

Risk Regionalization of Meteorological Calamities Based on GIS and Rough Set Theory

Fengchang Xue¹, Jin Hu² and Jin Wang³

^{1,2} Nanjing University of Information Science & Technology

³ School of Information Engineering, Yangzhou University

Abstract

The calculation of risk regionalization of meteorological calamities, usually determines the weight factors involved in the calculation by a subjective evaluation and objective calculation method, and the information should be more accurate or perfect. As the determination of the weight factors of meteorological disasters have limitations by subjective evaluation method and the objective calculation method, since the acquisition of information with multiple dimensions, has uncertainty. We propose a calculation method of meteorological disaster risk zoning of a rough set theory with GIS technology, and use this technology to evaluate the division of spatial unit and spatial factors index. The spatial unit of conditional attributes and decision attributes determine the weight of influence factors based on meteorological disasters. We can also determine the regionalization of the risk of meteorological calamities based on GIS spatial technology. The result indicated that the application of rough sets theory, combined with GIS technology can fully describe the relationship of spatial data though the information is still incomplete and uncertain. Moreover, it can be better to solve the practical problems in the calculation of meteorological disaster risk zoning.

Keywords: Rough set; GIS; Risk regionalization; Meteorological calamities

1. Introduction

By the calculation of regionalization of the risk of meteorological calamities, we usually adopts the following two methods to determine the weight factors involved in the calculation (Liu M *et al.* 2012) : (1) Subjective method: according to subjective degree of attention with factors participate in the evaluation, and then identified by experts according to their experience, such as the Delphi expert determination, the paired comparison method, queuing method and so on; (2) Objective method :according to certain objective rules to compute the relative importance of the alternative factors in the collection, commonly used with the method of principal component analysis, entropy method, multi objective programming method, etc. However, as the method above are based on the certain or perfect information, and on the other hand the meteorological calamities risk assessment involves multi-source: various types of spatial data and the process of the data acquisition and processing exist with huge uncertainties. Under such circumstances, the weight value is to calculate with incomplete information and conditions of uncertainty, so whether it is subjective or objective method, it presents its limitations in the calculation of risk regionalization of meteorological calamities.

Rough set is a new mathematical tool to deal with the geographic information that was provided with uncertainty and imprecision, and has obvious advantages in expressing and revealing the multi level (or size) knowledge(Wang G Y *et al.*2009; Lin W H 2013) ; GIS (Geographic Information System) is the analysis and processing of geospatial information computer system, which widely used in resource and environment, analysis and evaluation of natural disaster(Tang G A, Yang X.2012). The combination of GIS and

rough set theory, which can fully describe the relationship of the spatial data in the incomplete information and uncertain conditions. Above all, it is the better method to solve the practical problems in the calculation of risk regionalization of meteorological disaster.

The essence of meteorological disaster risk zoning is in the selected variety attribute based on regional, which process is depending on a specific target by using GIS composite calculation of multi-source, multi dimensional spatial data. The evaluation factors are attributes as a continuous variable in this process, on the basis of certain rules, transforming it into a specific digital interval (its realization process is called the standardization of a factor). Then make the combination of computing the standardized factor according to a definite weight, and get the evaluation value of the unit disaster risk index, then realize the quantitative evaluation of disaster risk.

According to the basic idea of rough set theory, knowledge is the subject's ability to classify objects that in the domain, factors that affecting the classification ability of different importance, some of these factors play a decisive role (Hu M L, Liu 2007; He Y Q, Hu S S. 2004). Rough set theory is based on the principle of the critical degree, which can realize the weight of evaluating factors risk zoning of meteorological disasters in the calculation. And then use the GIS spatial data composite calculation ability to realize a distribution calculation of regional spatial meteorological disaster risk.

2. Basic Concepts of the Rough Sets Theory

2.1 Information System

Formally, an information system S , can be seen as a system $S = \langle U, A, V, f \rangle$, where U is the universe (a finite set of objects, $U = \{x_1, x_2, \dots, x_n\}$) and $A = C \cup D = \{a_1, a_2, \dots, a_m\}$ is the set of attributes (features, variables). Each attribute $a \in A$ (attribute a belonging to the considered set of attributes A) defines an information function $f_a: U \rightarrow V_a$, where V_a is the set of values of a , called the domain of attribute a (Yang X J. 2012).

Suppose we are given an information system S , where $U = \{x_1, x_2, x_3, x_4, x_5\}$ is the universe, $A = \{a, b, c, d\}$ is the set of attributes, $C = \{a, b, c\}$ is the set of condition attributes and the decision attributes set is $D = \{d\}$. The range of V_a, V_b, V_c and V_d is $\{1, 2, 3\}, \{1, 2, 3\}, \{1, 2\}$ and $\{1, 2\}$ respectively. Then, the information system of S can be used to represent the information table, as showed in Table 1.

Table 1. Information Table

U	condition attribute C			decision attributes D
	a	b	c	d
X ₁	1	2	2	1
X ₂	1	2	2	2
X ₃	2	3	1	1
X ₄	2	1	2	2
X ₅	3	2	1	1

2.2 The Attribute Dependence of Rough Set

The information system of S , if all attribute values of the attributes set D only determined by values of attribute sets C , then D is called completely dependent on C . And if all attribute values of the attributes set D partly determined by values of attribute sets C ,

that is to say D is partially dependent on C. The degree of dependence between rough set attribute used to indicate the dependence of attribute:

$$\gamma(C, D) = \frac{\sum_{x \in U/D} |C(x)|}{|U|}$$

Where $\gamma(C, D)$ represents classification of attributes D depend on attributes C, $|U|$ is the

number of individuals in the U and $\sum |C(x)|$ represents the positive region

$C(x)$ according to the division of the value of attributes D, which value determined by the value of attributes C (according to a subset of known information can determine the individual ownership of the U in the number of individuals) individuals . In the example

table: $U = \{x_1, x_2, x_3, x_4, x_5\}, |U| = 5$; In the attributes C, U can be divided into distinguish categories as follows based on the attributes a、 b、 c:

$$U_a = \{(x_1, x_2), (x_3, x_4), x_5\},$$

$$U_b = \{(x_1, x_2, x_5), x_3, x_4\},$$

$$U_c = \{(x_1, x_2, x_4), (x_3, x_5)\}.$$

On the basis of $\sum_{x \in U/D} |C(x)| = 3$ attributes common of C to divide U into:

$$U_{abc} = U_a \cap U_b \cap U_c = \{(x_1, x_2), x_3, x_4, x_5\}$$

Based on the decision attributes D to divide U into:

$$U_D = \{(x_1, x_3, x_5), (x_2, x_4)\}$$

And then on the basis of condition attributes C and decision attributes D jointly to divide U into:

$$C(x) = U_{abc} \cap U_D = \{x_3, x_4, x_5\},$$

Hence,
$$\gamma(C, D) = \frac{\sum |C(x)|}{|U|} = \frac{3}{5} = 0.6 .$$

That is, the attribute D classification depends on the C is 0.6.

2.3 Attribute Importance of Rough Set

The importance of attribute in S information system is determined through the impact on the classification ability of S to remove an attribute $a \in C$ from attributes in C, through the comparison between $\gamma(C, D)$ and $\gamma(C - \{a\}, D)$ to evaluate the change.

In the example in the information table, attribute importance of (a) can be expressed as:

$$\sigma_{(C,D)}(a) = \frac{\gamma(C, D) - \gamma(C - \{a\}, D)}{\gamma(C, D)}$$

$$C_{-(C-\{a\})}(X) = U_b \cap U_c \cap U_d = \{x_3, x_4\},$$

$$C_{-(C-\{a\})}(X) = U_b \cap U_c \cap U_d = \{x_3, x_4\}$$

can obtained in section 2.1

$$\gamma(C, D) = 0.6$$

$$\sum_{x \in U/D} \left| C_{-(C-\{a\})}(X) \right| = 2, \gamma(C - \{a\}, D) = 0.4,$$

$$\text{and } \sigma_{(C,D)}(a) = \frac{\gamma(C, D) - \gamma(C - \{a\}, D)}{\gamma(C, D)} = 1 - \frac{0.4}{0.6} = 0.33 \quad .$$

That is, as for the attribute (a), the attribute D classification depends on the C is 0.33.

Attribute importance of rough set is to express the effect of attribute on decision under the current data environment, reflects various factors to occupy in the evaluation and decision-making process of the position or role, and then to determine the objective weight of every attribute in decision making.

3. Implementation of Meteorological Disaster Risk Zoning Based on GIS and Rough Set Theory

3.1 Data Processing Method

3.1.1 Space Division

Space partitioning technique is an essential foundation of spatial data of composite computation, which the basic goal is to identify the assessment unit, so as to determine the elementary unit of GIS space analysis and decision evaluation. In the GIS environment, space partition must have the following two properties (capabilities) [5] : (1) can produce infinite repeat mode, so that it can be applied to any spatial scale, (2) can be decomposed into infinite fine pattern so as to generate arbitrary resolution expression of spatial elements of different grade and level by the utilization of these models. In order to facilitate the calculation of space by using GIS, the space division method is commonly used grid method. GIS space partitioning technique through spatial unit control grid size to form a series of internal external homogeneous and different spatial units. These constitute the domain of U rough sets.

3.1.2 Spatial The Attribute Value of Evaluation Factors

The evaluation factors reflected the properties of spatial elements attributes in a form a point, line or area feature which affects a broad range of more around areas, reflecting the impact of the elements' situation of the corresponding index value to achieve.. Through the realization of evaluation factors, attributes discretization in domain U with GIS space technology, namely the evaluation factors attribute values will be given to the spatial unit after discretisation. Because of the spatial element attribute values are distinguished by thousands of ways in the presence of forms and sources. So its standardization in spatial processing must also be taken into consideration before being used. Standardized method was sum of several methods (standardization, standard deviation, maximum range standardization, standardization).

(1) Points of space

Spatial interpolation technology is the main method of space point factors of attribute value. Generally, the spatial interpolation can be expressed by the formula:

$$F(r_j) = z_j \quad j=1,2,\dots,N$$

Among them, z_j is the d dimension space discrete point measurement values of $r_j=(x^1_j, \dots, x^d_j)$. And the above function to a specific target for additional conditions, common spatial additional condition is variation function, polynomial function and so on.

(2) Space of linear features

Linear elements generally refer to road, river and so on. These elements of space through computing the evaluation of linear elements unit to achieve the effect of density, however, spatial range is needed in meteorological disaster risk zoning in the evaluation of these factors. Therefore it generally takes the first set up of the linear elements buffer, and then calculates the evaluation unit of linear elements to achieve linear buffer area elements of the density calculation. Finally, the spatial surface features. .

The spatial surface feature element generally refers to vegetation, land use types, which by computing the evaluation unit areal feature coverage is achieved, that the calculation of unit area of surface elements in grid coverage area, the function formula shown as below:

$$D = \frac{L}{A} = \frac{nl}{na} = \frac{l}{a}$$

Where D represents the area feature coverage, n is the number of grid network, L represents the total area covered by planar elements, A represents the total area of the grid, and L denotes the grid unit area and unit grid covering planar elements area.

3.2 Establish the Evaluation Factors Attribute Weights

By using GIS spatial partitioning technique, evaluation factors attribute value space technology for data preprocessing and then forming a decision table as showed in Table 2:

Table 2. Calculated of Meteorological Disaster Risk Decision

Unit Number	Local attribute	Condition attribute			Decision attribute
1	(x1, y1)	A1	B1	E1
2	(x2, y2)	A2	B2	E2
...

On the basis of Table 2, using rough set attribute importance to establish the principle of evaluation factors attributes weighting system.

3.3 Meteorological Disaster Risk Composite Index Calculated Based on GIS

Layer superposition is the basic function of GIS, and logic, probability of a laminated, fuzzy and superimposed method is commonly used. The spatial data digital overlay refers to as a continuous variable property of spatial objects, according to certain rules, transforming it into a specific digital interval arithmetic combination and forming property of the new space object. Space partition is dividing the continuous space object for the evaluation of the relative homogeneous unit, then spatial attribute data of the spatial variables of continuous distribution into discrete values. When the above data stored in the database data table, you can digitally start corresponding spatial data and computation. It can be described as the following mathematical model for calculation of meteorological disaster risk index:

$$S = \sum_i^n D_i \times P_i$$

where, S representative quantifies meteorological disaster risk index, Di quantifies the impact of meteorological disasters factor index values, Pi representatives use rough set attribute importance to establish the principle of evaluation factors attribute weights.

4. Application

The above technique is applied to Jiading District, Shanghai to analysis the Regionalization rainstorm water logging disaster risk.

4.1 Survey, Evaluation and Selection of Study Area Factor

4.1.1 General Situation of Study Area

Jiading District is located in the northwest of Shanghai City, the South Bank of the Yangtze Estuary, the center position for the latitude 31°23' and east longitude 121° 15', and the west border is Kunshan, Jiangsu, the East border is the two adjacent area of Baoshan and Putuo, the north is Liuhe river, which adjacent to the Taicang city of Jiangsu province, the south is Suzhou River, across Minhang, Changning and Qingpu. The whole terrain is the flat and slightly higher in the northeast and lower in the southwest, which is the accumulation landform type, the total area is 464.2 square kilometers, mainly divided into lacustrine plain, coastal plain and alluvial plain. Lacustrine plain is located in the west of the district, with wide 4 kilometers northwest strip, flat terrain and elevation are 3.4 - 3.6 meters, which accounting for about 6.01% of the total area of the region. The coastal plain of the region accounted for about 80.84% of the total area and the alluvial plain is about 13.15%. City, district river Wen zaobang, Lian qitang, Lou tanghe flows eastward through Baoshan District through the Yangtze River and the Huangpu River. And the Yan tietang, Hengli, new Cuo pu runs through the north and the south, connected with the Wusong River, Liuhe. The channel length is 1800km and the average river network density is 4 kilometers per square kilometer.



Figure 1. Schematic Diagram of the Study Area

4.1.2 Selection of Evaluation Factors

Storm frequency extracted from 30 years of regional precipitation data as hazard risk factor, drainage density, elevation drop and the drainage capacity of disaster as environmental sensitivity factor join with population density, road density, farmland hazard bearing body density factor to study the extent of flood disaster risk assessment and grading.

4.2 Data Preprocessing

In the administrative area of Jiading District Space division by 200 meter grids, the formation of evaluation units, as showed in Figure 2.

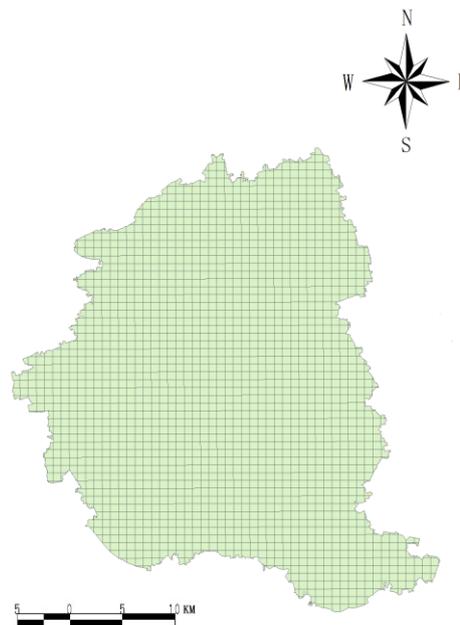


Figure 2. Spatiality of the Partition of the Grid Unit

Of the selected factors attribute value after standardization for space computation, each discrete attribute value gives the evaluation unit and the formation of an evaluation factor index value spatial distribution map, as showed in Figure 3.

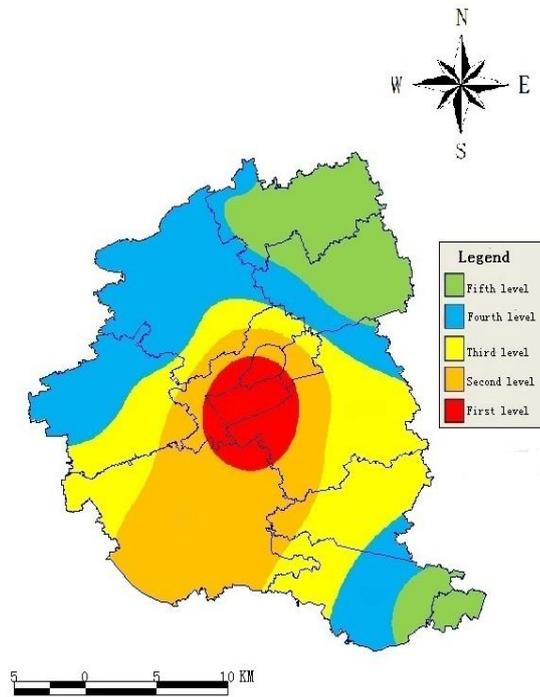


Figure 3. Spatiality of the Rainstorm Frequency

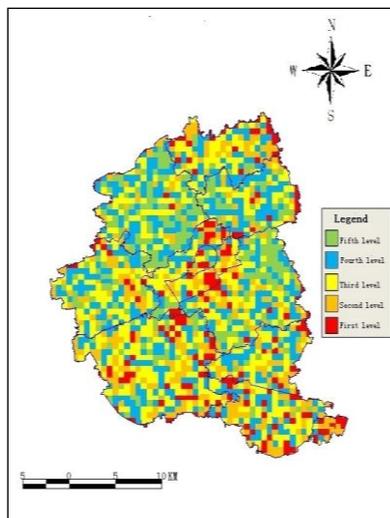


Figure 4. Spatiality of the River Network Density

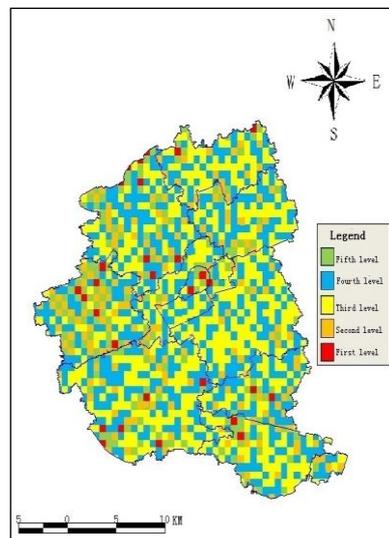


Figure 5. Spatiality of the Elevation Drop Height

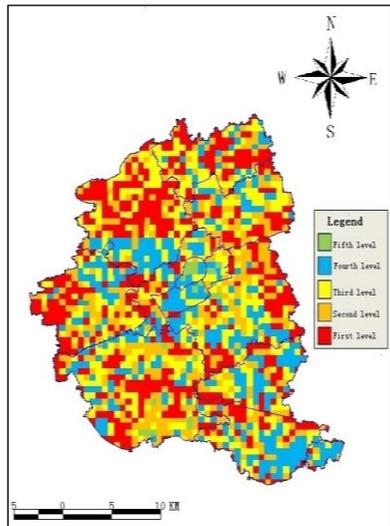


Figure 6. Spatiality of the Drainage Capacity

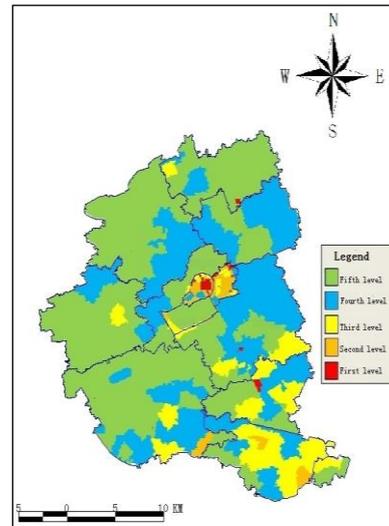


Figure 7. Spatiality of the Population Density

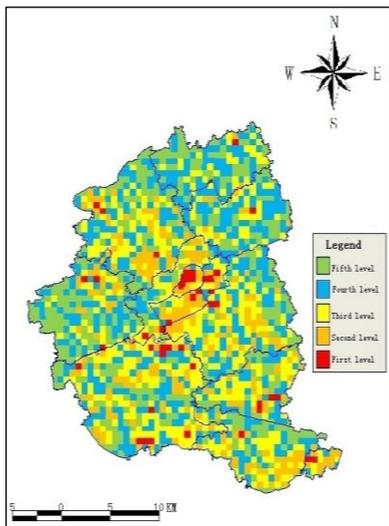


Figure 8. Spatiality of the Road Density

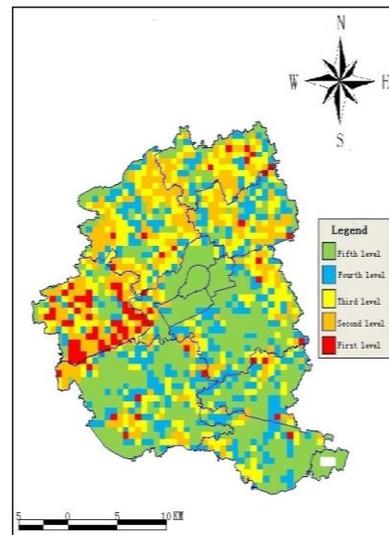


Figure 9. Spatiality of the Plowland Density

4.3 Determine The Weights of Evaluation Factors

This consists of choosing the area that has had heavy rain waterlogging disasters within the survey combined by the loss of choice with high, medium and low three regions in the Jiading District to establish standardized evaluation unit for each attribute value (condition attributes) and disaster losses (decision attribute). The decision table is as shown in Table 3.

Table 3. Decision Table

unit	rainstorm frequency	river network density	elevation drop height	drainage capacity	population density	plowland density	disaster damage
1	0.41	0.36	0.50	0.36	0.34		High
2	0.63	0.73	0.62	0.17	0.59		Medium
3	0.53	0.24	0.31	0.22	0.39		Low
4	0.28	0.11	0.29	0.32	.027		Medium
.....

Through the calculation of each factor, important degree of attributes is to obtain the weights of evaluation factors, as showed in Table 3:

Table 4. Evaluation Factors Weight Divisions of Rainstorm Water Logging Disaster Risk, Jiading District

Evaluation factor	Rainstorm frequency	River network density	Elevation drop height	Drainage capacity	Population density	Plowlands density
Weight	0.32	0.18	0.17	0.12	0.12	0.11

4.4 Risk Zoning Outcomes

Risk zoning in the evaluation unit data table is obtained, by using the weighted digital superposition calculation to obtain the comprehensive scores of each evaluation unit in the common effect of all factors under different evaluation unit and the frequency of occurrence of different scores through statistics evaluation unit and classifies it.

We can realize the risk grading of Jiading District rainstorm water logging disaster, as showed in Figure 10.

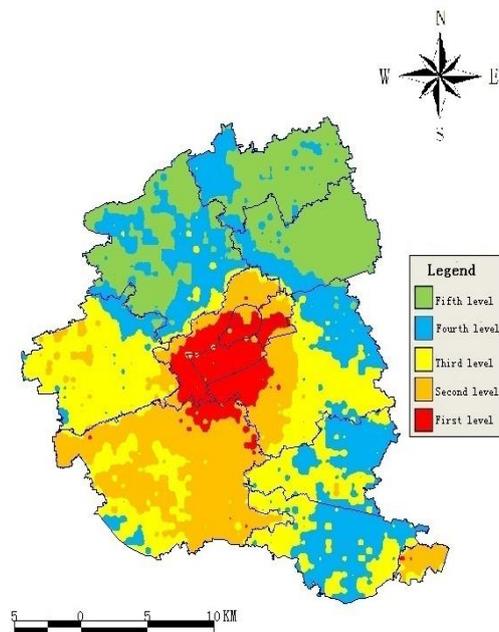


Figure 10. Rainstorm Water Logging Disaster Risk Zoning, Jiading District

5. Conclusion

Rough set theory can be more accurate and more comprehensive to characterize the relationship between the spatial data under incomplete information and conditions of uncertainty. Moreover, the use of GIS has good practical value for meteorological disaster risk zoning calculation to determine the relative importance of the different factors between improving the uncertainty of processing the spatial data and attribute data. GIS combined rough set theory can better evaluate all kinds of spatial data and social statistics meteorological disaster risk assessment, involved solving meteorological disaster risk. In addition, it can be better used to solving problem such when the data are imperfect and not sufficient or with gaps. Overall, it can come up with quantitative evaluation of meteorological disaster risk.

Acknowledgements

This work is supported by the postdoctoral scientific research fund of Jiangsu province (1101024B). It is also supported by the Industrial Strategic Technology Development Program (10041740) funded by the Ministry of Trade, Industry and Energy Korea.

References

- [1] M. Liu, R. S. Quan and S. Y. Xu, "Risk Assessment of Urban Rainstorm Waterlogging Disaster", Theory, Method and Practice, Science Press, (2012).
- [2] L. Nie, "Flooding Analysis of Urban Drainage Systems", Norwegian University of Science and Technology, (2003).
- [3] W. H. Lin, "The Comprehensive Evaluation Model Based On Rough Set Theory", Zhangzhou Normal University, (2013), pp. 4-8.
- [4] G. Y. Wang, Y. Y. Yao and H. Yu, "A Survey on Rough Set Theory and Applications", Chinese Journal of Computers, vol. 32, no. 7, (2009), pp. 1229-1246.
- [5] Y. L. Sang and Y. H. Qian, "A Granular Space Reduction Approach to Pessimistic Multi-Granulation Rough Sets", Pattern Recognition and Artificial Intelligence, vol. 25, no. 3, (2012), pp. 361-366.
- [6] X. J. Yang, "The Study Of Multi-Criteria Classification Decision-Making Problem Based On Rough Set Theory", Hebei University Of Technology, (2012), pp. 7-11.
- [7] M. L. Hu and S. F. Liu, "Rough Analysis Method Of Multi-Attribute Decision Making Base On Generalized Extended Dominance Relation", Control And Decision, vol. 22, no. 12, (2007), pp. 1347-1351.
- [8] Y. Q. He and S. S. Hu, "Rough Analysis Method Of Multi-Attribute Decision Making With Incomplete Information", Journal Of Systems Engineering, vol. 19, no. 2, (2004), pp. 117-120.
- [9] H. Zhu, Y. G. Fan and T. T. Wu, "Analysis Of Meteorological Disaster Risk Zoning Based on Open GIS", Meteorological and Environmental Sciences, vol. 35, no. 3, (2012), pp. 33-40.
- [10] G. A. Tang and X. Yang, "ArcGIS Experiments Courses Of GIS Spatial Analysis", Beijing: Science Press, (2012), pp. 8-10.
- [11] M. Han, J. J. Zhang and F. Peng, "Bayesian Rough Set Model Based On Multiple Decision Classes", Control and Decision, vol. 24, no. 11, (2009), pp. 1615-1619.
- [12] B. B. Yang, J. Y. Yang and C. Wu, "Rough Set Model Based on Characteristic Relations in Incomplete Information System", Pattern Recognition and Artificial, vol. 20, no. 4, (2007), pp. 450-457.
- [13] Y. Y. Xie, S. Q. Han and Y. L. H., "Risk Analysis Of Urban Rainfall Waterlogging In Tianjin City", Journal Of The Meteorological Sciences, vol. 24, no. 3, (2004), pp. 342-349.
- [14] D. Dubois and H. Prade, "Rough Fuzzy Sets And Fuzzy Rough Sets", International Journal Of General Systems, (1990), vol. 17, pp. 191-209.

Authors



Fengchang Xue, He is an associate professor of school of geography and remote sensing, Nanjing University of Information Science & Technology. He received the PhD degree in Cartography and Geographic Information Engineering from China University of Mining and Technology (2008). His research interests include Meteorological GIS (Geographic Information Systems), Surveying and mapping information.



Jin Hu, He is a Master in Meteorological IT from School of Atmospheric Science, Nanjing University of Information Science & Technology. He received the Bachelor degree in Computers with applications at Nanjing University of Information Science & Technology (2013). His research interests include Meteorological GIS (Geographic Information Systems), Risk Regionalization of Meteorological Calamities.