascal Rebreyend, Johan Håkansson

The road network is a necessary component in transportation. It facilitates spatial movements of people and goods, and it also influences the optimal locations of facilities that usually serve as destinations of the movements. To fulfill the transportation needs and to adapt to the facility development, the road network is often organized hierarchically and asymmetrically with various road levels and spatial structures. The complexity of the road network increases along with the increase of road levels and spatial structures. However, location models locate facilities on a given road network, usually the most complex one, and the influence from the complexity of road network in finding optimal locations is not well-studied. This paper aims to investigate how the complexity of a road network affects the optimal facility locations by applying the widely-applied p-median model. The main result indicates that an increase in road network complexity, up to a certain level, can obviously improve the solution, and the complexity beyond that level does not always lead to better solutions. Furthermore, the result is not sensitive to the choice of algorithms. In a specific case study, a detailed sensitivity analysis of algorithm and facility number further provides insight into computation complexity and location problems from intra-urban to inter-urban.

Key Words: Transportation system; Spatial optimization; Location models; Heuristics

1. Introduction

Movements of people and goods take place between origins and destinations on the road network; therefore, the road network is a necessary component in facilitating these spatial movements. A facility usually serves as a destination of the movements, thus, a facility optimally located on the road network is a desired goal in many decision-making processes since it minimizes travel distance (e.g. Carling et al., 2012). However, the road network is often organized hierarchically and asymmetrically with various road levels and spatial structures in order to fulfill transportation needs and to adapt to facility development. The complexity of the road network increases along with the increase of road levels and spatial structures, thus finding the optimal facility locations could become computationally complicated and time-consuming.

* Xiaoyun Zhao is a PhD-student in Micro-data analysis and corresponding author: xzh@du.se, phone: +46 23-778509. Pascal Rebreyend is an associate professor in Computer Science, and Johan Håkansson is a professor in Geography. The authors are all at the School of Technology and Business Studies at Dalarna University, Sweden
In locating facilities, the road network is usually given (Melkote and Daskin, 2001). Peeters and Thomas (1995) claimed that the more complex a road network, the higher probability of identifying a more optimal solution. Following which, many studies always work with a most complex road network and focus on aggregating the demand points and using a simple distance measure, such as the Euclidean distance, to simplify the problem and make it computationally treatable (e.g. Schilling et al., 2000; Francis et al., 2009). In addition, the given road network in these studies is simulated, and the number of facility locations is limited to represent various scenarios. In using a real road network for finding optimal locations in rural areas, Carling et al. (2012, 2015) demonstrate that the level of aggregation needs to be more sophisticated and the distance measure needs to be as accurate as possible in order to avoid suboptimal locations in areas where demands and road network are asymmetrically distributed. Few studies use a real road network with changing complexity, along with the road levels and the spatial structures, in finding optimal facility locations in both intra-urban and inter-urban scenarios.

In a pre-test conducted by Rebreyend et al. (2013), the complexity of the road network is found influential in finding the optimal location in rural areas. However, the trade-off among more optimal solutions, computational complexity and spatial distributions of facilities in intra-urban and inter-urban scenarios has not yet been studied.

This paper aims to investigate how the changes in road network complexity influence the optimal facility locations by applying the widely used p-median model. In addition, to provide further insight into computation complexity and location problems from intra-urban to inter-urban, a detailed sensitivity analysis of four algorithms and various facility numbers is conducted.

The four algorithms are: greedy search (Kuehn and Hamburger, 1963); CPLEX; simulated annealing (Al-Khedhairi, 2008); and imp-Genetic algorithm (Rebreyend et al., 2015). Each algorithm works in a different way in solving the p-median model. The facility numbers vary from 5 to 50 (with a common difference of 5) to simulate intra-urban and inter-urban location distributions in the case of Dalarna province, Sweden. The size of Dalarna, as well as its structure, is similar to the regions of Vermont and New Hampshire in the US.

The entire road network database provided by the official National Road Database (NVDB)¹

¹ www.nvdb.se
in Sweden consists of 2.6 million roads and 30 million nodes, where 0.15 million roads and 1.8 million nodes are distributed in the Dalarna region. The road network in the Dalarna region has a hierarchical structure composed of ten levels of road. The European highways are at the biggest level, representing the simplest road networks, where there is a limited number of candidate nodes for locating facilities. The local and private streets are at the smallest level and represent the most complex road network, with a huge number of candidate nodes for locating facilities.

The remaining part of the paper is organized as follows. In Section 2, we present the \( p \)-median model and some related previous studies that guide us in the model choice. Section 3 describes the data applied in the case study and settings for the sensitivity check. Section 4 illustrates the main results, and Section 5 concludes the paper.

2. The \( p \)-median model

The \( p \)-median model was first introduced by Hakimi (1964), and it has since then been widely used in transportation and location-allocation studies (Farahani, et al., 2012). Given that demand nodes are fixed in the network, the model finds the optimal facility locations that minimize the total travel distance for all demand points to the closest facility. Specifically, the objective function is to minimize the sum of weighted distances between demand points and their respective nearest facilities.

Minimize \( f = \sum_{i=1}^{Q} \sum_{j=1}^{N} w_i \cdot d_{ij} \cdot X_{ij} \) \hspace{1cm} (1)

Subject to:

\[ \sum_{j=1}^{N} X_{ij} = 1, \forall i \in (1,2, ..., Q) \] \hspace{1cm} (2)

\[ \sum_{j=1}^{N} Y_j = p, p \in N \] \hspace{1cm} (3)

Decision variables:

\[ X_{ij} = \begin{cases} 1 & \text{if demand at node } i \text{ is allocated to facility } j \\ 0 & \text{if not} \end{cases} \] \hspace{1cm} (4)

\[ Y_j = \begin{cases} 1 & \text{if a facility is located at candidate site } j \\ 0 & \text{if not} \end{cases} \] \hspace{1cm} (5)

\[ 0 \leq X_{ij} \leq Y_j, \forall i \in (1,2, ..., Q), j \in (1,2, ..., N) \] \hspace{1cm} (6)

\[ X_{ij} \in \{0,1\}, \forall i \in (1,2, ..., Q), j \in (1,2, ..., N) \] \hspace{1cm} (7)
\[ Y_j \in \{0,1\}, \forall j \in (1,2,\ldots,N) \] 

(8)

\[ Q = \text{Total number of demand points in the space of interest} \]

\[ N = \text{Total number of candidates for locating facilities} \]

\[ p = \text{Total number of potential facilities} \]

\[ w_i = \text{The weight associate to each demand node } i \]

\[ d_{ij} = \text{The distance between demand node } i \text{ and potential facility } j \]

The constraint (2) requires that each demand point is assigned to exactly one facility. The constraint (3) ensures that exactly \( p \) facility locations are to be chosen among the \( N \) candidates. The constraint (6) links the location variables and the allocation variables. Finally, constraints (7) and (8), insure that the location variables\( (X) \) and allocation variables\( (Y) \) are binary.

To find the optimal location for \( p \) facilities to fulfill the demand using the \( p \)-median model is NP-hard (Kariv and Hakimi, 1979), and optimal solutions to large problems are difficult to obtain (Al-Khedhairi, 2008). Francis et al. (2009) conducted a literature review of the \( p \)-median model, in which about half of the 40 reviewed articles are studies based on real data. The largest number of candidate nodes in the reviewed studies was some 70,000 candidate nodes. However, \( p \) was relatively small (<20) in those studies. The review also showed that almost all distance measures are Euclidean distance and rectilinear distance.

Many recent studies have been developed to illustrate the greater efficiency of network distance versus Euclidean distance, in the analysis of network-constrained objects or phenomena (Yu et al., 2015). Schilling et al. (2000) examined the Euclidean distance, network distance and a randomly generated network distance. Both Euclidean distance and network distance were found to have high computational efficiency and yield better solution quality. The problem scale in their study was, however, small and they did not study a network with different levels of candidate nodes. Although Euclidean distance is most widely used, the network distance in most cases is more accurate in measuring the travel distance between two points, since the Euclidean distance leads to suboptimal solutions in an unpredictable way. Xie and Yan (2008), and Shiode (2011), found that measuring distance by connecting the straight line between locations could possibly overestimate the clustering tendency of the network. The network distance is more appropriate for spatial phenomena or activities constrained by transportation networks, especially in the field...
of microscopic analysis (Ai et al., 2015). Peeters and Thomas (1995) examined the performance of the $p$-median model in different network topologies by changing the nature of the links. They found that there is a difference in optimal solutions when the links are changed, but they did not check the computational efforts in finding the optimal solutions. Following them, the topologies of the network have been studied, taking the computational effort into account (Peeters, D., et al., 1998; Melkote and Daskin, 2001; Bigotte et al., 2010; Rahmaniani and Ghaderi, 2013). However, on the one hand, these studies are mainly based on fully connected Euclidean and rectilinear networks or networks with varying numbers of radial and rectilinear arcs. On the other hand, the number of candidate nodes and links are limited, which means that the simulation cannot represent the real road network well.

The complexity of a real road network varies along with the road levels and the spatial structures, which cannot be simply represented by nodes and links in topology. Apart from Rebreyend et al. (2013), very few studies have examined the impact of varying the complexities of road network on the optimal facility locations and their spatial distributions.

3. Data and settings for sensitivity check

3.1 Data

Dalarna is a province located in the middle of Sweden with an area of 28,189 km$^2$. The population of Dalarna amounts to 231,934, in December 2015 (Statistic Sweden$^2$). The population is geo-coded and registered on 250 meters by 250 meters squares. The center of each registry square represents one demand point. Each demand point is assigned a weight, corresponding to the population number in that square. There are 15,729 weighted demand points representing the whole population in this region. Figure 1 shows that the population in the studied region is highly asymmetrically distributed, with the majority living in the southeast part.

---

$^2$ [www.scb.se](http://www.scb.se)
Figure 1. Map of the Dalarna region showing 250-by-250 meter squares of inhabitants

Figure 2 shows the complete digitalized representation of the road network in Dalarna, with its hierarchical structure corresponding to 10 road levels. Figure 2(a) shows that the main structure of the road network is composed of levels 0-5, which are European highways, national and regional roads with a total length of 5,479 km. The spatial structure of the road network of these big roads corresponds to the spatial distribution of the population, as illustrated in Figure 1. The European highway forms the simplest road network with a limited number of candidate nodes for locating facilities. While the most complex road network includes all the roads up to the smallest level of the local and private streets, where there is a huge number of candidate nodes for locating facilities. Figure 2(b) shows the small roads at levels 6-9, with a total length of 33,975 km.
Figure 2. (a) Main structure of the road network corresponding to road levels 0-5 in Dalarna, Sweden (b) Local dense road network corresponding to road levels 6-9 in Dalarna, Sweden

Specifically, there are 1,797,939 nodes (the start and the end of the roads), and 1,964,801 road segments in the whole road network in Dalarna. Approximately, one node can be found asymmetrically distributed every 20 meters on the road network. In order to have a valid measure to make comparisons between different network densities, the whole network is used to calculate all the network distances between the demand points and the candidate nodes, which can also ensure that the variation of candidate points from different network densities does not affect the distance measure. However, considering the computation complexity in solving the NP-hard $p$-median model, the whole network with all candidate nodes is not practical for use with most locations problems. A 500 by 500 meter grid aggregation on all density level roads was conducted for mainly two reasons: one reason is to enable all the density levels computationally feasible, and the other reason is to practically stick to the real world situation, since it is very rare to locate a facility every 20 meters.
In order to calculate the network distances, the connectivity of the nodes in the network is first checked. There are 9020 not connected and they are removed. All remaining nodes are tagged to the closest demand origin, or to the closest intersection node. In each grid, at the most, one node is kept as a potential candidate node for optimal location by applying three criteria sequentially. First, we choose the node connects with most road segments. If there are none, we then select the node at the highest density level. If there are none there, we at last pick the node which is closest to the center of the grid. Table 1 summarizes the numbers of nodes according to the complexity of the road network; the number of nodes increases when the complexity of the road network increases by including more levels of roads. There are 452 nodes in the simplest road network, and 67,020 nodes in the most complex road network.

**Table 1.** Number of nodes in different road network complexities that correspond to the included road levels under the grid aggregation level of 500m by 500m in Dalarna.

<table>
<thead>
<tr>
<th>Road Densities</th>
<th>Density Level</th>
<th>Number of Nodes</th>
<th>Length (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>European highway</td>
<td>452</td>
<td>167</td>
</tr>
<tr>
<td>0 – 1</td>
<td>+Primary national highway</td>
<td>1,994</td>
<td>883</td>
</tr>
<tr>
<td>0 – 2</td>
<td>+Primary national road</td>
<td>2,909</td>
<td>1,299</td>
</tr>
<tr>
<td>0 – 3</td>
<td>+Secondary national road</td>
<td>3,926</td>
<td>1,725</td>
</tr>
<tr>
<td>0 – 4</td>
<td>+Primary regional road</td>
<td>6,735</td>
<td>2,923</td>
</tr>
<tr>
<td>0 – 5</td>
<td>+Secondary regional road</td>
<td>12,417</td>
<td>5,479</td>
</tr>
<tr>
<td>0 – 6</td>
<td>+Local arterial road</td>
<td>12,552</td>
<td>5,631</td>
</tr>
<tr>
<td>0 – 7</td>
<td>+Local collector road</td>
<td>20,718</td>
<td>10,964</td>
</tr>
<tr>
<td>0 – 8</td>
<td>+Local rural arterial road</td>
<td>45,336</td>
<td>23,086</td>
</tr>
<tr>
<td>0 – 9</td>
<td>+Local rural collector road</td>
<td>67,020</td>
<td>39,454</td>
</tr>
</tbody>
</table>

As the coordinates of the residents (demand points) do not perfectly coincide with the nodes in the road network, correspondingly, we use the nearest node in the network to represent the location of residents. The average distance between the residents’ node and the nearest network node is 62 meters, which means the approximation does not introduce much difference to the final results.

The travel distance between the demand point and the nearest facility is one key variable that is in the objective function of the $p$-median model, therefore, the distance measure influences the solutions for optimal facility locations. Carling et al. (2012, 2015) investigated, empirically, the consequences of different distance measures for the optimal location of multiple service centers in rural areas. They stated that the shortest travel time or minimal cost along an existing network
intuitively seems to be the most accurate measure for most settings, yet it is infrequently employed. One reason for this is the difficulty and cost associated with collecting data on travel time. Another reason is the complication which arises in modelling the inherent variation in travel time. This paper employs network distance as the distance measure.

### 3.2 Settings for sensitivity check

Heuristics and approximation algorithms are the predominant techniques used for solving the $p$-median location problem, as described and explored in the recent literature (Tansil et al., 1983; Rees, 2006; Mladenović et al., 2007; Varnamkhasti, 2012). Four algorithms that work rather differently in solving the $p$-median model to find the solutions of optimal facility locations are chosen: CPLEX; greedy search (Kuehn and Hamburger, 1963); simulated annealing (Al-Khedhairi, 2008); and imp-Genetic algorithm (Rebreyend et al., 2015).

If the $p$-median model is formulated as a 0-1 binary programming problem it can then be solved by a Mixed Integer Problem (MIP) solver by using a branch-and-cut approach. The CPLEX from IBM is a commonly used software package for solving optimization problems. Following Rebreyend et al. (2015), some parameters of the solver have been tuned in order to adapt CPLEX to work on large problem instances, specifically, removing default computation time limits, allowing intermediate data storage, and tuning branch-and-cut search tree strategies, according to the manual 3.

The standard greedy algorithm for the $p$-median model was studied by Cornuejols et al. (1977). Resende and Werneck (2004) conducted a constructive greedy algorithm to perform the most profitable move among candidates to get the best local minimum in the path. The Greedy algorithm follows the problem-solving heuristic of finding the locally optimal choice at each stage, with the hope of getting a global optimum. It always chooses the optimal choice at the current stage, rather than considering all other conditions to ensure finding a global optimal. This can be characterized as being ‘short sighted’, but it is easy to implement and can achieve acceptable results within a short time.

Simulated annealing (SA) is one commonly used heuristics for solving the $p$-median model. Murray and Church (1996) proposed a basic SA algorithm for the $p$-median model. Levanova and Loresh (2004) studied the SA heuristic that used the 1-interchange neighborhood structure.

---

Carling et al. (2012, 2015), and Han et al. (2013) used tuned SA to solve specific $p$-median model in a real road network context. The basic idea of SA is not only accepting all the better results in the search process, but also accepting some worse results based on certain probabilities. It is simple to implement and can provide high quality solutions to many problems. The performances of SA are sensitive to the values of control parameters. In this study, we employ the same parameters of SA to the same real world network data as employed by Han et al. (2013). The specific parameter settings and the implementation of SA follow Zhao et al. (2013). This includes their dynamic scheme to update the temperature, which allows the algorithm to have setting for efficiency and accuracy, regardless of the size of the input. After testing with various parameter settings, we found that the scheme used to increase and decrease the temperature works well to avoid the search being trapped in the local optimal for a long time, and finally provides satisfying results.

Genetic algorithms (GAs) are another commonly used heuristics that are designed based on mimicking the evolution process. New solutions are based on previous solutions in ways that are reminiscent of the interaction of genes. Most previous studies used a classical string representation, in which each chromosome is represented as a single string of length $p$, embedding the index of the selected facilities or nodes. Thorough treatments of GAs can be found in (Davis, 1987, 1991; Reeves, 1993; Dowsland, 1996). Bozkaya et al. (2002) found GAs could produce solutions that are better than exchange algorithms; however, the convergence is very slow. Alp et al. (2003) proposed a GA which is simpler and produces good solutions faster. Following Correa et al. (2001), Rebreyend et al. (2015) proposed an improved genetic algorithm, called imp-GA, to solve large-scale $p$-median problems. The imp-GA used in this paper follows the description and settings as used by Rebreyend et al. (2015).

The facility numbers vary from 5 to 50 (with a common difference of 5) to simulate the intra-urban to inter-urban location distributions. We run CPLEX and Greedy once, due to the deterministic property of the methods. Since SA is sensitive to its starting point, we randomly select 3 different initial configurations to conduct the SA, and keep the solution with the minimum objective function value (OFV); each run contains 20,000 iterations. As for imp-GA, we set 3 runs with 100 iterations, due to its higher requirement for computation time and computer memory. Max running time is 48 hours for each $p$ and network density level. In this paper, the CPLEX version 12.6, Linux 64 bits is used. All the programs are coded with C and
compiled using GCC version 4.8.2, and they are launched under a system of Linux (Kernel 3.11-2-amd64). The computer has a memory of 32 G, CPU of Intel Core i7-3770.

4. Results

4.1 Sensitivity to complexity levels

Figure 3 shows how worse the optimal solutions found at each complexity level of the road network compared to the best solution found for a selection of $p$ (results for all tested $p$ can be found in Appendix, Figures A1-A3). Specifically, the improvement of the solution is measured by checking the difference between the OFV in the current complexity level, and the optimal OFV found at the same/other level ($\frac{|\text{current solution} - \text{best solution}|}{\text{current solution}} \times 100\%$).

**Figure 3.** Variations from algorithms Greedy search, SA and imp-GA in excess distances (in percent), compared to the best solutions for an increased density level. The x-axis shows the density level. The y-axis shows the difference in percentage between the best solution and the current.

The figure reveals that the solutions found from candidate nodes on the simplest road network...
network, namely, the European highways (level 0), are 85 to 95 percent worse than the optimal solutions that can be found. However, the figure also shows that there is a limit to how complex the road network needs to be to find the best solutions. In general, when the complexity of the road network continues to increase after including road level 5, the solutions barely improve, and even deteriorate.

4.2 Sensitivity to algorithms

As is shown in Figure 3, in general, the algorithms behave in a similar pattern of obvious improvement when the complexity increases from the simplest level 0 to levels up to, and including, 0-5 (ca. 12,500 nodes). When the complexity level increases from 0-5, the OFVs from SA deteriorate, the OFVs from the Greedy search neither improve nor deteriorate, and the OFVs from imp-GA show some very small improvements (also see Table 2).

Table 2. The optimal objective function value found on different road network densities by tested algorithms for various p.

<table>
<thead>
<tr>
<th>p</th>
<th>Greedy</th>
<th>SA</th>
<th>imp-GA</th>
<th>CPLEX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Optimal solution in meters</td>
<td>Included complexity level</td>
<td>Optimal solution in meters</td>
<td>Included complexity level</td>
</tr>
<tr>
<td>5</td>
<td>19685.104</td>
<td>0-5</td>
<td>19715.020</td>
<td>0-2</td>
</tr>
<tr>
<td>10</td>
<td>11134.182</td>
<td>0-5</td>
<td>11259.282</td>
<td>0-2</td>
</tr>
<tr>
<td>15</td>
<td>8316.884</td>
<td>0-5</td>
<td>8625.458</td>
<td>0-4</td>
</tr>
<tr>
<td>20</td>
<td>6698.967</td>
<td>0-5</td>
<td>7094.882</td>
<td>0-6</td>
</tr>
<tr>
<td>25</td>
<td>5804.399</td>
<td>0-5</td>
<td>6155.251</td>
<td>0-5</td>
</tr>
<tr>
<td>30</td>
<td>5122.226</td>
<td>0-5</td>
<td>5465.033</td>
<td>0-4</td>
</tr>
<tr>
<td>35</td>
<td>4708.024</td>
<td>0-5</td>
<td>5013.738</td>
<td>0-6</td>
</tr>
<tr>
<td>40</td>
<td>4349.791</td>
<td>0-5</td>
<td>4611.287</td>
<td>0-4</td>
</tr>
<tr>
<td>45</td>
<td>4049.786</td>
<td>0-5</td>
<td>4306.700</td>
<td>0-4</td>
</tr>
<tr>
<td>50</td>
<td>3796.301</td>
<td>0-5</td>
<td>4066.063</td>
<td>0-4</td>
</tr>
</tbody>
</table>

Table 2 further shows how the best OFVs derived from the four algorithms vary among these algorithms. Note that the CPLEX algorithm that gives the exact solutions fails to produce a result higher than density level 0. The reason for this is that the location problem addressed in this study is too complex, even when the representation of the road network is at the simplest level (level 0). Imp-GA gives the best solutions. However, for the same facility number, the improvements of the OFVs are rather small compared to Greedy and SA, and the largest improvement is only 2.78%, i.e. approximately 122 meters shorter in network travel distance. Table 2 also shows that the
algorithms show a different requirement of network complexity for finding better solutions by including the road levels. However, the differences between the Greedy and imp-GA OFVs are just 1-2 percent. For p=20, or larger, the SA gives an OFV that is 8 to 9 percent worse than the OFV from the imp-GA. Whereas, the computation time and computation effort of imp-GA become tremendously large, so that the results cannot be derived for the most complex of the road networks within the time threshold of 48 hours, which is not the case for Greedy and SA.

4.3 Sensitivity to facility numbers

Based on Figure 3 and Table 2, the changes of facility numbers do not introduce obvious changes either in the complexity level or among the algorithms. Figure 4 visualizes the spatial distribution of the optimal locations found by imp-GA.

![Figure 4. Distribution of optimal locations (p from 5 to 40) found by imp-GA on the study area Dalarna](image)
In general, the optimal locations found when \( p \) is low remain as optimal locations when \( p \) increases. The locations are mostly on the major roads in the road network and correspond very well with the complexity structure, as shown in Figure 2(a). This further indicates that the increase in complexity of the road network above level 0-5 does have a limited influence on the spatial distribution of the optimal facility locations. To make the location-allocation practical in transportation and land-use planning, a location that decreases the travel distance a few hundred meters is not appealing, compared to the risk of losing the attraction of customers due to poor accessibility on a denser road network, especially when the road network is asymmetrically distributed.

5. Concluding discussion

This paper aims to investigate how the changes of road network complexity influence the optimal facility locations by applying the widely used p-median model. In addition, to provide further insights on computation complexity and location problems from intra-urban to inter-urban, a detailed sensitivity analysis of four algorithms and various facility numbers is conducted on a specific case study, namely, Dalarna province, Sweden.

The road network has hierarchical road levels and an asymmetrical structure, where the European highway forms the simplest road network with 452 candidate nodes, while the most complex road network includes all the roads up to the smallest level of the local and private streets, with 67,020 candidate nodes.

The main result indicates that there is a limit to how complex the road network needs to be to find the best solutions. When the complexity of the road network continues to increase after a certain level the solutions barely improve, but even deteriorate. We found that the choice of algorithms and facility number \( p \) has little to no impact on the results, although due to differences in the mechanics of the algorithm, there are some variations among the algorithms. The most complex level of road network (including local streets) is not preferable, especially when the number of the facility is less than 20. The spatial distributions of the optimal locations show that optimal locations found for smaller \( p \) remain as part of the optimal locations when \( p \) increases.

The specific complexity level that is found efficient here might be case-specific, and additional empirical analysis on other cases is needed. In the \( p \)-median model, people are assumed to always choose the closest facility and neglect the multiple purpose travel, or the
heterogeneous preferences on facilities. Therefore, using the gravity $p$-median model (Drezner and Drezner, 2007) would provide more insight into the impact of road network complexity on optimal facility locations for aiding the decision-making in location and transportation planning.

References


Inc., Los Altos, CA.


Appendix

**Figure A1.** Variations from Greedy in excess distances (in percent) compared to the best solutions for an increased density level. The x-axis shows the density level. The y-axis shows the difference in percentage between the best solution and the current solution in accordance to \( \frac{|\text{current solution} - \text{best solution}|}{\text{current solution}} \times 100\% \).
Figure A2. Variations from SA in excess distances (in percent) compared to the best solutions for an increased density level. The x-axis shows the density level. The y-axis shows the difference in percentage between the best solution and the current solution in accordance to \( \frac{|\text{current solution} - \text{best solution}|}{\text{current solution}} \times 100\% \).

Figure A3. Variations from imp-GA in excess distances (in percent) compared to the best solutions for an increased density level. The x-axis shows the density level. The y-axis shows the difference in percentage between the best solution and the current solution in accordance to \( \frac{|\text{current solution} - \text{best solution}|}{\text{current solution}} \times 100\% \).