Subdivision of Polygon Parcel in Land-Use Data Generalization

Al Tinghua

1 Introduction

Multi-scale representation of GIS data is the basis of seamless data navigation, progressive web transfer, self-adaptable visualization and other applications. Recently it gains considerable concern in such fields as map generalization and spatial data mining.

In order to obtain geographic data representation suitable for multiple resolutions, one has to build hierarchical spatial partitioning and clustering. The GIS user hopes to obtain more details during continuous zoom-in rather than just the exaggeration of graphic symbols. For category polygon map, such as land-use, vegetation, soil etc., the progressive zoom-out requires to remove insignificant parcels. To resolve this problem, two questions need to be answered: Which is unimportant or less important? and which is the most compatible neighbor of unimportant polygon? The former question is resolved in accordance with object attribute and geometrical properties, e.g. size, and the latter is based on the computation of topological neighbor relationship and semantic similarity.

How to remove the insignificant parcel? Traditionally, the generalization method is to completely assign the small parcel to one of its neighbors. This study tries to improve the generalization by separating the insignificant parcel into parts around the weighted skeleton and assigning these parts to different neighbors. The distribution of the weighted skeleton depends on the compatibility between the removed object and its neighbor, which considers not only topological relationship but also distance relationship and semantic similarity. This process is based on the Delaunay triangulation model. This paper gives the detailed geometric algorithms for this operation.
has to be removed. The higher the collapse value of a neighbor is, the more it pushes away the skeletons (and the larger piece of the object will be assigned to this neighbor). The interpolated constrained Delaunay triangulation is used as a basis for computing the adjusted skeleton.

2 Neighborhood analysis

Each object in GIS has a semantic description (at least a classification or feature category) and a spatial description. Selecting all objects from a single feature category results in a sub-set of the objects within the planar partition. To represent and compute the neighborhood for one category of objects, the topological model which has been embedded in database will not be strong enough and a new model has to be built. The Voronoi diagram has the geometric properties of equally separating space between the objects and can be used to support disjoint polygonal cluster analysis.

One polygon object, on the one hand, has topological neighbors with possible other category in whole space. On the other hand, it has geometric neighbors with the same category in sub-category-space. As far as the adjacency degree is concerned, the former is spatially closer but semantically farther, the latter is semantically closer but spatially farther. Which is the most compatible neighborhood object? Van Oosterom (1991, 1994) developed the compatibility formula to resolve this question. But the impact of distance should be also included.

The geometric neighborhood representation assigns those objects with distance to each other less than the tolerance (buffer zone) into a set called geometric neighborhood set. If we expect to further distinguish the elements of the neighborhood set in distance relationship representation, the tolerance distance, which is responsible for the neighborhood determination, can be adjusted to smaller. This principle implies that the topological neighbor has no privilege in the neighborhood representation, although it has the least distance, zero to the referenced object. On the basis of this idea, the topological and the geometric neighborhood are unified. Two polygons sharing the same boundary can be thought of having zero distance to each other. Next we consider the semantic contribution in neighborhood identification. We define the following principle; within a neighborhood cluster, the closer the semantic relationship between object elements is, the closer the neighborhood relationship to each other is. In Fig. 1, an example of a neighborhood cluster representation is given: parcel a has neighbors d, c, f, b, h. When comparing two candidate parcels b and f with a, we find that parcel b has the same category with parcel a, so it has closer neighborhood than parcel f to parcel a, although parcel f touches parcel a. This is just an intuitive judgment and the strict judgment depends on compatibility computation.

In the neighborhood definition, the representation resolution has to be considered. The above determination of tolerance distance is the behavior of spatial resolution. The semantic closeness is also decided by the semantic resolution of the hierarchical classification tree. For example, in land-use hierarchical category, each super-category contains several basic categories. In a higher resolution representation, every basic category is able to be distinguished from each other and the semantic neighborhood is strict in the feature identification. But in a lower resolution representation, only super-category is distinguishable and the basic categories under it are assigned into
one super-category. The large tolerance distance implies that the neighborhood identification is loose in geometry, and the neighborhood strength is mainly decided by the semantic resolution. The low semantic resolution implies that geometric distance and the area between polygons mainly determine the neighborhood. We can stress one of two facts in the compatibility formula to obtain the neighborhood that we need in a specific application.

In the neighborhood definition for disjoint objects, another consideration is the context influence. For a large neighborhood tolerance distance, two objects having less distance are regarded as strong neighbors. But when the third object of another category not from the "other land" (or "unclassified" or "empty space") type appears between them and breaks their connection, they are no longer having neighborhood relationship even they remains the original position. It means that we can not simply use the metric distance to find neighbor, as the buffer operation in GIS software. In the polygonal cluster distribution, each element competes to release its influence region to obtain parts of the empty space and the competition result can be represented as the equally partitioning geometric construction, the Voronoi diagram. The neighborhood with the consideration of the corresponding context can be based on the topological touch relationship between partitioning polygons in the Voronoi diagram.

On the basis of the analysis above, we give the following neighborhood definition: Objects have the neighborhood relationship to each other if they satisfy one of the following conditions:

1. having topological touch relationship to each other;

2. among sub-category space, their Voronoi cell polygon shares an identical boundary and the distance between objects is less than the predefined tolerance.

Furthermore, the neighborhood strength is decided by the semantic adjacency and the length of the shared boundary, which is not only between objects themselves but also between the touching cell polygons in the Voronoi diagram for disjoint objects.

According to this definition, in Fig. 1, parcel $a$ has neighbors $b$, $c$, $f$, $d$ and this order also represents the neighborhood strength. Parcel $g$ is of the same category but is far away from parcel $a$, so it can not act as the neighbor of $a$. Although parcel $h$ is of the same category with $a$ and locates within the buffer region of parcel $a$ (visualized as dash line), the appearance of parcel $b$ makes its corresponding Voronoi diagram partitioning polygons untouched and, therefore, $h$ is not the neighbor of parcel $a$.

3 Subdivision of insignificant polygon

On the basis of the analysis above, we can obtain the neighbors of each polygon parcel in the polygon map. How to process the relationship between current parcel and its neighbors in the map generalization? In the traditional generalization of polygon cluster, we apply the progressive method which removes the most insignificant polygon object among the current remaining objects step by step. As a result, the space will be filled by its most compatible neighbor. This process can also be understood as the amalgamation operation in the case of two semantically different objects. In GAP-tree method, this operation is based on the topological structure and semantic compatibility computation, and the insignificant object will completely be assigned to one of neighbors. We can not always find an obviously strong neighbor to completely receive the insignificant object. From the point of view of error adjustment, separating the insignificant polygon around its skeleton and assigning different parts respectively into different neighbors is better than assigning the whole polygon into one of the neighbors. Bader and Weibel (1997) presented an idea to
separate insignificant parcels in such category map generalization based on the constrained Delaunay triangulation model [33]. But if the insignificant object has not an obvious linear stretch distribution, the method cannot guarantee that the terminal point of skeleton exactly meets one of intersection points between two neighbors. The adjustment of skeleton is required. Next we will provide the detailed method of this amalgamation operation with the geometric algorithm.

The basic idea of this method is to generate skeleton within the insignificant polygon and then to organize the polygon composed of some skeleton edges and one boundary edge. The generated polygons are respectively assigned to the neighbor polygon on the other side of the boundary edge. We use Delaunay triangulation to support this partitioning. In order to seamlessly partition the insignificant polygon, we expect that the terminal of each skeleton edge meets the node of the boundary edge. Supposing in the polygon-edge-node topological structure, there is no pseudo node, it means each node must be related to at least three edges. Obviously, directly extracting skeleton on the basis of all triangles within one polygon cannot guarantee the above conditions (Fig. 2(b)). We use the following two steps to adjust the skeleton edges.

All triangles linking the same boundary edge (that is, with all three points on the same boundary of one neighbor polygon) are removed, as shown in Fig. 2(c). According to the part of triangles within the polygon constructing the new skeleton edge, we now can guarantee that each terminal exactly meets one node, not being located in the middle of the boundary edge. In Fig. 2(c), the grey area in which triangles are removed serves as the concave part of one neighbor polygon. Since this part is strongly surrounded by one neighbor polygon, in the partitioning it should not act as a competition region for different neighbors to separate. According to the “nearest connection” of the Delaunay triangulation, the darker area covered by the remained triangles belong to the competition region. It has to be separated by the skeleton edge between the neighbors connected by the triangles.

The skeleton edge begins from the triangle with one neighbour, which appears in some places of boundary which is not smooth. If the angle between two boundary edges in some nodes approximates to 180°, there is perhaps no this kind of triangle, (see \( E_1, E_2, E_1, E_4 \) in Fig. 3). It implies that no skeleton edge starts from these nodes according to the normal skeleton generation. As in Fig. 3, we can just obtain one skeleton edge from \( E_0 \) to \( E_5 \). Additional skeleton edges are needed. From this kind of node, search the shortest link to the point in the existing skeleton edge along the triangle edge. Some nodes relate to just one triangle edge intersecting with the existing skeleton edge, such as \( E_1, E_4 \) in Fig. 3. Others may be related to more than one, such as \( E_1, E_4 \) in Fig. 3. Select the shortest link (half of one triangle edge) and let it act as the added skeleton edge. The original skeleton edge needs to be clipped as two parts in the new added link node when a new skeleton edge is added (Fig. 3).
If one neighbor wants to receive part of an insignificant polygon, it must share the common boundary with some length. For example, in Fig. 3, parcel c shares just one point E3 with the insignificant parcel a, it will not collapse with the part of parcel a either between E3 and E4 or E3 and E2, though the spatial relationship parcel c is the neighbor of parcel a. Through these operations, for the matching relationship between boundary nodes and skeleton terminals, we can exactly guarantee the conditions discussed in the beginning of this subsection. Select all the boundary edges of the insignificant polygon and all its skeleton edges to organize the polygons. Each boundary edge will make up one polygon with other skeleton edges. And the generated partitioning polygon will be assigned to the neighbor of the boundary edge on other side. This amalgamation maintains the topological consistency, not resulting in gap regions or overlapped regions. Fig. 4 is an example of this amalgamation. In the left, the parcels surrounded by the wide dark boundaries are unimportant (small size) and will be removed. The right graphic is the result from subdividing the insignificant parcels. Different parts have been assigned to different neighbors and visualized with the same color as the corresponding neighbor.

The skeleton generation above is based on the supposition that the neighbors have the same collapse abilities to separate the insignificant object and the skeleton edge partitions left/right region in an equal way. Further improvement can be performed when considering geometric relationship and semantic similarity. We use function $\text{collapse}(a, b)$ to compute the collapse ability between object a (to be removed) and its neighbor b. This function has to consider the following aspects:

1. compatibility of the types: $\text{type_compat}(a, b)$;
2. the length of the shared edge (either real shared edge or shared edge as part of the skeleton within buffer distance): $\text{length}(a, b)$;
3. the weighted distance between a and b: $\text{dist}(a, b)$;
4. the importance of b: $\text{Imp}(b)$.

The function can be written as:

$$\text{collapse}(a, b) = \text{type_compat}(a, b) \times \text{length}(a, b) \times \text{Imp}(b) / (1 + \text{dist}(a, b))$$
Earlier researches have made the discussion of \( \text{type}_\text{compat}(a,b) \) and \( \text{imp}(b) \) \(^{2,4}\). We call this kind of skeleton weighted skeleton, since it considers the collapse difference between regions competing for extending area.

In the geometric algorithm, the weighted skeleton can be adjusted on the basis of the normal skeleton. In Fig. 5, the position \( P_i \) needs to be computed according to two segment ratios based on the collapse values, rather than the middle position unless the two collapse values are the same. For example, in Fig. 5, parcel \( c \) has collapse value 3 and \( b \) has 7. For the region between parcel \( c \) and \( b \), 30 percent will be assigned to parcel \( c \) and 70 percent will be assigned to parcel \( b \).

![Fig. 5 Generation of weighted skeleton](image)

4 Conclusions

On the basis of the skeleton handling, this paper presents the improvement of amalgamation between different category objects. The extensions use the Delaunay triangulation. This strategy is very suitable for progressively transferring polygon map in Internet. Further research work in this field should include:

(1) The improvement of neighborhood definition, simultaneously considering topological, geometric, semantic and Gestalt nature’s impacts. The compatibility formula between two neighbors has been defined. There is need to further study the way to compute the parameters as well as the way to find the better definition.

(2) The application of this method to a web transferring environment, especially in a server-client framework, to realize the transfer of vector map in the way from coarse to refine-

References

1 van Oosterom P (1991) The reactive-tree: a storage structure for a seamless, scaleless geographic database, Auto-Carto 10.
2 van Oosterom P (1994) Reactive data structure for geographic information systems. Oxford: Oxford University Press.
3 Badar M, Weibel R (1997) Detecting and resolving size and proximity conflicts in the generalization of polygonal maps. The 18th ICC, Stockholm, Sweden.
4 van Putten J, van Oosterom P (1998) New results with generalized area partitions. The 8th International Symposium on Spatial Data Handling, Ottawa.
5 Hernandez D, Clementini E (1995) Qualitative distance, COSIT'95, Semmerring, Austria.
6 Jones C B, Bundy G L, Ware J M (1995) Map generalization with a triangulated data structure, Cartography and GIS, 22(4):317-331
7 Li C, Chen J, Li Z L (1999) A raster-based algorithm for computing Voronoi diagrams of spatial objects using dynamic distance transformation, International Journal of Geographic Information Sciences, 13(3):209-225
8 van Oosterom P (1995) The GAP-tree, an approach to on-the-fly map generalization of an area partitioning. In: Muller J C, Lagrange J P, Weibel R, eds, GIS and Generalization: Methodology and Practice, London: Taylor & Francis.
9 Ai T H, Guo R Z, Liu Y L (2000) A binary tree representation of bend hierarchical structure based on Gestalt principles. The 9th International Symposium on Spatial Data Handling, Beijing.
10 Ai T H, Guo R Z (2000) A constrained Delaunay partitioning of areal objects to support map generalization. Journal of Wuhan Technical University of Surveying and Mapping, 25(1):35-41 (in Chinese)
11 Peng W, Muller J C (1996) A dynamic decision tree structure supporting urban road network automated generation, The Cartographic Journal, 33(1):5-10
12 Peng W (1995) Automatic generalization in GIS. ITC Publication Series.
13 Poorten P, Jones C B (1999) Customisable line generalization using Delaunay triangulation. The 19th ICC, Ottawa.
14 Preparata F P, Shamos M I (1985) Computational geometry an introduction. New York: Springer-Verlag.