Article

A Spatial Optimization Approach for Simultaneously Districting Precincts and Locating Polling Places

Kamyoung Kim

Department of Geography Education, Kyungpook National University, Daegu 41566, Korea;
kamyoungkim@knu.ac.kr; Tel.: +82-53-950-5861

Abstract: Voting is the most basic form of political participation. The agencies that are responsible for voting must delineate precincts and designate a polling place for each precinct. This spatial decision-making requires a strategic approach for several reasons. First, changes in the location of polling places induce transportation and search costs from the perspective of voters. Second, improving accessibility to polling places can increase turnout. Third, differences in the population sizes of precincts may produce biased voting results. Spatial optimization approaches can be a strategic method for delimiting precincts and siting polling places. The purpose of this paper is to develop a spatial optimization model, namely, the capacitated double p-median problem with preference (CDPMP-P), which simultaneously delimits boundaries of precincts and selects potential facilities in terms of mixed integer programming (MIP). The CDPMP-P explicitly includes realistic requirements, such as population balance, the spatial continuity of precincts, the preferences of potential facilities where polling places can be installed, and the possibility of allocating multiple polling places in one facility.

Keywords: accessibility; population balance; preference; spatial contiguity; double location-allocation process; capacitated double p-median problem with preference (CDPMP-P)

1. Introduction

Voting is the most basic form of political participation in current democratic societies. Every time a presidential or parliamentary election or a referendum occurs, a government agency that is in charge of election administration and management delineates precincts and establishes polling places. Polling places are physical locations where eligible voters cast their votes, and precincts are the smallest electoral unit [1], which are geographically continuous areas for grouping residents to assign them to a polling place [2] (p. 138). The terms *precinct* and *constituency* are used to refer to electoral districts. To clarify the meaning of a precinct, it is necessary to distinguish the two terms. A constituency (*seongeo-gu* in Korea) is referred to as a unit that elects a representative(s), and a precinct (*tupyo-gu* in Korea) as a subdivision of planning units, such as a county, town, city, or ward, to manage and operate voting. Generally, a constituency consists of a number of precincts. In Korea, voters on Election Day must vote at the designated polling places based on their residential address, while those who opt to pre-vote can vote at any pre-polling place irrespective of their residence. The former is a traditional precinct-voting method, while the latter is a non-precinct-voting approach [3,4]. The agency that is responsible for the electoral administration of jurisdictions delineates precincts based on the voters' residence and then notifies them individually where they will vote on Election Day. Because population changes continually over time, the process of designating precincts and polling places should be conducted whenever there is an election. Thus, partitioning a continuous geographical voting zone based on voters' residence and informing voters where they can vote are common tasks in electoral administration.

Precincts and polling places are used temporarily for the convenience of voters and electoral administration on Election Day. Delineating precincts differs in many ways from political districting problems. Generally, population equality, contiguity, and compactness are considered key modeling concerns when districting constituencies. This type of political districting may have the problem of gerrymandering, which adjusts districts for specific political purposes, and political interests are sensitive to the outcome of redistricting [5]. The reason why politically sensitive and societal issues arise is that the results of elections can vary depending on how many and what types of people are included in constituencies; if the population imbalance among constituencies is severe, the equivalence of the voting might be damaged. Unlike constituency redistricting, which can cause gerrymandering due to political interests, delineating precincts and determining polling places target the convenience of voting.

Why should we strategically approach the delineation of precincts and the siting of polling places? Polling places may induce two types of costs for voters. One is transportation costs that are incurred by moving from voters' residence to the designated polling place, and the other is the search costs arising from the redistricting of precincts and the repositioning of polling places [6]. Several studies showed that the distance to polling places affects turnout rate [3,6–9]. Furthermore, the turnout rate can be increased through the relocation of polling places [10]. In addition, some claimed that the population size of the precincts (especially larger ones) could produce biased results [11]. For these reasons, it is necessary to partition precincts with a suitable population and increase access to polling places. Spatial optimization approaches can significantly contribute to these spatial decisions [12–14]. On the other hand, with many countries reducing the number of polling places to cut election costs [6,9,13,14], spatial optimization approaches can provide the alternative solutions to minimize accessibility deterioration.

Despite the practical usefulness of spatial optimization approaches, not many optimization studies that relate to demarcating precincts and locating polling places have been conducted. References [13] and [14] developed a multi-objective genetic algorithm that can reduce the number of polling places while increasing turnout. However, these authors did not present a mixed integer programming optimization model. Reference [15] formulated a generalized assignment problem [16] that minimizes the cost of moving to polling places while balancing the population between precincts. This model finds the optimal precinct boundaries by assigning demand spatial units to a predetermined number of polling places to minimize the population-weighted distance under capacity constraints. This is an allocation model because the location and number of polling places have already been determined. On the other hand, [12] proposed a capacitated plant location problem that allows multiple polling places at the same site. In this model, multiple polling places can be installed on one site by adding a new index to track multiple places on it. In this study, the authors pre-defined the number of facilities that can be installed per site. This model can determine the locations of the polling places, but may not demarcate a spatially continuous precinct. For example, suppose that two polling places are located on a site. Because the distance between the two polling places and the allocated demand are the same, even if the objective function minimizes the sum of distance, the demand that is allocated to the two polling places may be spatially mixed. Similarly, [17] suggested a spatial optimization model that allows the assignment of multiple polling places to the same site by relaxing the binary integer requirement of the location decision variables in a capacitated *p*-median problem. The authors defined an objective function with a utility cost, which is a function with a combination of the traveling cost and the preference of potential sites. Their computational results showed that one can co-locate multiple polling places on the same site by emphasizing a preference factor in the objective function. However, this approach also cannot delineate geographically continuous precincts using allocation information on the same site. This paper proposes a way to overcome the limitations of previous studies.

The purpose of this paper is to develop a spatial optimization model for simultaneously redistricting precincts and locating polling places in terms of mixed integer programming (MIP). During modeling, the population balance, spatial contiguity, preference (or attractiveness as a polling

place) of potential facilities where polling places can be installed, and the possibility of installing multiple polling places on the same facility are explicitly considered.

2. Modeling Concerns

This section describes the aforementioned considerations in more detail. Related to the location of polling places, the most important consideration might be the ease of finding and accessing polling places. Many studies empirically demonstrated that accessibility to polling places can significantly affect turnout or voter participation [3,6–9]. Improving the accessibility to polling places is important both prescriptively and practically. Prescriptively, a lack of accessibility to polling places may be a type of suffrage violation. Improved accessibility to polling places might provide more equal political participation opportunities to voters. Practically, delimiting precincts and installing polling places are common election-management tasks for all elections, and factors such as accessibility can be easily reflected in this decision-making process [9]. If these practical actions can increase turnout or improve the convenience of voters, they become very valuable.

The second modeling concern is population balance among polling places (or precincts). The balance of population between precincts might be important when redistricting precincts because the voting time on Election Day and the manpower that is available for polling places are restricted. In addition, the population size of the precincts can cause large precinct bias, which creates an advantage for specific candidates as the precinct size increases above a designated total precinct vote count [11].

The third consideration is the preference of facilities where polling places can be installed. Many studies combined consumers' preferences for facilities with optimization models, such as p-median problems [18–20]. Generally, optimization models with consumers' preferences were defined as bi-level models that consisted of two components; that is, minimizing the location cost, which is estimated based on distance, and optimizing the consumers' preferences. The preference for facilities is predefined in an order form for each individual consumer. When finding solutions for bi-level optimization models, the preference and distance interact with each other; but, these two values are measured independently. In practice, distance to a facility, alongside the characteristics of the facility, such as size, affects the preference of the facility [21]. To reflect this reality, a utility cost is defined in this study as a function of distance and the preference, which is evaluated from the characteristics of the facilities where polling places are installed [17]. The number of precincts depends on the size of the population in the jurisdiction. Because the population continually changes, the number of precincts and the location of the polling places also may variate each time. In most cases, the same sites are often used as polling places, but some sites can be designated as a new polling place for each election. In terms of election administration, designating public facilities as polling places is easier than private facilities. In addition, it is possible to install multiple polling places at larger facilities. Therefore, larger public facilities may be preferred. Meanwhile, voters would prefer to vote at a place they had previously voted, at a known facility, or at a geographically close facility.

Fourth, installing multiple polling places at one facility can be a modeling issue. In Korea, several spatially separated spaces of a facility may commonly be used as polling places for different precincts, which is similar to Italy [12]. Several approaches exist to model the installation of multiple polling places at one facility or one site. The first method is to mitigate the integrality constraints of the location decision variables into a positive integer in optimization models [17]. In this method, however, if multiple polling places are installed on the same site, we cannot know to which precinct the assigned voters into the facility belong. The second method is to add a new index that points to different places in a facility, such as in [12]. This method can distinguish demand that is allocated to different polling places on the same site. However, the demand that is allocated to a polling place may not be spatially contiguous. Furthermore, the addition of a new index increases the complexity of a model ($n \times t \times m$, where n is the index indicating potential facilities, t is the index indicating the different places of a facility, and m is the index indicating demand). Thus, the new index makes finding solutions to practical problems with MIP more difficult. In addition, adding indices to track different places in one facility does not guarantee that multiple polling places will be installed on a

site. If demand is dispersed and potential facilities are distributed more evenly over a study area, the objective function that minimizes distance and fixed costs will likely not result in multiple polling places being located on the same site. Conversely, if the demand is concentrated in a specific area and the number of potential facilities is small, the co-location of the polling places may be possible. Therefore, an alternative approach is required to locate multiple polling places on a site, regardless of the distribution of demand and potential facilities.

The last consideration is the spatial contiguity of the delineated precincts. If demand is represented as points, one can delimit continuous precincts by grouping demand points that are allocated into the same polling places. In this case, the partitioned precincts are relatively free from contiguity constraints. Because the objective function minimizes the sum of the distance to polling places, one can distinguish the allocated demand spatially and exclusively. However, demand is generally represented as areal spatial units in districting problems, and these spatial units are grouped into larger districts under predefined conditions [22–24]. When demand is represented as polygons, the distance from the centroid of a polygon to the center of the districts or facilities is measured, rather than the distance between the polygons themselves. Therefore, depending on the shape or arrangement of the polygons, the allocated polygons into a district can be discontinuous [22,24–26]. Additionally, the basic spatial units may be allocated to more distant facilities than the closest one when population balance constraints are included. In this case, discontinuous districts are more likely to be delineated. In this study, geographically contiguous precincts are delimited by using a flow-based contiguity constraint that enforces unit flows from individual spatial units to the centers of districts [22,24].

3. Capacitated Double *p*-median Problem with Preference

To model the five aforementioned considerations, this paper proposes a new model, called the capacitated double *p*-median problem with preference (CDPMP-P), by extending the capacitated *p*-median problem [27]. This model is a type of districting problem because it includes the delineation of precincts. Demand is represented as polygons and potential facilities are represented as points. Unlike *p*-median problems [17,28], general assignment problems [15], or capacitated plant location problems [12], which have a single location-allocation process, the CDPMP-P consists of double location-allocation processes. The first process involves selecting a given number of basic spatial units as centers and allocating basic spatial units to the selected centers, and the second process comprises selecting polling facilities and allocating the selected centers to them. The latter process of determining candidate facilities and assigning delineated districts is the opposite of the regionally constrained *p*-median problem, which allocates facilities while complying with the minimum and maximum number of facilities in a predefined area [29–31]. Figure 1 illustrates the double location-allocation process of the CDPMP-P. In this figure, three precincts were delineated by assigning eleven basic spatial units to three centers (*x3*, *x4*, and *x8*). In the second phase, the three centers were assigned to two selected facilities (*y2* and *y5*). The basic spatial units are assigned to selected centers to minimize the sum of the population-weighted distance (first term in the objective function, see Equation (1)), so the selected units are the population centers of delineated precincts. These population centers are, in turn, assigned to facilities to minimize the sum of the population-weighted distance (second term in the objective function, see Equation (1)). Thus, access to polling places is indirectly modeled. Directly assigning basic spatial units to facilities to model voters' accessibility to polling places may be desirable, but several practical reasons exist for adopting this indirect approach.
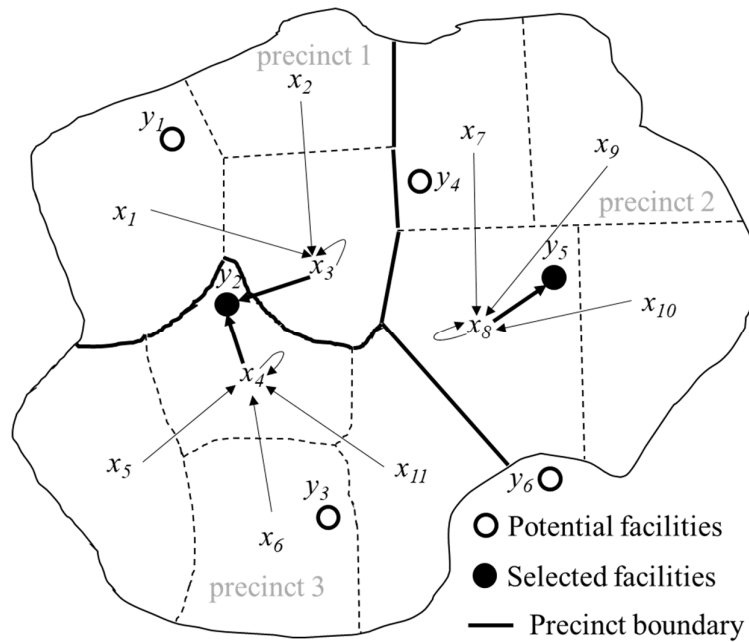
Figure 1. Double location-allocation process of the capacitated double *p*-median problem with preference (CDPMP-P).

The first reason to introduce the double location-allocation process is to reflect realistic needs. According to [32], if no suitable places are available within a precinct, it is possible to site a polling place in an adjacent precinct. If demand is directly allocated to polling places or potential facilities, selected polling places must be located in precincts because an objective function minimizes the sum of distance in location-allocation problems. However, no adequate facilities may be available to install a polling place in a precinct. In such cases, polling places must inevitably be located outside precincts. The second reason is to obtain spatially separated precincts that correspond to each polling place when multiple polling places are installed in one facility.

Here, we must explain the terms and notation that are used to clearly describe the proposed model. As mentioned earlier, the *center* is the population center of a precinct, which serves as a seed for grouping basic spatial units. The *sink*, which is used to model the continuity of the delineated precincts, is an imaginary point where unit flows that occur between adjacent basic spatial units within the same precinct are gathered. A *facility* is a location that has a physical space in which to set up polling places, such as community centers, schools, and gyms. On the other hand, in districting problems with continuity constraints, *i*, *j*, and *k* indices are needed to represent the adjacency relationship between the basic spatial units and to track the basic spatial units that are selected as the centers of the districts [22,24]. The notation that is required to mathematically define the CDPMP-P is as follows:

i, j, k = index of basic spatial units, $i, j, k = 1, \ldots, n$,

l = index of potential facilities, $l = 1, \ldots, m$,

a_i = demand of basic spatial unit i,

d_{ik} = distance between basic spatial unit i and k selected as the center of a precinct,

d_{kl} = distance between basic spatial unit selected as a center k and a facility l,

pr_l = preference of facility l,

β = preference impact factor,

p = number of precincts to be delineated,

C_{min} = lower bound of population,

$N_i = \{j | c_{ij} = 1\}$, a set of spatial units that are adjacent to i,

$c_{ij} = \begin{cases} 1, \text{basic spatial units } i \text{ and } j \text{ are adjacent} \\ \quad\quad 0, \text{otherwise} \end{cases}$,

f_{ijk} = the amount of unit flow from i to j in a precinct centered on k,

$x_{ik} = \begin{cases} 1, \text{if spatial unit } i \text{ is assigned to center } k \\ \quad\quad 0, \text{otherwise} \end{cases}$,

$w_{ik} = \begin{cases} 1, \text{if spatial unit } i \text{ is selected as a sink of precinct centered on spatial unit } k \\ \quad\quad 0, \text{otherwise} \end{cases}$,

$y_{ikl} = \begin{cases} 1, \text{if precinct with sink } i \text{ and center } k \text{ is assigned to facility } l \\ \quad\quad 0, \text{otherwise} \end{cases}$,

$z_l = \begin{cases} 1, \text{if facility } l \text{ is selected} \\ \quad 0, \text{otherwise} \end{cases}$.

Capacitated double p-median problem with preference (CDPMP-P):

$$Minimize\ Z = \sum_i \sum_k a_i d_{ik} x_{ik} + \sum_i \sum_k \sum_l \frac{d_{kl}}{pr_l^\beta} y_{ikl} \tag{1}$$

Subject to

$$\sum_k x_{ik} = 1, \forall i \tag{2}$$

$$\sum_i \sum_k w_{ik} = p, \tag{3}$$

$$\sum_j a_j x_{jk} \geq C_{min} w_{ik}, \forall i, k \tag{4}$$

$$\sum_l y_{ikl} = w_{ik}, \forall i, k \tag{5}$$

$$y_{ikl} \leq z_l, \forall i, k, l \tag{6}$$

$$f_{ijk} \leq x_{ik}(n - p), \forall i, j \in N_i, k \tag{7}$$

$$f_{ijk} \leq x_{jk}(n - p), \forall i, j \in N_i, k \tag{8}$$

$$\sum_{j \in N_i} f_{ijk} - \sum_{j \in N_i} f_{jik} \geq x_{ik} - (n - p)w_{ik}, \forall i, k \tag{9}$$

$$w_{ik} = \{0, 1\}, \forall i, k \tag{10}$$

$$x_{ik} = \{0, 1\}, \forall i, k \tag{11}$$

$$y_{ikl} = \{0, 1\}, \forall i, k, l \tag{12}$$

$$z_l = \{0, 1\}, \forall l \tag{13}$$

$$f_{ijk} \geq 0, \forall i, j \in N_i, k \tag{14}$$

The objective function (Equation (1)) consists of two terms: the first term minimizes the sum of the population-weighted distance from the basic spatial units to the centers of the precincts, and the second term minimizes the sum of the utility cost when allocating the selected centers to potential facilities. The utility cost is defined as a function of distance between the center of a precinct and the facilities and the preference of a facility as a polling place [17].

If β equals 0, then double allocations proceed only based on distance. If $\beta > 0$, the preference is evaluated only by the characteristics of facilities when assigning the centers of the precincts to potential facilities. When strongly emphasizing the preference in this utility cost ($\beta > 1$), the likelihood of choosing more preferred facilities is increased, and multiple precinct centers can be assigned because the denominator of the second term is much smaller. Because the number of basic spatial units is larger than those of the precinct centers and potential facilities (the contribution from the first term to the objective function is much greater than that from the second term), the emphasis on the preference affects the selection of potential facilities but has relatively little effect on the locations of the centers. Constraint (2) means that all basic spatial units should be allocated to only one center. Constraint (3) enforces p precincts to be delineated. Constraint (4) limits the minimum population size of the delineated precincts. According to Constraint (3), w_{ik} is only 1 for the given number p and 0 for all other cases. Thus, w_{ik} is a subset of x_{ik} ($x_{ik} \geq w_{ik}$); that is, w_{ik} is 1 out of x_{ik} = 1. If we limit the upper bound of the population in Constraint (4), then the right side will often have a value greater than 0, while the left side will be mostly 0. Therefore, no solutions satisfy this type of constraint. Strictly speaking, the lower bound, similar to Constraint (4), is a lower threshold of the central place theory [33] rather than capacity. Constraint (5) ensures that all the selected centers should be assigned to one potential facility; that is, only the basic spatial unit i that is selected as the center of the precinct k ($w_{ik} = 1$) should be assigned to one potential facility l once. Constraint (6) enables us to assign the centers only to the selected facilities. Constraints (7), (8), and (9) are contiguity requirements that enforce unit flows between the spatial units i and j, which are adjacent to each other and assigned to the same precinct k [22,24–26]. Constraints (10), (11), (12), and (13) impose

integer restrictions on the location and allocation decision variables. Finally, Constraint (14) means that f_{ijk} is a non-negative decision variable.

Generally, MIP models with contiguity constraints are computationally intractable [22,24,26]. y_{ikl} and f_{ijk} can each produce $n^2 \times m$ variables, so finding the exact solution of a problem instance with MIP is difficult. If the contiguity constraints are removed from the CDPMP-P, the complexity of the model is greatly reduced. After removing the contiguity constraints, we can re-state the notation, decision variables, and the reduced CDPMP-P as follows:

i, k = index of basic spatial units, $i, k = 1, \dots, n$,

l = index of potential facilities, $l = 1, \dots, m$,

a_i = population of basic spatial unit i,

d_{ik} = distance between basic spatial units i and k selected as the center of a precinct,

d_{kl} = distance between basic spatial unit selected as a center k and facility l,

pr_l = preference of facility l,

β = preference impact factor,

p = number of precincts to be delineated,

C_{min} = lower bound of population,

$x_{ik} = \begin{cases} 1, \text{if spatial unit } i \text{ is assigned to basic spatial unit } k \text{ selected as a center} \\ 0, \text{otherwise} \end{cases}$,

$w_k = \begin{cases} 1, \text{if spatial unit } k \text{ is selected as a center} \\ 0, \text{otherwise} \end{cases}$,

$y_{kl} = \begin{cases} 1, \text{if selected center } k \text{ is assigned to facility } l \\ 0, \text{otherwise} \end{cases}$,

$z_l = \begin{cases} 1, \text{if facility } l \text{ is selected} \\ 0, \text{otherwise} \end{cases}$.

Reduced CDPMP-P:

$$Minimize\ Z = \sum_i \sum_k a_i d_{ik} x_{ik} + \sum_k \sum_l \frac{d_{kl}}{pr_l^\beta} y_{kl} \tag{15}$$

Subject to

$$\sum_k x_{ik} = 1, \forall i \tag{16}$$
$$x_{ik} \leq w_k, \forall i, k \tag{17}$$
$$\sum_i a_i x_{ik} \geq C_{min} w_k, \forall k \tag{18}$$
$$\sum_k w_k = p, \tag{19}$$
$$\sum_l y_{kl} = w_k, \forall k \tag{20}$$
$$y_{kl} \leq z_l, \forall k, l \tag{21}$$
$$w_k = \{0, 1\}, \forall k \tag{22}$$
$$x_{ik} = \{0, 1\}, \forall i, k \tag{23}$$
$$y_{kl} = \{0, 1\}, \forall k, l \tag{24}$$
$$z_l = \{0, 1\}, \forall l \tag{25}$$

In the reduced CDPMP-P, the meaning of the objective function and constraints is the same as that of the full version except for Constraint (17). These constraints mean that basic spatial units can only be assigned to the central spatial unit of a precinct. By excluding the contiguity constraints, the total number of constraints is greatly reduced from $(2n^3 + n^2(m + 3) + n + 1)$ to $(n^2 + n(m + 3) + 1)$.

4. Computational Results

The proposed models were evaluated by using two datasets. To test the full version of the CDPMP-P, we used a very small dataset with 13 basic spatial units and six potential facilities. To evaluate the reduced CDPMP-P, we used a relatively large dataset with 200 basic spatial units and 38 potential facilities. The minimum number of voters in the basic spatial units was zero, the maximum was 563, and the average was 106. The minimum area of the basic spatial units was 0.002 km², the maximum was 0.052 km², and the average was 0.008 km². The Euclidean distance between the centroids of the basic spatial units and between the basic spatial units and potential facilities were used for analysis. Using network distance may be more realistic. However, the spatial extent of the case area was small (the maximum distance between the basic spatial units was 2990 m and the

average distance was 751 m), demand was aggregated into polygons, and most people walk to polling places. For these reasons, alongside the ease of calculation, we used the linear distance. Figure 2 shows the distribution of the voters and the locations of potential facilities within an administrative unit (Seokyo-dong, Mapo-gu, Seoul) that can be itself a single precinct or can be divided into several precincts in Korea. In this figure, the preference of potential facilities was evaluated by considering their histories as polling places in previous elections, the sizes of facilities, and whether facilities were public. The preferences in this map were normalized to values between one and two. As of April 1, 2016, 24,128 voters were in this area. In previous elections, the area was divided into seven precincts and seven polling places were established. The spatial distribution of the voters was represented by reallocating the voters of each precinct proportionally to the number of households in basic spatial units. This approach is similar to the dasymetric technique of estimating the values of small spatial units [34].
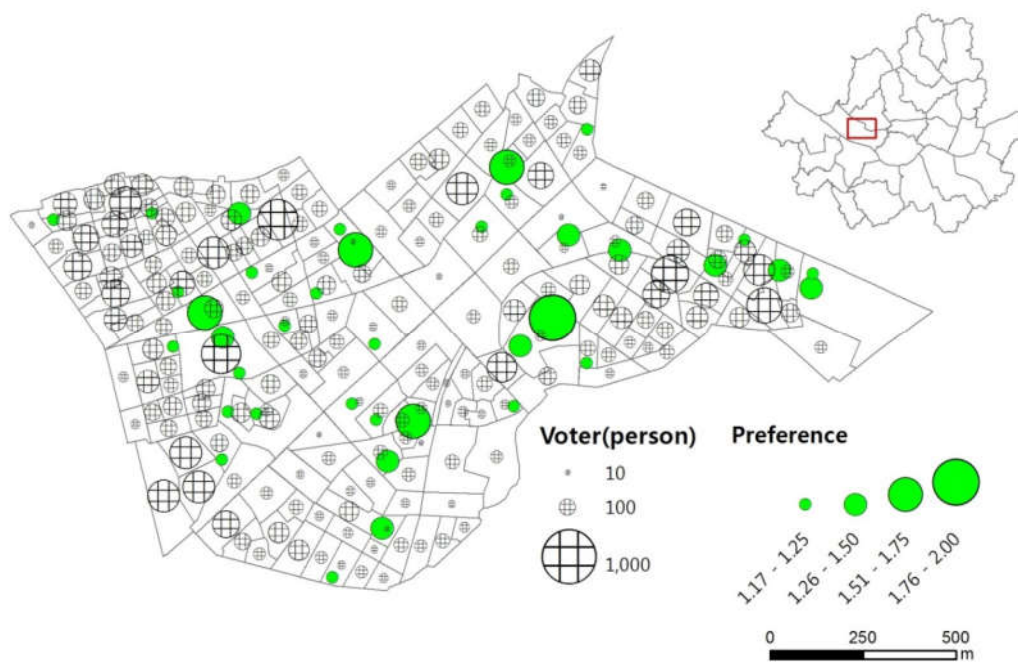

Figure 2. Spatial distribution of the voters and potential facilities.

ILOG CPLEX Optimization Studio (version 12.5.1) was used to find the optimal solutions to the problem instances. The solver was run on a computer with the following specifications: Windows 7 Enterprise K (64-bit OS), Intel(R) Core(TM) i5-6500 CPU @ 3.20 GHz, and 8.00 GB RAM. ESRI ArcGIS (version 10.2) was used to spatially represent demand, generate input data for the proposed spatial optimization, and visualize results.

First, let us examine why contiguity constraints are required in districting problems, such as precinct delineation. Figure 3a,b show three delineated precincts from the small dataset when using the reduced and full versions of the CDPMP-P, respectively. In Figure 3a, the delineated precincts from the reduced model were spatially non-contiguous. The Euclidean distance from Spatial Unit 8 to Spatial Unit 5 (selected as the center of Precinct 1) was closer than that from Spatial Unit 8 to Spatial Unit 6 (selected as the center of Precinct 2), so this unit was allocated to Center 5 because no contiguity constraints were used. On the other hand, the full version delineated contiguous precincts. Because of the contiguity constraints, Spatial Unit 8 was assigned to Center 6 in Figure 3b. This example shows that the contiguity constraints worked correctly in the proposed full model.
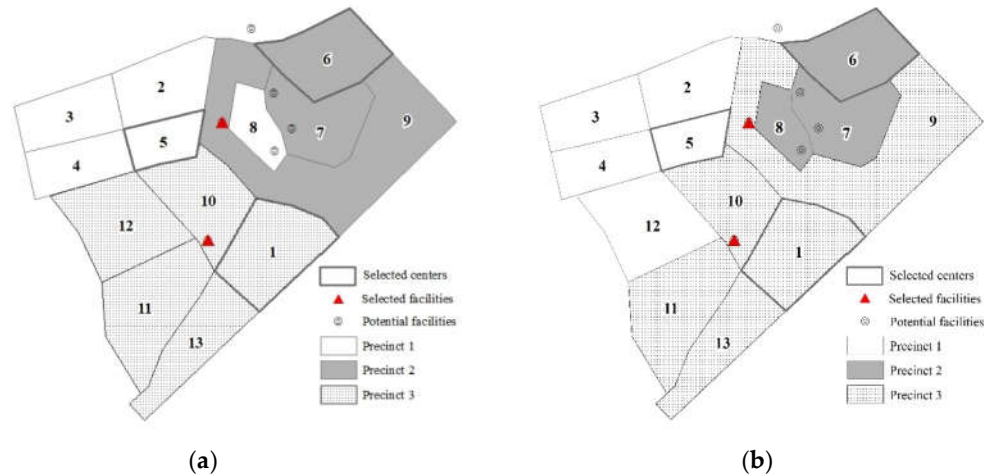
Figure 3. Role of the contiguity constraints in the CDPMP-P: (**a**) without contiguity constraints; (**b**) with contiguity constraints.

Table 1 summarizes the computational results for the small dataset. When including the contiguity constraints, Spatial Unit 8 was not assigned to the nearest center because of them, so the objective function's value increased. In addition, the computational time, the number of nodes, and iterations significantly increased because the number of cases in which the basic spatial units could be allocated due to the continuity constraints increased. When solving the practical problem instance with 200 spatial units and 38 potential facilities, the solution procedure of the MIP solver reached the out-of-memory limit without finding solutions. Therefore, the reduced model was applied to solve the large dataset and check whether the contiguity was violated.

Table 1. Computational results for the small dataset ($n = 13$, $p = 3$).

Contiguity constraints	β	C_{min}	Objective Value	Time (sec)	Node	Iteration	IDs of Selected Facilities
Without	3	750	95.601	0.12	0	46	2, 6
With	3	750	97.151	2.54	1467	69,818	2, 6

Table 2 summarizes the computational results when applying the reduced version to the large dataset. This table shows some interesting points. First, the tighter population balance constraints tended to increase the number of nodes, the number of iterations, and the time that is required to find solutions. By emphasizing population balance, basic spatial units were allocated to centers that were located farther than the nearest center to balance the population among the delimited precincts, which increased the number of allocation cases. Thus, the number of sub-problems or nodes to be solved and checked for integrality increased. Secondly, 11 problem instances among the 30 cases violated the contiguity requirements. In particular, contiguity violations were more likely to occur when the population balance constraint was applied more tightly. As the population balance constraint became tighter, basic spatial units had to be allocated to centers other than the nearest center, which increased the likelihood that dis-contiguous districts are demarcated. Districting problems that use the minimized sum of distance between centers and the assigned spatial units as an objective function implicitly assume that this objective function produces a contiguous and compact district [26]. Emphasizing population balance increases the likelihood that the model violates this assumption. This result demonstrates that contiguity conditions must be explicitly modeled in districting problems, particularly when models include population balance constraints. Thirdly, the modeling results shows that the co-location of polling places is possible if the preference was emphasized. When $\beta = 3$ or 4, multiple centers of precincts were assigned to one potential facility.

From these results, in addition, it can be seen that population balance as well as preference can affect the co-location of polling places.

Table 2. Computational results for the large dataset (n = 200, m = 38, p = 7).

$ß$	C_{min}	Objective Value	Time (sec)	Node	Iteration	Contiguity Violation	IDs of Selected Polling Facilities
0	2500	3645.116	8.82	71	3,250	No	15, 18, 24, 26, 28, 32, 38
0	2600	3657.363	7.58	2	2,603	No	15, 18, 24, 26, 28, 32, 38
0	2700	3679.259	4.98	0	337	No	15, 18, 24, 26, 28, 32, 38
0	2800	3714.938	10.05	149	12,479	No	15, 18, 24, 26, 28, 32, 38
0	2900	3756.439	17.72	129	34,111	Yes	15, 18, 24, 26, 28, 32, 38
0	3000	3863.127	3007.76	26,212	3,975,306	Yes	15, 18, 24, 26, 28, 32, 38
1	2500	3631.458	11.12	193	11,411	Yes	15, 18, 24, 26, 28, 31, 32
1	2600	3644.737	7.91	18	1,095	No	15, 18, 24, 26, 28, 32, 38
1	2700	3666.633	6.33	0	487	No	15, 18, 24, 26, 28, 32, 38
1	2800	3702.053	9.92	161	9,086	No	15, 18, 24, 26, 28, 32, 38
1	2900	3743.593	21.56	283	58,893	Yes	15, 18, 24, 26, 28, 32, 38
1	3000	3846.118	1999.23	13,909	2,325,943	Yes	15, 18, 24, 26, 28, 32, 38
2	2500	3616.215	10.14	84	6,127	No	15, 18, 24, 26, 28, 31, 32
2	2600	3634.916	9.39	59	4,945	No	15, 18, 24, 26, 28, 32, 38
2	2700	3656.812	8.53	17	1,341	No	15, 18, 24, 26, 28, 32, 38
2	2800	3691.677	11.22	190	10,177	No	3, 15, 18, 24, 26, 28, 32
2	2900	3731.337	30.42	465	57,041	Yes	3, 15, 18, 24, 26, 28, 32
2	3000	3883.495	3883.54	26,220	4,289,912	Yes	3, 15, 18, 24, 26, 28, 32
3	2500	3600.918	8.33	65	3,864	No	3, 15, 18, 24, 26, 31, 32
3	2600	3622.543	10.39	114	9,520	No	3, 15, 18, 24, 26, 31, 32
3	2700	3648.490	11.97	136	18,020	No	3(2), 15, 18, 24, 26, 32
3	2800	3682.106	11.15	96	6,631	No	3(2), 15, 18, 24, 26, 32
3	2900	3718.859	31.68	512	93,063	Yes	3(2), 15, 18, 24, 26, 32
3	3000	3822.775	3181.55	20,225	4,089,337	Yes	3(2), 15, 18, 24, 26, 32
4	2500	3589.139	9.67	235	11,585	No	3(2), 15, 18, 26, 31, 32
4	2600	3608.441	9.59	93	4,713	No	3(2), 15, 18, 26, 31, 32
4	2700	3635.920	12.75	226	14,650	Yes	3(2), 15, 18, 26, 31, 32
4	2800	3671.420	16.49	303	27,371	No	3(3), 15, 18, 26, 32
4	2900	3706.331	16.54	230	36,502	Yes	3(3), 15, 18, 26, 32
4	3000	3810.764	2025.94	18,905	2,923,362	No	3(3), 15, 17, 18, 32

Figures 4 and 5 show changes in the boundaries of the precincts and the locations of the selected facilities when emphasizing the preference and population balance. When $ß$ = 0 (Figure 4), emphasizing population balance shifted the allocation of the basic spatial units at the boundaries of the precincts, thereby slightly shifting the locations of the selected centers and facilities; but, the changes were not noticeable. All the selected facilities were located within the precincts. When the lower bound of the population balance was greater than or equal to 2900, the disjointed precincts were delineated, with spatial units assigned to a more distant center to meet the tight population balance. However, facilities with a high preference outside precincts were selected when emphasizing the preference for potential facilities (Figure 5). As the $ß$ value increased in the second term of the objective function, the distance friction decreased, and the selected precinct centers can be assigned to a more distant but more favorable potential facilities. Under the same $ß$ value, emphasizing population balance resulted in the assignment of multiple centers to a high-preference potential facility with changes in the location of the precinct centers. In Figure 5, the centers of the three precincts around a high-preference potential facility were assigned to it when C_{min} was greater than or equal to 2800.
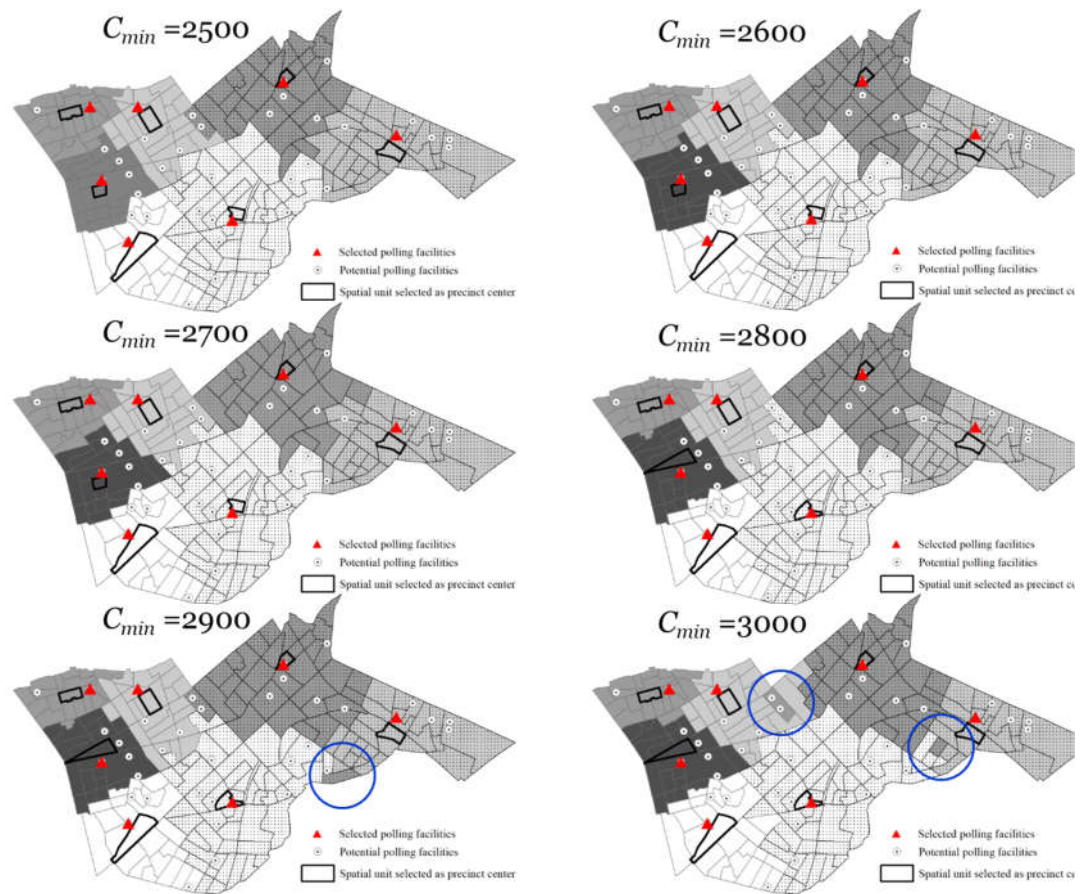
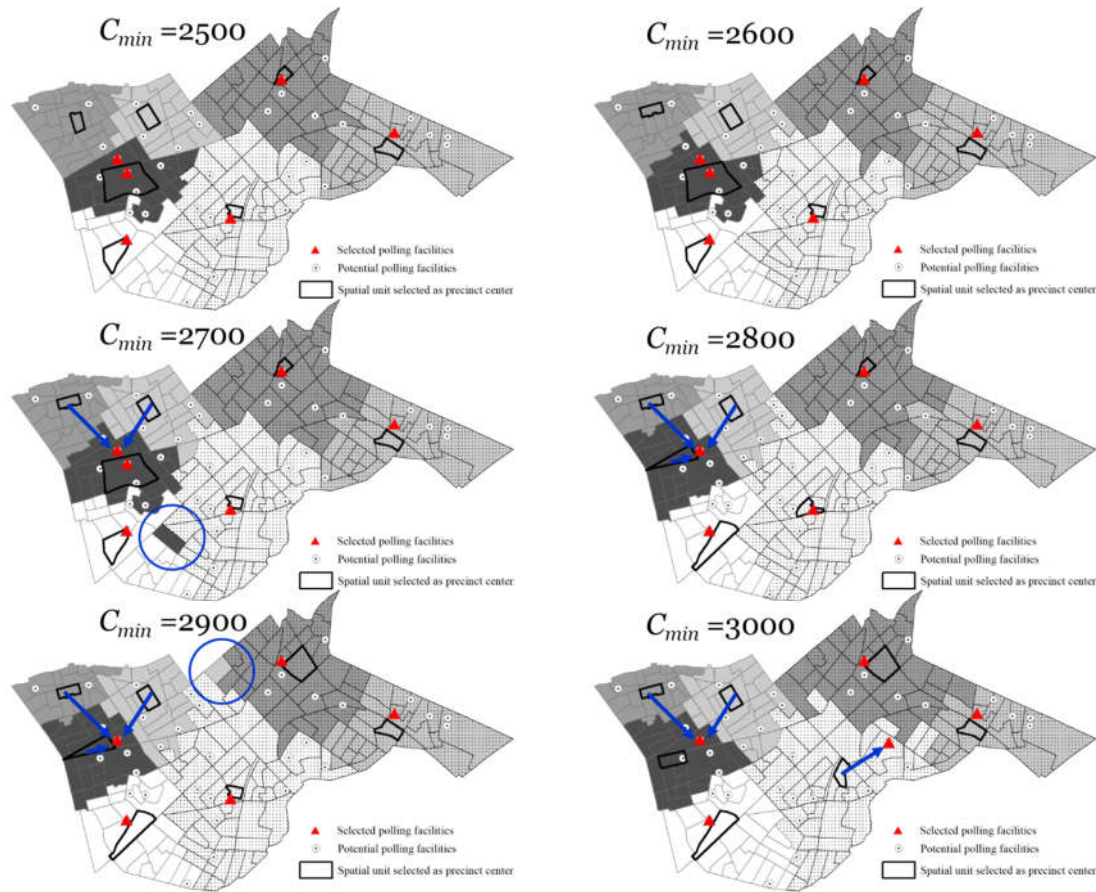**Figure 4.** Delineated precincts and selected potential facilities when $\beta = 0$.

Figure 5. Delineated precincts and selected potential facilities when *ß* = 4.

Under the tightest population balance (Figure 6), contiguity violations occurred when *ß* ranged from 0 to 3, but geographically contiguous precincts were delineated when *ß* = 4. However, the delineation of contiguous precincts when *ß* = 4 in Figure 6 seemed to be a coincidence. We must also stress the necessity of simultaneously delineating precincts and determining the locations of polling facilities. In Figure 6, you can compare the maps for when *ß* was 3 and 4. By emphasizing the preference of potential facilities, the distance friction from the center of Precinct 5 to **A** was reduced. Thus, **A** was chosen as a polling facility for this precinct instead of **B**, which was selected from other *ß* values. That is, when the preference was emphasized, the location of the selected polling facilities could be changed, which could affect the centers of precincts and the allocation of basic spatial units. This change was possible because the reduction in the second term in the minimization objective function was greater than the increase in the first term, which was caused by the allocation of basic spatial units to the non-closest center. These results show that determining the locations of polling facilities and delineating precincts can interact with each other. Therefore, both tasks should be processed simultaneously rather than independently or sequentially.
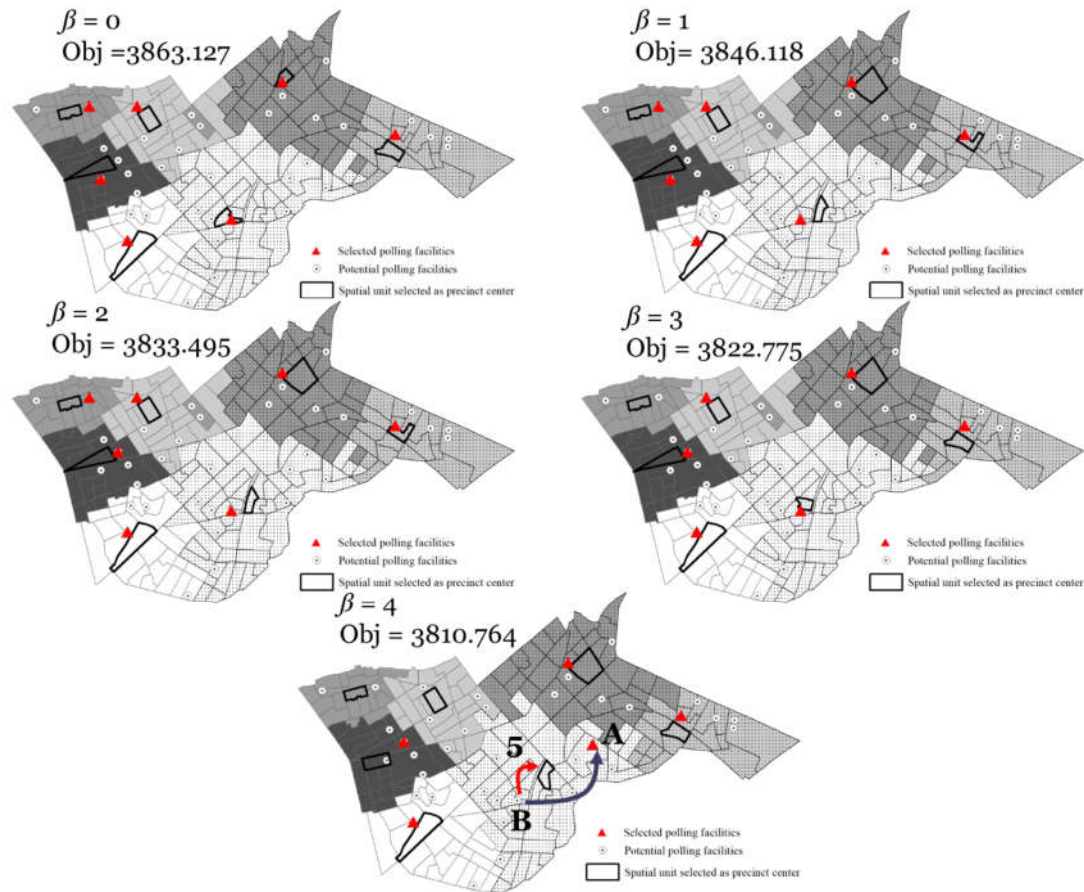
Figure 6. Delineated precincts when C_{min} = 3000.

5. Discussion

There are several discussions regarding optimization modeling and analysis results. First, it is necessary to evaluate whether the model developed provides improved outcomes theoretically and practically. Previous studies that were related to delineating precincts or determining polling places were based on a single location-allocation model, whether MIP or heuristic [12–15,17], while the CDPMP-P has a double location-allocation structure. Therefore, directly comparing the performance of the model with the results of previous studies is difficult. However, it is possible to compare the results obtained through this study with the existing system. In the existing system, the study area is divided into seven precincts, and different polling places are allocated to each of them. The population of the precincts varied from 2453 to 4874 and the average distance traveled by voters was 250 m. For comparison, the population range and average distance were calculated for a solution obtained from β = 2 and C_{min} = 2800 (in Table 2, seven falling places were selected, spatial contiguity was not violated, and C_{min} was the largest). As a result, the population range of the precincts ranged from 2800 to 4119. The population deviation was significantly reduced compared to the existing system. In addition, the average distance to the polling places was reduced to 185 m compared to the existing 250 m, resulting in improved access to the polling places. These results suggest that the spatial optimization model developed could support better decision-making regarding election administration.

Secondly, the fact that the results of the reduced version without contiguity constraints are often spatially non-contiguous, alongside the computational intractability of the full version, suggests that an alternative approach is required to find solutions to the CDPMP-P. One of the most important considerations when developing a heuristic algorithm may be how to effectively check the continuity

of the partitioned precincts. In previous studies that involved districting problems, techniques such as the connectivity matrix multiplication [35], the switching point method [36,37], and the spanning tree method [38] have been suggested. Another important consideration when developing heuristic algorithms is how to extend the search space beyond local optima so as to increase the quality of solutions. Alternatives may include tolerating poor-quality solutions, such as simulated annealing [39] or TABU search [40], or increasing the number of alternative solutions, such as genetic algorithms [41]. In addition, disassembling and dividing districts may be an alternative to increase the search space [38]. The final consideration is related to the feasible solution sets of potential facilities. In the proposed CDPMP-P, the number of precincts to be delineated (p) is given in advance, but the number of facilities to establish polling places is not fixed. Depending on the preference of facilities, several precincts may be assigned to one facility. Therefore, all solutions with a range of potential facilities greater than 1 and less than or equal to p should be evaluated in the second location-allocation process of the model.

The need to develop a heuristic algorithm for the CDPMP-P is related to how spatial optimization can be integrated with GIS. This is the last discussion. In this study, an optimization solver was used to find solutions for CDPMP-P. This is an example where spatial optimization and GIS are "loosely" combined [42]. Here, GIS was used to spatially represent demand and facilities, to create input data to be used in the developed optimization model, such as adjacency and distance matrices, and to visualize the modeling results. Even with these roles, GIS has sufficient value in the implementation of the spatial optimization model. In such a loose coupling, however, whenever a dataset is changed, the input data must be regenerated each time, the values specifying the constraints in the optimization model must be reset for the data, and the solutions must be visualized through additional processing. The development of a heuristic algorithm makes it possible to combine spatial optimization and GIS more closely [43]. If a heuristic algorithm and user interface are developed using a script language like Python provided by GIS, many hassles encountered in the loose coupling can be reduced, and spatial decision-making to meet users' needs can be supported more effectively.

6. Summary and Conclusions

Voting is an important channel of political participation. The agencies that are responsible for voting must delineate precincts and designate a polling place for each precinct. A strategic approach is required for the spatial decision-making for several reasons. First, the locations of and changes in polling places induce transportation and search costs from the perspective of voters. Secondly, improving accessibility to polling places can increase turnout. Thirdly, differences in the population sizes of the precincts may produce biased voting results. Spatial optimization approaches can be a strategic method for delimiting precincts and siting polling places. In this study, we developed a spatial optimization model, i.e., the capacitated double p-median problem with preference (CDPMP-P), which simultaneously delineates the boundaries of precincts and selects potential polling facilities with mixed integer programming (MIP). The CDPMP-P explicitly includes realistic requirements, such as population balance, the spatial continuity of precincts, the preference of potential facilities where polling places can be installed, and the possibility of allocating multiple polling places in one facility. The CDPMP-P consists of a double location-allocation process: determining the centers of the precincts and allocating basic spatial units to them, as well as determining the locations of the potential facilities and assigning the selected centers of the precincts. This double location-allocation process enables us to install polling places outside the boundaries of the precincts and multiple polling places in one facility.

For optimization models that involve continuity constraints, finding solutions becomes difficult when the problem size increases. Therefore, solutions to the application data were obtained by using the reduced CDPMP-P, which excluded the continuity constraints from the full version. The main results are as follows. First, when the population balance constraints were applied more strictly, the time required to find solutions increased because the basic spatial units were allocated to remote centers rather than the nearest centers to satisfy the conditions. In addition, the probability that

discontinuous precincts were partitioned increased as the basic spatial units were allocated to more geographically distant centers. These results disproved the need to explicitly consider continuity constraints for districting problems with population balance. Secondly, emphasizing the preference for potential facilities where polling places were installed enabled us to assign multiple polling places to a facility because the distance friction from the centers of precincts to the potential facilities was reduced. Finally, the analysis results confirmed that determining the locations of polling places and delineating precincts could influence each other, demonstrating the need for both tasks to be performed simultaneously and not independently or sequentially. The model proposed in this study may be useful in supporting spatial decision-making because it reflects realistic requirements that are related to the delineation of precincts and the installation of polling places.

This study proposed a new optimization model that can deal with mutually affecting districting and location problems simultaneously. If the developed model can be applied to electoral administration in practice, it would be possible to obtain solutions improving voter access to polling places while reducing the population variation between precincts. However, in order for the developed model to be utilized in reality, it is necessary to develop a heuristic algorithm that can find good solutions quickly, as well as an application with a user-friendly interface that can be easily used by users who are unfamiliar with GIS and spatial optimization.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kinsella, C.; McTague, C.; Raleigh, K.N. Unmasking geographic polarization and clustering: A micro-scalar analysis of partisan voting behavior. *Appl. Geogr.* **2015**, *62*, 404–419, doi:10.1016/j.apgeog.2015.04.022.

2. Shambon, L.; Abouchar, K. Trapped by Precincts? The Help America Vote Act's Provisional Ballots and the Problem of Precincts. *NYUJ Legis. & Pub. Pol'y*. **2006**, 10, 133–194.

3. Dyck, J.J.; Gimpel, J.G. Distance, turnout, and the convenience of voting. *Soc. Sci. Q.* **2005**, *86*, 531–548, doi:10.1111/j.0038-4941.2005.00316.x.

4. Stein, R.M.; Vonnahme, G. Voting at non-precinct polling places: A review and research agenda. *Elect. Law J.* **2011**, *10*, 307–311, doi:10.1089/elj.2011.1036.

5. Williams, J.C. Political redistricting: A review. *Pap. Reg. Sci.* **1995**, *74*, 13–40, doi:10.1111/j.1435-5597.1995.tb00626.x.

6. Brady, H.E.; McNulty, J.E. Turning out to vote: The costs of finding and getting to the polling place. *Am. Political Sci. Rev.* **2011**, *105*, 115–134, doi:10.1017/S0003055410000596.

7. Haspel, M.; Knotts, H.G. Location, location, location: Precinct placement and the costs of voting. *J. Politics* **2005**, *67*, 560–573, doi:10.1111/j.1468-2508.2005.00329.x.

8. McNulty, J.E.; Dowling, C.M.; Ariotti, M.H. Driving saints to sin: How increasing the difficulty of voting dissuades even the most motivated voters. *Political Anal.* **2009**, *17*, 435–455, doi:10.1093/pan/mpp014.

9. Bhatti, Y. Distance and voting: Evidence from Danish municipalities. *Scand. Political Stud.* **2012**, *35*, 141–158, doi:10.1111/j.1467-9477.2011.00283.x.

10. Orford, S.; Railings, C.; Thrasher, M.; Borisyuk, G. Changes in the probability of voter turnout when resiting polling stations: A case study in Brent, UK. *Environ. Plan.* **2011**, *29*, 149–169.

11. Webb, G. Precinct Size Matters-The Large Precinct Bias in US Presidential Elections. *arXiv Prepr.* **2014**, arXiv:1410.8868.

12. Ghiani, G.; Guerriero, F.; Musmanno, R. The capacitated plant location problem with multiple facilities in the same site. *Comput. Oper. Res.* **2002**, *29*, 1903–1912, doi:10.1016/S0305-0548(01)00065-X.

13. Konishi, K. Voting simulation for improving voter turnout and reducing the number of polling places. *J. of Japan Soc. Fuzzy Theory Intell. Inform.* **2010**, *22*, 203–210.

14. Murata, T.; Konishi, K. Making a practical policy proposal for polling place assignment using voting simulation tool. *SICE J. Control Meas. Syst. Integr.* **2013**, *6*, 124–130.

15. Cantoni, E. A precinct too far: Turnout and voting costs. *Am. Econ. J.* **2020**, *12*, 61–85, doi:10.1257/app.20180306.

16. Ross, G.T.; Soland, R.M. A branch and bound algorithm for the generalized assignment problem. *Math. Program.* **1975**, *8*, 91–103.

17. Kim, H.; Kim, K. Spatial Optimization Problem for Locating Polling Facilities and Stations and Policy Implications. In *Optimal Districting and Territory Design*; Ríos-Mercado, R.Z. Ed.; Springer: Cham, Switzerland, 2020; pp. 173–190.

18. Hanjoul, P.; Peeters, D. A facility location problem with clients' preference orderings. *Reg. Sci. Urban Econ.* **1987**, *17*, 451–473, doi:10.1016/0166-0462(87)90011-1.

19. Vasilyev, I.L.; Klimentova, K.B. The branch and cut method for the facility location problem with client's preferences. *J. Appl. Ind. Math.* **2010**, *4*, 441–454, doi:10.1134/S1990478910030178.

20. Camacho-Vallejo, J.F.; Casas-Ramírez, M.S.; Miranda, P. The p-Median bilevel problem under preferences of the customers. In *Recent Advances in Theory, Methods and Practice of Operations Research*, Ríos-Mercado, R., Camacho-Vallejo, J.-F., Gonzalez-Velarde, J., Laguna, M., Eds.; UANL: Monterrey, Mexico, 2014; pp.121–127.

21. Brunner, J.A.; Mason, J.L. The influence of driving time upon shopping center preference. *J. Mark.* **1968**, *32*, 57–61, doi:10.1177/002224296803200209.

22. Duque, J.C.; Anselin, L.; Rey, S.J. The max-p-regions problem. *J. Reg. Sci.* **2012**, *52*, 397–419, doi:10.1111/j.1467-9787.2011.00743.x.

23. Kalcsics, J. Districting Problems. In *Location Science*; Laporte, G., Nickel, S., Saldanha da Gama, F., Eds.; Springer: Cham, Switzerland, 2015; pp. 595–622.

24. Kim, H.; Chun, Y.; Kim, K. Delimitation of Functional Regions Using a *p*-Regions Problem Approach. *Int. Reg. Sci. Rev.* **2015**, *38*, 235–263, doi:10.1177/0160017613484929.

25. Shirabe, T. A model of contiguity for spatial unit allocation. *Geogr. Anal.* **2005**, *37*, 2–16, doi:10.1111/j.1538-4632.2005.00605.x.

26. Shirabe, T. Districting modeling with exact contiguity constraints. *Environ. Plan. B Plan. and Des.* **2009**, *36*, 1053–1066, doi:10.1068/b34104.

27. Pirkul, H.; Schilling, D.A. The siting of emergency service facilities with workload capacities and backup service. *Manag. Sci.* **1988**, *34*, 896–908, doi:10.1287/mnsc.34.7.896.

28. Hakimi, S.L. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Oper. Res.* **1965**, *13*, 462–475, doi:10.1287/opre.13.3.462.

29. Church, R.L. The Regionally Constrained p-Median Problem. *Geogr. Anal.* **1990**, *22*, 22–32, doi:10.1111/j.1538-4632.1990.tb00194.x.

30. Gerrard, R.A.; Church, R.L. A general construct for the zonally constrained p-median problem. *Environ. Plan. B* **1995**, *22*, 213–236, doi:10.1068/b220213.

31. Murray, A.T.; Gerrard, R.A. Capacitated service and regional constraints in location-allocation modeling. *Locat. Sci.* **1997**, *5*, 103–118, doi:10.1016/S0966-8349(97)00016-8.

32. Public Official Election Act. Available online: https://elaw.klri.re.kr/ (accessed on 7 October 2019).

33. Christaller, W. *Die Zentralen Orte in Süddeutschland*. *Central Places in Southern Germany*; Translated by Carlisle, W. Baskin; Prentice-Hall: Englewood Cliffs, NJ, USA, 1966.

34. Pavía, J.M.; Cantarino, I. Dasymetric distribution of votes in a dense city. *Appl. Geogr.* **2017**, *86*, 22–31, doi:10.1016/j.apgeog.2017.06.021.

35. Openshaw, S.; Rao, L. Algorithms for reengineering 1991 Census geography. *Environ. Plan. A* **1995**, *27*, 425–446, doi:10.1068/a270425.

36. Macmillan, W.; Pierce, T. Optimization modelling in a GIS framework: The problem of political redistricting. In *Spatial analysis and GIS*; Fotheringham, S., Rogerson, P., Eds.; Taylor & Francis: London, UK, 1994; pp. 221–246.

37. Macmillan, W. Redistricting in a GIS environment: An optimisation algorithm using switching-points. *J. Geogr. Syst.* **2001**, *3*, 167–180, doi:10.1007/PL00011473.

38. Kim, K.; Dean, D.J.; Kim, H.; Chun, Y. Spatial optimization for regionalization problems with spatial interaction: A heuristic approach. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 451–473, doi:10.1080/13658816.2015.1031671.

39. D'Amico, S.J.; Wang, S.J.; Batta, R.; Rump, C.M. A simulated annealing approach to police district design. *Comput. Oper. Res.* **2002**, *29*, 667–684, doi:10.1016/S0305-0548(01)00056-9.

40. Bozkaya, B.; Erkut, E.; Laporte, G. A tabu search heuristic and adaptive memory procedure for political districting. *Eur. J. Oper. Res.* **2003**, *144*, 12–26, doi:10.1016/S0377-2217(01)00380-0.

41.	Bacao, F.; Lobo, V.; Painho, M. Applying genetic algorithms to zone design. *Soft Comput.* **2005**, *9*, 341–348, doi:10.1007/s00500-004-0413-4.

42.	Murray, A.T. Advances in location modeling: GIS linkages and contributions. *J. Geogr. Syst.* **2010**, *12*, 335–354, doi:10.1007/s10109-009-0105-9.

43.	Caro, F.; Shirabe, T.; Guignard, M.; Weintraub, A. School redistricting: Embedding GIS tools with integer programming. *J. Oper. Res. Soc.* **2004**, *55*, 836–849, doi:10.1057/palgrave.jors.2601729.