Analysis of the spatial relations among point distributions on a discrete space

Yukio Sadahiro
Department of Urban Engineering, University of Tokyo
sada@ua.u.t.u-tokyo.ac.jp

Abstract: This paper proposes a method of analyzing spatial relations among point distributions on a discrete space. Spatial topology and proximity are discussed in an integrated framework at both local and global scales. Local relations are described by geographical representations while global relations are visualized by graph representations. The latter also provides a means of classifying the point distributions. The proposed method is applied to school location planning in Japan. The results reveal the appealing properties of the method and provide empirical findings.

Keywords: spatial topology, spatial proximity, geographical representations, graph representations

1. Introduction

Spatial relations among point distributions have drawn considerable attention in geography, ecology, epidemiology, and other academic fields related to spatial phenomena. Maps of John Snow’s London (Cliff and Haggett, 1988) clearly visualizes the clustering of cholera cases around a water well. Pielou (1961) analyzes the segregation and symmetry in the distributions of Douglas firs and pines. Christaller (1933) discusses the hierarchical relation among urban settlements in terms of their location and function.

Among spatial relations, spatial proximity has been the most important topic in point pattern analysis. Statistical methods have been developed to analyze the spatial proximity between two different types of points. Given two sets of points, nearest-neighbor contingency table (Pielou, 1961; Dixon, 1994; Ceyhan, 2009) summarizes the relations between base points and their nearest neighbors into four categories. Statistical test is performed by comparing the table with that expected from the null hypothesis where point sets are distributed independent with each other. Bivariate \( J \)-function (Lieshout & Baddeley, 1999) also evaluates the closeness of points between distributions. Cross \( K \)-function (Ripley, 1981) counts the number of points of one type located within a certain distance from points of the other type.
Since the latter two are both functions of arbitrary distance, interpretation of results needs careful consideration to avoid the problem of multiple testing.

For more than two types of points, the choice of approach depends on the objective of analysis. If the focus is on the relation between pairs of point distributions, one option is to perform the above methods for all the distribution pairs. Spatial data mining is effective especially for a large set of points (Koperski and Han, 1995; Morimoto, 2001; Huang et al., 2004; Wan and Zhou, 2009). It is a powerful tool for detecting spatial association between point distributions, which provides research hypothesis in exploratory spatial data analysis. If the interest lies in the relation among more than two types of points, nearest-neighbor contingency table and multivariate J-function are useful to test whether the distributions are mutually correlated.

As seen above, statistical and computational methods are available for analyzing spatial proximity among point distributions. Spatial relations, however, are not limited to spatial proximity. Spatial hierarchy has been one of the main topics in quantitative geography since Christaller's pioneering works on Central Place Theory (King, 1984; Mulligan, 1984). Hierarchical structure can be found in the location of administrative offices, educational facilities, and retail outlets (Okabe and Sadahiro, 1996). Heterogeneity in spatial phenomena yields a wide variety of local patterns in point distributions (Anselin, 1995; Kulldorff, 1997; Lloyd, 2006). It sometimes happens that two distributions are correlated positively in some areas while negatively in other areas.

To treat a wider variety of spatial relations, this paper aims to provide a general framework for analyzing the relations among point distributions. As well as spatial proximity, spatial topology including hierarchical relations is discussed in an integrated framework. It provides a means of describing the relations among point distributions and classifying them into similar groups. Another emphasis is on the local analysis of point distributions. Spatial topology and proximity are considered at various scales, from local to global.

The focus of this paper is on points distributed on a discrete space, i.e., a limited number of locations. Urban facilities, both public and private ones, often fall into this category, where land blocks are firmly predetermined by a land ownership system. Since they are hard to change, new facilities choose their location from a limited set of available lots.

Methods are proposed in the following five sections. Section 2 defines basic topological relations between point distributions and those between locations. Using these relations, Sections 3 and 4 propose geographical and graph representations of the relations among point distributions. They are used in local and global analysis of
the relations, respectively. Section 5 introduces numerical measures that also describe the global relations among point distributions. Section 6 extends the proposed method to treat a wider variety of situations. Section 7 applies the method to school location planning in Japan. Section 8 summarizes the conclusions with a discussion.

2. Topological relations between point distributions and those between locations

The basic topological relations between point distributions and those between locations serve as a basis for the analysis and visualization of more complicated relations among point distributions.

2.1 Relations between a pair of point distributions

Consider a set of locations \( \Lambda = \{ Q_i, \ i \in [M] \} \) in a two-dimensional region \( S \), where \([M]=\{1, 2, ..., M\}\). A set of non-empty point distributions \( \Omega = \{ \Omega_i, \ i \in [N] \} \) (\([N]=\{1, 2, ..., N\}\)) are defined on \( \Lambda \), each of which consists of \( m_i \) points denoted by their locations \( \{ Q_{i1}, Q_{i2}, ..., Q_{imi} \} \). Points do not overlap with each other within the same distribution.

Two distributions \( \Omega_i \) and \( \Omega_j \) are called equal and denoted as \( \Omega_i = \Omega_j \) if they consist of the same set of locations (Figure 1a). Two distributions are inclusive if one set is a subset of but not equal to the other (see Figure 1b where \( \Omega_4 \subset \Omega_3 \)). If \( \Omega_i \subset \Omega_j \), \( \Omega_j \) is a higher level distribution of \( \Omega_i \), while \( \Omega_i \) is a lower level distribution of \( \Omega_j \). If \( \Omega_i \cap \Omega_j = \emptyset \), they are called exclusive (Figure 1c). They do not share any location in their distributions. Two distributions are overlapping (Figure 1d) if they are neither inclusive nor exclusive. These four relations are mutually exclusive.

Another set of relations is defined by considering \( \Lambda \), the universal set of locations. Two sets \( \Omega_i \) and \( \Omega_j \) are complete if \( \Omega_i \cup \Omega_j = \Lambda \) (Figure 1d), while they are incomplete if \( \Omega_i \cup \Omega_j \neq \Lambda \) (Figure 1e). Since these two relations are independent of the previous four relations, eight relations (i.e., combinations of four distribution related ones and two location related ones) are defined between \( \Omega_i \) and \( \Omega_j \). For instance, \( \Omega_i \) and \( \Omega_j \) are complete and overlapping if \( \Omega_i \cup \Omega_j = \Lambda \) and \( \Omega_i \cap \Omega_j \neq \emptyset \) (Figure 1e). Two distributions are called complementary if they are complete and exclusive (Figure 1f).

2.2 Relations between a pair of locations
The above discussion also applies to the relations between a pair of locations. Locations $Q_i$ and $Q_j$ are called equal (Figure 2a) if any distribution in $\Omega$ contains either both or none of them. If $Q_i$ and $Q_j$ are not equal but every distribution containing $Q_j$ also contains $Q_i$, or vice versa, they are called inclusive. In the former case $Q_i$ is a higher level location of $Q_j$, while $Q_j$ is a lower level location of $Q_i$ in the latter case. In Figure 2b, $Q_3$ is a higher level location of $Q_4$. Locations $Q_i$ and $Q_j$ are exclusive (Figure 2c) if any distribution does not contain both $Q_i$ and $Q_j$ simultaneously. If $Q_i$ and $Q_j$ are neither equal, inclusive, nor exclusive, they are called overlapping (Figure 2d).

Consideration of the universal set of distributions $\Omega$ gives another independent set of relations. Locations $Q_i$ and $Q_j$ are complete (Figure 2d) if any distribution contains at least $Q_i$ or $Q_j$. On the other hand, locations $Q_i$ and $Q_j$ are incomplete (Figure 2e) if a distribution exists that containing neither $Q_i$ nor $Q_j$. As well as point distributions, locations can be described by a pair of independent relations such as incomplete and overlapping (Figure 2e). If locations $Q_i$ and $Q_j$ are complete and exclusive, they are called complementary (Figure 2f).

3. Geographical representations of the relations among point distributions

This section proposes geographical representations of the relations among point distributions. They serve as graphical tools for analyzing the relations at a local scale.

To this end, the relations defined for a pair of locations are extended to the cases of more than two locations. A set of locations are called complete if any distribution contains at least one of the locations. In a complete set of locations, if every distribution contains exactly one location, the locations are called complementary.

A $K$-core is a complete set of $K$ locations any smaller subset of which is not complete. For instance, $Q_i$ is a 1-core if it is shared by all the distributions. Locations $Q_i$ and $Q_j$ form a 2-core if they are complete and neither is 1-core (Figure 3). Locations $Q_i$, $Q_j$ and $Q_k$ are a 3-core if any two of the three is not complete.

$K$-cores are classified into complementary and non-complementary ones denoted by $K_C$-cores and $K_O$-cores, respectively. In a $3_C$-core every distribution has exactly one point while a $3_O$-core contains at least one location shared by more than one distribution. Since a $K$-core does not contain smaller $K$-cores, a $K$-core contains at least one location shared by only one distribution.

$K$-cores indicate local similarity in composition of points among distributions. All the distributions are locally equal at a 1-core in the sense that they share the same
location. In a 2-core at least one point is found in every distribution, which can be interpreted as a local similarity among point distributions.

Visualization of $K$-cores gives us a useful information about the local similarity among distributions. However, since $K$-cores are permitted to share the same locations, numerous $K$-cores can exist simultaneously. It may be impossible to enumerate all the $K$-cores, and their visualization may yield a very complicated map.

To resolve this problem, this paper considers the desirable properties of $K$-cores to visualize in three aspects: 1) simplicity of representation, 2) spatial compactness, and 3) usefulness of information.

The first property claims that $K$-cores of fewer locations are more desirable than larger ones. Simpler $K$-cores reveal the structure of point distributions more clearly and effectively, as seen by comparing 1-cores, 2-cores and 3-cores.

The second property is that smaller $K$-cores should be chosen before larger ones. This is critical when a focus is on the spatial aspects of relations among point distributions. It is consistent with the first law of geography “near things are more related than distant things (Tobler, 1970).” This property assures that $K$-cores indicate local similarity in location among point distributions, which will be shown later. Proximity of locations is measured by, for instance, the radius of their circumcircle.

The third property claims that $K$-cores should be visualized in such a way that they convey as much information about point distributions as possible. The amount of information is evaluated by a decrease in entropy, which is a measure of uncertainty associated with stochastic phenomena (Shannon, 1948). Consider two locations $Q_i$ and $Q_j$ in $\Omega$, a set of $N$ point distributions. If no information is available about whether the locations are contained in each distribution in $\Omega$, $4^N$ cases can equally occur. Entropy of this situation is

$$-\log_2 \frac{1}{4^N} = 2N, \quad (1)$$

If $Q_i$ and $Q_j$ are known to be a 2-core, the number of possible cases reduces to $3^N - (2^{N+1} - 1)$ (subtraction of $2^{N+1} - 1$ represents that a 2-core does not contain 1-cores). For large $N$, entropy reduces to

$$-\log_2 \frac{1}{3^N - (2^N - 1)} \approx \left( \log_2 3 \right)N, \quad (2)$$

Consequently, average amount of information per location is

$$J(2\text{-core}) = \frac{2 - 1.585}{2} N, \quad \approx 0.208N, \quad (3)$$
where $J(K$-core) is the average amount of information per location given by indicating a $K$-core.

If $Q_i$ and $Q_j$ are a $2_c$-core, the entropy reduces accordingly to $2^N-2$. Consequently,

$$J(2_c\text{-core}) = \frac{1}{2} \left( 2N - \log_2 \frac{1}{2^N-2} \right).$$

$$\approx 0.500N$$

The amount of information given by a 1-core is

$$J(1\text{-core}) = -\log_2 \frac{1}{2^N} - \left( \log_2 \frac{1}{1} \right).$$

$$= N$$

For $3_c$- and $3_o$-cores, the average information is approximately $0.472N$ and $0.066N$, respectively (for details, see Sadahiro (2009)). These measures give us a means of ordering $K$-cores to visualize. If $2_c$- and 2-cores are overlapping, the former should be visualized because it conveys more information. If $3_c$, $2_c$- and $2_o$-cores share a single location, the $2_c$-core should be chosen while the others are omitted. The priority order of 1-, 2- and 3-cores is

$$1\text{-core} > 2_c\text{-core} > 3_c\text{-core} > 2_o\text{-core} > 3_o\text{-core}.$$  \hspace{1cm} (6)

Note that the three properties do not specify a particular procedure of choosing $K$-cores to visualize. They are not strict rules to obey, but desirable properties to satisfy. They are not always compatible and even not comparable with each other. Consequently, in practice, a general principle should be determined by analysts in light of the objective of analysis. One option is to choose $K$-cores that convey the largest amount of information without overlapping. Since it is a combinatorial problem, heuristic methods would be effective to obtain a reasonable result. Section 7 will illustrate another example of implementation.

**4. Graph representations of the relations among point distributions**

From local analysis of the relations among point distributions, this section shifts the focus to global analysis. Graph representations are introduced to describe the global structure of the relations.

**4.1 Lattice representation of the relations among point distributions**
Consider the power set of $\Lambda$, denoted by $2^\Lambda=\{\Psi_1, \Psi_2, ..., \Psi_P\}$. This contains all the point distributions in $\Omega$. Since they are defined as sets of locations, Boolean operations such as intersection, union, and complement can be applied. Consequently, the algebraic structure $(2^\Lambda, \cap, \cup)$ is a lattice, where the smallest and largest elements are $\emptyset$ and $\Lambda$, respectively.

A lattice as a partially ordered set can be visualized as a Hasse diagram (Davey and Priestley, 2002; Pemmaraju and Skiena, 2003). Hasse diagram is a simple representation of a partially ordered set that permits us to evaluate the relations among elements in $2^\Lambda$. In the Hasse diagram, nodes and links represent point distributions in $\Psi$ and inclusion relation between them, respectively. Binary operations applied to point distributions are given by tracing links either upward or downward from the point distributions. The intersection and union of $\Psi_i$ and $\Psi_j$ are the point distributions at the lowest and highest levels connected either directly or indirectly to both $\Psi_i$ and $\Psi_j$, respectively.

Hasse diagram can also be used to visualize the closeness of relations between point distributions. To this end, two distances are defined between $\Psi_i$ and $\Psi_j$: Size distance is the difference in the number of elements in $\Psi_i$ and $\Psi_j$:

$$d_s(\Psi_i, \Psi_j) = |n(\Psi_i) - n(\Psi_j)|,$$

It is zero if two distributions consist of the same number of points.

Inclusion distance evaluates the separation from inclusive relation. Consider four distributions $\Psi_1=\{Q_1, Q_2, Q_3, Q_4\}$, $\Psi_2=\{Q_1, Q_2, Q_3\}$, $\Psi_3=\{Q_1, Q_2, Q_5\}$, and $\Psi_4=\{Q_1, Q_5, Q_6\}$. Distribution $\Psi_1$ contains all the three elements in $\Psi_2$, two of three elements in $\Psi_3$, and only one element in $\Psi_4$. The relation between $\Psi_1$ and $\Psi_2$ is inclusive while $\Psi_1$ and $\Psi_3$ are overlapping. The relation between $\Psi_1$ and $\Psi_4$ is also overlapping but the relation is weaker than that between $\Psi_1$ and $\Psi_3$. Inclusion distance evaluates such closeness of relations between distributions:

$$d_i(\Psi_i, \Psi_j) = n(\Psi_i \cup \Psi_j) - \max\left(n(\Psi_i), n(\Psi_j)\right)$$

$$= \min\left(n(\Psi_i), n(\Psi_j)\right) - n(\Psi_j \cap \Psi_j),$$

It is zero if $\Psi_i$ and $\Psi_j$ are in inclusion relation, and increases as their intersection becomes smaller.

Taking $n(\Psi_i)$ as the vertical axis, Hasse diagram tells us the closeness of relations between two distributions (Figure 4). Inclusion distance $d_i(\Psi_i, \Psi_j)$ is represented as the length of the shorter link connected to $\Psi_i \cup \Psi_j$. If $\Psi_i$ and $\Psi_j$ are inclusive, $d_i(\Psi_i, \Psi_j)=0$ and thus $\Psi_i$ and $\Psi_j$ are connected directly by a single link. If the relation is mostly inclusive, the shorter link becomes so short that $\Psi_i$ and $\Psi_j$ look like connected directly by a single link. Size distance is the difference in length of two
links connected to $\Psi_i \cup \Psi_j$. If $\Psi_i$ and $\Psi_j$ consist of the same number of elements, they are connected indirectly by two links of the same length. Four links connecting $\Psi_i$, $\Psi_j$, $\Psi_i \cup \Psi_j$, and $\Psi_i \cap \Psi_j$ form a rhombus.

### 4.2 Tree representation of the relations among point distributions

Though Hasse diagram is a useful tool for visualizing the relations among point distributions, it is not effective for a large set of distributions because it becomes too complicated. Hasse diagram is complemented by intersection tree, a visual representation of a stepwise process that groups point distributions into one according to the similarity between distributions (see also Sadahiro and Sasaya (2008)). It is constructed in a way similar to hierarchical methods in cluster analysis. Choice of distance measure depends on the objective of analysis. On option is to use both the size and inclusion distances, say, $d(\Omega_i, \Omega_j)=d_s(\Omega_i, \Omega_j) + d_i(\Omega_i, \Omega_j)$ or $d(\Omega_i, \Omega_j)=\frac{d_s(\Omega_i, \Omega_j) + d_i(\Omega_i, \Omega_j)}{n(\Omega_i \cap \Omega_j)}$. The measure is calculated for all the pairs of point distributions in $\Omega$, from which the most similar pair is chosen. They are replaced by their intersection so that the elements in $\Omega$ decrease by one. This process continues until $\Omega$ consists of only one element. Figure 5 illustrates an intersection tree constructed from three distributions of points. The tree grows downward from original distributions.

Using union instead of intersection yields another tree called a union tree. It is generated by replacing the most similar pair of point distributions with their union. A union tree extends upward from original distributions.

Both intersection and union trees consist of at most $N-1$ links. While they inherit the properties of Hasse diagram, they are not so complicated as to allow us intuitive understanding of the relations among point distributions.

Intersection and union trees are processes of gathering similar distributions into groups. Consequently, they can also serve as hierarchical methods of clustering. Their partial trees naturally define groups of similar point distributions. Note, however, that intersection and union trees are different from dendrograms used in cluster analysis. The vertical axis of the trees indicates $d_s(\Omega_i, \Omega_j)$, which is not necessarily used as a distance measure in clustering process. Since the trees do not directly indicate the order of clustering, the process of classification should be visualized by using dendrograms.
5. Numerical measures of the properties of point distributions

As well as graphical representations proposed above, numerical measures are also effective for understanding the global relations among point distributions. The mean and variance of the number of points are the most basic statistics that describe the properties of point distributions. These measures, however, are defined based only on the number of points so that they do not distinguish different compositions of point distributions. To complement them, this paper introduces a measure called \( \gamma \), which evaluates the difference in point composition among distributions:

\[
\gamma(\Omega) = 1 - \frac{N n\left(\bigcap_{i} \Omega_i\right)}{\sum_{i} m_i}, \tag{9}
\]

where \( n() \) is an operator that gives the number of elements in a set. This is the ratio of locations not shared by all the distributions. For instance, if \( \{\Omega_1, \Omega_2, \Omega_3\} = \{\{Q_1, Q_2, Q_3\}, \{Q_1, Q_3\}, \{Q_2, Q_3\}\} \), \( Q_1 \) and \( Q_2 \) are not shared by all the distributions and thus the diversity is \( 2/3 \). If \( \Omega_3 = \{Q_1, Q_2, Q_3\} \), \( \Omega_1 \cap \Omega_2 \cap \Omega_3 = \{Q_1, Q_3\} \) and \( \gamma(\Omega) = 1/3 \). The measure \( \gamma(\Omega) \) tends to be small when all the distributions consist of similar locations and increases with the variation of elements in the distributions.

Since diversity only counts the number of locations shared by all the distributions, it may not work for a large set of distributions. To treat such cases, \textit{weighted diversity} relaxes the definition of \( \gamma(\Omega) \):

\[
\gamma_w(\Omega) = 1 - \frac{N \sum_{i} m_i \sum_{j} w(\Omega, Q_j)}{\sum_{j} w(\Omega, Q_j)}, \tag{10}
\]

where \( w(\Omega, Q_j) \) is a measure of sharing \( Q_j \) in \( \Omega \). A possible definition of \( w(\Omega, Q_j) \) is

\[
w(\Omega, Q_j) = \left[ \frac{1}{N} \sum_{i} n(\Omega_i \cap Q_j) \right]^2, \tag{11}
\]

where the right hand side is the square of the ratio of distributions in \( \Omega \) that contains \( Q_j \).

6. Extension of the method: smoothing of point distributions

The theory of the proposed method heavily relies on the concept of intersection and union of point distributions. This may cause a practical problem when a large number of distributions are analyzed because their intersection often
becomes empty. To resolve this problem, smoothing operation is applied to point distributions that transforms points into a surface. Surface function of point distribution $\Omega_i$ at $Q_j$ is defined as

$$s(\Omega_i, Q_j) = \frac{\sum_k \rho(Q_j, Q_k)n(\Omega_i \cap Q_k)}{\sum_k \rho(Q_j, Q_k)}.$$  

(12)

where $\rho(Q_j, Q_k)$ is spatial proximity between $Q_j$ and $Q_k$. Spatial proximity is typically given by $\rho(Q_j, Q_k)=\exp(-\alpha d_E(Q_j, Q_k))$, where $d_E(Q_j, Q_k)$ is the Euclidean distance between $Q_j$ and $Q_k$. The surface function is continuous ranging from 0 to 1 defined on discrete locations $\Lambda$. It shows a large value if points are densely distributed around $Q_j$ in $\Omega_i$.

Transformation of points into a surface involves redefinition of spatial objects, operations and variables. Distribution $\Omega_j$ is represented by a set of continuous values:

$$S(\Omega_j) = \{s(\Omega_i, Q_j), j \in [M]\}.$$  

(13)

The number of elements in $\Omega_i$ is given by

$$n_s(\Omega_i) = \sum_j s(\Omega_i, Q_j).$$  

(14)

Intersection of sets is also a set of continuous values:

$$\eta(S(\Omega)) = \left\{\min_{i \in [N]} (s(\Omega_i, Q_j)), j \in [M]\right\}.$$  

(15)

Intersection is usually positive so that the problem mentioned earlier can be avoided.

The mean and variance of the number of points remain the same because the surface function is standardized. Diversity is given by

$$\gamma_s(\Omega) = 1 - \frac{N \sum_{j \in [N]} \min_{i \in [N]} (s(\Omega_i, Q_j))}{\sum_i m_i}.$$  

(16)

Intersection and union trees can still be constructed in the same way. On the other hand, $K$-cores need careful consideration because they can be redefined in a wide variety of ways. Due to the limitation of space, this paper just suggests two options: one is to visualize the value of $s(\Omega_i, Q_j)$ at $\Lambda$, and the other is to show the $K$-cores extracted from original point distributions.

7. Application

This section applies the proposed method to school location planning in Inage and Wakaba wards in Chiba City, Japan (Figure 6).
There were 16 and 20 public elementary schools in 2008 in these wards, respectively (Figure 7). With a rapid decrease in birth rate, however, children of elementary schools have been decreasing since 1981. School reduction has been being discussed to improve educational environment of schools and economic efficiency of school operation.

Future plans of school location were derived as the solutions of spatial optimization problem. The minimum numbers of schools were calculated that satisfied the given constraints including the maximum distance from home to school, the minimum and maximum capacity of schools, and so forth.

The solutions were 11 and 15 schools in Inage and Wakaba wards, respectively. Since the constraints were not so restrictive, the solutions could be achieved by numerous combinations of 11 and 15 schools chosen from existing ones. Fifty and a hundred alternatives were extracted in Inage and Wakaba wards, respectively, in the ascending order of average distance from home to school.

The proposed method was utilized to devise a few desirable plans from the alternatives. The alternatives were classified into several groups and their geographical properties were described by $K$-cores. The results were used for discussing which schools should remain in future.

Analysis was performed based on the original point distributions and surfaces generated by smoothing operation. The spatial proximity was defined by

$$\rho(Q_i, Q_j) = \begin{cases} 1 & \text{if } D(Q_i, Q_j) \leq D_{\max}, \\ 0 & \text{otherwise} \end{cases}$$

where $D(Q_i, Q_j)$ was the graph distance (minimum length of the shortest path; see Diesel, 1997) between $Q_i$ and $Q_j$ on Delaunay triangulation generated from $\Lambda$. The maximum distance $D_{\max}$ was set to 1 or 2.

Table 1 shows the summary statistics of the results. Diversity of original distributions is one in both wards indicating that there is no school shared by all the plans. Weighted diversity shows similar values to those of ordinary measures, while spatial smoothing greatly reduces diversity among point distributions. Though weighting and smoothing are both extensions of the ordinary diversity measure, they are different in that the former evaluates points at different locations while the latter considers points in other distributions. In this empirical study, it is shown that smoothing is more effective to obtain a significant result.

School plans were then classified into several groups by using intersection tree. The distance between plans was evaluated by

$$d(\Psi_i, \Psi_j) = d_s\left(\Psi_i, \Psi_i \cap \Psi_j\right) + d_s\left(\Psi_j, \Psi_i \cap \Psi_j\right) \frac{n(\Psi_i \cap \Psi_j)}{n(\Psi_i) + n(\Psi_j)}.$$
Table 2 shows the summary of classification results where four and nine groups were obtained. As seen in the table, the number of school plans in each group considerably varies among three cases in both wards. In cluster analysis, it is not desirable to obtain one large and many small groups because it does not tell much about the property shared in the same group. In this sense, four-group classification seems better in Inage ward because nine groups include many small ones. In Wakaba ward, on the other hand, nine-group classification is better where school plans are more uniformly classified.

Figure 8 shows the dendrograms of school plans in Wakaba ward. In Figure 8a, groups A, B, C, and D are connected by dashed lines because their intersection is empty. Among the three cases, classification based on the original distributions seems most effective because of the following two reasons. First, it yields plan groups whose variation is small in size. This applies to classifications into both four and nine groups. Second, plan groups are clearly separated from each other. This is confirmed by the fact that the distance between plan groups is larger in Figure 8a than those in Figure 8b and 9c. A comb-like dendrogram as seen in Figure 8c implies difficulties in distinguishing groups with each other.

To describe the geographical properties of school plans shared in each group, 1-, 2- and 3-cores were extracted as follows. Delaunay triangulation was generated from school locations to extract the pairs and triplets of schools located within 800m connected directly by links. After extracting all the 1-cores, the 2- or 3-core having the largest amount of information was chosen from all the pairs and triplets. The next $K$-core was chosen from those not overlapping with the first one. This process continued until no $K$-core remains.

Figures 10-12 show the $K$-cores of each group in Wakaba ward. Plans were classified into nine groups, among which those consisting of more than nine plans are presented.

$K$-cores indicate school locations shared by many plans within each group, i.e., local similarities among school plans. Groups containing many $K$-cores consist of school plans similar in school location. In Figure 10, for instance, groups a, e and f have only six or seven $K$-cores while groups b, g, and c have ten $K$-cores. This implies that larger groups have greater variation in school plans within each group. It is quite reasonable, which is also confirmed by the fact that larger groups contain fewer 1-cores than smaller groups. Among $K$-cores, those consisting of fewer locations indicate greater similarity in school location.

Concerning 1-cores, they are observed in Figure 9 more frequently than in Figures 11 and 12. School plans share more locations within each group in Figure 9. This is mainly because spatial smoothing permits a difference in location. Since 1-
cores are defined by points exactly located at the same location, they hardly exist in Figures 11 and 12 in which groups are obtained based on smoothed distributions.

From Figure 8-12, it is concluded that classification based on the original point distribution gives the best result among three alternatives. It classified school plans into groups of similar size each of which shares many school locations.

In actual school location planning, the above result can be utilized as follows. In Japan, educational planning is usually discussed by local administrators, teachers, and parents of students. It is a collaborative decision making, where many options are discussed from various points of view. Since the school plans derived as the solutions of spatial optimization problem reflect a limited aspect of educational environment, they are not considered as the best solutions to be chosen. Nevertheless, they are very helpful for understanding the spatial conditions imposed on school locations, say, school at 1-cores are indispensable to keep the physical environment of students. Consequently, in Wakaba ward, schools represented by 1-cores may be extracted first in each group shown in Figure 9. Then, one school is chosen from every 2\(_c\)- and 3\(_c\)-core. Finally, schools are chosen from 2\(_o\)- and 3\(_o\)-cores and others up to 16 in such a way that the given constraints are satisfied. The results are compared among different groups in other aspects such as the racial composition, minority distribution, and school quality to make a final decision. Thought this process is not completely logical, its flexibility is often effective for reaching the agreement.

8. Conclusion

This paper has proposed a new method of analyzing the relations among point distributions on a discrete space. Spatial topology and proximity are discussed in an integrated framework at various scales, from local to global. Geographical representations such as \(K\)-cores permit us to analyze the similarity among distributions at a local scale, while Hasse diagram and intersection tree are useful for grasping the global structure of the relation among distributions. The latter also provides a means of classifying the point distributions into similar groups. The method was applied to school location planning in Chiba City, Japan. The result revealed the properties of the method as well as illustrated its utilization in school location planning.

We finally discuss some limitations and extensions of the paper for future research. First, this paper adopts a hierarchical method to construct intersection trees. It is primarily because implementation is straightforward and result is easy to interpret. One drawback is that such a hierarchical method does not assure the global optimality of result. Other non-hierarchical and heuristic methods such as \(k\)-means can also be
considered. Second, this paper focuses on points distributed on a discrete space. This is a reasonable setting in urban environment where land blocks are predetermined by land ownership system. In ecology and epidemiology, however, it is more realistic to assume a continuous space in point distributions. Extension in this direction is indispensable to treat a wider variety of point distributions. Third, the proposed method does not discuss the statistical significance of the result of analysis. One option to solve this problem is compare the observed result with those obtained from a Monte Carlo simulation that randomizes the point locations as the realization of null hypothesis. However, its computational cost is not negligible, especially when generating intersection trees and classifying point distributions. Efficient algorithms as well as analytical approaches should be sought.
References


Table 1 Summary statistics of school plans in Inage and Wakaba wards.

<table>
<thead>
<tr>
<th>Ward</th>
<th>Number of plans</th>
<th>Number of schools</th>
<th>Point distributions</th>
<th>Diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ordinary</td>
</tr>
<tr>
<td>Inage</td>
<td>50</td>
<td>11</td>
<td>Original</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Smoothed ($D_{max}=1$)</td>
<td>0.593</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Smoothed ($D_{max}=2$)</td>
<td>0.173</td>
</tr>
<tr>
<td>Wakaba</td>
<td>100</td>
<td>15</td>
<td>Original</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Smoothed ($D_{max}=1$)</td>
<td>0.521</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Smoothed ($D_{max}=2$)</td>
<td>0.214</td>
</tr>
</tbody>
</table>
### Table 2 Classification of school plans in Inage and Wakaba wards.

<table>
<thead>
<tr>
<th>Ward</th>
<th>Distribution</th>
<th>Classification into 4 groups</th>
<th>Classification into 9 groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number of school plans in each group</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classification into 4 groups</td>
<td>Classification into 9 groups</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18, 13, 11, 8</td>
<td>9, 8, 7, 7, 5, 5, 4, 3, 2</td>
</tr>
<tr>
<td></td>
<td>Smoothed (D_{max}=1)</td>
<td>39, 5, 3, 3</td>
<td>29, 7, 4, 3, 2, 2, 1, 1, 1</td>
</tr>
<tr>
<td></td>
<td>Smoothed (D_{max}=2)</td>
<td>21, 16, 12, 1</td>
<td>14, 11, 9, 6, 5, 2, 1, 1, 1</td>
</tr>
<tr>
<td>Inage</td>
<td>Original</td>
<td>54, 19, 16, 11</td>
<td>24, 13, 13, 11, 11, 9, 8, 6, 5</td>
</tr>
<tr>
<td></td>
<td>Smoothed (D_{max}=1)</td>
<td>56, 31, 12, 1</td>
<td>29, 18, 13, 12, 11, 10, 5, 1, 1</td>
</tr>
<tr>
<td></td>
<td>Smoothed (D_{max}=2)</td>
<td>77, 15, 6, 2</td>
<td>52, 15, 13, 6, 5, 5, 2, 1, 1</td>
</tr>
<tr>
<td>Wakaba</td>
<td>Original</td>
<td>54, 19, 16, 11</td>
<td>24, 13, 13, 11, 11, 9, 8, 6, 5</td>
</tr>
<tr>
<td></td>
<td>Smoothed (D_{max}=1)</td>
<td>56, 31, 12, 1</td>
<td>29, 18, 13, 12, 11, 10, 5, 1, 1</td>
</tr>
<tr>
<td></td>
<td>Smoothed (D_{max}=2)</td>
<td>77, 15, 6, 2</td>
<td>52, 15, 13, 6, 5, 5, 2, 1, 1</td>
</tr>
</tbody>
</table>
Figure 1 Point distributions in a one-dimensional space. Black dots aligned on horizontal lines indicate the location of points in each distribution from $\Omega_1$ to $\Omega_{12}$. Black dots on the same vertical lines share the same location. Relations between a pair of point distributions is (a) equal, (b) inclusive, (c) exclusive, and (d) overlapping. These four relations are mutually exclusive. Distributions in (d) are also complete, while those in (e) are incomplete. Distributions in (f) are called complementary because they are both complete and exclusive.
Figure 2: Point distributions in a one-dimensional space. Black dots aligned on horizontal lines indicate the location of points in each distribution from $\Omega_1$ to $\Omega_{12}$. Black dots on the same vertical lines share the same location. Relations between a pair of locations is (a) equal, (b) inclusive, (c) exclusive, and (d) overlapping. Locations in (d) are also complete, while locations in (e) are incomplete. Locations in (f) are called complementary because they are both complete and exclusive.
Figure 3 Point distributions in a one-dimensional space. Black dots aligned on horizontal lines indicate the location of points in each distribution from $\Omega_1$ to $\Omega_7$. Black dots on the same vertical lines share the same location. Sets of points above are either 1-, 2- or 3-core in $\Lambda$. 
Figure 4 Relationship among the number of elements in sets, distance measures and binary operations.
Figure 5 An intersection tree constructed from $\Omega_1$, $\Omega_2$ and $\Omega_3$. White and black circles indicate original distributions and intersections replacing the most similar pair of distributions, respectively. Links represent intersection operation applied to point distributions. The similarity between distributions is measured by $d_S(\Omega_i, \Omega_j) + d_I(\Omega_i, \Omega_j)$. Distributions $\Omega_1$ and $\Omega_2$ are chosen first to generate $\Omega_1 \cap \Omega_2$, and then $\Omega_3$ and $\Omega_1 \cap \Omega_2$ yield their intersection $\Omega_1 \cap \Omega_2 \cap \Omega_3$. 
Figure 6 Inage and Wakaba wards in Chiba City. They are located 40-60 minutes from the central area of Tokyo. Residents are working in Tokyo and Chiba downtown.
Figure 7 Public elementary schools and density distribution of children aged 6-12. Children concentrate around Chiba downtown and railway stations.
Figure 8 Dendrograms of school plans in Wakaba ward. School plans were classified based on (a) original distribution, (b) smoothed distribution ($D_{\text{max}}=1$), and (c) smoothed distribution ($D_{\text{max}}=2$). $n_p$ is the number of plans in each group. Black and white circles indicate groups of school plans where a hundred plans are classified into four and nine groups, respectively. The vertical axis $d(\Psi_i, \Psi_j)$ is the distance between sets of distributions when clustered into one group. In dendrogram (a), groups A, B, C, and D are connected by dashed lines because there is no point shared by all the distributions.
Figure 9 $K$-cores in school plans in Wakaba ward. Classification is based on original distributions. Groups consisting of more than nine plans are presented. Labels under the figures such as (a) and (c) correspond to those representing nine groups from (a) to (i) in Figure 9a. Large black dots indicate 1-cores. Small black dots encircled by solid lines are $2_c$- and $3_c$-cores. Small black dots encircled by dashed lines are $2_o$- and $3_o$-cores.
Figure 10 $K$-cores of school plans in Wakaba ward. Classification is based on smoothed distributions ($D_{\text{max}} = 1$). Labels under the figures correspond to those representing nine groups in Figure 9b. The same symbols are used as in Figure 9.
Figure 11 $K$-cores in school plans in Wakaba ward. Classification is based on smoothed distributions ($D_{max}=2$). Labels under the figures correspond to those representing nine groups in Figure 9c. The same symbols are used as in Figure 9.