Journal of the Korean Geographical Society, Vol. 46, No. 6, 2011(724~737)

An Optimization Approach for a Spanning Tree-Based Compactness Measure in Contiguous Land Acquisition Problems

Myung Jin Kim* · Ningchuan Xiao**

토지 획득 문제에서 공간적 밀집도 측정을 위한 최적화 연구

김명진* · Ningchuan Xiao**

Abstract : The goal of solving a contiguous land acquisition problem is to find an optimal cluster of land parcels so that one can move from an acquired parcel to another without leaving the cluster. In many urban and regional planning applications, criteria such as acquisition cost and compactness of acquired parcels are important. Recent research has demonstrated that spatial contiguity can be formulated in a mixed integer programming framework. Spatial compactness, however, is typically formulated in an approximate manner using parameters such as external border length or other shape indices of acquired land parcels. This paper first develops an alternative measure of spatial compactness utilizing the characteristics of the internal structure of a contiguous set of land parcels and then incorporates this new measure into a mixed integer program of contiguous land acquisition problems. A set of computational experiments are designed to demonstrate the use of our model in a land acquisition context.
Key Words : compactness, land acquisition problem, optimization problems, exact methods

요약 : 토지 획득 문제(land acquisition problems)란 일련의 목적에 맞게 서로 인접하고 있는 최적의 토지 필지들을 찾는 것이다. 이 문제는 도시 및 지역 계획과 각종 구획 문제 등에서 사회적 활용도가 높은 분야로서, 공간적 요소인 인접성(contiguity)와 밀집도(compactness)는 중요한 제약요소로 다루어 지고 있다. 그렇지만, 공간적 밀집도(spatial compactness)는 완벽한 측정방법이 존재하지 않고, 획득된 필지들의 둘레를 제거나, 모양을 측정하는 등의 여러 가지 방법으로 측정되고 있다. 그리하여 이 논문에서는 공간적 밀집도를 측정하는 새로운 방법을 제시하고자 한다. 인접한 토지 필지간의 내부적인 구조적 특징을 바탕으로 proximity degree라고 불리는 공간적 밀집도를 측정하는 최적화 연구모델(optimization model)을 발전시켰다. 일련의 실험을 통해 proximity degree에 따라 다양한 공간적 밀집도를 가진 모습을 확인할 수 있다.
주요어 : 공간적 밀집도, 토지 획득 문제, 최적화 연구모델

1. Introduction

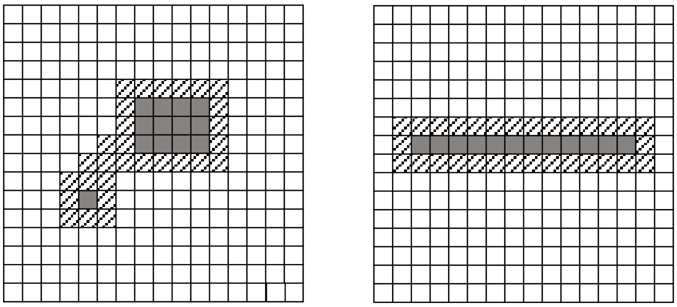
Land acquisition models have been used in a variety of application domains, including nature reserve design (Williams and ReVelle 1996; McDonnell *et al.* 2002; Nalle *et al.* 2002; Fisher and Church 2003), residential development (Baerwald 1981; Cocks and Baird 1989; Diamond and Wright 1989), recreational facility location (ReVelle and Swain 1970; Hopkins 1977; Wright et al. 1983; Gilbert *et al.* 1985), and landfill siting (Minor and Jacobs 1994). Though a common

* Part-time lecturer, Department of Geoinformatics, The university of Seoul
** Associate Professor, Department of Geography, Ohio State University, 1036 Derby Hall, 154 North Oval Mall, Columbus OH 43210, xiao.37@osu.edu

objective is to minimize acquisition cost, to effectively solve a land acquisition problem, many additional criteria must also be considered. These criteria include spatial contiguity, compactness, and other environmental and economic factors. Among these criteria, spatial contiguity and compactness are often critical when acquired land parcels are required to be adjacent and close to each other.

A cluster of acquired land parcels is contiguous if one can move from an acquired parcel to another without leaving the cluster. Recent research has demonstrated that spatial contiguity can be formulated in a mixed integer programming framework. Cova and Church (2000), for example, developed a set of contiguity constraints based on finding the shortest path between a land parcel to a pre-selected root parcel (Cova and Church 2000). Williams (2002) developed a contiguity model using the concept of a minimum spanning tree on a planar graph. More recently, Shirabe (2005) presented an alternative contiguity model based on network flows. Shirabe (2009) and Duque *et al.* (2011) developed exact optimization models for districting problems considering contiguity constraints.

Compactness, however, does not have a perfect measure and is often more difficult to measure and formulate, mainly because of the lack of an objective definition (Wright *et al.* 1983; Young 1988; Altman 1998). In a common sense, compactness means the closeness or proximity of the acquired land parcels to each other (Aerts *et al.* 2003). It is generally accepted that a region is compact if it exhibits a circular or square shape. In the context of land acquisition problems, Wright *et al.* (1983) developed a compactness measure based on the length of the external border that separates acquired and non-acquired parcels. Based on this measure, a given area is maximally compact if the border that encloses the acquired area is a circle; the border of a non-circular shape has a total length that is greater than its circular counterpart. A compact set of acquired land parcels can therefore be obtained by minimizing the length of external borders. To account for variation in land parcel size, especially for cases when irregular land parcels are used, Minor and Jacobs (1994) and Diamond and Wright (1989) used the ratio between the external perimeter and the total area of the acquired parcels. These measures, though effective, may not distinguish a variety of parcel collections that exhibit the same external border length or perimeter-area ratio but have different



(a) Core = 13, buffer = 29 (b) Core = 12, buffer = 30

Figure 1. Two sets of example acquired land parcels

The parcels filled with patterns represent the buffer zones and the grey parcels enclosed by the buffer parcels consist of the core area. Both examples have the same size of all acquired land parcels (42) and the same length of external border (34).

shapes and compactness (Young 1988; Altman 1998).

A second type of compactness measure can be used to address some of the limitations discussed above by delineating acquired parcels into core and buffer zones (Williams and ReVelle 1996; Aerts et al. 2003). A compact set of parcels can be obtained by maximizing the core area. Spatial compactness is promoted when required surrounding buffer zones are minimally generated. A drawback of this approach, however, is that acquired parcels with a larger core area may not necessarily be more compact than parcels with the same area size but a smaller core (Figure 1).

The purpose of this paper is to develop a new compactness measure (called proximity degree) that can be used to account for the internal structure of a contiguous set of land parcels based on a spanning tree. The new compactness measure is basically based on the common sense definition regarding compactness, which is proximity or closeness of the acquired land parcels and will be incorporated into the exact optimization model. In Section 2, we introduce the definition of the new measure and then formulate it in an integer programming framework for contiguous land acquisition problems. In Section 3, a set of computational experiments are designed to demonstrate the use of our compactness measure in a land acquisition context. Finally, in Section 4, we conclude the paper by discussing the applications and limitations of the measure.

2. Methodology and Problem Formulation

A graph consists of two components: vertices that represent objects (e.g. locations or land parcels) and edges that represent relationships or connections among vertices. Formally, a graph is a pair $G=(V, E)$ where V is a set of all vertices and E is a set of all edges. A planar graph is constructed in the Cartesian plane or on the surface of a sphere in a way such that no two edges intersect, except at vertices. Figure 2a illustrates the representation of land parcels using a graph. Note that the use of graph should not be confined by the regularity of the parcels, though a set of regular land parcels are used here for the sake of illustration.

A tree is a connected graph with a simple path and is acyclic. A spanning tree of a graph is a tree that connects all the n vertices of the graph with n-1 edges (Figure 2b) (Evans and Minieka 1992, 9). A subtree of a spanning tree consists of p vertices ($p \leq n$) which are a subset of the vertices in the spanning tree with $p-1$ edges and p contiguous land parcels represented by the vertices on a subtree from a contiguous set (see Figures 2c and d). Williams (2002) utilized this feature to develop a contiguity model for land acquisition problems (see below in section called contiguity model).

Given a subtree, we define its corresponding subgraph of G as $G'=(V', E')$, where V' is a set that contains vertices of the subtree, and E' is a subset of E containing all edges amongst vertices in V'. For example, p ($p \leq n$) vertices and their connections among p vertices can be formed in the subgraph G'. The number of edges in G' may be greater than the number of edges used to construct the subtree (see Figures 2d). More importantly, the difference between these two numbers (the number of edges in the subgraph and that of edges in the subtree) can be used to indicate the compactness of the vertices in the subtree. A subtree is least compact if all the edges in the subgraph are used to construct the subtree that exhibits a linear feature and the number of edges of a subtree is same as the number of

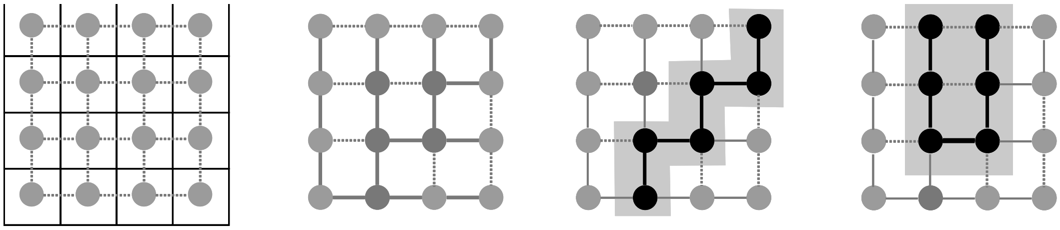


Figure 2. A graph representation of land parcels

(a) Each vertex is used to represent a regular land parcel and the edges (dashed lines) represent the connectivity between adjacent parcels. (b) A spanning tree where the solid lines represent edges on the spanning tree. (c) A subtree of the spanning tree and its corresponding subgraph (included in the shaded area). Dark dots represent the vertices in the subtree and dark solid lines represent the edges on the subtree. The proximity degree of the subtree is 0. (d) Another subtree and corresponding subgraph of six vertices where the proximity degree is 2, indicating a more compact shape than the subtree in (c).

edges of its subgraph (Figure 2c). The shape becomes compact if the length of edges in the subgraph but not in the subtree decreases and the number of edges of a subtree is less than the number of edges of its subgraph (Figure 2d). In Figure 2c, the number of vertices in G' is 6, the number of edges in G' is 5 and the number of edges in the subtree is also same as 5.The difference between the number of edges in the subgraph and the subtree is 0. Figure 2c indicates the least compact because all the edges in the subgraph (G') are used to construct the subtree. However, In Figure 2d, the number of vertices in G' is 6 and the number of edges in the subtree is 5 but the number of edges in G' is 7. This difference between the number of edges in the subgraph and the subtree is 2 which can lead to measure proximity degree.

Here, we informally define a new compactness measure, called the raw proximity degree (c'), as the difference between the length of edges that are in the subgraph but not in the subtree and the length of edges used in the subtree. A high value can be used to indicate a relatively compact shape, while a low value reflects a relatively un-compact shape. It should be noted that in many cases the length of edges in the subgraph but not in the subtree is smaller than the length in the subtree, which yields a negative value (see Figures 2c and 2d). This is the reason c' is called raw proximity and we discuss the normalization of c' in below section "Modeling compactness" after we introduce the contiguity model developed by Williams (2002).

1) Contiguity model

A significant feature of a planar graph is that it has a dual graph, created by placing a new vertex in each region enclosed by edges of the original graph and connecting these new vertices by new edges that cross all original edges (Figure 3a). The dual graph is also a planar graph and the dual of a dual graph is always the original graph (Evans and Minieka 1992, 6). We call the original graph a primal graph. The edges in the primal graph are intersecting with the edges in the dual graph. Note that the spanning trees can be constructed for both the primal and dual graphs. The primal and dual spanning trees are complementary if the edges in the two trees do not intersect (Figure 3b).

Williams (2002) developed a contiguity model by maintaining the complementarity of the primal

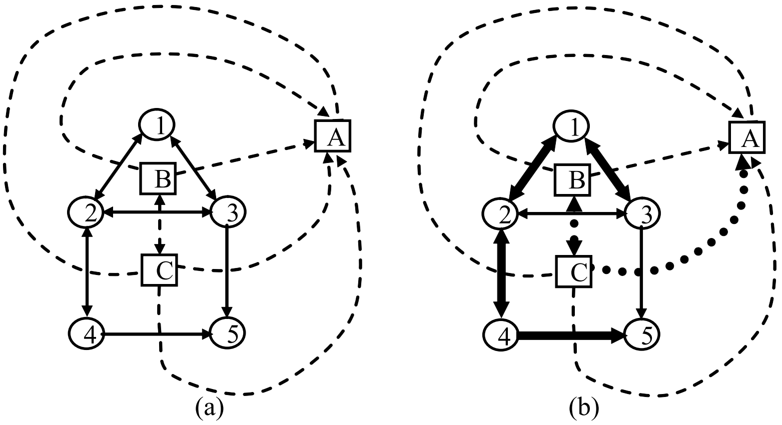


Figure 3. Primal and dual planar graphs (a), and their complementary spanning trees (b)

Here, circles represent vertices in the primal graph and vertex 5 is specified as the terminal vertex; squares represent vertices in the dual graph and vertex A is assigned to be the terminal vertex. Thick lines in (b) represent the arcs used to construct complementary spanning trees.

and dual spanning trees. In this way, cycling can be prevented and the valid spanning trees can be subsequently ensured because a cycle in any of the two trees will violate the complementarity. For each graph (primal and dual), a vertex can be arbitrarily selected as a terminal vertex. A total of n-1 directed arcs are required to build a spanning tree in the primal graph of n vertices and, similarly, m-1 direct arcs are needed for the spanning tree in the dual graph of m vertices. A subtree that consists of p vertices in the primal spanning tree represents a contiguous set of p land parcels. The goal of the program is to search for the subtree that yields minimal acquisition cost given such cost for each vertex.

The following indices and parameters are used in the formulation of this model:

i, j , I=the indices and set of primal vertices, where i, j=1, …, n;

k, l, K=the indices and set of dual vertices, where k, l=1, …, m;

D_i=the set of primal vertices that are adjacent to primal vertex i;

D_k=the set of dual vertices that are adjacent to dual vertex k;

A_i=the acquisition cost of primal vertex i; and

p=the number of land parcels (vertices) to be selected.

The decision variables are defined below:

$$X_{ij}= \begin{cases} 1 \text{ if direct arc } (i, j) \text{ in the primal graph is selected for the spanning tree and is also selected for the subtree} \\ 0 \text{ otherwise} \end{cases}$$

$$Y_{ij}= \begin{cases} 1 \text{ if direct are } (i, j) \text{ in the primal graph is selected for the primal spanning tree but is not selected for subtree} \\ 0 \text{ otherwise} \end{cases}$$

$$Z_{kl}= \begin{cases} 1 \text{ if direct arc } (i, j) \text{ in the dual graph is selected for the complemenary dual spanning tree} \\ 0 \text{ otherwise} \end{cases}$$

$$U_i= \begin{cases} 1 \text{ if a primal vertex } i \text{ is selected for acquisition} \\ 0 \text{ otherwise} \end{cases}$$

With the above variables defined, the contiguity model can be presented as:

Minimize: $\sum_{i \in I} A_i U_i$ (1)

Subject to: $\sum_{j \in D_i} X_{ij} + \sum_{j \in D_i} Y_{ij} = 1$

for all primal vertices $i = 1, 2, \ldots, n$-1 (2)

$\sum_{j \in D_k} Z_{kl} = 1$

for all dual vertices k=1, …, m-1 (3)

$X_{ij} + Y_{ij} + X_{ji} + Y_{ji} + Z_{kl} + Z_{lk} = 1$

for all intersecting primal-dual arc pairs (4)

$$\begin{cases} X_{ij} + X_{ji} \leq U_i \\ X_{ij} + X_{ij} \leq U_i \end{cases}$$

for all primal arcs (i, j), $i < j$ (5)

$\sum_{j \in D_i} X_{ij} \leq U_i$

for all primal vertices $i = 1, \ldots, n$-1 (6)

$\sum_{i \in I} U_i = p$ (7)

$\sum_{i \in I} \sum_{j \in D_i} X_{ij} = p$-I (8)

(9)

$X_{ij}, Y_{ij}, Z_{kl}, U_i \in \{0, 1\}$

The details of the above model have been discussed by Williams (2002); here we focus on the essence of this model. The objective (1) is to minimize the total acquisition cost. Constraints (2) through (4) are used to create complementary primal and dual spanning trees. Constraints (2) specify that only one vertex is selected from the primal graph. Constraints (3) indicate that only one vertex is selected from the dual graph. Constraints (4) are used to choose only one direct arc between the primal graph and the dual graph. Constraints (5) through (8) guarantee the creation of a connected subtree of p vertices in the primal spanning tree. Constraints (5) and Constraints (6) ensure contiguity in selected parcels. Constraints (7) specify p parcels to be chosen. Constraints (8) indicate p-1 edges to construct a spanning tree. Constraints (9) specify the binary decision variables.

2) Modeling compactness

This now develops a mixed integer program to model the proximity degree (our compactness measure) of the vertices in the subtree and first introduce a new set of decision variables:

$$W_{ij} = \begin{cases} 1 \text{ if edge between vertices } i \text{ and } j \text{ in the primal graph is in the subgraph but not in the subtree} \\ 0 \text{ otherwise} \end{cases}$$

With this definition, this paper further defines a raw proximity degree which refers to the closeness between selected vertices. Proximity degree can be measured by considering the length (or distance) of edges that are in the subgraph but not in the subtree. That is, the shorter the length (or distance) of edges in the subgraph but not in the subtree, the more compact shapes. There are three fundamental conditions of proximity degree. First conditions is that the subtree with p parcels is a minimum spanning tree, second condition is that the number of edges in the subgraph but not in the subtree is more and more, and third condition is that the length of edges in the subgraph but not in the subtree is shorter and shorter. A raw proximity degree (c') is also mathematically explained as $c' = \sum_i \sum_{j \in D_i} \frac{W_{ij}}{d_{ij}} - \sum_{j \in D_i} \sum_i d_{ij} X_{ij}$, where d_{ij} is the distance between vertices i and j. The use of d_{ij} is necessary for irregular graphs since edges may different lengths. Lastly, this paper defines

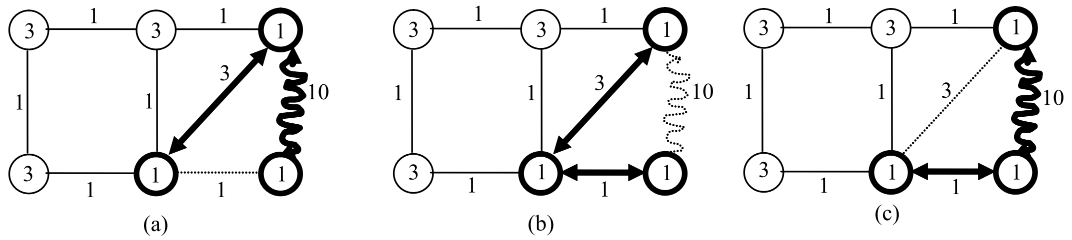


Figure 4. Graphical illustrations of the model of Williams (2002)

when p = 3. Circles represent vertices in the primal graph, costs in circles are land acquisition cost, and each arc has its length.

proximity degree (c) as a normalized measure that is intuitive to use in practice:

$$c=\frac{c'-c_{min}}{c_{max}-c_{min}} \quad (10)$$

where c_{min} and c_{max} are the lower and upper bounds of the c' measure, respectively. These bounds are constants and the value of c ranges from 0 (least compact) to 1 (most compact). These proximity degree measures successfully not only investigate internal structure among selected parcels, but also give normalized index from 0 (least compact) to 1(most compact).

The idea of proximity degree, the closeness between selected vertices, is totally different from the original model of Williams (2002). Williams (2002) finds p contiguous parcels with the minimum land costs by finding a spanning tree and p subtrees in the spanning tree. The model of Williams (2002) currently does not find MST (minimum spanning tree) which does not consider the length of edges, but only find a spanning tree. The model of Williams (2002) finds p contiguous parcels with minimum land costs without consideration of the length of edges. There are several possible solutions when p=3 with minimization of land acquisition cost (Figure 4). However, a new compactness model wants to find a compact shape in p contiguous parcels by considering the length (or distance) of edges called proximity degree which refers to the closeness between vertices. Then, Figure 4a shows high proximity degree because length of arcs (or edges) in not in the subgraph only in the subtree is short as 1 while Figure 4b shows low proximity degree because it is very long as 10. The consideration of arc lengths (or edge lengths) can be useful to measure the closeness between selected vertices.

Based on the above definitions, we formulate the following constraints:

$$c \geq c_0 \quad (11)$$

$$U_i+U_j-(X_{ij}+X_{ji})-1 \leq W_{ij}+W_{ji}$$
for all primal arcs (i, j), $i<j$ (12)

$$\frac{1}{2}(U_i+U_j)-(X_{ij}+X_{ji}) \leq W_{ij}+W_{ji}$$
for all primal arcs (i, j), $i<j$ (13)

$$W_{ij} \in \{0, 1\} \quad (14)$$

Constraint (11) specifies that the compactness of the shape must be greater than or equals to a user-specified value c_0 ($0 \leq c_0 \leq 1$). Constraints (12) and (13) control all four possible cases of the relationship between vertices and direct arcs that are needed to compute compactness. The first case occurs when two adjacent vertices (i and j) are selected, direct arc (i, j) or (j, i) is in the subtree and there is no edges in the subgrah but not in the subtree. In the first case, we have U_i+U_j=2, X_{ij}+X_{ji}=1, and W_{ij}+W_{ji}=0. The second case is when two adjacent vertices (i and j) are selected, and arc (i, j) or (j, i) is only in the subgraph but not in the subtree. In the second case, we have U_i+U_j=2, X_{ij}+X_{ji}=0, and W_{ij}+W_{ji}=1. The third case takes place when only one vertex of i and j is selected. The third case clearly leads to the fact that there is no directed arc (i, j) or (j, i) in the subtree and there is no edges in the subgrah but not in the subtree. In the third case, we have U_i+U_j=1, X_{ij}+X_{ji}=0, and W_{ij}+W_{ji}=0. The last case represents the situation when i and j are not selected, and we have U_i+U_j=0, X_{ij}+X_{ij}=0, and W_{ij}+W_{ji}=0. Note that W_{ij}+W_{ji} is always 0 when one of the directed arcs between two adjacent vertices is included in the subtree (X_{ij}+X_{ji}=1). Constraints (12) stipulate that all the above four possible cases are controlled. However, constraint (12) may produce the only one exception case, that is when there exists a directed arc between vertices i and j in the subgraph but not in the subtree even though verteices i and j are not selected (i.e., U_i+U_j=0, X_{ij}+X_{ji}=0 and W_{ij}+W_{ji}=1). Constraints (13) successfully prevent the possibility, the one exception case. Constraints (14) require W_{ij} to be a non-negative binary decision variable.

In summary, the mixed integer program to model the compactness for a contiguous land acquisition problem can be formulated as follows:

Model 1

Minimize: $\sum_{i \in I} A_i U_i$

Subject to: Equations (2) to (9), and (11) to (14).

Solving Model 1 requires the knowledge about c_{min} and c_{max} as specified in Equation (11). For a regular graph based on a grid of cells (see, for example, Figure 2a), we can analytically derive the upper bound of its proximity degree given the value of p (see Appendix). The lower bound occurs when all vertices in the subgraph are adjacent to each other as in a line (see Figure 2c) and therefore cmin is always -(p-1). For an irregular graph, however, the lower and upper bounds of the proximity degree may not be analytically derived because of the irregularity of the graph. In this case, the upper bound can be obtained by solving the following problem:

Model 2

Maximize: $\sum_i \sum_{j \in D_i} \frac{W_{ij}}{d_{ij}} - \sum_i \sum_{j \in D_i} d_{ij} X_{ij}$

Subject to: Equations (2) to (9), and (12) to (14).

Similarly, the lower bound can be obtained by solving the following problem:

Model 3

Minimize: $\sum_i \sum_{j \in D_i} \frac{W_{ij}}{d_{ij}} - \sum_i \sum_{j \in D_i} d_{ij} X_{ij}$

Subject to: Equations (2) to (9), and (12) to (14).

3. Computational Experiments

A set of experiments are designed to demonstrate the use of a compactness measure for land acquisition problems. This model incorporates the compactness constraints described above into the contiguity model developed by Williams (2002) and then investigates the model performance with different proximity degrees.

The experiments were conducted on a Xeon

2.8 GHz computer with 8 GB memory. A solver called CPLEX (version 11) is used to solve the problem. There are two types of test data used to demonstrate the model. The first type of test data is a regular data. The surface of acquisition costs developed by Williams (2002) is used here in order to see the difference between models. On this surface, acquisition costs are created on a grid with 10 rows and 10 columns. The cost at each cell (A_i) is a random number taken from a uniform distribution with a range of [0.2, 1.8], with a step size of 0.1. In order to demonstrate the impact of our compactness model on the results of land acquisition, the p value is decided to set to 30, which presented a significant computational challenge in the previous work by Williams (2002). This p value is sufficient to demonstrate the use of my new compactness. The second type of test data is an irregular data, Iowa 99 counties. The acquisition cost for each county is the population in 2000 for each county and p value is decided to set to 30.

For computational experiments using regular data, the lower bound of proximity degree for the case of p=30 is -29. Using the analytical procedure provided in the Appendix, the upper bound of proximity degree (c_{max}) is -9 for the case of p=30. I set c_0 to a value between 0 and 1, with an increment of 0.05. Table 1 reports the proximity degree requirement (c_0), the actual proximity degree of the result (c), the raw proximity degree (c') of the result, the objective function value, the number of iterations, the number of branch-and-bound nodes, and the computing time (in seconds) for each c_0 value. Model 2 and 3 are also solved in order to examine the computational complexity of these two models. The spatial configurations of the results are illustrated in Figure 5 and the computational results are listed in the bottom two rows of Table 1. The results also show that Model 2 and 3 can be solved efficiently, which implies that it is feasible to obtain the lower and upper bounds of proximity degree. Figure 6 shows the experiment solutions with different proximity degree requirements. The objective function value increases from 13.90 to 23.10 when c_0 increases from 0 to 1. The table also includes the performance of the original model by Williams (2002).

It can be concluded that owing to the addition of compactness constraints (11) to (14), the model show different computing time depending on c_0 specified by the user. In general, the model can be efficiently solved (in less than a minute) for most of the cases. The highest computation challenge occurs when c_0 is set to 0.5, which refers to something in-between in terms of proximity degree. It can be noted that the identical solutions are returned when c_0 is set

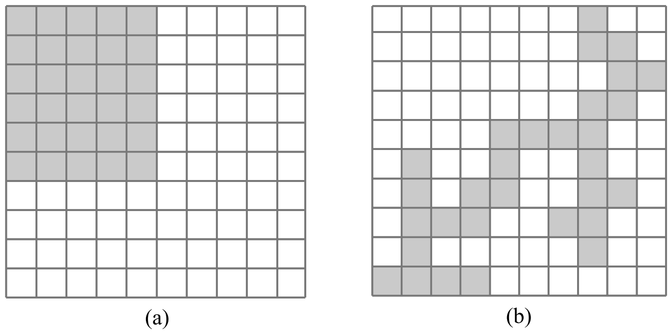


Figure 5. Spatial configurations of results from Model 2 (a) and Model 3 (b)

Table 1. Results for computational experiments (p=30)

c_0	c	c'	Objective	Iteration	Node	Time (s)
1.00	1.00	-9	23.10	65039	375	16.94
0.95	0.95	-10	20.60	9630	46	3.97
0.90	0.90	-11	19.90	5297	16	2.15
0.85	0.85	-12	19.60	37613	565	18.21
0.80	0.80	-13	19.20	89843	1307	32.05
0.75	0.75	-14	18.80	81368	1099	27.94
0.70	0.70	-15	18.10	29166	411	9.99
0.65	0.65	-16	17.70	217030	3979	61.29
0.60	0.60	-17	17.30	363619	7050	99.13
0.55	0.55	-18	16.80	395579	7021	101.71
0.50	0.50	-19	16.30	2924969	61298	660.01
0.45	0.45	-20	15.80	1113029	26761	241.47
0.40	0.40	-21	15.10	257086	5761	83.12
0.35	0.35	-22	14.70	53021	910	21.08
0.30	0.30	-23	14.40	199588	4169	57.42
0.25	0.25	-24	14.30	127886	2990	38.51
0.20	0.20	-25	14.00	77810	1788	28.67
0.15	0.15	-26	13.90	42166	1128	19.09
0.10	0.15	-26	13.90	35592	1063	16.5
0.05	0.15	-26	13.90	46643	1222	15.92
0.00	0.15	-26	13.90	50547	1261	19.51
*	0.15	-26	13.90	44394	1792	12.82
Model 2	1.00	-9		72655	633	34.16
Model 3	0.00	-29		11065	61	7.00

* The performance of the original contiguity model by Williams (2002).

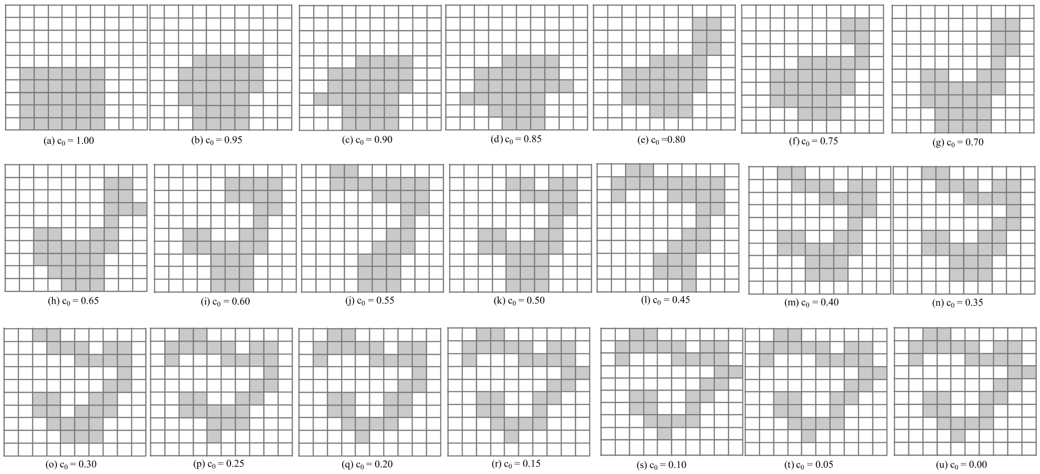


Figure 6. Spatial configurations of results using different c_0 values

The shaded symbols represent the acquired land parcels.

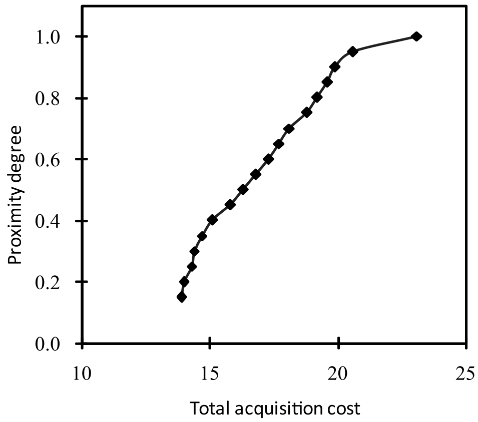
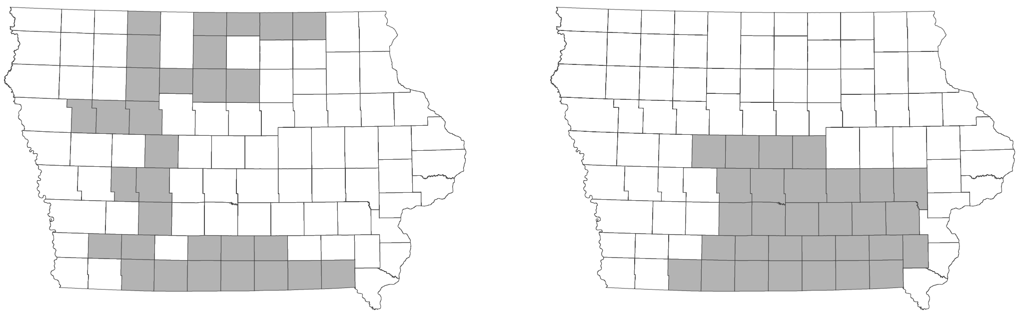


Figure 7. The tradeoff between total acquisition cost and compactness (measured by proximity degree)

from 0 to 0.15; these solutions are also as same as the solution to the original model by Williams (2002). This is because there is no solution that yields a smaller acquisition cost and exhibits a more compact shape than c_0 (as specified by Equation 11).

Finally, the tradeoff between total acquisition cost and the compactness is shown in Figure 7, where the dots, starting from the rightmost, correspond to solutions illustrated in Figures 6a to u and Table 1 respectively. It can be noted the concave shape of the tradeoff curve. As proximity degree is close to 1 which indicates the most compact areas, the land acquisition cost becomes high and *vice versa*. This observation is useful to develop multiobjective models in many applications that require to minimize the cost and to maximize the compactness simultaneously. In this paper, we solve the problem by converting the compactness objective to a constraint (Equation 11). Some commonly used multiobjective solution approaches (e.g., weighted sum) require the convexity of the tradeoff curve (see Cohon 1978), which however may not be satisfied in many applications (such as the one presented in this paper).

For computational experiments using irregular data, Model 2 is used here to demonstrate the usefulness of the model because it is impossible to analytically derive the upper and the lower bounds of the proximity degree. Model 2 produces the most compact shapes by maximizing proximity degree as a result. Figure 8 shows spatial configuration results using irregular data when p=30. For comparison, the original contiguity model by Williams (2002) is also presented. Computational results show that Model 2 finds the most compact and contiguity area while the original model by Williams (2002) finds contiguous area.



(a) The original contiguity model by Williams (2002) (b) Model 2 with compactness objective

Figure 8. Spatial configuration using Irregular data (Iowa) when p=30

4. Concluding Discussion

We developed a spanning tree-based compactness measure termed proximity degree that can be used in modeling contiguous land acquisition problems. Proximity degree is alternative new measure to capture compact area. Different from many existing compactness measures, proximity degree is designed to capture the internal structure acquired land parcels using the closeness of vertices by taking into account the length of edges used to construct the spanning tree of the graph. From several computational experiments, the new measure called proximity degree can be applied to both regular and irregular cases, and depending on different proximity degrees, diverse compact land parcels can be founded. It should be noted that proximity degree is not an exception to the common problems with compactness measures such as subjectivity (Young 1988; Altman 1998). However, we argue that this new measure is especially suitable for applications for the purpose of searching for a single contiguous cluster of land parcels. Though we only demonstrated the use of this measure in the context of a particular formulation of land acquisition problems, it will be relatively straightforward to implement the measure in other formulations that utilize graph theory.

Several variations or extensions are possible for future research. The first extension is the development of multiobjective optimization approaches (exact or heuristic methods) in order to find lots of compromised solutions. Among them, the best compromised solution can be selected. For heuristic methods, it is necessary to implement more effective and efficient heuristic. Another extension is the application to other related problems such as political redistricting, school districting, emergency service territories and electrical power districting. Compactness is also an essential factor in these problems. The idea to measure compactness in this paper can used to develop new optimization models in these areas.

Appendix

For a regular graph, the length between two arcs i and j is always 1 (i. e., d_{ij}=1). The upper bound of proximity degree can be derived given the value of p:

$$c_{min}\begin{cases}(m-1)^2+t(k)-(p-1) & \text{if } p=m^2+k(0\leq k\leq m-1)\\ m(m-1)+t(k)-(p-1) & \text{if } p=m(m+1)+k(0\leq k\leq m-1)\end{cases} \quad \text{(A2)}$$

where m is any positive integer, k is an integer between 0 and m-1, and $t(k)$ is function defined as:

$$t(k)\begin{cases}0 \text{ if } k=0,\ 1\\ k-1 \text{ if } k\geq 2\end{cases} \quad \text{(A2)}$$

We first prove the correctness of equation A1 for the cases when p equals m^2 or $m(m+1)$ and k is 0. For the case of $p=m^2$, the graph is the most compact when the vertices are arranged into a block of m rows and m columns. In this case, the total number of edges in the subgraph is $2[(m-1)^2+m-1]$. Because a spanning tree of such a subgraph contains m^2-1 edges, according to the definition of proximity degree, we have $\sum_i\sum_{j\in D_i}W_{ij}=2[(m-1)^2+m-1]-(m^2-1)=(m-1)^2$. $\sum_i\sum_{j\in D_i}X_{ij}$ is always p-1. Therefore, we have $c_{max}=(m-1)^2-(p-1)$. Similarly, when $p=m(m+1)$, the subgraph is the most compact when the vertices are arranged into a block with m rows and m+1 columns. In this case, the total number of edges in the subgraph is $2m^2$-1. Because a spanning tree of such a graph consists m^2+m-1 of edges, we have $\sum_i\sum_{j\in D_i}W_{ij}=2m^2-1-(m^2+m-1)=m(m-1)$. Again, $\sum_i\sum_{j\in D_i}X_{ij}$ is always p-1. Therefore, we have $c_{max}=m(m-1)-(p-1)$.

Compactness cases can be derived using extensions of the above two cases by increasing k vertices($0\leq k\leq m-1$). For $p=m^2+k$, when one parcel is added (k=1), the total number of edges in the subgraph becomes $2[(m-1)^2+m-1]+1$. Because a spanning tree of such a subgraph contains $(m^2+1)-1$ edges, we have $\sum_i\sum_{j\in D_i}W_{ij}=2[(m-1)^2+m-1]+1-$

$[(m^2+1)-1]=(m-1)^2$. $\sum_i \sum_{j\in D_i} X_{ij}$ always is p-1. So, we have $c_{max}=(m-1)^2-(p-1)$. When $k\geq 2$, adding one vertex to the subgraph will increase the value of $\sum_i \sum_{j\in D_i} W_{ij}$ by 1 (Figure Ala). Hence, we have $\sum_i \sum_{j\in D_i} W_{ij}=(m-1)^2+k-1$. Here, $\sum_i \sum_{j\in D_i} X_{ij}$ is always p-1. Therefore, we have $c_{max}=[(m-1)^2+k-1]-(p-1)$. Similarly, for the case of $p=m(m+1)+k$, when one parcel is added (k=1), the total number of edges is increasing with $(2m^2-1)+1$. Because a spanning tree of such a graph contains $[m(m+1)+1]-1$ edges, we have $\sum_i \sum_{j\in D_i} W_{ij}=[(2m^2-1)+1]-[(m(m+1)+1-1]=m(m-1)$. When $k\geq 2$, adding one vertex to the subgraph will increase the value of $\sum_i \sum_{j\in D_i} W_{ij}$ by 1 (Figure Alb). So, we have $\sum_i \sum_{j\in D_i} W_{ij}=m(m-1)+k-1$. Here, $\sum_i \sum_{j\in D_i} X_{ij}$ is always p-1. Therefore, we have $c_{max}=[m(m-1)+k-1]-(p-1)$.

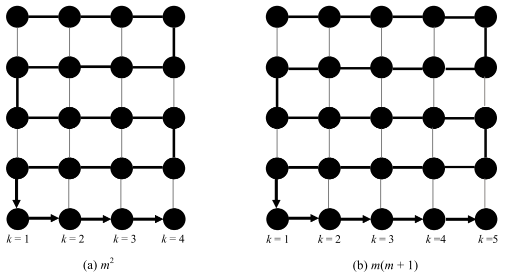


Figure A1. Adding k vertices to a block of (a) m^2 or (b) $m(m+1)$ vertices.

A dashed line represents an edge in the subgraph but not in the subtree, and a thick solid line represents an edge in the subtree.

References

Aerts, J. C. J. H., E. Eisinger, G. B. M. Heuvelink and T. J. Stewart, 2003, Using linear integer programming for multi-site land-use allocation, *Geographical Analysis*, 35(2), 148-169.

Altman, M. ,1998, Modeling the effect of mandatory district compactness on partisan gerrymanders, *Political Geography*, 17(8), 989-1012.

Baerwald, T. J., 1981, The site selection process of suburban residential builders, *Urban Geography*, 22, 339-357.

Cocks, K. D. and I. A. Baird, 1989, Using mathematical programming to address the multiple reserve selection problem: an example from the Eyre Peninsula, South Australia, *Biological Conservation*, 49, 113-130.

Cohon, J. L., 1978, *Multiobjective Programming and Planning*, New York: Academic Press.

Diamond, J. T. and J. R. Wright, 1989, Efficient land allocation, *Journal of Urban Planning and Development*, 115, 81-96.

Duque, J.C., Church, R.L., and Middleton, R.S., 2011, The p-Regions Problem, *Geographical Analysis*, 43, 104-126.

Evans, J. R. and E. Minieka, 1992, *Optimization algorithms for networks and graphs*, 2nd Ed. New York: Marcel Dekker, Inc.

Fisher, D. and R. L. Church, 2003, Clustering and compactness in reserve site selection: an extension of the biodiversity management area selection model, *Forest Science*, 49, 555-565.

Gilbert, K. C., D. D. Holmes and R. E. Rosenthal, 1985, A multiobjective discrete optimization model for land allocation, *Management Science* 31(12), 1509-1522.

Hopkins, L. D., 1977, Land-use design-quadratic assignment and central facility models, *Environment and Planning A*, 9, 625-642.

Love, R. F. and J. G. Morris, 1979, Mathematical models of road travel distances, *Management Science*, 25(2), 130-139.

McDonnell, M. D., H. P. Possingham, I. R. Ball and E. A. Cousins, 2002, Mathematical methods for spatially cohesive reserve design. *Environmental Modeling and Assessment*, 7, 107-114.

Minor, S. D. and T. L. Jacobs, 1994, Optimal land allocation for solid and hazardous waste landfill siting, *Journal of Environmental Engineering*, 120,1095-1108.

Nalle, D. J., J. L. Arthure and J. Sessions, 2002, Designing compact and contiguous reserve networks with a hybrid heuristic algorithm, *Forest Science*, 48(1), 59-68.

ReVelle, C. S. and R. Swain, 1970, Central facilities location, *Geographical Analysis*, 2, 30-42.

Shirabe, T., 2005, A model of contiguity for spatial unit allocation, *Geographical Analysis*, 37, 2-16.

Shirabe, T., 2009, Districting modeling with exact contiguity constraints, *Environment and Planning B: Planning and Design*, 36(6), 1053-1066.

Skiena, S., 1990, I*mplementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica*, New York: Addition-Wesley.

Williams, J. C., 2002, A zero-one programming model for contiguous land Acquisition, *Geographical Analysis*, 34, 330-349.

Williams, J. C. and C. S. ReVelle, 1996, A 0-1 programming approach to delineating protected reserves, *Environment and Planning B*, 23, 607-624.

Williams, J. C. and C. S. ReVelle, 1998, Reserve assemblage of critical areas: A zero-one programming approach, *European Journal of Operational Research*, 104, 497-509.

Wright, J. R., C. S. ReVelle and J. L. Cohon, 1983, A multiobjective integer programming model for the land acquisition problem, *Regional Science and Urban Economics*, 13, 31-53.

Young, H. P., 1988, Measuring the compactness of legislative districts, *Legislative Studies Quarterly*, 13(1), 105-115.

Correspondence: Myung Jin Kim, Department of Geoinformatics, The university of Seoul, 163 Siripdaero, Dongdaemun-gu, Seoul 130-743 Korea (e-mail: wizlove51@gmail.com, phone: +82-10-7155-8250)

교신: 김명진, 130-743 서울특별시 동대문구 서울시립대로 163(이메일: wizlove51@gmail.com, 전화: 010-7155-8250)

Recieved September 1, 2011

Revised December 19, 2011

Accepted December 22, 2011