A comparison of using Euclidean and road distance to solve facility location problems

J Pittayarugsarit and Y Song
Department of System Management, Fukuoka Institute of Technology, Fukuoka, Japan
mjm18201@bene.fit.jp and song@fit.ac.jp

Abstract. To solve any facility location problem, the distance between facility locations and demand nodes is a crucial factor. In previous studies, the distance used most frequently is the Euclidean distance, despite that the road distance should be used when we travel by car or on foot. In this paper, we focus on the difference of the results by using the Euclidean distance and road distance to solve facility location problems. We apply the data of Fukuoka City to the p–median model and the p–center model, and compare the optimal solutions with Euclidean distance and road distance. Numerical results suggest that wide difference exists in some cases.

1. Introduction
Facility location decisions are an important factor in strategic planning for achieving commercial success and competitive advantages. Location decisions also arise in a wide range of public and private sectors. Since the cost of property holding and facility construction are high, facility location or relocation projects are a long-term investment. Thus, poor location decisions may lead to increased costs and decreased competitiveness. In other words, the success or failure of any facility depends in part on the location that was chosen. The decision makers must choose facility location that will not only perform well according to the current circumstance, but will continue to be profitable throughout the facility’s lifetime, even as environmental factors, populations, and market trends change [1]. Hence, for finding beneficial or optimal facility locations, the facility location problem is raised.

For solving the facility location problem, the distance between facility locations and a demand node (e.g., residential communities, industrial estates) is one of the crucial factors. Currently, the distance can be obtained by measuring the length of the Euclidean distance and the road distance between locations. Note that the road distance is always greater than or equal to the Euclidean distance. The Euclidean distance is a length of a straight line between two locations, and the road distance is the length of the road from one location to another location that is used in real life, i.e. the exact traveling distance by car. Even so, on related works, concerning the location problem, the Euclidean distance is used most frequently [2].

For that reason, this study concentrates on the different results obtained between using the Euclidean distance and the road distance to solve the facility location problem. In this paper, the P–median and P–center problem, classical types of the facility location problem, are considered to solve the problem. This study uses Gurobi optimizer to demonstrate the simulation results.

The rest of this paper is divided in three main parts as follows: Section 2 illustrates the methods and formulation of the p–median and p–center problem. Then, the simulation results tested in reallocation are presented in Section 3. Finally, the conclusion and future work are discussed in Section 4.
2. Problem formulation

The \( p \)-median and \( p \)-center problem, both have the problem of selecting the specified number (\( p \)) of optimal location from various candidate locations to locate a facility without considering the area that the facility can cover. However, the method of these two problems is different.

In order to formulate these problems, we set \( i \in I \) to represent a demand node, and the facility can be located at candidate location \( j \in J \). Introducing the following notations:

- \( c_{ij} \) = Distance from demand node \( i \) to facility location \( j \)
- \( p \) = Number of facilities to locate
- \( x_{ij} = \begin{cases} 1, & \text{if demands at node } i \text{ are served by the facility at site } j \\ 0, & \text{otherwise} \end{cases} \)
- \( y_j = \begin{cases} 1, & \text{if the facility is located at candidate location } j \\ 0, & \text{otherwise} \end{cases} \)

2.1. \( p \)-median problem formulation

The method to select the optimal location of the median problem is by minimizing the total distance from each demand node to the nearest facility which can serve those demands.

With the notations mentioned above, the P–median problem model can be formulated as follows:

\[
\begin{align*}
\text{minimize} & & \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \\
\text{subject to} & & \sum_{j \in J} x_{ij} = 1 \quad \forall i \in I \\
& & \sum_{j \in J} y_j = p \\
& & x_{ij} \leq y_j \quad \forall i \in I; j \in J \\
& & x_{ij} \in \{0,1\} \quad \forall i \in I; j \in J \\
& & y_j \in \{0,1\} \quad \forall j \in J 
\end{align*}
\]

The objective function (1) minimizes the total distance between each demand node and the facility which it serves. Constraint (2) states that all of the demand at node \( i \) must be served by the facility at some candidate site \( j \). Constraint (3) ensures that only \( p \) facilities can be located. Constraint (4) confines that demands at node \( i \) cannot be served by the facility at node \( j \) unless the facility is located at candidate site \( j \). Constraints (5) and (6) are the binary constraints.

2.2. \( p \)-center problem formulation

The center problem selects the location by minimizing the maximum distance which is the distance from a demand node to the nearest facility.

With the notations mentioned above, the P–center problem model may be formulated as follows:

\[
\begin{align*}
\text{minimize} & & Z \\
\text{subject to} & & \sum_{j \in J} x_{ij} = 1 \quad \forall i \in I \\
& & \sum_{j \in J} y_j = p \\
& & x_{ij} \leq y_j \quad \forall i \in I; j \in J \\
& & \sum_{j \in J} x_{ij} c_{ij} \leq Z \quad \forall i \in I \\
& & x_{ij} \in \{0,1\} \quad \forall i \in I; j \in J \\
& & y_j \in \{0,1\} \quad \forall j \in J 
\end{align*}
\]
3. Numerical results.

This section presents two simulations of using the road distance and Euclidean distance applied to the $p$-median and $p$-center problems. The optimal solutions of the simulation were performed with Gurobi Optimizer.

The simulations test with a real location in Japan. The first simulation test with Chuo Ward, Fukuoka City in Fukuoka Prefecture, Japan, by considered 12 primary schools as demand nodes and 22 gas stations as candidates of facility locations. The second simulation tests with a larger area, Fukuoka City in Fukuoka Prefecture, with 146 primary schools and 178 gas stations as demand nodes and candidates of facility locations as well. The road and Euclidean distance data are obtained from Google Map [3] and Benricho website [4].

To indicate how the result using the road distance and using the Euclidean distance is different numerically, we considered following these steps:

1. Obtain an optimal location using the road and Euclidean distance, denoted by $R$ and $E$, respectively.
2. Optimal location of using road and Euclidean distances, denoted by $V(R)$ and $V(E)$.
3. Compute the objective function value of Euclidean distance’s optimal location but using road distance, denoted to $V(R)$.
4. Calculate the percentage difference (%Diff) between $V(R)$ and $V_R(E)$.

Note that the objective function value of using road distance is equal to $V(R)$.

The simulation results of applied various lengths in the $p$–median problem and the $p$-center problem is shown in table 1 for Chuo Ward area, and table 2 for Fukuoka City area. The example of located facilities and its covered demand node in Chuo Ward is shown in figure 1 for the $p$-median problem result, and figure 2 for the $p$-center problem result, where it was determined that $p = 3$.

### Table 1. The result of applied the $p$-median and $p$-center problem with Chuo Ward.

| $p$ | Optimal Value (km) | $V_R(E)$ (km) | %Diff | Optimal Value (km) | $V_R(E)$ (km) | %Diff |
|-----|--------------------|--------------|-------|--------------------|--------------|-------|
| 1   | 25.25              | 19.6         |       | 27.35              | 23.75        | 8.30% |
| 2   | 17                 | 12.9         |       | 18.15              | 15.55        | 6.00% |
| 3   | 14                 | 10.4         |       | 15.45              | 12.35        | 10.40%|

### Table 2. The result of applied the $p$-median and $p$-center problem with Fukuoka city.

| $p$ | Optimal Value (km) | $V_R(E)$ (km) | %Diff | Optimal Value (km) | $V_R(E)$ (km) | %Diff |
|-----|--------------------|--------------|-------|--------------------|--------------|-------|
| 1   | 1103.11            | 882.91       |       | 1103.11            | 882.91       | 0.00% |
| 2   | 852.31             | 662.79       |       | 868.71             | 685.71       | 1.92% |
| 3   | 634.79             | 492.62       |       | 671.25             | 570.25       | 5.74% |
| 4   | 572.04             | 439.59       |       | 609.98             | 542.98       | 6.63% |
| 5   | 526.78             | 397.02       |       | 581.41             | 514.41       | 10.37%|
From the results, the percentage difference of using road distance and using Euclidean distance is rather high. Moreover, it not only causes a difference in optimal location but also causes the number of the covered demand nodes of each facility too. To describe the reason why this %Diff is rather high, we decided to compute to the correlation coefficient and the average of %Diff of all road distances and Euclidean distances as shown in table 3. The correlation coefficients are positive and close to 1, meaning that there is high correlation between the two types of distances. That is, if the road distance gets larger the Euclidean distance will get larger too. Furthermore, from the average of %Diff of the Euclidean and road distance, it is indicated that the length of road distance in these simulations, are greater than the length of Euclidean distances by an average of 24.6% for Chuo Ward area, and 21.3% for Fukuoka city area.

| Table 3. The correlation coefficient and the average difference percentage of the road and Euclidean distances in various area. |
|---------------------------------------------------------------|
| Correlation Coefficient | 0.975 | 0.945 |
| Average %Diff. | 24.6% | 21.3% |

4. Conclusion and future work
This paper compared the differences in optimal solutions of facility location problems by using the road distance and the Euclidean distance. We applied the data of Fukuoka City to two classical facility location models: the $p$-median and $p$-center problems. The numerical results have shown that the percentage difference between using road distance and Euclidean distance is rather high in some cases, especially for the $p$-center problem. For future work, we intend to apply this idea to check other location models.

References
[1] Owen S H and Daskin M S 1998 Strategic facility location: A review European Journal of Operational Research 111 423–47
[2] Liu L and Zhou J 2016 A road-like demand location problem based on squared-Euclidean distance 2016 International Conference on Logistics Informatics and Service Sciences 16673386
[3] Google 2019 Google map Online: https://www.google.co.jp
[4] Everybody's knowledge 2019 Measuring the distance between two points Online: https://www.benricho.org
[5] Daskin M S 2013 Network and Discrete Location 2nd Edition (New Jersey: John Wiley & Sons)
[6] Kubo M, Pedroso J P, Muramatsu M and Rais A 2016 Mathematical Optimization: Solving Problems using Gurobi and Python 1st Edition (Japan: Kindaiagaku)