www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

GIS FOR ARCHEOLOGICAL DATA

CAO Ligang and WANG Xuben
Key Laboratory of Earth Exploration & Information Techniques of Ministry of Education
Chengdu University of Technology
Chengdu, Sichuan, P.R.China 610059
Emails: caoligang08@cdut.cn

Submitted: Mar. 12, 2014 ***Accepted: July 15, 2014*** ***Published: Sep. 1, 2014***

Abstract- The digital system of Archaeology includes multi-scale non-destructive detection (NDD) of archaeological methods, data mining technologies and the GIS of archaeological detect. Heritage preservation is not just to protect the cultural relics which have been excavated. NDD method could detect archaeological artifacts and clarify the statues of cultural relics buried underground. Due to data mining algorithm we can obtain archaeological information via detection data. Finally, apply GIS technology to achieve the data management and data mining methods integrating. A whole new system of digital archaeology is built on the GIS platform by applying of data mining technology to realize a detection method and archaeological information mapping. The system has provided a complete technical support for the digital archaeology

Index terms*: Archaeology explore, Geographic information system, Data mining*

I. PREFACE

The digital system of Archaeology includes multi-scale non-destructive detection system of archaeological methods, data mining system and the data of archaeological GIS. Apply the non-destructive method to detect archaeological relics buried underground, we can gather data; after that data mining technology will provide information extraction from detected archaeological data; at last through the archaeological GIS data processing, to display the data mining results.

The method for archaeological exploration can trace back to single method, and now archaeological exploration always contains multi-method and multi-scale. In 2013, application of multi-frequency electromagnetic profiling in studying the distribution of bronze in Jinsha ruin worship, proving that multi-frequency electromagnetic profiling method is an effective and nondestructive means of assessing the distribution of bronze ware in archaeological ruins [1]. In 2012, analysis of land use change characteristics based on remote sensing and GIS in the Jiuxiang river watershed, proving the effectiveness of spatial data management by GIS [2]. In 1982, the Archaeological Institute of Shaanxi Province applied magnetics, electric and geochemical exploration methods to prove the boundary of the Qin Shi Huang Mausoleum where is an ancient Tomb existing large quantities of mercury inside[6]; in 1987, Shaanxi Provincial Center for Remote Sensing and Aerial Survey cooperated together with Zhao-ling Museum and conducted "Zhao-ling ancient burial sites and positioning of Remote Sensing Research"[10]; in 2004, Chengdu University of Technology, Chengdu archaeological institute and Sichuan archaeological institute together researched both in Jinsha and Sanxingdui ruins also applied electric, magnetic, remote sensing and other geophysical methods to achieve archaeological exploration[4-5].

The data mining method has been applied to the field of archaeological research in recent years. Archaeological data mining is currently and mainly focus on archaeological relic where historical relic clusting and classification be fulfilled. Such as in Shaanxi Lintong Jiangzhai Site it was conducted an archaeological stratigraphic section scan of Spatial Data mining research [11]; the settlement of Jiangzhai archaeological applicated clustering algorithm [3]. This paper studies a digital system of archaeology based on GIS, specifically the detection of data for data mining and archaeological information extraction, and ultimately the use of GIS platform for data management and results display.

II. THE KEY TECHNOLOGY OF DIGITAL SYSTEM OF ARCHAEOLOGICAL

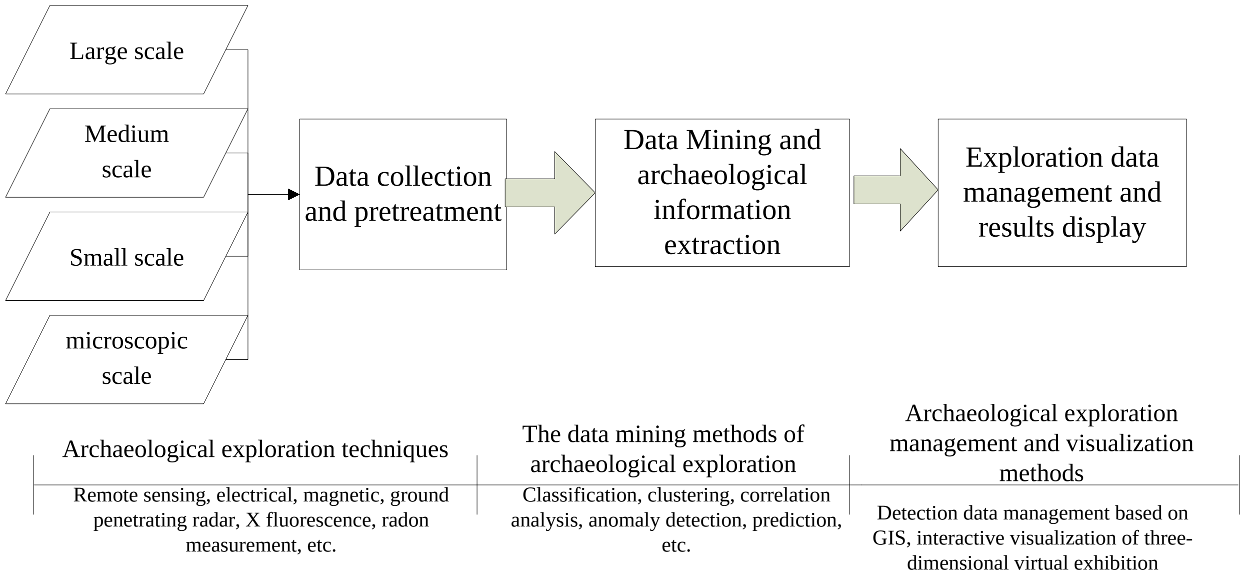

Figure.1. Archaeological exploration flow

The digital system of archaeological exploration includes archaeological detection techniques, data mining methods, management of archaeological exploration and visualization methods. Namely the realization of the archaeological exploration data acquisition requires data analysis and visualization of archaeological result.

(1) The technology of archaeological detection

Archaeological detection methods including ground, underground, aquatic relics detecting of human history ruins, tombs, artifacts, etc., which mainly consist of land-based geophysical survey and aquatic archaeological geophysical survey. The land-based geophysical detect methods include remote sensing, electrical, magnetic, seismic, gravity measurements, and geochemical surveys. Provide each method also includes a number of geophysical methods, such as remote sensing i.e. aerial remote sensing, space remote sensing, microwave remote sensing and infrared remote sensing. Electrical methods includes electrical resistivity method, ground penetrating radar, the natural electric field method, direct-current induced polarization method and multi-frequency electromagnetic method, etc.. Magnetic method is mainly high-precision magnetic measurement method. And gravimetry mainly contains microgravity measurement, geochemical surveys, including mercury measurement, gas measurements, X fluorescence measurements, isotopic measurements and C14 test methods. Aquatic archaeological-geophysical methods include magnetic, seismic and acoustic wave method.

(2) The data mining methods of archaeological exploration

The data mining methods of Archaeological exploration include classification, clustering, correlation analysis, anomaly detection and prediction those methods. [9,12]
Classification must have a clear definition on each class firstly, but also has a series of classified examples, classification process is to build model(s), and then the model will be used to classify unknown data. The classification method includes: decision tree classification, rule-based classification, neural networks, support vector machines, Bayesian classifier and genetic algorithm.
Clustering is to vary the individual, divided it into more sub-group similarity or cluster work. The biggest difference from classification is that the clustering does not rely on a priori knowledge of pre-established categories. Clustering methods include: Based on the level of approach, based on the division of methods, density-based methods, model-based methods and grid-based methods. [8]
Correlation analysis is used to determine which matter(s) is interrelated and analysis should be considered together. Correlation analysis methods include: frequent pattern mining, sequential pattern mining and structural pattern mining and so on.
Anomaly detection is mainly used to figure out the different object from most audiences. In the clustering process, the anomalous data points as noise filtering, in some application scenarios, we need to look for these outliers. Anomaly detection methods include: statistical methods, anomaly detection based on the nearby degree, density-based anomaly detection and anomaly detection based on clustering and so on.
Prediction mainly refers to the future behavior prediction or estimate of future value, it is the historical data used to build the model and explain the current observed data or behavior. To reuse this model, applies it to draw on the future behavior prediction. Prediction methods include: linear regression, nonlinear regression and other regression-based methods.

Most of these methods are not completely separated; a variety of methods can be used in conjunction with each other. In fact, the issue of data mining can be used to solve a number of ways also.

(3) Archaeological exploration management and visualization methods

There are a great amount of space data from the multi-scale archaeological detect methods in the ancient ruins, which data is necessary for geographic information systems technology management and analysis. After the data mining, the archaeological information will combine with the virtual reality technology, and be embedded into the GIS platform to display the results. GIS platform is a digital system key of archaeological, the data of archaeological exploration and data mining algorithm modules are integrated into the platform to achieve integrated management and analysis.

III. THE DATA MINING METHODS OF ARCHAEOLOGICAL EXPLORATION

"We are inundated by information, but lack of knowledge" (Rutherford D. Roger, 1985), Data Mining from the large amounts of data are extracted, or "mining" knowledge as well. Data Mining Technology combines dynastic history of multi-disciplinary knowledge, including the database, artificial intelligence technology, knowledge of statistics, machine learning technology, high-performance computing and data visualization technology. We focused on non-destructive archaeological data for data mining algorithms and applied research.

The main purpose of the archaeological exploration is to answer two questions: Where is heritage? What is the heritage status?

In order to understand the whole distribution in archaeological site, it can be applied that remote sensing or ground the electrical, magnetic and other geophysical methods. These methods of data mining mainly relay on the cluster, such as ancient river borders extraction, the direction of the ancient city walls, ancient lakes and other border objects. In the clustering algorithm, based on the division of the method and the density-based methods we can consider it.

IV. DIGITIZED ARCHAEOLOGICAL EXPERIMENT

Our archaeological exploration missions are mainly concentrated in the ancient worship ruins, based on previous archaeological experience, our ancestors often held a ritual nearby the river, and we hope to be able to find Paleochannel flows to determine the laying of the sacred area, guiding archaeological excavation or site protection. So, during a non-destructive detection using high density resistivity method to track the flow of the Paleochannel, the approximate range of the sacred area maybe appears nearby riparian. The main study is using clustering algorithm to high-density electrical method data in recognition of the Paleochannel flow and its change, including using K-means algorithm to cluster analysis in electrical method data for different depths in order to identify the Paleochannel.

(1) Processing flow

The result of Paleochannel toward by using K-means algorithm cluster analysis is from the data of the archaeological sites of high-density resistivity method, which is consistent with interpret by the experts and which proves the K-means algorithm can get the information from the data of high density resistivity method and can be promoted to other ruins. The data-mining process of high density resistivity method is shown in Figure 2:

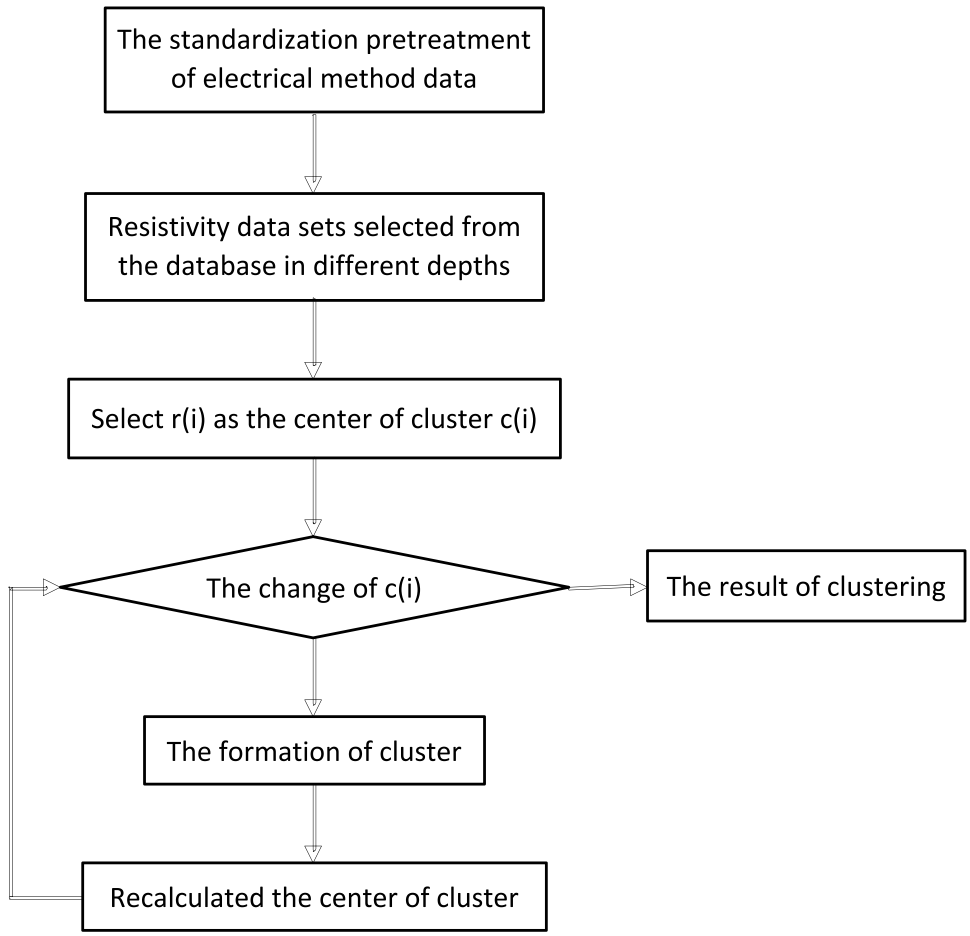

Fig. 2 The data-mining process of high density resistivity method

First unify the resistivity value of the property to the range of [0, 1], by the data standardization of high density resistivity method, and get the results of cluster by using the K-means algorithm. Below is the detailed process of cluster:

(2)The preprocessing of high-density resistivity data

During the paleochannel ruins clustering analysis, according to the Chengdu Plain part rock, soil resistivity values of sample data, the data of high-density resistivity needs pretreatment; Electrical method data is a profile data, during the ancient river ruins be outlined in the horizontal plane sketch, there would be different probing depths during electrical method data preprocessing, in line with the requirements of data mining. Sample data can be obtained via the high density resistivity method, the measured resistivity in the archaeological sites are shown in Table 1:

Table 1 The resistivity of rock, soil measured in the sacred area

Horizon	Lithology	Resistivity (Ω·m)
H2313	Silt layer	17.6
H2313	Silt layer	25.22
43	Clay	37.5
43	Clay	28.7
38	Gray-green clay	32.6
39A	Gray-green clay	37.0
39B	Gray-green clay	37.1
39A • B	Gray-green loam	26.3
H2313	Gray-green loam	24.8
41B	Gray-green loam	24.9
40	Gray-green loam	39.9
39B	Gray-green loam	37.2
35-37	Brownish yellow loam	33.2
35-37	Brownish yellow loam	33.6
The top surface of	Silt folder pebble	52.6

The basic characteristic of the rock and soil conductivity from the resistivity values in the Table1:

1) The range of varies soil resistivity are 24.9-39.9, which is a low resistivity layer;

2) The range of the silt layer resistivity is 17.6-25.22, which is the conductive best low resistance layer;

3) The range of gravelly sand and muddy gravel layer resistivity is about 70-114, which is the high impedance layer;

4) The resistivity of the gravel layer value is greater than 220, which is a high resistivity layer;

According to the requirements of data mining, we should create a database of high density electrical field like table 2, in which the X, Y, Z are spatial data of detection point, the value is particular point resistivity value's logarithm, and lineno is the measured line number of the point. To trace Paleochannel, we utilize a total of 55 survey lines and 29,735 data points.

Table 2 The database fields of electric method

Name	Data types	Length
X	float	8
Y	float	8
Z	float	8
value	float	8
lineno	varchar	255

(3)The data feature extraction of High-density resistivity method

The data of high density resistivity method including spatial coordinates (X, Y, Z) and resistivity values V and measured line number. The resistivity values can be used for cluster analysis.
Use minimum - maximum standardized formula for data standardization in resistivity value data preprocessing, set minA and maxA as the minimum and maximum values of the resistivity V.

$$v' = \frac{v - \min A}{\max A - \min A}(new_{\max A} - new_{\min A}) + new\_\min_{A}$$

Resistivity value of V is mapped to the interval [newmaxA, newminA] of V’. The mapping interval is [0, 1].

(4)Experimental results and analysis

1) The paleochannel geological inversion interpretation
According to resistivity inversion cross-section of the Jinsha ruins and the findings of field geological conditions corresponding to lithology and resistivity values, here below is corresponding relationship in Table 3.

Table 3 Chengdu plain rock, soil resistivity (ρ) parameter tables

Lithology	Resistivity values(Ω·m)
argillaceous sand gravel	18-47
gravel pebble gravel	220-770
Mud gravel gravel	117-372
boulder-clay gravel	70-114

No.37 measured line resistivity inversion (Figure 3). From Table 3 correspondence between the geological interpretations maps shown in Figure 4. Which geological inversion map is divided into a total of five, from bottom to top they are pebble bed, gravel layer, sand bed, and sediment layer and soil layer respectively. According to the geological interpretation map, it can be more intuitive to draw geological interpretation of the ancient river.

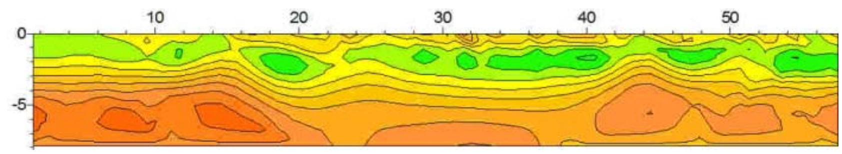

Figure 3. No.37 measured line resistivity inversion

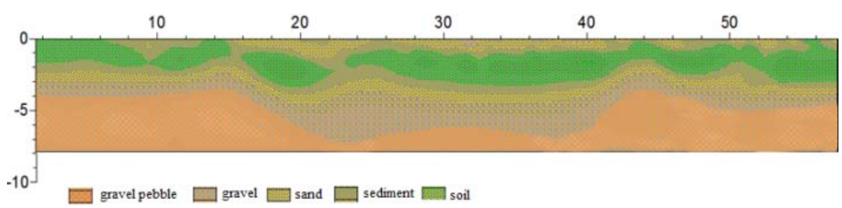

Figure 4. No.37 geological interpretation of the survey line diagram

2) The move of paleochannel

By conclusion of the density map of resistivity cluster result including 0-1 meters, 1-2 meters and 2-3 meters: the gravel layer in north is shallow, while deeper in south, it shows that the river is shallow in north and deeper in south, it can be inferred that the paleochannel migration was from southeast to northwest.

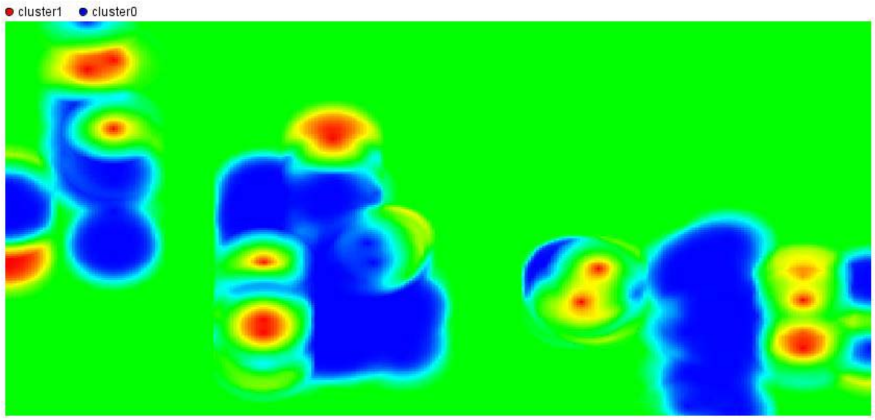

Figure 5. The clustering density map of high-density resistivity depth 0-1 meters

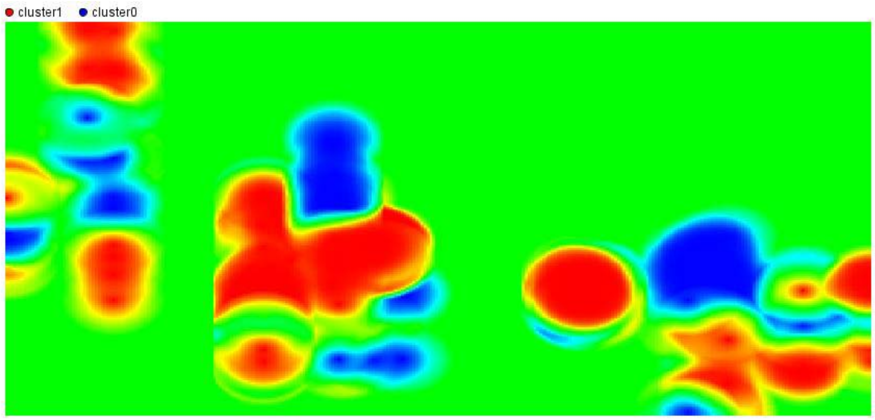

Figure 6. The clustering density map of high-density resistivity depth 1-2 meters

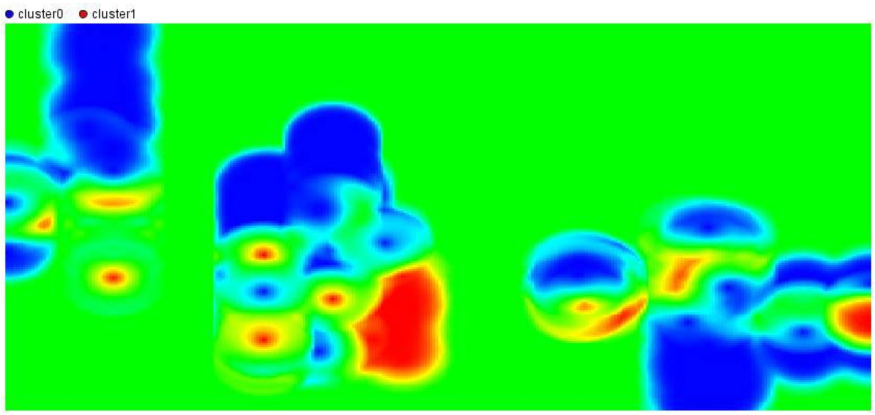

Figure 7. The clustering density map of high-density resistivity depth 2-3 meters

We can obtain the migration and location of the paleochannel in the archaeological site, by using K-means algorithm of high density resistivity method in which different depth data clustering analyzed of the result. This result is consistent with the results obtained by the using of traditional geophysical interpretation methods. K-means clustering algorithm can automatically extract data for high-density resistivity method provided a new technical support which can greatly improve the efficiency of the archaeological probing interpretation and provide a new technical means of data processing.

V. GIS MANAGEMENT

The data of archaeological exploration has space properties. The location of archaeological sites and boundary can be expressed by spatial data. The ancient artifacts such as pottery, ivory and so on, are always in specific position. Accurate detect data can be used to express their spatial location. Traditional archaeological research generally uses words description or chart. This approach is very simple, but cannot express all archaeological research, and haven't analysis and simulation. Archaeological GIS can conduct multi-spatial analysis and modeling, manage detect data and integrate the data mining results easily. [5]

The function modules of GIS are mainly divided into data management, layer management, spatial analysis and output functions. Data management is mainly responsible for detecting data input and editing and so on. Layer management is for different data layers to superimpose and edit. Spatial analysis modules applies to inquiry detect line and some unusual buffer analysis and output module for thematic map print. Archaeological exploration GIS system functions into modules shown in Figure 8.

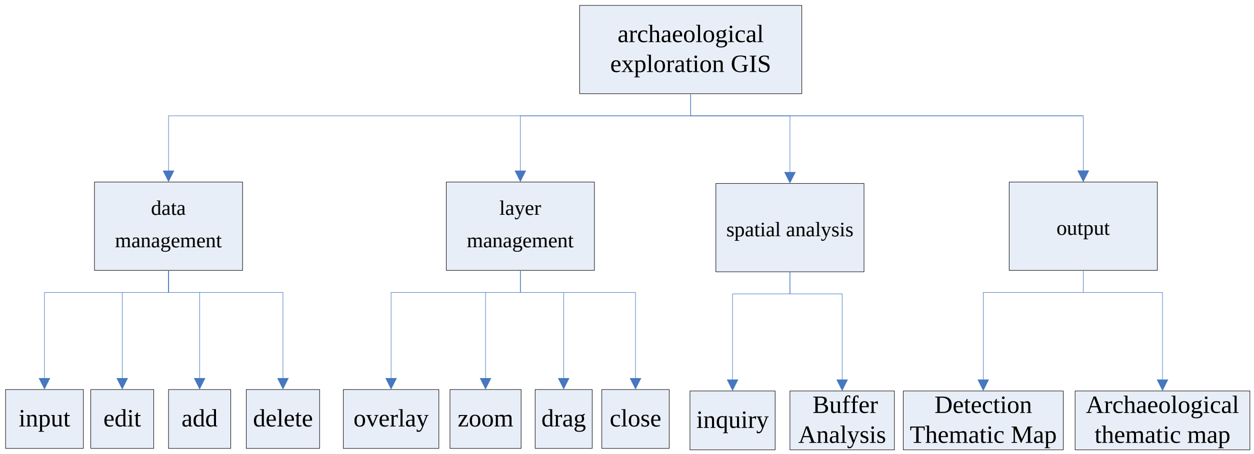

Figure 8. Archaeological exploration GIS system functions

The database manages Jinsha base map data, archaeological exploration pit data, detection work area data, electrical data, magnetic data, radar data, as well as artefacts and unusual detected data management, detect data in the database structure shown in Figure 9 as follows:

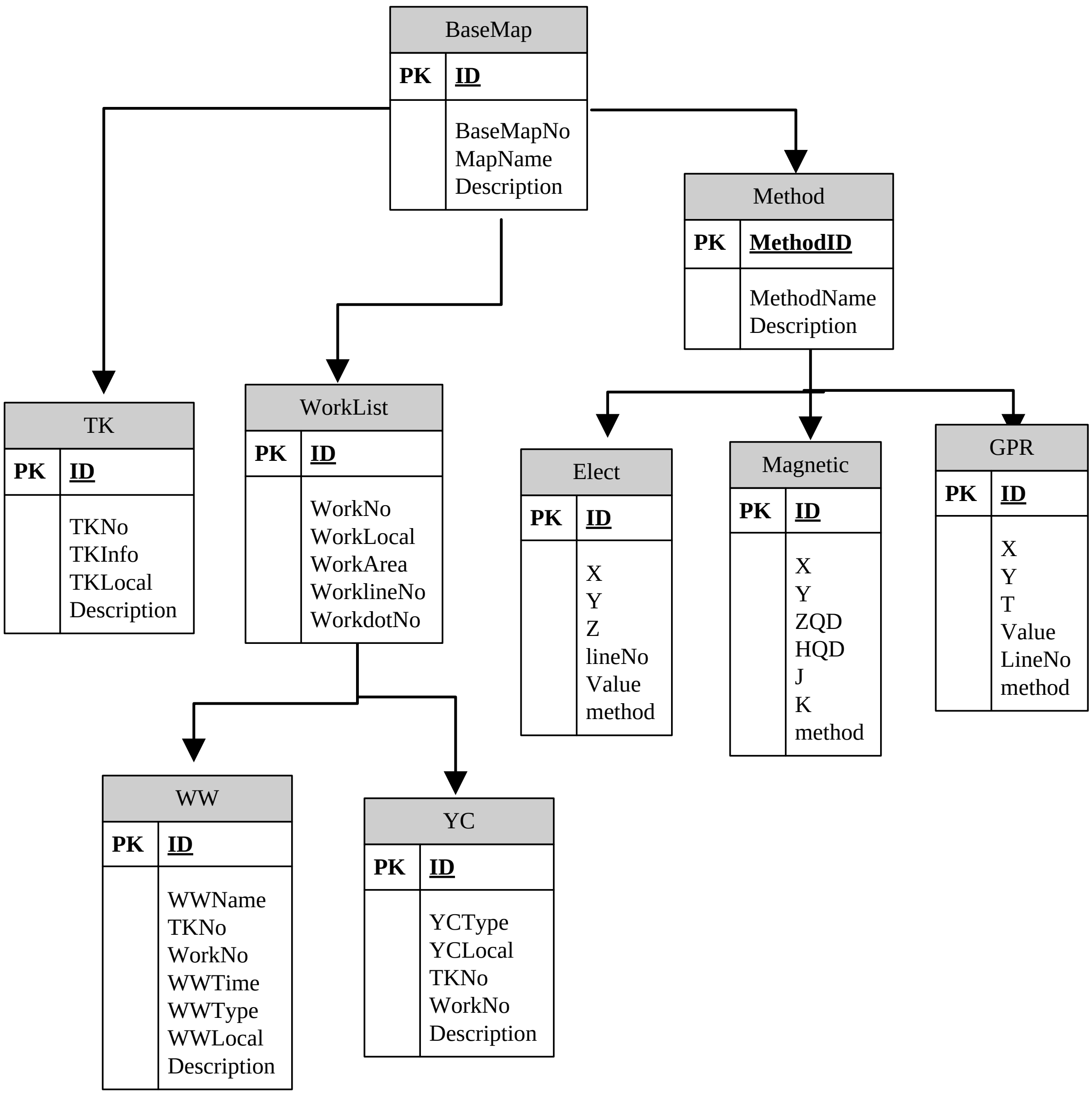

Figure 9 archeological exploration database structure diagrams

Archaeological exploration GIS system development using C/S framework, use Visual Studio. NET is as a system development language and Intergraph's Geomedia Professional 6.0 is as a software development platform. According to system design function to achieve map operation, such as zoom, move, output, etc.; spatial analysis, such as query and analysis, imagery analysis; and layers of management and interactive interpretation of development. Figure 10 for high-density electrical method to detect the GIS Manager screenshot, Figure 11 for the electrical data mining management screenshot, and Figure 12 for the magnetic data mining management screenshot. [4-5, 7]

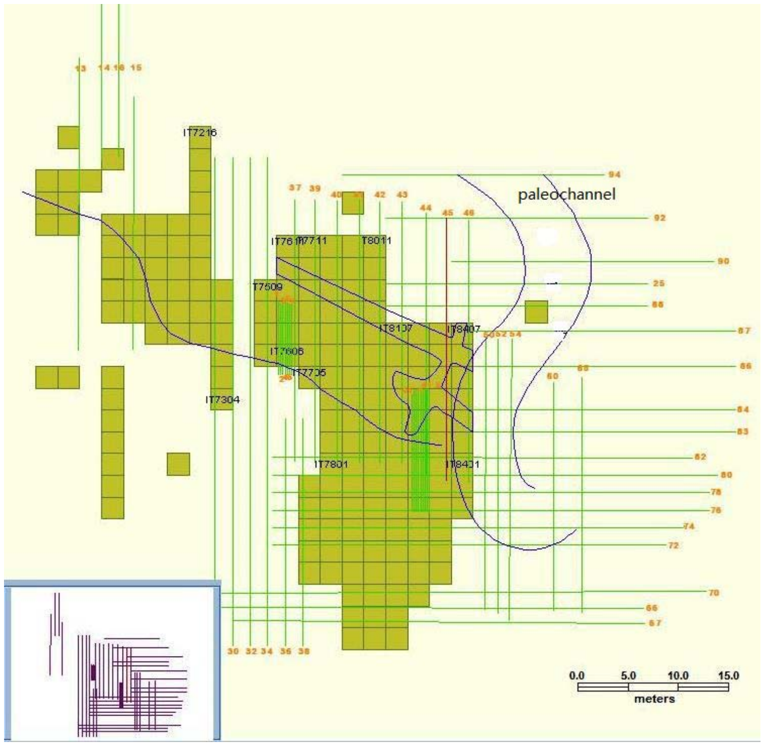

Figure 10. High-density electrical method management

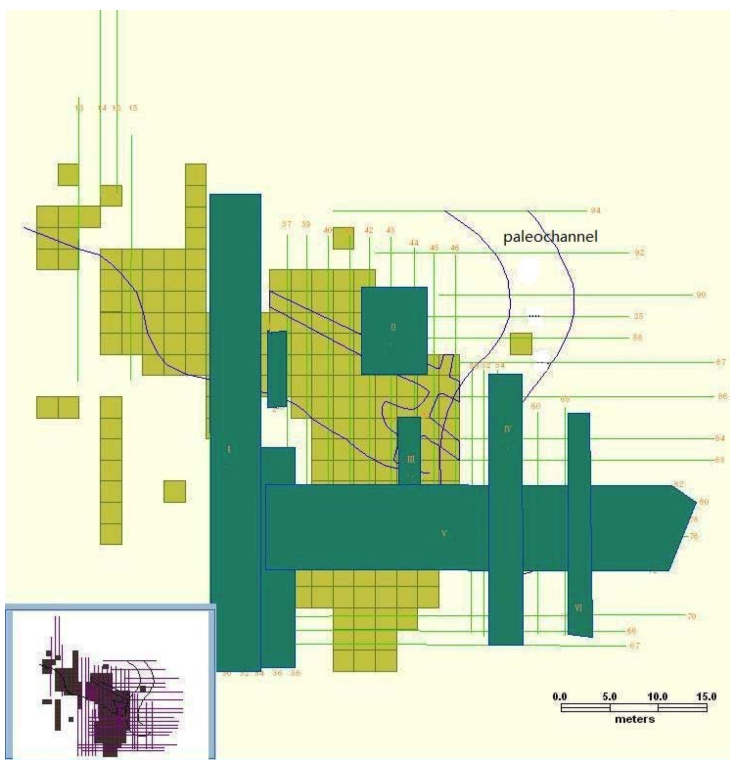

Figure 11. The electrical data mining management

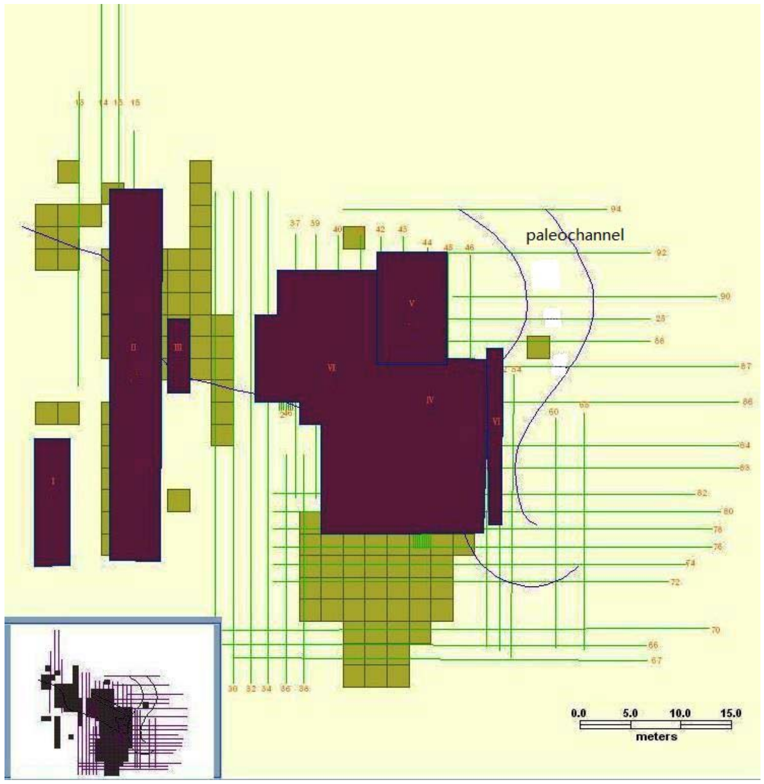

Figure 12. The magnetic data mining management

VI. CONCLUSION

The digital system of archaeological is based on that GIS integrates archaeological detection technology, data mining technology and geographic information systems technology. Using non-destructive detection of archaeological artifacts we can greatly improve work efficiency without excavation. We will be able to conduct archaeological research when excavation conditions are not fulfilled. Popularization of this technology system will provide a broad sense in the heritage conservation.

VIII. ACKNOWLEDGEMENT

Many thanks for the great support from the Key Laboratory of Earth Exploration and Information Technology of Ministry of education in Chengdu University of Technology.

REFERENCES

[1]Wenfeng Zheng, Xiaolu Li and Nina Lam et al, "Application of multi-frequency electromagnetic profiling in studying the distribution of bronze in jinsha ruin worship", International journal on smart sensing and intelligent systems,Vol.6, No.2, 2013, pp.733-751.

[2]He-Bing Hu, Hong-Yu Liu and Jing-Feng Hao et al, "Analysis of land use change characteristics based on remote sensing and GIS in the Jiuxiang river watershed", International journal on smart sensing and intelligent systems, Vol.5, No.4, 2013, pp.811-823.

[3]Bi S. B., Pei A. P., Chen J. M. and Lv G. Y., "Clustering algorithm application in first period culture of jiangzhai settlement", Computer Engineering, Vol.32, No.8, 2006, pp.89-91.

[4]Cao,L G and Wang X B, "Spatial data transformation and archaeological detection webgis", Computer Engineering,Vol. 33, No.18, 2007, pp.273-275.

[5]Cao, L G and Wang X B, "Research on interactive 3D scene in digital archaeology", Computer Engineering & Design, Vol.29, No.2, 2008, 490-495.

[6]Chang Y. and Li T., "A preliminary study of the mercury buried in the Qin Shi Huang Mausoleum", Archaeology, Vol.7, 1983, pp.659-663.

[7]Li H. R., "Integrated interpretation of archaeological detection system support by GIS", Chengdu: Chengdu University of Technology, 2006, pp. 41-48.

[8]Micheline K. and Han J.W., “Data Mining: Concepts and Techniques", Beijing: Machinery Industry Press, 2005, 263-267.

[9]Tan P. N., Michael S. and Vipin K. , "Introduction to Data Mining", Beijing: People's Posts and Telecommunications Press, 2005, pp.13-16.

[10]Song D. W., Rao S. H. and Qiu W. D., "Zhaoling ancient burial sites interpretation of remote sensing and positioning", cultural, vol.7, 1992, pp.87-92.

[11]Yang L., Lv G. Y., Bi S. B., Pei A. P., Shen Y. H., Chen J. M. and Sun Y. Q. "Data Mining of Field archaeology Stratum based on geographical information system database:a case study of jiangzhai site", Geography and Geo-Information Science, Vol. 21, No.2, pp.28-31.

[12]Wang Guanghong and Jiang Ping, "Survey of Data Mining", Journal of Tongji Unversity, Vol.32, No.2, pp.246-252.