

Geographical information services classification based on FCA - a case study in vector spatial statistic analysis

Yumin Chen¹, Jingyang Wu^{1,a}, Fei Zeng^{1,2}, Xiang Gao¹ and Xiaomei Bi¹

¹School of Resource and Environment Science, Wuhan University, Wuhan, China

²61363 Troops, Xi an, China

^a Corresponding author: ^awjytcren@163.com

Keywords: Geographical information services, Classification, Formal concept analysis

Abstract. The classification of geographical information services is very important to the effective management and sharing of geographical spatial information, especially in a web-based GIS system. However, most current classification criteria focus on an abstract level. Discussions on the detailed services classification are still insufficient. This paper proposes a new thought of constructing these diverse services using the Formal Concept Analysis (FCA). After briefly discussing the current services classification systems and basic knowledge of FCA, the paper takes vector data spatial statistic analysis as an example to perform the classification with the method of FCA. The resultant categories prove to be acceptable and practical.

Introduction

A standard, comprehensive and reasonable geographical information service taxonomy criterion is very important to the effective management and sharing of geographical spatial information, especially in a web-based GIS solution. At present, several classification systems are already established. For example, the ISO 19119 criteria divided the whole system into six categories[1]. And the OGC (Open GIS Consortium) put forward the OpenGIS service framework[2]. All these classifications are built in a coarse and abstract level. While the detailed categories are necessary for improving the resource sharing and interoperability between different geographical service systems, yet there is few relevant researches concerned with the classification of the specific and diverse geographical services on a fine scale. This paper comes up with a new thought of classifying these complex geographical services with the method of Formal Concept Analysis considering the traditional empirical classification may be incomplete and unreasonable. In the following sections, after briefly introducing basic conceptions of FCA, the paper discusses the vector spatial statistic analysis classification with the FCA method.

Formal Concept Analysis

Formal Concept Analysis which is designed for data analysis, knowledge representation and information management was first proposed by R.Wille and B. Ganter[3]. This analysis supplies a method to automatically and effectively dig the hidden concepts and their relationships from the existed formal concepts. A formal concept is defined as a set of objects, attributes and their relationships. Here the objects and attributes are not limit to their literature meanings. These formal concepts together form a formal context[4]. In the method of FCA, all these concepts in a formal context are organized into a single framework, namely the concept lattices. A concept lattice explicitly displays the hierarchical relationships between different concepts. Moreover, these formal concepts often represent related conceptual categories. Consequently, FCA offers a new way to classify and manage these complex geographical information services.

Vector Spatial Statistic Analysis Services Classification Based on FCA

Determine the Objects and Attributes. In this paper, 25 most basic vector spatial statistic services which can no longer be subdivided are taken as the objects[5].

They include equal interval classification, user-defined interval classification, equal element classification, standard deviation classification, natural break classification, manual classification, five-elements classification based on box-plot, histogram, scatter plot, parallel coordinate plot, quadrat analysis, nearest-neighbor distance analysis, centroid identification, inversed distance weight, kriging interpolation, first order trend surface, K-means aggregation, ISODATA, line density distribution, nearest-neighbor analysis, line element orientation, Moran’ s I, Getis’ G, Geary’ s C.

For the convenience of depiction, these objects are encoded with g1 to g25.

By comprehensively considering these objects ,the attributes possessed by these objects could be summarized as the below 12 kinds: point, line, polygon, classified by attribute, aggregated by attribute, Exploratory Data Analysis, Pattern Recognition, interpolation, line distribution pattern, global autocorrelation analysis, local autocorrelation analysis and vector data

These attributes are encoded with m1 to m12.

Establish the Formal Context. Usually, a formal context could be expressed by a matrix. Firstly, put the objects in the columns and attributes in the rows. Then if one object and one attribute have a relationship, a cross sign will be put in the corresponding cell[6]. With this operation, the formal context is built up as shown in Table 1.

Table 1. The matrix of objects (g1-g25) and related attributes(m1-m12)

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12
g1	X		X	X								X
g2	X		X	X								X
g3	X		X	X								X
g4	X		X	X								X
g5	X		X	X								X
g6	X		X	X								X
g7	X		X	X								X
g8	X		X		X							X
g9	X		X		X							X
g10	X		X		X							X
g11	X		X		X							X
g12	X					X						X
g13	X					X						X
g14	X					X						X
g15	X						X					X
g16	X						X					X
g17	X						X					X
g18	X		X					X				X
g19	X		X					X				X
g20		X							X			X
g21		X							X			X
g22		X							X			X
g23			X							X	X	X
g24			X							X	X	X
g25			X							X		X

Identify New Concepts in the Formal Context. New concepts are the combinations of those attributes (identified by the “X” sign) along rows in the matrix[7]. For example, a combination of four attributes m1, m3, m8, m12 is present in g18 and g19. Thus, a new formal concept stands for this combination can be described as Ci= (g18, g19, m1, m3, m8, m12).

In this way, all the formal concepts are derived and shown in Table 2. The example listed above is C9= (g18, g19, m1, m3, m8, m12).

Construct a Concept Lattice. Constructing a concept lattice with these formal concepts includes two steps: (1) examine the inclusion relationships between these concepts and (2) determine the hierarchical relationships between them. Many researches about the FCA constructing algorithms have been proposed[8] and in this paper we use the software *Concept Explorer* to help construct the concept lattice.

Fig.1 lists the resultant concept lattice. In this lattice, a concept in a lower level is considered as a sub-concept of an upper level one if they are connected directly with lines. Moreover, the sub-concept inherits all attributes from its super-concept and possess its own attributes.

Table 2. The list of 13 concepts derived from the matrix in table 2. In the list, C=concepts

-
- C1=(g1,g2,g3,g4,g5,g6,g7,g8,g9,g10,g11,g12,g13,g14,g15,g16,g17,g18,g19,g20,g21,g22,g23,g24,g25,m12)
 - C2=(g1,g2,g3,g4,g5,g6,g7,g8,g9,g10,g11,g12,g13,g14,g15,g16,g17,g18,g19,m1,m12)
 - C3=(g1,g2,g3,g4,g5,g6,g7,g8,g9,g10,g11,g18,g19,g23,g24,g25,m3,m12)
 - C4=(g20,g21,g22,m2,m9,m12)
 - C5=(g25,m3,m10,m12)
 - C6=(g1,g2,g3,g4,g5,g6,g7,g8,g9,g10,g11,g18,g19,m1,m3,m12)
 - C7=(g23,g24,m3,m10,m11,m12)
 - C8=(g1,g2,g3,g4,g5,g6,g7,m1,m3,m4,m12)
 - C9=(g18,g19,m1,m3,m8,m12)
 - C10=(g8,g9,g10,g11, m1,m3,m5,m12)
 - C11=(g15,g16,g17, m1,m7,m12)
 - C12=(g12,g13,g14,m1,m6,m12)
 - C13=(m1,m2,m3,m4,m5,m6,m7,m8,m9,m10,m11,m12)
-

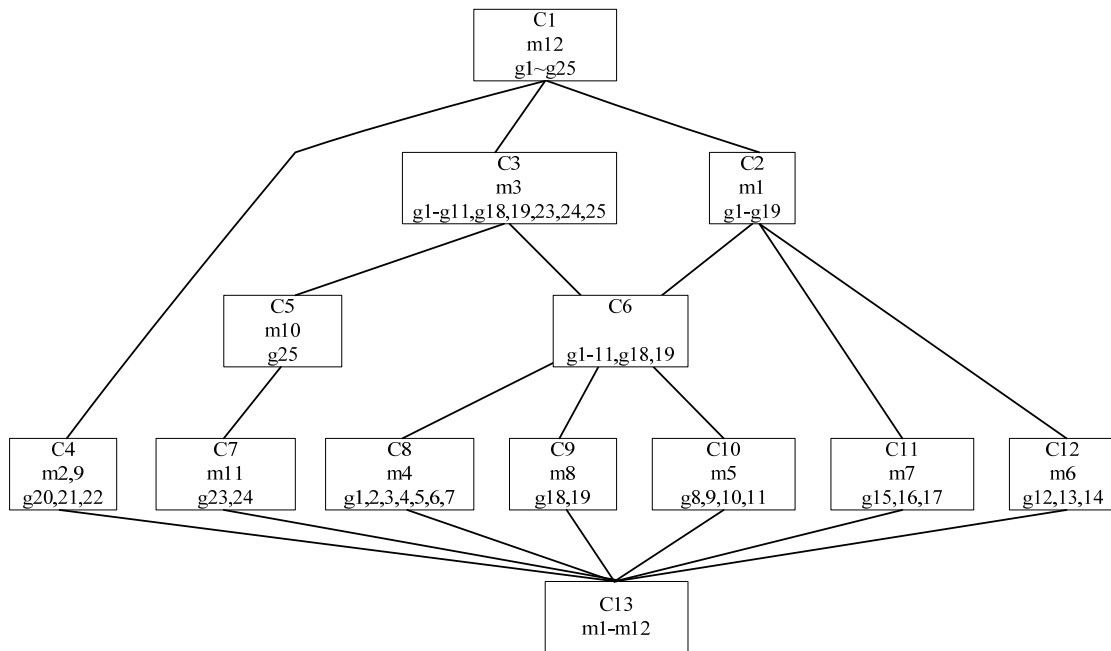


Fig. 1

Figure 1. The graphic form of the concept lattice shown in table 2. Notice that in a lower level sub-concept, the inherited attributes from its super-concept are not listed. So in C6 no attribute is shown.

Classification Results. As is mentioned above, each of these concepts in the concept lattice stands for one kind of statistic analysis service which can be identified by the included attributes. Furthermore, concepts between levels are in a ‘is-a’ relationship. Utilizing these features, the classification is easy to define. For example, the top concept C1 has three sub-concepts (C2, C3, C4) connected with it, which means that “vector data analysis” can be categorized into “point analysis”, “line distribution pattern analysis” and “polygon analysis” according to the concerned attributes and objects. In this way the final classification system is established in Fig. 2.

It is clearly shown that this result categorize those services depending on the dimension of geographical features. Compared with the categories of vector spatial statistic analyses in some prevailing geographical statistic softwares such as ArcGis Desktop and Geoda, this classification system is more explicit and more suitable from the view of cognition habits.

Conclusion

This paper presents a FCA-based method to establish the classification system for vector spatial

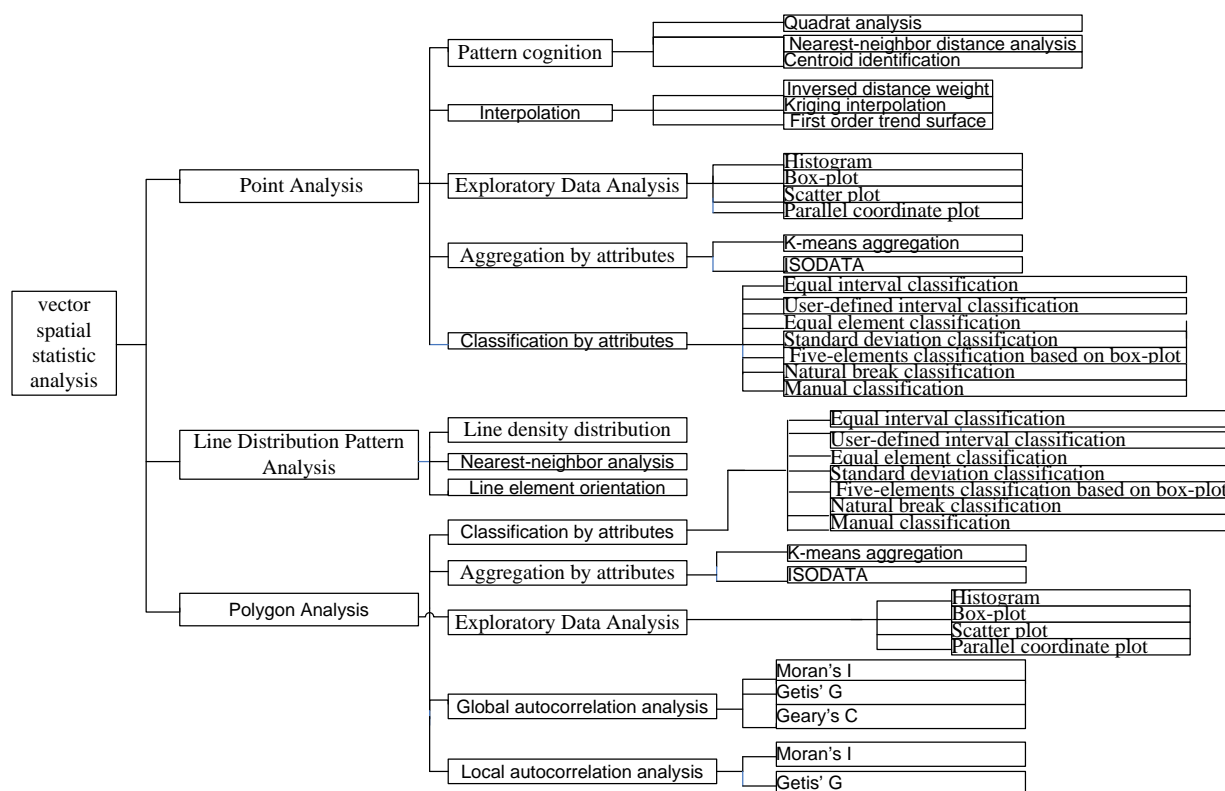


Fig. 2

Figure. 2 The derived classification of vector spatial statistic analyses from the concept lattice in figure 2.

statistic analysis. Compared with the categories in some business softwares, the result is reasonable and acceptable empirically and semantically. It proves that FCA is an effective and feasible method in classification, especially for a large number of complex services while it may be difficult and incomplete doing such a classification manually considering the similarities and heterogeneities. This method could be promoted to set up other geographical services classification in order to improve the interoperability and sharing between different web-based GIS systems.

References

- [1] ISO 19119-Geographic information services, first ed. 2005.
- [2] Information on <http://www.opengis.org/>.
- [3] B. Ganter, R. Wille, Formal Concept Analysis: Mathematical Foundations, M. Berlin: Springer,1999.
- [4] Shixiong Hua, Ling Bian, Interoperability of functions in environmental models - a case study in hydrological modeling, J. International Journal of Geographical Information Science. 23(2009), 657-681.
- [5] Information on <http://resources.arcgis.com/en/help/>.
- [6] Ling Bian, Shixiong Hua, Identifying components for interoperable process models using concept lattice and semantic reference system, J. International Journal of Geographical Information Science. 21(2007), 1009-1032.
- [7] M. Koklar, M. Kavouras, Spatial concept lattice: an integration method in model generalization, J. Cartographic Perspectives. 34(1999), 23 - 38.
- [8] Keyun Hu, Yuchang Lu, Chunyi Shi, Advances in concept lattice and its application, J. Tsinghua Univ(Sci & Tech). 40(2000), 77-81.