The research of rural land loss based on data mining technology

Xusen Zhu[1, 2], Zhongcheng Wang[3*], Jiqiang Niu[4]

[1]*Southwest University, Chongqing, 400715, China*

[2]*Chongqing Normal University, Chongqing, 401331, China*

[3]*Central South University of Forestry and Technology, Changsha, Hunan, 410004, China*

[4]*College of Urban Environmental Science, Xinyang Normal University, Xinyang, 464000, China*

Received 1 October 2014, www.cmnt.lv

Abstract

As for Chinese society, land has bearing thousands of years of culture. Land change not only pulls social development, but also relates to the life of hundreds of millions of farmers. With the development of urbanization and industrialization, land resource is facing with the dilemma of constant loss. Thus a kind of technology is eagerly needed to excavate and utilize our land territory. This paper proposed the objective function and constraint system of land use regionalization, and designed the system framework of land use regionalization based on spatial data mining. This paper selectively analyzed the content and implementation strategy of data layer, knowledge layer and spatial data mining layer in system, and realized land use regionalization coupled and integrated by GIS and application analysis model.

Keywords: land loss, land use regionalization, data mining, GIS

1 Introduction

Modernization firstly lies in believing that people are capable of changing nature and social environment through rational behavior. No matter it is political modernization or economical modernization, "reason" cannot be missing. Economical modernization means to not excessively and disorderly develop and utilize resource, but to seek sustainnable development in the long run. However, in current land resource protection, the intensification of governmental responsibility is increasingly impendent and significant. The significance of studying governmental responsibility in rural collective land loss is: as for land protection, discuss the governmental responsibility of rural land loss. It has fundamental theoretical value and practical significance for seeking effective strategy of land resource protection and realizing the sustainable development of land resource. As for government, it is benefit for enhancing administrative capacity, strengthening idea of responsibility, completing the construction of governmental responsibility system, protecting the legitimate rights of peasants and promoting the establish of political democratization. As for society, it is benefit for establishing the restriction mechanism of land protection, efficient land utilization and the sustainable development of social economy.

In recent years, in spatial information technology field, a large number of data has been collected from the multiple applications of RS, GIS and GPS with the rapid development of earth observation technology, database technology and network technology. Thus, plenty of data and materials from land utilization have been obtained [2]. The complexity and quantity of these data are far more beyond human's analysis. Spatial database has the capacity of preserving the spatial object represented by the relation of spatial data type and the spatial relationship of objects, but users cannot analyze all the data in details and extract interesting spatial knowledge and model. Data mining will be an effective tool [3], which provides opportunity for solving land loss. Spatial Data Mining (SDM) refers to the process of extracting spatial or non-spatial general knowledge rules that implicated in it, beforehand unknown and potentially useful. Thus, we should closely integrate spatial data mining model and land evaluation specific model so as to solve the problem of land use regionalization.

2 Goal constraint of land use regionalization

In order to transform land use regionalization to the goal and constraint system of land use regionalization constructed by multiple goal optimizations, land loss is regarded as the problem of multiple goal optimizations. Such kind of system can better prevent the occurrence of land loss.

2.1 BENEFIT GOAL

The benefit of land use is the concentrated expression of land use regionalization. The higher optimal degree the land use regionalization is, the better the land use benefit

* ***Corresponding author's*** e-mail: wzcwangzc@163.com

is 1. Land use is a complex system that involves many influencing factors and stresses comprehensive benefits, that is, to well coordinate the economical, social and ecological benefit of land [4].

In the optimization model of land use structure, economic benefit function is often used as the objective unction of optimization model [5]. The calculation of economic benefit value is mainly the land use scale of various kinds multiply by certain economical coefficient and then performs linear accumulation. Each kind of land use regionalization is determined as acting as the rigid constraint condition of the upper level project, which guarantees the float of economical benefit within certain control area. However, the difference of use designation causes the difference of land spatial arrangement, which will directly influence the size of project cost. Therefore, the reduction of project cost is one of goals of land use regionalization model.

2.2 CLUSTERING GOAL

Clustering goal is also known as concentrated goals or layout optimization goal. Generally, compared with long and narrow land use area, the relatively compact form of land use arrangement (such as present as form of square or round) is more convenient in management and higher in utilization efficiency. Thus, we can introduce the form index δ_k of type cluster, or the form index δ_k of land use type *k*. It is calculated by the follow formula: $\delta_k = \sum_{k}^{K}\sum_{c=1}^{C}\frac{Y_{kc}}{\sqrt{A_{kc}}}$ and in $A_{kc} \geq S_k$, S_k refers to the prescribed minimum value of each use cluster, that is the minimal scale of limitation of clustering so as to avoid the participation of scattered object unit in calculation.

The spatial form formed by land use regionalization is relatively compact, that is, the land use of same type is as mush as possible concentrated which is benefit for obtaining larger benefit. Therefore, the form index value δ_k for representing the compact degree of land use is tending to small. The simple expression is: $Min\ \delta_k$.

2.3 SUITABILITY CONSTRAINT

Land use regionalization based on land suitability is restrained on the basis of regional land suitability evaluation, which relatively easy to determine. Present situation of land use can be evaluated as multi-suitability: suitable for cultivation, garden, forestation, farm, etc. We should adopt relevant constraint to conduct land use regionalization when performing optimal planning layout on agricultural land. This kind of constraint also reflects the maintain of reasonable current situation of land use, changes unreasonable idea of land use and encourages land use regionalization to develop in beneficial direction, which conforms to the basic principle of land use regionalization. In order to apply these constraint conditions to the model, we adopt spatial data mining to reflect these constraints in the form of quantization index value.

2.4 LAND USE TRANSFORMATION CONSTRAINT

The alteration and transformation of land use is influenced by many factors. The designation of land use regionalization is mainly performed based on the present situation of land use. Transformation constraint of land use is mainly used to express the condition possibility or probability needed for transforming land from certain type to the other type, or the cost or price needed to pay in transforming. Therefore, the analysis of land use transformation constraint within planning region has significance on performing land use regionalization. This paper adopted spatial data mining to have the application of some of land use units cannot be changed. For example, urban construction land generally cannot transforms to other land. Given the rules of conversion of land use, the existing urban and rural construction lands are directly designated to city and town construction land area. If not conduct special handling on these land units, then we cannot guarantee to reasonably designate these units to certain application land. Therefore, before using intelligent algorithm, we need to conduct certain pretreatment on the figure spot data in database. In addiction, forestry lands like forest land, open forest land and not forest land are divided to forestry area. In order to apply these constraint conditions to model, we can use P_{cf} to represent the possibility rate or conversion coefficient of present land use type c converting to land use regionalization type f. It was the constant among 0 to 1, of which 0 refers to unconvertible (such as: generally, other used land can not convert to unused land), and 1 refers to optimal conversion method.

3 System framework design of land use regionalization

The system framework of land use regionalization based on SDM is divided into data layer, knowledge layer, excavation layer and human-machine interaction layer. Through the integration of land use data and relevant professional field planning data, this framework conducts extraction, transformation and spatial conversion on data of land use regionalization. It constitutes database of land use regionalization and spatial data cube and makes data preparation for land use regionalization (Figure 1).

3.1 DATA LAYER

Data layer includes the multi-scale and multi-type raw database of each relevant department (spatial database and attribute database). In order to obtain intermediate data needed in the process of spatial data mining, data are integrated and form database. These can provide functions like index, query and optimization for land use regionalization. Moreover, these are the data origins of

obtaining and extracting the relevant knowledge in land evaluation field. According to its function in land use regionalization, data is mainly divided into basis geographic data, present situation of and use data, land use suitability evaluation data, important construction specific planning data and land supply and demand balance data.

1) Spatial database mainly includes basic geographic data and data of present situation of land use

2) Specific spatial database mainly includes important construction specific database, regional specific planning database and land use suitability evaluation.

3) Other relevant data mainly includes land supply and demand balance data, regional natural condition and social economy data.

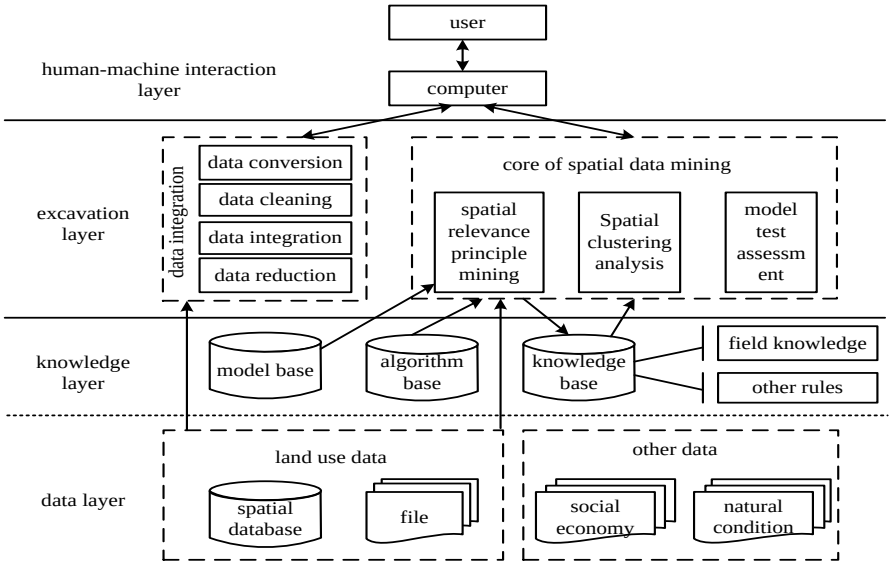

FIGURE 1 System architecture of land use regionalization based on spatial data mining

3.2 KNOWLEDGE LAYER

Knowledge layer refers to perform organization and management on domain knowledge and rules obtained by spatial analysis or knowledge engineering. Model base and method base are also parts of knowledge [6]. Combined with knowledge of land use regionalization obtained by data layer which mainly including the domain knowledge of land use planning, various of parameters settled during mining and model collection mined out by spatial data mining algorithm and extracted various of knowledge. Therefore, knowledge layer mainly contains model base, algorithm base and knowledge base.

1) Model base consists of two parts: deposited model dynamic link library based on file form and model dictionary for the convenient of retrieval service of base file management.

2) Algorithm base refers to all kinds of algorithms used in mining process, which mainly provide algorithm support for model base. It includes some of standard algorithms and basic methods of spatial analysis.

3) Knowledge base includes fact base and rule base.

3.3 SPATIAL DATA MINING LAYER

Spatial data mining layer conducts conversion, integration and compilation on the participated data waiting for excavation, thus to make it become data that suitable for excavation. Then according to the excavation purpose, proper algorithm is elected to finish excavation task. Excavation layer is closely related with data layer, knowledge layer and human-machine interaction layer, which mainly including data integration, spatial association rule mining, spatial clustering, and model test and evaluation.

1) Data integration module can guarantee system to obtain better land use regionalization scheme, and reduce the workload of data mining core. It includes unit treatment and data pre-treatment [7].

2) Spatial association rule mining module mainly adopts model base and algorithm to land use database and other data, so as to discover the spatial association rule hiding in mass data.

3) Based on the obtainment of data integration and spatial knowledge, spatial clustering module realizes the association between attribute data and spatial data through spatial data engine.

4) Model test and evaluation mainly refers to conduct measurement on spatial data mining by using corresponding standard, thus to obtain scientific and reasonable land use regionalization scheme.

3.4 HUMAN-MACHINE INTERACTION LAYER

Human-machine interaction layer is the only layer of whole system to interact with user, which mainly is used for the connector of user to inter communicate with computer. It can makes computer to operate according to user's purpose, and also can display the knowledge and

results of spatial data mining to user through many understandable ways, like map, form, etc. Users also can adopt many ways to manage the excavated knowledge and result, or even obtain satisfying land use regionalization results through human-machine inter operation and feed back to users in the form of friendship and visualization for understand, analysis and evaluation. Therefore, the merits of user interface has the direct bearing on whether system can work in efficient, and which is also the foundation of smoothly performing land use regionalization and obtaining satisfying results.

4 System implementations

In the face of land use regionalization, spatial clustering needs to equip with the input, display, calculation and output of spatial data, and at the same time, it also needs to provide data regionalization model for the alterative offer of decision support. In order to realize these functions and guarantee development efficiency, system adopts the current relatively advanced Microsoft Visual and NET 2005 as the development platform. The secondary development is realized based on Arc Engine 9.2. The function of GIS is custom made. The function of professional model is realized by adopting edit code.

5 Conclusion

This paper proposed the objective function and constraint system of land use regionalization, and conducted design on the system framework of land use regionalization based on spatial data mining, which realized the land use regionalization system integrated by the close coupling of GIS and applied analysis model. It can provide intuitive expression of spatial information for land use planning, and at the same time provide an interactive environment and half intelligentized aid decision capable for decision maker. Therefore, intelligent land use regionalization system is formed, which provide theoretical basis for the solution of land loss.

Acknowledgement

The Project Affiliated to the Education Department of Chongqing Municipality (12SKE18). The Project Affiliated to Chongqing Normal University (10XWQ10).

References

[1] Liu Y L, Zhao X, Liu D F 2014 A Parallel Decision Support System for Land-Use Allocation Optimization Based on Artificial Immune System *Geomatics and Information Science of Wuhan University* **39**(2) 166-71 *(in Chinese)*
[2] Xu H X, Jing X D 2013 Study on Land Use Rapid Investigation Method Using Unmanned Aerial Vehicles Remote Sening and GIS Technology *Geomatics & Spatial Information Technology* **36**(9) 11-4 *(in Chinese)*
[3] Zhang X R, Chen X D, Ding L 2013 Self-service Product Innovation of User-oriented Based on Data Mining Technology *Soft Science* **26**(12) 10-3 *(in Chinese)*
[4] Li D R, Wang S L, Li D Y 2006 Spatial Data Mining Theories and Application *Beijing Science Press* 337-49 *(in Chinese)*
[5] Li X, Ou M H, Liu J S 2014 Regional land use structure optimization under uncertain theory *Transactions of the Chinese Society of Agricultural Engineering* **30**(4) 176-84 *(in Chinese)*
[6] Zhang J, Tong L G, Dong J J 2013 Land Use and Coverage Change *Journal of Inner Mongolia University (Science Edition)* **3** 20 *(in Chinese)*
[7] Xun Z J, Zhao G F 2011 Design and Realization of Land Use – Based Spatial Data Mining System *Journal of Anhui Agricultural Sciences* **39**(7) 4214-6 *(in Chinese)*

Authors

Xusen Zhu, 1979, Zhejiang Province, China.

Current position, grades: lecturer.
University studies: master's degree was earned in major of human geography, Chongqing Normal University in 2005.
Scientific interests: land management.

Zhongcheng Wang, 1970, Zhejiang Province, China.

Current position, grades: associate professor.
University studies: master's degree was earned in major of music, Xiamen University in 2012.
Scientific interests: land use and evaluation.

Xusen Zhu, 1979, Zhejiang Province, China.

Current position, grades: lecturer.
University studies: master's degree was earned in major of human geography, Chongqing Normal University in 2005.
Scientific interests: land management.