Available online at www.humg.edu.vn

ESASGD 2016

GIS-IDEAS (2016)

Conference Title: International Conference on GeoInformatics for Spatial-Infrastructure Development in Earth & Allied Sciences (GIS-IDEAS)

Fixing The Topology Rule Error of "Polygon Contains Point" in Processing Spatial Data

Quynh-An Tran[a], Quang-Minh Nguyen[b], Thi-Luong Le[a]

[a] *Hanoi University of Mining and Geology, Ha Noi, Vietnam*
[b] *Survey and Aerial Mapping One Member Limited Liability Company, Ha Noi, Vietnam*

Abstract

Establishing topology relationship among spatial data is an essential part in improving spatial data processing. The topology rule that a polygon must contain at least one point is one of the basic topology rules laid out by ESRI in ArcGIS platform. In case that a polygon object violates this topology rule, ArcGIS provides an auto-error correction tool. This article reviews the efficiency of the topology error correcting tool "polygon contains point" of ArcGIS, and proposes a solution to auto correct the error of "polygon contains point" on the basis of automatic recognition of the various types of polygons and outlines how to fix topology for each corresponding polygon type, resulting in higher efficiency in terms of map aesthetics compared with method used by ArcGIS.

Keywords: Spatial data processing; topology rule; polygon contains point.

1. Context

In a Geographic Information System, regional geographic entities are digitalized in a database as polygon type. During the process of establishing topology relationship, it is important to automatically generate a point object inside a polygon object. This is useful when every polygon should have at least one associated point, such as to assign address points for parcels, to identify the geographic entity, to connect a polygon with its attributes, to automatically display label, etc.

The inside point of a polygon needs to be at the geometric center of the polygon in order to meet the requirements of map aesthetics. The current algorithms of finding an inside point for a polygon only ensure that the point falls within the polygon, but does not guarantee that the point will be at the geometric center of the polygon, e.g. trianglizing algorithm (DING Yao 2008), (GAO Zhi-jun and WANG Ji-yao 2007), the algorithm to calculate inside points based on minimum bounding rectangle (MBR) (Cai Shaohuan, Din zhiyuan and Zhu Tao 1998), improvements algorithm MBR (LU Hao et al. 2013), which leads to poor visual effects.

ArcGIS is a software platform which can work with maps and geographic information. It also provides relatively adequacy of functionalities to create and fix the topology relationships. One of the topology rules that ArcGIS platform requires is that each polygon must contain at least one point from another feature class. The point must fall within the boundary of the polygon. In case of a polygon does not contain any point, it would be an error. The automatic error-correcting function in ArcGIS will fix this error using the tool "Create Feature". This tool will create a new point that lies inside the error-polygon. The newly created point is guaranteed to be

* Corresponding author. Tel.:+84-964 -889 -458
E-mail address: tranquynhanmdc@gmail.com.

inside the polygon feature. This process can be applied to one or many error-polygons (ESRI). However, similar to the existing algorithms, the algorithm used in ArcGIS, in some cases, such as a concave polygon, can create a point which is not at the geometric center of the polygon, thus failing to best meet the requirements of map aesthetics. Especially, in the case of hole-polygons, ArcGIS platform is unable to fix the error. In this regard, this article proposes a new method to solve the issue of creating point for polygon to bring higher aesthetic efficiency, suitable for all types of polygons. This method is based on the principle of automatic recognition of different types of polygon and applies suitable algorithm of finding inside point for each type of polygon.

2. Polygon types and auto-recognition methods

In terms of geometry, there are two ways of classifying polygons based on whether a polygon is convex or concave and whether a polygon contains holes or not. A polygon can be used to describe a parcel, a house, etc. A hole-polygon can be used to describe a parcel of land which contains a pond inside, or a rounding house with a courtyard in the middle, etc.

Therefore, in combination, there are four types of polygons: convex polygon without hole, concave polygon without hole, convex polygon with hole, concave polygon with hole. With existing algorithms, interpolating an inside point of a convex polygon is very simple by directly using the centroid point of a polygon as the inside point of the polygon. However, with the concave polygon or those containing holes, centroid points, in many cases, fall outside polygons. Finding an inside point of a polygon using different algorithms give different results, but can not resolve the aesthetic efficiency issue. It can be seen that with each different type of polygons, different methods need to be chosen to find out the best inside points. Consequently, before automatically calculating an inside point for a polygon, it is needed to automatically identify the type of the polygon.

2.1. Automatic identification of convex and concave polygons

The existing algorithms that automatically recognize convex and concave polygon include: angle magnitude method, topo mapping method, area splitting method, comparison method based on vector boundary deviation... (Wu Lixin and Shi Wenzhong 2003).

The algorithm of identifying convex or concave polygon that is used in this research is: Firstly, to calculate the area of the original polygon S_1, then eliminate a vertex of the polygon, calculate the area of the new polygon S_2, compare S_1 and S_2, if $S_2 > S_1$ then this is a concave vertex, i.e. the original polygon is concave; if $S_2 < S_1$, we remove the next vertex from the original polygon and follow the steps of calculation and comparison above. If any vertex is concave, immediately the polygon is proved to be concave polygon. If all the vertexes are convex, that is a convex polygon.

2.2. Automatic identification of hole-polygon

A polygon without hole will contain an inner zone which is completely interconnected, i.e. any point within the polygon boundary belongs to the polygon. Hole-polygon has more than one boundary, including internal and external boundaries. None-hole polygon and hole-polygon require different methods of finding inside points. For a none-hole polygon, an inside point is obviously located inside the polygon. However, for a hole polygon which has two or more boundaries, an inside point can lie inside external boundary but also inside the internal boundary, i.e. the hole. Therefore, the first task is to automatically recognize the hole-polygon, and then use suitable algorithm to find the inside points automatically.

ArcGIS is able to convert hole-polygon data from other formats, for example, convert the data format from Mirostation SE software to ArcGIS format keeping the original structure of the hole-polygon; In addition, ArcGIS is also able to digitalize multi-part polygon objects such as hole-polygons. However, in the process of creating topology rules for polygons, ArcGIS can only recognize none-hole polygon, and it also modifies the original data structure of hole-polygon to none-hole polygon, as illustrated in figure 1. This proves that ArcGIS is only able to create inside points for normal none-hole polygon, but can not handle hole-polygon.

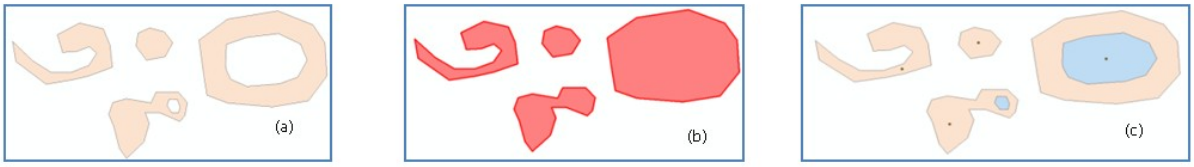

Fig. 1. (a) ArcGIS software can digitalize and display correctly the polygon data types converted from other softwares; (b) ArcGIS is unable to recognize hole polygon when creating topology rules and modifies the original input polygons; leading to (c) after error correction by ArcGIS, the newly created point is still outside the polygon (the top-right polygon: new point lies inside the hole).

In terms of data structure, there are three types of data structure: point, polyline and polygon. A point is stored independently, a polyline is formed by points, a polygon is formed by a closed polyline. In ArcGIS, a closed polyline is called Ring object, with the starting point and ending point are the same. Polygon objects can be formed by one or more rings, and a ring can be an inner or outer ring. The algorithm of recognizing hole-polygon used in this research bases on querying the number of inner rings in a polygon. If the returned result is zero, i.e. this is none-hole polygon; otherwise, it is the hole-polygon.

3. Auto-fix "polygon contains point" topology rule

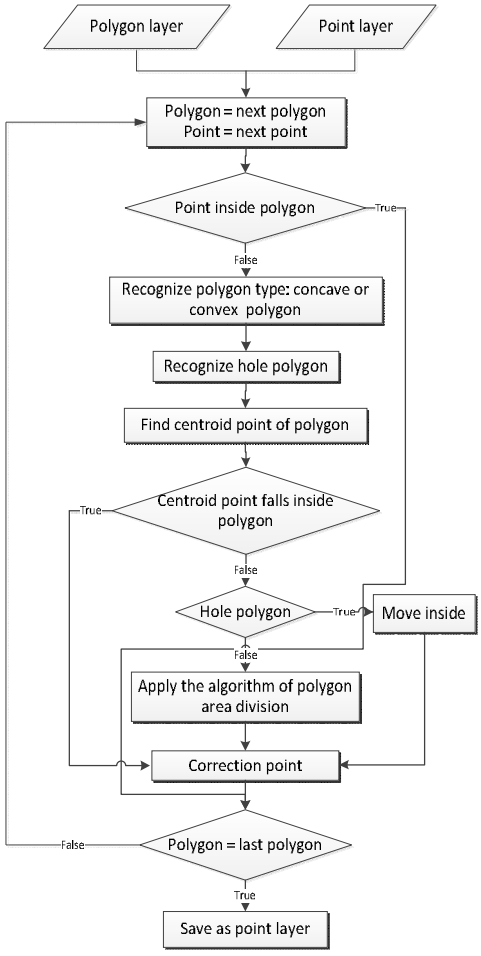

Fig. 2. Flowchart of fixing the error topology rule "polygon contains point"

In this article, the topology rule of "polygon contains point" is fixed by creating a new point layer in which the new points are located inside and at the geometric center of the violating polygons.

With four types of polygon mentioned above, using the centroid point of a polygon as the inside point is only suitable for none-hole convex polygon type. With the other polygon types like concave polygon or hole-polygon, the centroid point does not always fall within the polygon. Figure 1(c) shows that ArcGIS is not really capable of thoroughly fix this error for various types of polygon. Therefore, to ensure an inside point is within the geometric center of a polygon when viewed by readers, it requires that each type of polygon above should

have a corresponding method of generating inside point.

With convex polygon type: Directly use polygon's centroid point as the correction point.

With concave polygon type: Calculate the position of the centroid point of the polygon. If the centroid point of the polygon falls inside the polygon then use it as the point of correction. If the centroid point does not fall inside the polygon, apply the principle of area splitting, draw scanning lines that go through the centroid point and cut the polygon, choose the line that split the polygon into two parts with equal areas. On the segment that cut polygon into two equal areas, select the midpoint of the segment as a point of correction.

With hole polygon type: First, calculate the centroid point of the polygon, if that point falls inside the polygon then use it as the correction point. Otherwise, base on the principles of displacement provided in reference (Hong Fan and QuynhAn Tran 2013) to move the centroid point to the nearest position inside the polygon. Figure 2 shows the flowchart of how to fix the topology rule error of "polygon contains point".

4. Experimental results and discussion

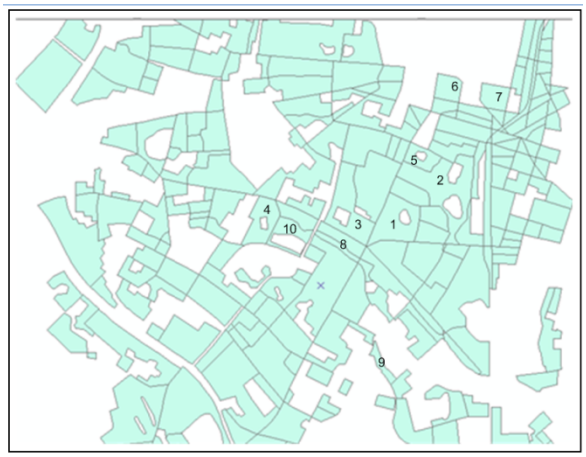

Fig. 3. The input data for experiment

The figure 3 showing the real data to test the research result, in this data, there have all types of polygon mentioned in this article, for example, the polygons 1, 2, 3, 4, 5, 10 are hole polygons; polygons 6, 7 are convex polygons and polygons 8,9 are concave polygons.

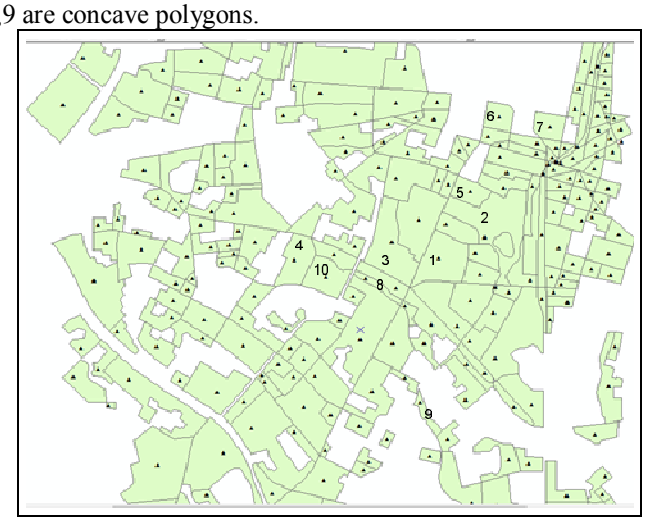

Fig. 4. Polygon features violated the rule of "polygon contains point" are corrected by ArcGIS

Fig. 5. Overlapping newly created point layer by ArcGIS on the initial input polygon layer

Figure 4 shows the result of correcting topology rule error of "polygon contains point" by ArcGIS software, the newly created points all lies inside polygons. However, the process of fixing the error also modified input data, hole polygons became none-hole polygon e.g. polygons 1, 2, 3, 4, 5, 10. Figure 5, when the newly created point layer is put on top of the initial polygon layer, there are still some points lie inside interior ring (not belong to the hole-polygon) e.g. polygons 5, 10.

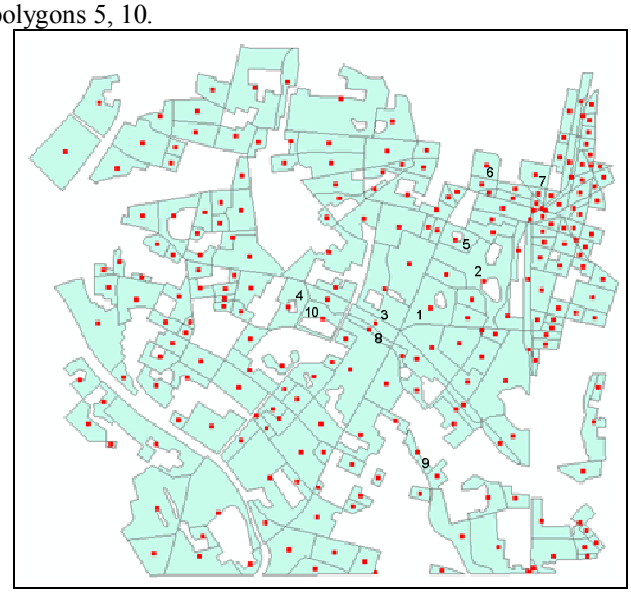

Fig. 6. Polygon features violated the rule of "polygon contains point" are fixed using the solution proposed by the authors

Figure 6 shows the result of fixing topology rule error of "polygon contains point" using the software developed by the authors, the correction points completely fall inside polygons, no matter that is a hole or none-hole polygon, the process of correction does not modify the input data. Regarding to the processing of concave polygon e.g. polygons 8, 9, the result shows that the correction points created by the software developed by the authors lie at the geometric center of polygons, bring higher aesthetic efficiency compared with the points created by ArcGIS. With convex polygons e.g. 6, 7 results from the two softwares are similar.

Looking at all aspects, the software developed by the authors can solve the shortcomings of topology rule error correction that exist in ArcGIS, and the feasibility and aesthetic efficiency of the algorithm proposed by the authors are also proven.

5. Conclusion

This article studies and proposes solutions of automatic recognizing polygon types available in describing geospatial data, pointing out shortcomings in identifying hole polygons which leads to errors in fixing the topology rule of "polygon contains point" in the current ArcGIS platform. Furthermore, the article proposes solutions to fix the topology rule error "polygon contains point". The experimental results demonstrate the

feasibility and visual effectiveness of the proposed solutions.

Acknowledgements

This work was supported by grants from the Foundation for Science and Technology of University of Hanoi Mining and Geology (No. T16-22).

References

Cai Shaohuan, Din zhiyuan & Zhu Tao, 1998. New Method to Automatically Create Polygon Node Using MBR. *Engineering of surveying and mapping* 7(2), p61.

DING Yao, 2008. New Method to Automatically Create Polygon Node Based on Triangulation Theory. *Journal of Chongqing Institute of Technology (Natural Science),* 22(3), p13.

ESRI. *The help of ArcGIS 10.2.*

GAO Zhi-jun & WANG Ji-yao, 2007. Study on Polygonous Node Automatic Creation Based on Triangle. *Journal of Jiangsu Geology,* 31(3), p251.

Hong Fan & QuynhAn Tran, 2013. Cartography Automatic Generalization. *The Journal of Science & Engineering of Hanoi Mining and Geology,* 44, p23.

LU Hao, ZHONG Er-shun, WANG Tian-bao & WANG Shao-hua, 2013. An Improved Polygon Data Interior Point Automatic Generation Algorithm. *Computer Engineering,* 39(1), p41.

Wu Lixin & Shi Wenzhong, 2003. *Geographic Information System Principle and Algorithm.* . Beijing: Science Press.