CO-182

A CONSISTENCY MAINTENANCE OF SHARED BOUNDARY AFTER POLYGON GENERALIZATION

AI T.(1), ZHANG W.(2)

(1) Wuhan University, WUHAN CITY, CHINA ; (2) Zhongnan University of Economics and Law, WUHAN CITY, CHINA

ABSTRACT

For area spatial data, such as land-use, vegetation and soil class, the boundary should be consistent and all polygons should cover the whole region with neither gap nor overlapped area. However the map data processing such as map generalization usually leads to the inconsistency due to the independent simplification of polygon boundaries. A post-correction is required to keep the spatial relation, to maintain the consistency between neighbouring polygons. This study aims at this question investigating three inconsistent cases resulted from polygon generalization, which are intersection, separation and interlacing inconsistency. Considering the different importance of two neighbor polygons locating beside the corrected boundary, the correction is developed by two approaches: less important polygon to snap its neighbor, and equal polygons to move together to adjust their shared boundary. So 3*2 methods have been built in this study for three different inconsistent questions. The algorithms apply Delaunay triangulation model to detect inconsistent area among shared boundaries, and further use triangulation skeleton to produce common boundary.

KEYWORDS

Polygon map, map generalization, spatial consistency, Delaunay triangulation

1. INTRODUCTION

After more and more spatial data infrastructures are completed, the integration of heterogeneous databases becomes an important issue. In the applications of spatial data update, multi-scale database creation, interoperation and web deliver, the integration of spatial data plays a significant role not only to derive combined abstracted information but also to extract differences between various data. The up-to-date data is usually integrated to the existed database to update the existed database by map generalization.

Database integration and transformation results in the matching question requiring some works at the schema level and some works at the data level, especially in the context of geographic data(Revell and Antoine, 2009). The inconsistency detection and adjustment is such an important work when integrating and matching different data. During the data integration, the data from different sources may contradict to each other in geometric representation, topological relationship or semantic description (Safra et al 2010). The same object in the real word may be represented in quite different ways, such as geometric dimensions, abstraction levels, semantic hierarchies and other properties. The consistency of shared boundaries of neighboring spatial entities may be destroyed when integrating different data.

On the other hand, the logical consistency is one of five aspects of spatial data quality (Goodchild 1991). To preserve the consistency becomes an important maintenance in database construction. A few methods have been developed to detect and adjust inconsistency when data matching in the field of spatial data handling. Among them the matching of different scale spatial data attracts more interests to build the associations between less detailed data and more detailed data (Revell and Antoine 2009, Safra et al 2010, Volz 2006). The road network is especially an active study focus in this domain(Zhang and Meng 2007, Mustiere and Devogele 2008).

The inconsistency results from different reasons, e.g. the different representations (the river is represented by a shallow polygon or a collapsed axis line)(Liu and Chen 2008), the database construction at different time by different agencies, or the cognition from different viewpoints. Correspondingly the consistency maintenance has to be performed aiming at different data integration situations, including the integration of (1) different scale data in same region, (2) different semantic description under same environment, (3) data from different domains, (4) associated data of different features, and so on.

In this study we focus on the consistency of shared boundaries of neighboring spatial entities investigating the spatial relation maintenance after polygon data generalization. For area spatial data, such as land-use, vegetation and soil class, the boundary should be consistent and all polygons should cover the whole region with neither gap nor overlapped area(Ai and Oosterom 2002). However the map data processing such as map generalization usually leads to the inconsistency (Muller and Wang 1992). A common question is that in the process of adjacent polygon generalization and integration, neighboring polygon overlay is not able to strictly guarantee the topological consistency but generate series of fragment area and gap area in overlap region. To correct the inconsistency between polygon boundaries, this study presents different solutions based on spatial neighbor relation analysis. Three cases of inconsistency are distinguished: intersecting, separating and interlacing. Intersecting inconsistency causes small fragment area and separating inconsistency causes small gap area. Since the correction of inconsistent shared boundary is to find single line which can approximate original two boundaries with high location precise, we develop the algorithms by Delaunay triangulation to generate such consistent lines.

The remainder of paper is organized as follows. Section 2 discusses the reason of inconsistency generation among neighboring polygons and presents the classification of inconsistencies. Section 3 investigates the correction methods aiming at three different inconsistent situations. The applications of the correction approaches are given in conclusion, section 4.

2. INCONSISTENCY ANALYSIS FOR POLYGON GENERALIZATION

Map generalization aims at the generation of abstracted spatial data by transformation from original large scale to target small scale to remove less important information. For area spatial data, such as land-use, vegetation and soil class, the generalization on one hand reduces the semantic resolution to abstract hierarchical levels, on the other hand simplifies the geometric shape and structure. The operations usually applied in polygon data generalization include the aggregation of polygon parcel at object levels and the simplification of boundary at geometric characteristic level(Ai and Oosterom 2002). The decision on which parcel is unimportant and which one combine the neighboring unimportant parcels need to consider the semantic difference, neighboring relationship and geographic context. The polygon aggregation, such as the GAP method (Oosterom 1995) does not result in inconsistent question. However the boundary simplification may generate the gap area or overlap area between neighboring polygons destroying original consistency. The main reason exists in the generalization aiming at independent polygons beside the shared boundary.

2.1 The inconsistency generation

The polygon data integration and generalization can generate the boundary inconsistency by different ways:

<u>*1) Independent polygon boundary simplification*</u>

Map generalization includes the decisions and operations at three levels: 1) selection of generalization operators, 2) choose of algorithms and 3) adjustment of parameters (Shea & McMaster, 1991). The above decisions and operations depend on the feature class, geometric characteristic of polygon data to be generalized. The shared boundary relates to two different feature class polygons, for example building class on one side and vegetation class on other side in land-use cover. As the building class belongs to the man-made feature with orthogonal geometric properties, while vegetation class belongs to nature feature with smooth boundary properties. Usually the boundary simplification has to select two different algorithms for two features. The result is that the generalized shared boundary does not keep the consistency leading to gap or overlapped fragment area between building and vegetation polygons. On the other hand, for different land-use class, we need to setup different generalization parameters to make important class with high accuracy and unimportant class with low accuracy. The setup of different parameters also generates the inconsistency question.

One polygon object, on the one hand, has topological neighbors with possibly other category in whole space. On the other hand, it has geometric neighbors with the same category in sub-category-space. In the topological data structure, one polygon is represented by series of arcs and each arc relates to one left polygon and one right polygon. In area data generalization, the simplification object is the complete polygon or the closed boundary rather than part of boundary arc. It means the shared boundary participates the simplification as part of role in two different closed boundary through two different times of operation. The result of independent simplification is that the shared boundary is not able to keep consistent as original as shown in figure 1. Alternatively, we can extract the arcs from the closed boundary by topological structure and simplify the shared arc once to keep the consistency. But this process strategy does not consider the properties of the whole polygon, such as the area size, the shape pattern, the context and others.

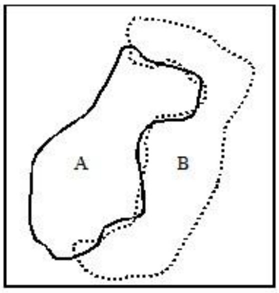


Fig 1. Boundary inconsistency caused by polygon simplification.

2) Narrow polygon collapse

In polygon data such as land-use, all polygons should cover the whole region with neither gap nor overlapped area. However, with the scale change those narrow parcels can not be represented by polygon geometry. From the perspective of map legibility, the polygon with too small width cann't be recognized. Alternatively, we need to transform the polygon representation to center-line representation by map generalization collapse. Once the narrow polygon is collapsed from two dimension area to one dimension line, the original place has a hole. It implies there is a gap area between the original polygons adjacent to the narrow polygon. In figure 2, the original road polygon is collapsed to center-line representation and two adjacent polygons becomes face to face with a gap connection. This situation of polygon cover should be modified to keep the original consistency.

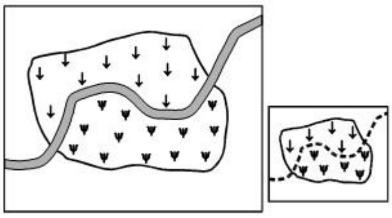


Fig 2. The collapse of road leads to the gap area between neighboring vegetation polygon in land-use data generalization

2.2 The inconsistency cases resulted from generalization

The generalization on neighboring polygons may result in the following three inconsistency cases:

1) Intersection inconsistency. After generalization two neighboring polygon generates the overlapped polygon with only two intersection points as shown in figure 3A.

2) Separation inconsistency. The generalization separates the original two neighboring polygons with the gap distance less than legibility tolerance as shown in figure 3B.

3) Interlacing inconsistency. After generalization two neighboring polygons intersect each other with more than two intersection points like a wrapped rope.

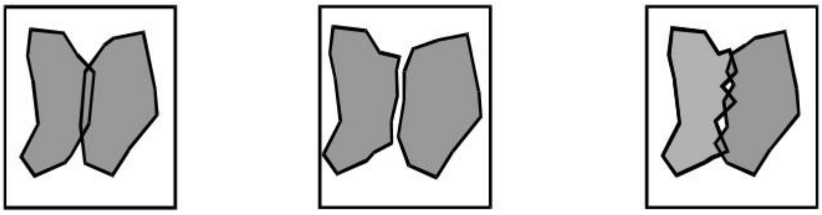
The intersection inconsistency generates overlapped fragments. The separation inconsistency generates gap area and the interlacing inconsistency results in both gap and overlapped fragments. In the integration or generalization of polygons, the resulted overlapped fragment has some properties which can be used to identify the fragments. Goodchild(1991) summarized as following:

___ Small size;

___ Long shape, the rate of perimeter to the root of area is big;

___ Only two arcs to compose the fragment and the arc associates with different polygons;

___ Each node relates to four arcs.



A. Intersection case　　B. separation case　　C. interlacing case

Fig 3. Three cases of polygon boundary inconsistency

3. CORRECTION METHODS

Since the generalization of polygon data can't guarantee the consistency of shared boundary between neighboring polygons generating gap or overlapped fragments, the post correction is necessary to maintain the spatial relationship. The correction of inconsistent shared boundary is to find single line which can approximate original two boundaries with high location accuracy. Considering the different importance of two neighbor polygons locating beside corrected boundary, we distinguish two correction strategies: less important polygon to snap its neighbor, and equal polygons to move together to adjust their shared boundary. We call the previous correction snap correction and the latter the adjustment correction. Aiming at three inconsistency cases, we build 6 correction methods based on two strategies.

3.1 Snap correction for intersection case

Two polygons are of different importance and the boundary of one polygon has to be kept without any change. This kind of correction is easy to carry out by polygon overlay operation. Suppose the important polygon is polygon *A* and unimportant *B*, the snap correction is to let *A* unchanged and *B* change to be *B-A* with part of polygon *A*'s boundary acting as the new shared boundary.

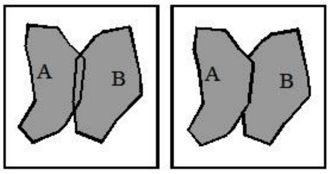


Fig 4. Snap correction for intersecting inconsistency between polygons.

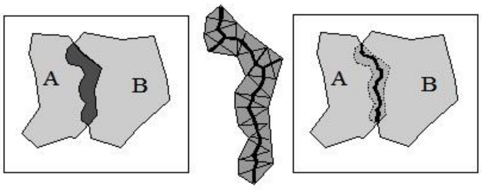


Fig 5. Adjustment correction for intersecting inconsistency between polygons.

3.2 Adjustment correction for intersection case

In the situation that two polygons are of equal importance, the corrected boundary should be equally displaced by an adjustment line. For polygons *A* and *B*, the intersection part $A \cap B$ can be extracted by polygon overlay operation. We let the skeleton of intersection $A \cap B$ act as the adjustment line. The skeleton extraction is based on Delaunay triangulation and for the detailed method see (Ai and Oosterom 2002). The skeleton directly extracted by Delaunay triangulation usually has some branches or hairs due to not smooth boundary of region $A \cap B$. To guarantee the single stretch line without branches, we remove some triangles before the skeleton generation only to remain the triangles locating in main stretch direction as shown in figure 6.

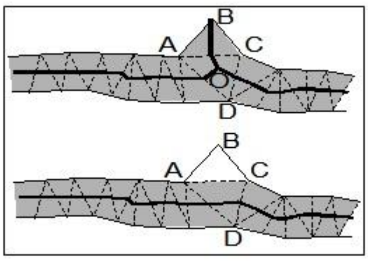


Fig 6. Analysis of hair skeleton producing.

3.3 Snap correction for separation case

The separation inconsistency between neighboring polygons has a gap area required to be detected. We use Delaunay triangulation again to find such adjacency area(Ware and Jones 1997, Ai et al 2000) . Construct Delaunay triangulation between two polygons as shown in figure 7. From outside remove those triangles with the triangle edge length longer than the connection tolerance bridging two polygons. In figure 7 the removal stops when two connections generates, say P_1P_1’, P_2P_2’, and then the adjacency region between two polygons is extracted. The snap correction for separation inconsistency can be conducted by the union of extracted adjacency region and the unimportant polygon, say polygon *B*, as illustrated in figure 8. It means in this correction the important polygon keep its boundary unchanged.

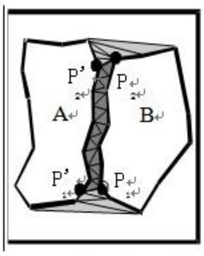


Fig 7. Extraction of adjacency area by Delaunay triangulation

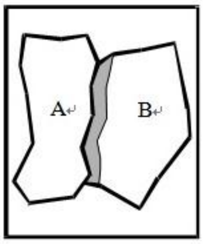


Fig 8. Snap correction for separation inconsistency

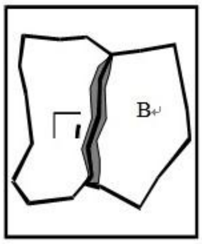


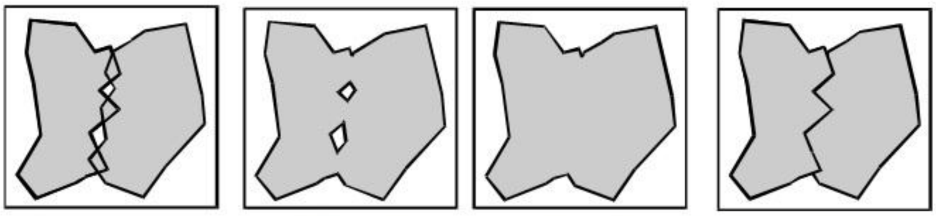
Fig 9. Adjustment correction for separation inconsistency

3.4 Adjustment correction for separation case

First the adjacency region extraction is similar to that of section 3.3. Since two polygons are of same importance. We can apply the skeleton of adjacency region acting as the correction boundary as shown in figure 9.

3.5 Snap correction for interlacing case

The interlacing inconsistency is the common case after polygon generalization. As for the snap correction, we first define a special union for polygon boolean operation: $A\cup^0 B = A\cup B\cup X$, where *X* is the islands of the union $A\cup B$. Then the correction can be performed by the difference operation $A\cup^0 B - A$. It means the important polygon A keeps its boundary unchanged while polygon *B* is clipped with common boundary with polygon *A*. The correction process is illustrated in figure 10.

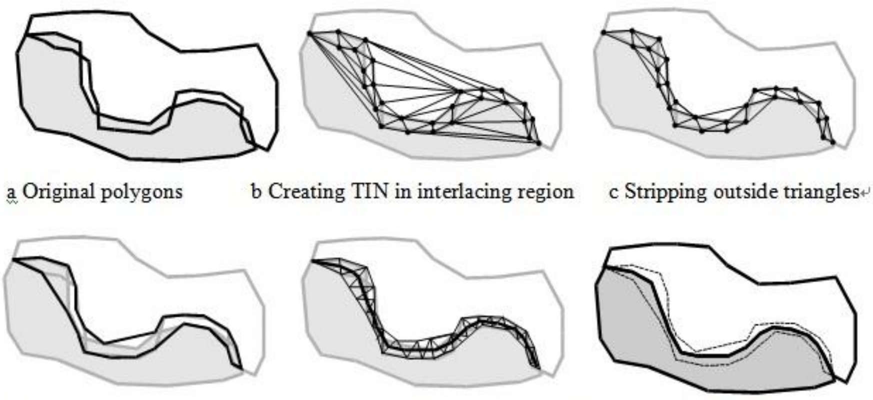


a. interlaced polygon A,B b. A∪ B c. $A\cup^0 B$ d. B changed to $A\cup^0 B - A$

Fig 10. Snap correction for interlacing shared boundary between polygons

3.6 Adjustment correction for interlacing case

Among six cases, this method is the most complex. First the contradict area needs to be extracted. Construct Delaunay triangulation in the two interlaced polygons as shown in figure 11a. Through triangle stripping operation, the inconsistent area can be extracted as set of triangles, like a sausage, as shown in figure 11c. Based on the triangle edge, the contradict area is generated as shown in figure 11d. Next use the same skeleton generation method as the above approaches to construct the center-line within the contradict sausage area. Finally use the center-line to replace the shared boundary between two polygons. The whole correction is illustrated with six steps in figure 11.



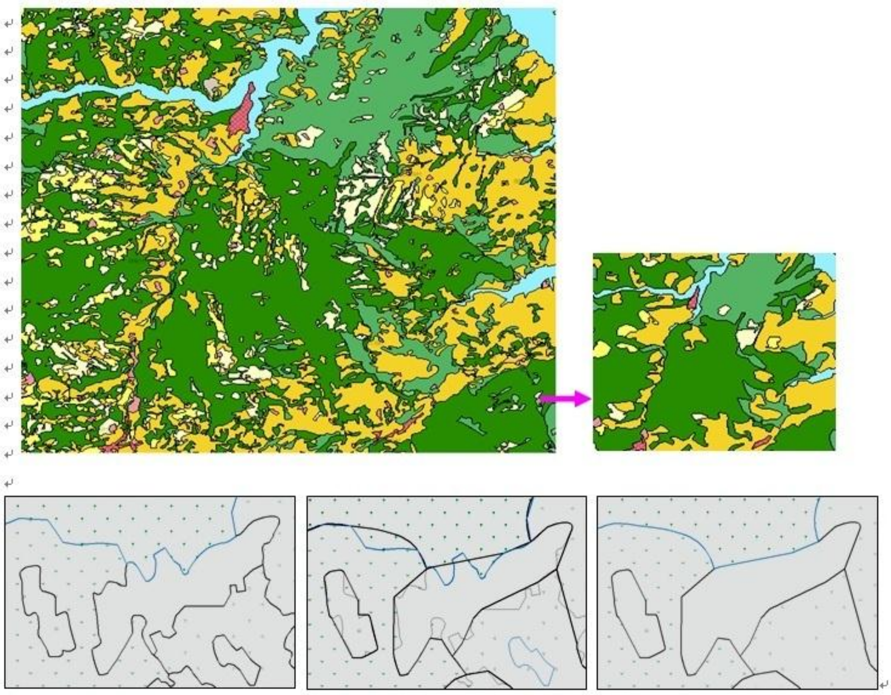
a Original polygons b Creating TIN in interlacing region c Stripping outside triangles

d. Extracting boundary e. Creating TIN again f. Generating skeleton of TIN

Fig 11. Adjustment correction for interlacing inconsistency between polygons

4. APPLICATIONS AND CONCLUSIONS

We apply the correction method in the maintenance works after the generalization of land-use data. The land-use map is a typical kind of categorical theme and has the properties covering the whole area with neither gap nor overlapping area. Usually it includes many categories in classification hierarchy. Its generalization has to consider geometric simplification, semantic hierarchy abstraction and statistical properties maintenance. The objects in this map are not all of the same importance. Some may be considered as foreground objects (important) while others may be considered as background objects (less important). Both snap correction and adjustment correction are necessary in land-use data generalization and consistency maintenance. Figure 12 is the illustration of land-use data generalization and consistency correction.



a. Original land-use polygon. b. After simplification of boundary c. After consistency correction

Fig 12. The application of consistency correction in land-use data generalization.

As one of important spatial data qualities, logical consistency plays a significant role in spatial analysis and decision. The consistency maintenance becomes a key step in map data processing. Several reasons in map data processing may destroy the consistency, for example map generalization. We can apply the generalization strategy which guarantees the consistency while the generalization is exercised at the same time. However, most consistency maintenances are required by post-process, such as the boundary consistency correction in land-use data generalization. To correct the inconsistency between polygon boundaries, this paper offers different solutions based on spatial neighbor relation analysis. Three cases of inconsistency are distinguished: intersection, separation and interlacing. Intersection inconsistency causes small fragment area and separation inconsistency causes short gap area. Correction of inconsistent shared boundary is to find single line which can approximate original two boundaries with high location precise. Considering the different importance of two neighbor polygons locating beside corrected boundary, we provide two correction approaches: less important polygon to snap its neighbor, and equal polygons evenly to adjust their shared boundary. In geometric operation, we use the polygon Boolean operation and Delaunay triangulation method to extract the skeleton to replace the shared boundary.

REFERENCES

1) Ai, T. and Oosterom, P. van. GAP-tree Extensions Based on Skeletons. In: Richardson D and Oosterom P van(eds) Advances in Spatial Data Handling, Springer-Verlag, Berlin, pp501-514, 2002.

2) Ai, T. Guo, R. and Liu, Y. A Binary Tree Representation of Bend Hierarchical Structure Based on Gestalt Principles, In: Proceedings of the 9th International Symposium on Spatial Data Handling, Beijing, pp2a30-43, 2000.

3) Goodchild M F. Issues of Quality and Uncertainty, in : Advances in Cartography, edited by Muller,J., published on behalf of the ICA by Elsevier Applied Science, 1991.

4) Jones C B, Bundy G L,Ware J M. Map Generalization with a Triangulated Data Structure ,Cartography and GIS, 1995, 22(4):317-331.

5) Liu W Z, Chen J, Deng K Z, Zhao R L, Cheng T. Detecting the Spatial Inconsistency between the Updated Rivers and Valleys. Journal of Image and Graphics,2008,13(5)::1003-1008.

6) Muller J C,Wang Z. Area-path Generalization: A Competitive Approach, The Cartographic Journal 1992, 29(2): 137-144.

7) Mustiere S, Devogele T. Matching Networks with different Levels of detail. GeoInformatic, 2008, 12(4)::435-453.

8) Oosterom P Van. 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. Taylor & Francis, London, 1995,pp 120-132

9) Revell P, Antoine B. Automated matching of building features of differing levels of detail: A case study. ICC. 2009.

10) Safra E, Kanza Y,Sagiv Y. Location-based algorithms for finding sets of corresponding objects over several geo-spatial data sets. International Journal of Geographical Information Science, 2010, 24 (1)::69-106.

11) Shea K S, McMaster R B. Cartographic Generalization in A Digital Environment: When and How to Generalize, Proceeding of AutoCarto , Baltimore, USA ,1991: 56-67.

12) Volz S. An iterative approach for matching multiple representations of street data. ISPRS workshop – multiple representation and interoperability of spatial data. 2006, Hanover, Germany.

13) Zhang M, Meng L. An iterative road-matching approach for the integration of postal data. Computer Environment and Urban Systems. 2007,597-615.