[AM-05-043] Location and Service Area Problems

Abstract

Many facilities exist to provide essential services in a city or region. The service area of a facility refers to a geographical area where the intended service of the facility can be received effectively. Service area delineation varies with the particular service a facility provides. This topic examines two types of service areas, one that can be defined based on a predetermined range such as travel distance/time and another based on the nearest facility available. Relevant location models are introduced to identify the best location(s) of one or multiple facilities to maximize service provision or minimize the system-wide cost. The delineation of service areas and structuring of a location model draw extensively on existing functions in a GIS. The topic represents an important area of GIS&T.

Keywords: coverage modeling, linear programming, location-allocation modeling, network analysis, service area delineation

Author & citation

Tong, D. (2020). Location and Service Area Problems. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2020 Edition), John P. Wilson (ed.).
DOI: 10.22224/gistbok/2020.1.10.

This Topic is also available in the following editions:

- DiBiase, D., DeMers, M., Johnson, A., Kemp, K., Luck, A. T., Plewe, B., and Wentz, E. (2006). Other classic network problems. The Geographic Information Science & Technology Body of Knowledge. Washington, DC: Association of American Geographers.

and

- DiBiase, D., DeMers, M., Johnson, A., Kemp, K., Luck, A. T., Plewe, B., and Wentz, E. (2006). Location-allocation modeling and p-median problems. The Geographic Information Science & Technology Body of Knowledge. Washington, DC: Association of American Geographers.

Explanation

1. Definitions
2. Service Area Delineation
3. Site Selection
4. Uncertainty Associated with Space Abstraction

1. Definitions

Service area: a geographical area where the intended service of a facility is effectively received

Location modeling: approach used to identify the best location(s) of one or multiple facilities for achieving one or multiple goals

LSCP (location set covering problem): location model constructed to find the minimum number of facilities for providing full service coverage to a region

MCLP (maximal covering location problem): location model constructed to find the best locations of a number of facilities for achieving the maximal service coverage

Location-allocation model: location model constructed to find the best locations of a number of facilities so that the total demand weighted travel is minimal

2. Service Area Delineation

Many facilities exist to provide essential services in a city or region. Such services include retailing, communication, banking, transportation, health care, emergency response, etc. The service area of a facility refers to a geographical area where the intended service can be received effectively. For example, the service area of a bus station includes an area within walking distance for people to use the transit service. This topic focuses on two common types of service areas: one can be defined using a predetermined range such as travel distance/time, and another is based on the nearest facility available. The delineation of both service area types can be greatly facilitated in a GIS.

For some facilities, a service range may exist for delineating the associated service area. Depending on the facility involved, the service range may be defined based on spatial proximity. For example, the service area of a warning siren or a cell phone tower can be delineated based on a distance range within which the signals can be received. In these applications, the service area of a facility is essentially a buffer zone with the radius equal to the service range (also see Figure 1a). In this case, the service range is defined based on the Euclidean/straight-line distance. When travel is involved, network distance is often relied upon. Figure 1b draws the service area of a fire station showing all the sites that can be reached within a 5-minute driving distance. Comparing Figure 1a with Figure 1b, while the straight-line distance metric gives circle-shaped service areas, the network distance metric may produce irregularly shaped zones depending on the network structure. More complicated service range and associated area delineation may be needed for some other applications. For example, the service area of a watchtower or a monitor camera may require a viewshed analysis to determine the service /visible area considering all line-of-sight obstructions, including trees as well as other building and non-building structures (also see Figure 1c).

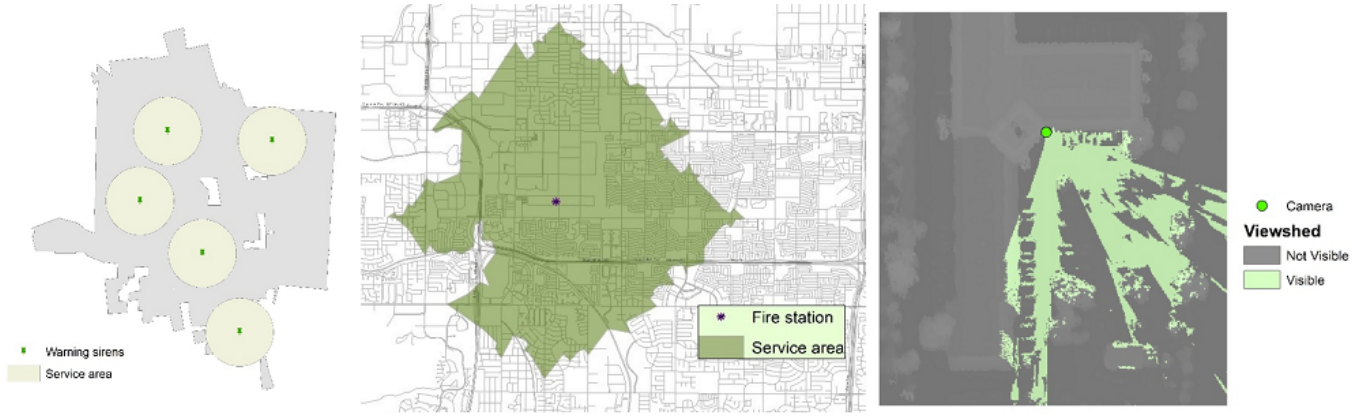


Figure 1a, 1b, and 1c: Service areas delineated based on service ranges: 1a (left): Euclidean-distance, 1b (center) 5-minute drive distance), 1c (right) viewshed. Source: author.

In some applications, there may not exist an obvious service range that can be used to delineate the service area of a facility. For example, people's visit choice of libraries, grocery stores, or postal offices may vary with the spatial distribution of service facilities. In these applications, it is more reasonable to assume that people visit the closest facility available. The associated service area delineation may also vary with the distance metric used. Assuming the Euclidean distance metric, Thiessen polygons can be used to generate the service areas for multiple facilities. In doing so, for any site inside a facility's service area the distance to the facility is smaller than that to any other facilities. When travel is involved, service areas can be drawn using the network distance metric accordingly. Figures 2a and 2b show the service areas delineated for four supercenters using the Euclidean distance and network distance, respectively.

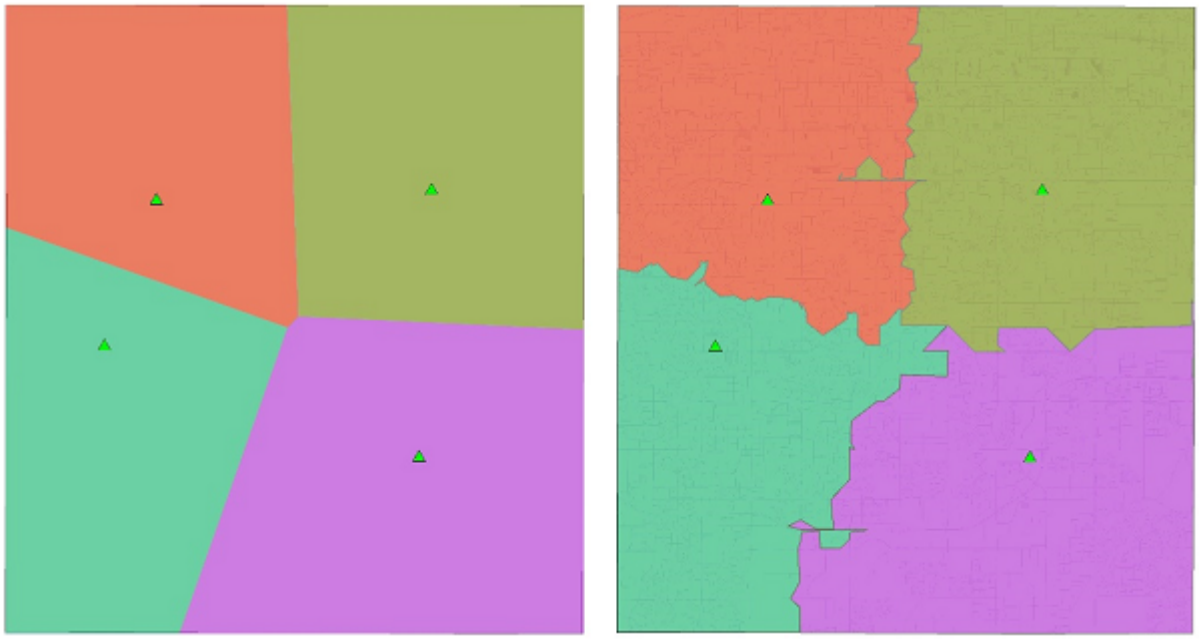


Figure 2a and Figure 2b. Service areas delineated on Euclidean distances (2a, left) and network distances (2b, right). Source: author.

3. Site Selection

Where to locate service facilities to maximize service provision or minimize the system-wide cost has been an important decision for both the private and public sectors. A range of location models have been developed to support these locational decisions. These models vary in their application contexts with different objective functions and constraining conditions. Also see Church and Murray (2009) for a comprehensive discussion on these models. In this topic, coverage problems and location-allocation problems are selected to illustrate the essential role of service areas in determining the best facility locations. The two set of problems differ in the way that services areas are defined and modeled: while in a covering problem service areas can be defined a priori, in a location-allocation problem service areas need to be determined in the modeling framework along with the facility site decisions. In addition, the two types of problems have been widely applied and are also accessible in many commercial GIS packages (Murray et al. 2019).

3.1 Coverage Models

3.1.1 Location Set Covering Problem

For applications where facilities have a fixed service range, coverage models have been widely applied to identify the best siting strategies. Also refer to Church and Murray (2019) for detailed discussion on model advancements and applications. In particular, two types of coverage models have been developed: the location set covering problem (LSCP) and the maximal covering location problem (MCLP). The LSCP aims to find the minimum number of facilities for providing full service coverage to an entire city or region. The problem was first introduced by Toregas et al. (1971). Consider the following notation:

i: index of demand sites

j: index of candidate facility sites

Ni = {j| facility when sited at j can cover/serve demand i}

Here set Ni defines the facility sites that can provide suitable coverage to demand i. In coverage models, Ni is constructed in advance based on the service area of a candidate facility. If the service area of a facility at site j contains i, site j is then included Ni.

The problem involves the locational decisions on whether a candidate facility site (j) is selected,

The LSCP can be formulated as:

Minimize $\sum_j x_j$ (1)

Subject to $\sum_{j \in N_i} x_j \geq 1 \qquad \forall i$ (2)

$x_j \in \{0,1\} \qquad \forall j$ (3)

Objective (1) aims to minimize the number of facilities sited. Constraints (2) ensure that each demand is covered. Constraints (3) impose binary conditions on the decision variables, being either covered or not covered.

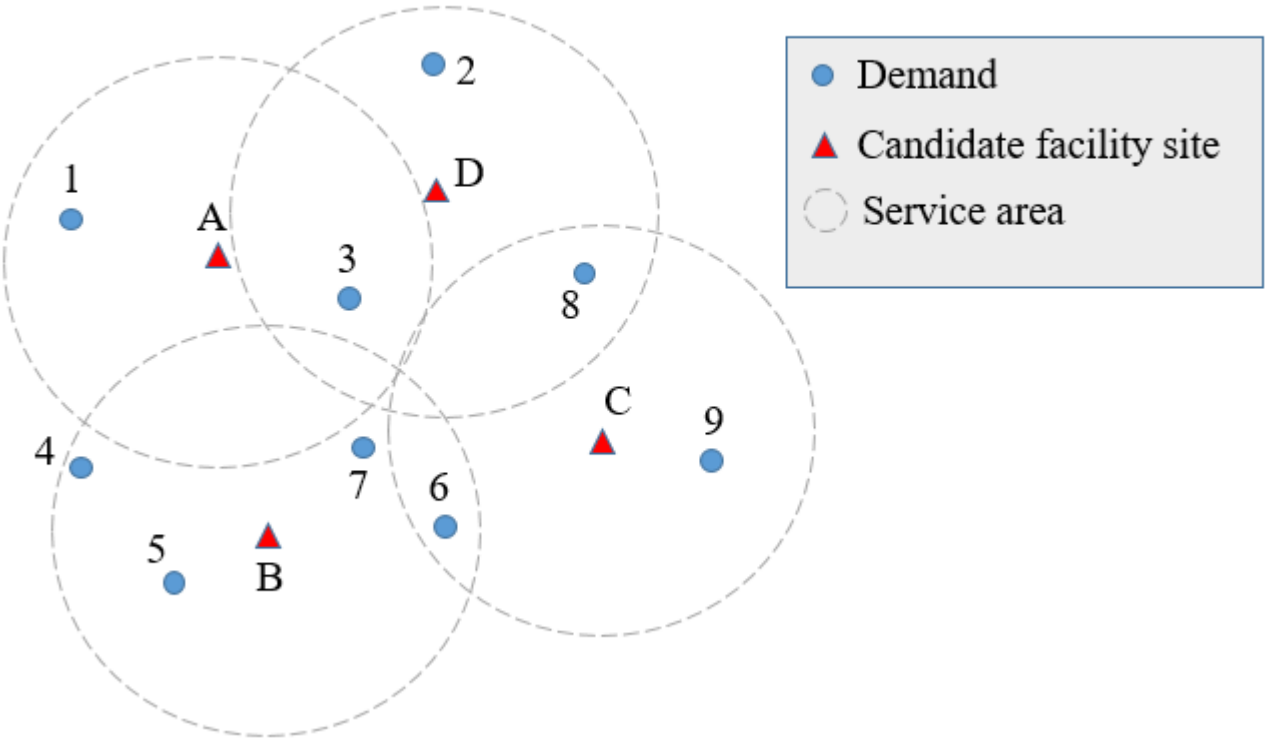


Figure 3. Coverage model illustrations. Source: author.

Figure 3 provides an example to illustrate how the service area of a facility is incorporated in the modeling process of the LSCP. In the example, a total of nine (i = 1, 2, …, 9) demands are to be covered. Given the four candidate facility sites (A, B, C, D), the LSCP finds the minimum number of facilities and the associated locations to provide service to all demands. For each demand i, set Ni identifies all the facilities that can suitably serve i based on the associated service areas. For example, for demand 3 (i = 3), N3={A, D} facilities sited at A and D can serve demand 3 given that the demand is inside the service areas of both facilities. According to Constraints (2), when i = 3, xA+xD ≥1 . That is, at least one of the two candidate locations, A and D, needs to be selected for siting to ensure demand 3 to be covered. Constraints (2) specify that such an evaluation needs to be

performed for each demand i. In doing so, when Constraints (2) are satisfied, the entire region will be served. The LSCP model can be expanded as:

Minimize $x_a + x_b + x_c + x_d$

Subject to

$$x_A \geq 1 \quad \text{(to ensure demand 1 is covered)}$$

$$x_D \geq 1 \quad \text{(to ensure demand 2 is covered)}$$

$$x_A + x_D \geq 1 \quad \text{(to ensure demand 3 is covered)}$$

$$x_B \geq 1 \quad \text{(to ensure demand 4 is covered)}$$

$$x_B \geq 1 \quad \text{(to ensure demand 5 is covered)}$$

$$x_B + x_C \geq 1 \quad \text{(to ensure demand 6 is covered)}$$

$$x_B \geq 1 \quad \text{(to ensure demand 7 is covered)}$$

$$x_C + x_D \geq 1 \quad \text{(to ensure demand 8 is covered)}$$

$$x_C \geq 1 \quad \text{(to ensure demand 9 is covered)}$$

$$x_A, x_B, x_C, x_D \in \{0, 1\}$$

Solving the problem gives the optimal solution: XA = 1, XB = 1, XC = 1, XD = 0. That is, to cover all demand in Figure 3, at least three facilities are required, and these facilities need to be sited at A, B, and C.

3.1.2 Maximal Covering Location Problem

When resources are insufficient for full coverage of a region, it is desirable to locate facilities to cover as much demand as possible. To address this, the maximal covering location problem (MCLP) has been proposed (Church and ReVelle 1971). Assume a limited budget allows only p facilities to be sited in a region. In this case, the limited resource may be insufficient to ensure coverage of the entire region. As a result, in the MCLP there is no guarantee that each demand (i) will be covered. To accurately assess the overall coverage achieved, a decision variable yi is used in the MCLP to track whether a demand is covered or not. With the following additional notation:

wi : weight (e.g., population) associated with demand site i

yi is 1 if demand i is covered, and 0 otherwise

The MCLP can be formulated as:

Maximize

$$\sum_i w_i y_i \quad (4)$$

Subject to

$$\sum_{j \in N_i} x_j \geq y_i \quad \forall i \quad (5)$$

$$\sum_{j \in N_i} x_j = p \quad (6)$$

$$x_j, y_i \in \{0, 1\} \quad \forall i, j \quad (7)$$

Objective (4) aims to maximize the overall coverage. Constraints (5) ensure that each demand is only covered when it is located within the service area of at least one facility. Constraint (6) specifies the number of facilities to be sited. Constraints (7) impose binary conditions on the decision variables.

In Figure 3, if we assume p = 2 and each demand site (i) has equal amount of demand, the MCLP can be expanded as,

Maximize $y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 + y_9$

Subject to

$$x_A \geq y_1$$

$$x_D \geq y_2$$

$$x_A + x_D \geq y_3$$

$$x_B \geq y_4$$

$$x_B \geq y_5$$

$$x_B + x_C \geq y_6$$

$$x_B \geq y_7$$

$$x_C + x_D \geq y_8$$

$$x_C \geq y_9$$

$$x_C \geq y_9$$

$$x_A + x_B + x_C + x_D = 2$$

$$x_A, x_B, x_C, x_D, y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9 \in \{0, 1\}$$

Solving the problem yields the optimal solution, xA = 0, xB = 1 , xC = 0 , xD = 1 with a total coverage of 7 (y1=0 , y2=1 , y3=1 , y4=1 , y5=1 , y6=1 , y7=1 , y8=1 , y9=0).

In both examples, service areas are evaluated using the Euclidean distance metric. For problems involving alternative distances (e.g., network distance) or other types of service range analysis (e.g., viewshed analysis), Ni needs to be evaluated accordingly.

3.2 Location-allocation Problem

Coverage problems are suitable for applications where facilities have a predefined service range so that the service area can be stipulated a priori. For applications where service areas need to be delineated based on nearest facilities, service areas need to be determined along with the siting decisions in the modeling framework. In this topic, the location-allocation problem is used to illustrate the process. The location-allocation model identifies the optimal locations of facilities so that the overall demand-weighted travel is minimal. When candidate facility sites can be pre-determined, the location-allocation problem can be formulated as:

Minimize

$$\sum_i \sum_j w_i d_{ij} z_{ij} \qquad (8)$$

Subject to

$$z_{ij} \leq x_j \quad \forall i,j \tag{9}$$

$$\sum_j z_{ij} = 1 \quad \forall i \tag{10}$$

$$\sum_j x_j = p \tag{11}$$

$$z_{ij} \in \{0,1\} \quad \forall i,j \tag{12}$$

$$x_j \in \{0,1\} \quad \forall j \tag{13}$$

With the following additional notation:

dij is the distance between demand i and facility j

zij is 1 if demand i is assigned to facility j , and 0 otherwise

Objective (8) minimizes the overall travel distance. Constraints (9) ensure that demand can only be assigned to a site when the site is selected. Constraints (10) state that each demand is assigned to only one facility. Similar to constraint (6), constraint (11) specifies the number of facilities to site. Constraints (12) and (13) impose binary conditions on the decision variables. According to constraints (10), a demand is assigned to only one facility. As the model imposes no constraints on the capacity of a facility, the objective function forces each demand to be assigned to the closest facility in order to achieve the minimal overall travel distance. Readers are referred to the topic Location-allocation model for more details about the model construction and problem solution.

The demand-to-facility assignments (zij) allocate demands to each sited facility. Both the allocation assignments along with the facility locational decisions need to be determined in the model when achieving the optimal system-wide efficiency. Figure 4 shows an example of selecting two facilities to serve nine demands. Different from the covering problem example in Figure 3, in a location-allocation problem each facility does not have a predefined service area. Instead, whether a demand can be served by a facility is determined by the allocation assignment. The solution illustrated in Figure 4 includes the selected facility sites and the associated service areas determined by the demand-to-facility assignments.

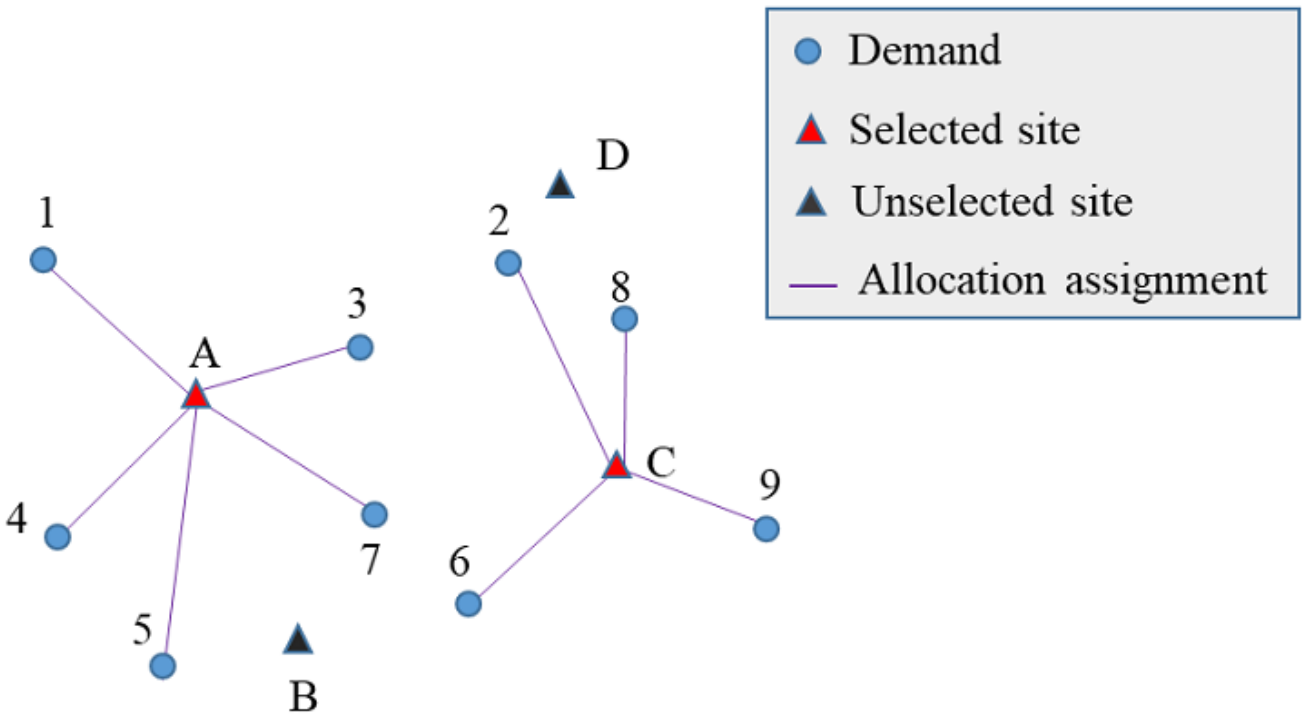


Figure 4. Illustration of the allocation assignments in a location-allocation problem. Source: author.

4. Uncertainty Associated with Space Abstraction

In location problem applications, points have been widely used to abstract and model demand, as is the case in the previous examples. In these applications, a point often represents demand inside a continuous area, such as using the population center to represent all people in a census tract. However, the point abstraction may result in inaccurate characterization of demand to be served. In the case of coverage models, coverage of points does not necessarily guarantee coverage of the associated areas. Figure 5 illustrates such an issue. In Figure 5, according to the point representation demand 3 is completely covered although the associated area is not, and demand 1 is not covered although part of the area receives coverage. To address the issue, studies have explored the use of spatial objects, such as polygons, to represent demand of interest. In doing so, the associated LSCP or MCLP can be directly applied with each demand i corresponding to an area/polygon. Such an approach can also be problematic. This is because, different from points that can be either covered or not covered at all, partial coverage of an area is possible and joint coverage can be provided by multiple facilities. As shown in Figure 5, part of areas 1 and 3 is able to receive coverage although neither demands can be covered completely. As for demand 2, the partial coverage provided by two facilities results in complete coverage of all demands. Refer to Tong and Church (2012) and Murray et al. (2010) for more discussion on uncertainty related to demand abstraction and strategies to solve the issue.

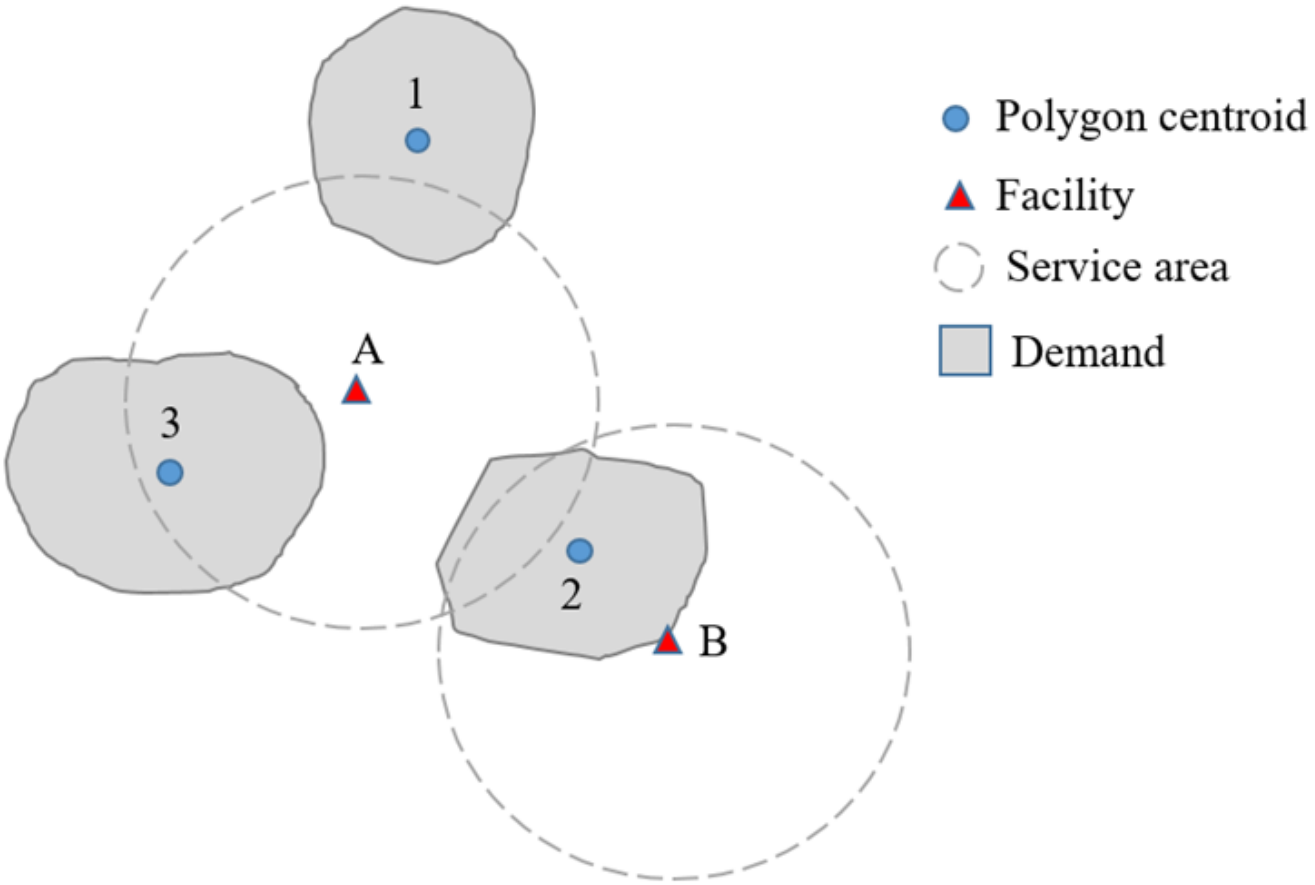


Figure 5. Issue associated with point representation. Source: author.

References

Church R. L. and Murray, A. T. (2009). Business Site Selection, Location Analysis, and GIS. Hoboken, NJ: John Wiley & Sons.

Church R.L. and Murray, A. T. (2019). Covering Models: History, Applications, and Advancements. Berlin: Springer.

Church, R. L., & ReVelle, C. S. (1974). The maximal covering location problem. Papers of the Regional Science Association, 32, 101–118.

Murray A. T., Xu, J., Wang, Z., and Church, R. L. (2019). Commercial GIS location analytics: capabilities and performance. International Journal of Geographical Information Science 33(5): 1106-1130.

Murray, A. T., Tong, D., and Kim, K. (2010). Enhancing Classic Coverage Location Models. International Regional Science Review 33(2): 115-133.

Tong, D. and Church, R.L. (2012). Aggregation in continuous space coverage modeling. International Journal of Geographical Information Science 26(5): 795-816.

Toregas C., Swain, R., ReVelle, C. S., and Bergman, L. (1971). The location of emergency service facilities. Operations Research 19(1):1363-1373.