Study on Forest Pests and Diseases Forecast based on the Rough Set Theory

HongHan Zhang[1] andYanRong Zhang[1]
[1]School of computer and information engineering
Harbin University of Commerce
Harbin, China
zhangyanrong_5@163.com

Abstract

This paper applies the rough set theory and artificial intelligence technology into the field of forestry diseases and pests. It proposed the model of forestry diseases and pests forecasting in the demonstration. Through the model validation and accuracy analysis, the accuracy of the result is high, and forecast results with the actual situation are match.. In addition, it proposed an improvement attribute reduction algorithm based on discernibility matrix. Reduced conditional attribute set of forestry diseases and pests forecasting based on the algorithm. Through the extraction and reduction of generation rules, obtained a new f forestry diseases and pests forecasting model based on rough set theory. By verifying, the model has achieved a good result.

Keywords: Rough Set Theory; Expert Systems, Pests and Diseases, Forecast

1. Introduction

As the most precious strategic resources of human society development, is constantly suffer from huge loss of various disasters. However in these major forest resources, the frequent biological disaster such as forest diseases and pests has been to limit the forestry development. With the artificial forestation area continues to grow, forest diseases and pests is always in upward trend. Therefore, we should take active measures to reduce the loss and avoid the damage caused by forest pest and disease.

Since forest pest forecasting expert system contains a greater amount of knowledge, in order to be able to quickly and accurately mine the data which has the potential value of the information from a chaotic mass of data and apply in forest pests forecasting the rough set theory is applied to the forest pest forecasting process in this paper. Through the forest pest forecasting data collection, completions and discrete, an improved discernibility matrix based attribute reduction algorithm is raised and the forest pest forecasting condition attribute set is reduced based on the algorithm, thereby extracts and reduces the generating rule, and obtains a new forest pest forecasting model based on rough set theory[1].

2. Rough Set Theory and the Application in Pests and Diseases Forecast

2.1. Information Indicates

In rough set theory, knowledge is a kind of object classification capabilities. Any objective thing is available to be described by some knowledge. According to these knowledge (of things different attributes or characteristics), things can be classified by different attributes or characteristics. About the object definition that refers to all the things around us and we can

imagine their concepts. The domain refers to knowledge and concrete or abstract things, which is the object associated with the context of this feature, it exists as a non-empty finite set of forms. Their definitions are as follows:

Definition 1: (Knowledge and concepts (categories or information granules)) let U be the object we are interested in the composition of non-empty finite set called a domain. Any subset of the domain U is called a concept or category, written $x \subseteq U$.

Definition 2: (Knowledge Base) given a domain U and U tuft on the equivalence relation S is called a tipple $K= (U, S)$ of a knowledge base about domain of or similar space.

Definition 3: (indiscernible relation (no clear relation)) Given a domain U and U tuft on the equivalence relation S, If $P \subseteq S$, and $P \neq Q$, then $\cap P$ (the intersection of all equivalence relations of P) is still the domain of an equivalence relation on U, called the indiscernible relation, denoted by *IND (P)*,

$$\forall x \in U \ [x]_{IND(P)} = [x]_P = \bigcap_{\forall R \in P} [x]_R \tag{1}$$

Classification of forest pests, for example, Given a forest pest species domain of $U = \{x_1, x_2, x_3, ..., x_{10}\}$ = {woodland pasture locust, grasshopper Si, dendrolimus sibiricus, dendrolimus punctatus, slug, Anoplophora glabripennis, Poplar tabaniformis, lapathi , gypsy moth, Adelges laricis }, These forest pests belong to different target pests (Orthoptera, Lepidoptera, Coleoptera, Hemiptera), science (Acridiidae, Any of Si Department, Lasiocampidae, Slimacodiidae, Cerambycidae, Aegeriidae, weevils, Lymantriidae, Adelgidae). Therefore, the classification of these pests can be orders, families and species that knowledge to describe. If we are under a particular attribute describes the classification of these pests situation, you can follow orders, families and other classified.

Head according to Category: Orthoptera = $\{x_1, x_2\}$, Lepidoptera = $\{x_3, x_4, x_5, x_7, x_9\}$, Coleoptera = $\{x_6, x_8\}$, Hemiptera = $\{x_{10}\}$.

In accordance with Section Category: Acrididae = $\{x_1\}$, Tettigoniidae = $\{x_2\}$, Lasiocampidae = $\{x_3, x_4\}$, Limacodidae = $\{x_5\}$, Cerambycidae = $\{x_6\}$, Aegeriidae = $\{x_7\}$, Curculionidae = $\{x_8\}$, Lymantriidae = $\{x_9\}$, Adelgidae = $\{x_{10}\}$.

In other words, two attributes define two equivalence relations: head R_1, Section R_2. Through these equivalence relation, we can get below domain of represented by a collection of the different divisions.

U/R_1= {{x_1, x_2}, {x_3, x_4, x_5, x_7, x_9}, {x_6, x_8}, {x_{10}}}

U/R_2= {{x_1}, {x_2}, {x_3, x_4}, {x_5}, {x_6}, {x_7}, {x_8}, {x_9}, {x_{10}}}

Knowledge of these equivalence classes form K = (U, {R_1, R_2}) in the elementary category. Basic category is determined by the intersection of elementary areas constituted [3].

Table 1. The Information Table of Forest Pests

U （Species）	R_1 （Order）	R_2 （Science）
x_1 （Locust）	Orthoptera	Acrididae
x_2 （Tettigonioidea）	Orthoptera	Tettigoniidae
x_3 （Dendrolimus）	Lepidoptera	Lasiocampidae
x_4 （Punctatus）	Lepidoptera	Lasiocampidae
x_5 （Limacodidae）	Lepidoptera	Limacodidae

x_6 （Glabripennis）	Coleoptera	Cerambycidae
x_7 （Paranthrene tabaniformis ）	Lepidoptera	Aegeriidae
x_8 （Lapathi）	Coleoptera	Curculionidae
x_9 （Gypsymoth）	Lepidoptera	Lymantriidae
x_{10} （Adelges laricis）	Homoptera	Adelgidae

2.2. Knowledge Reduction

Knowledge reduction is the core of rough set theory. The importance of the Knowledge in the general knowledge base has a big difference, and even some of that knowledge is unnecessary or redundant. The so-called knowledge reduction refers to the ability to maintain the knowledge base of classification under the same conditions, delete unnecessary knowledge.

During the forest pest forecasting process, real-time monitoring of its impact factor is a continuous process, so the accumulated information is growing. Therefore, in the forest pests forecasting, the knowledge minimalist become an important research direction. The main task of knowledge reduction is to reduce the amount of information, it will feature a number of irrelevant or redundant lose without affecting their original analysis and forecasting functions, That is, the n-dimensional information space $\{x_1,x_2,...,x_n\}$ is reduced to m-dimensional $\{x_1,x_2,...,x_m\}$ (m<n) but does not affect the original classification of knowledge. It is the nucleus' definition in the forest pest forecasting that is the dominant factor, based on which we can roughly predict the occurrence of forest pests and diseases or the occurrence of volume. Knowledge reduction will feature information after re-combined to create new decision-making rules, the premise of such decision rules may differ from the information and conclusions about the information prior to any one simple decision rules, but they can be obtained by reasoning identical or similar results.

3. Forest Pests Forecasting Model

3.1. Predict the Occurrence of Forest Pests

Using the effective temperature prediction method to predict the occurrence of dendrolimus, forest insect pests' each state to normal growth and development requires a certain threshold temperature, the completion of a developmental stage or complete an entire generation requires some effective temperature. For some forest pests, its developmental threshold temperature and effective accumulated temperature is constant, according to the value of the cumulative temperature, a sample of a group of insects, insect developmental state the number of days and the sample size is calculated. Effective temperature prediction formula is as (2) [5].

$$N = \frac{K \pm S_k}{T - \left(C \pm S_c\right)} \tag{2}$$

Under the various factors (including individual factors, environmental factors and marketing factors), Consumers come into being a certain sense of self and their lifestyle, specific self-consciousness and lifestyle can lead to appropriate consumer needs and

motivations. Consequently, this correspond the consumer behavior. And the realization of consumer behavior will impact the consumer characteristics of internal and the external environment.

Where N is the number of day's growth, K is the effective temperature, T is the average daily temperature, C is the developmental threshold temperature, S_k is the effective temperature standard deviation, S_c is the developmental threshold temperature standard deviation.

Effective temperature K, developmental threshold temperature C, effective temperature standard deviation S_k, S_c developmental threshold temperature standard deviation is calculated as (3), (4), (5), (6).

$$C = \frac{\sum f \sum fnT_{sum} - \sum fn \sum fT_{sum}}{\sum f \sum fn^2 - (\sum fn)^2} \tag{3}$$

$$K = \frac{\sum fn^2 \sum fT_{sum} - \sum fn \sum fnT_{sum}}{\sum f \sum fn^2 - (\sum fn)^2} \tag{4}$$

$$S_c = \sqrt{\frac{(\sum f)(\left|\sum fT_{sum}^2 - K \sum fT_{sum} - C \sum fnT_{sum}\right|)}{\left[\sum f \sum fn^2 - (\sum fn)^2\right](\sum f - 2)}} \tag{5}$$

$$S_k = \sqrt{\frac{(\sum fn^2)(\left|\sum fT_{sum}^2 - K \sum fT_{sum} - C \sum fnT_{sum}\right|)}{\left[\sum f \sum fn^2 - (\sum fn)^2\right](\sum f - 2)}} \tag{6}$$

Where T_{sum} is the accumulated temperature values, n is the Sampling a group of insect developmental instars of a number of days, f is the Sample size [7].

3.2. Forest Pest Forecast

3.2.1. The Method of Effective Base Projection: The amount of forest pests in their generations has a great association with the last generation population base. Smaller population base, and its number of occurrences of the next generation of the less; contrast, the larger population base, the higher the number the occurrence of the next generation. Therefore, the method of effective base prediction by forest pest population base, the next generation of its survival and the amount of forecast occurrence. This approach is more accurate to the year occurred in fewer generations of forest pests when the forest stand, climate, natural enemies and other circumstances are more stable.

Often in the number of forest pests fairly stable periods, such as after the winter and early spring when population base to investigate, to predict the next generation as a valid basis for generating capacity. And to grasp the impact of fluctuations in major forest pests factor changes, and conduct a comprehensive analysis to improve prediction accuracy.

According to a previous generation of forest pest insect population base to predict the next generation amount of forest pests, often using the formula (7) to calculate the amount of its next-generation occurrence [8].

$$P = P_0 \times e \times \frac{f}{m + f} \times (1 - a)(1 - b)(1 - c)(1 - d) \tag{7}$$

Where P is the volume of the occurrence of the next generation, P_0 is the previous generation population base, e is the average per head female fecundity, f is the number of females, m is the number of male,$(1-a)(1-b)(1-c)(1-d)$ is the namely egg, larva, pupa, adult reproductive viability before.

3.2.2. The Method of Climate Prediction: As the temperature and humidity of climate both have greater impact on the incidence of forest pests; some forest pests have certain selectivity about the temperature and humidity. In the appropriate temperature and humidity conditions, the occurrence of the amount will be larger, on the contrary, over the appropriate temperature and humidity range, the growth will be inhibited, and generation amount becomes smaller. These forest pests can be by drawing climate chart, graph analysis of its occurrence through climate volume and temperature, humidity, the relationship between the occurrences of the amount to its forecast.

Climate map is months or ten days total rainfall or relative humidity of the abscissa or ordinate, Or ten days, the average monthly temperature for the vertical axis or horizontal axis, the corresponding month or ten days the combination of temperature and relative humidity plotted as coordinate point, and then connect each coordinate point, thus forming a closed curve multilateral irregular. Draw the occurrence of pests' generations of climate diagram and some pest occurrence appropriate temperature range plotted in climate chart, you can compare the analysis of temperature and humidity and the amount of forest pests in the relationship.

In addition to climate map can be clearly seen in the forest pest outbreaks occurred in annual and small, but it is evident that different combinations of temperature and humidity have a different amount of forest insect pests. This can provide the basis for the forecast amount occurred. Climate diagram of a pest may also be suitable and appropriate temperature and humidity range of temperature and humidity range plotted in the figure, so that the analysis can be more clearly certain pests in the relationship between the amount of temperature and humidity. For example, the life history of Dendrolimus as shown in Table 2.

Table 2. Life History of Dendrolimus Superans

Mid-May	1~4 Early midmonth Late	5 Early midmonth Late	6 Early midmonth Late	7 Early midmonth Late	8 Early midmonth Late	9 Early midmonth Late	10~12 Early midmonth Late
First year				•	• • •		
				– – –	– – –	– – Δ	Δ Δ Δ
Second year	Δ Δ –	– – –	– – –	– – –	– – –	– – Δ	Δ Δ Δ
	Δ Δ –	– – –	– – –	– – –			
Third year			⊕ ⊕	⊕ ⊕ ⊕			
				+ +	+ + +		

The two-year in July, the middle and late larvae of Dendrolimus average temperature during the occurrence of relative humidity and atmospheric climate chart compared with each other, we can see these two years the incidence of Dendrolimus different as shown in Figure 1.
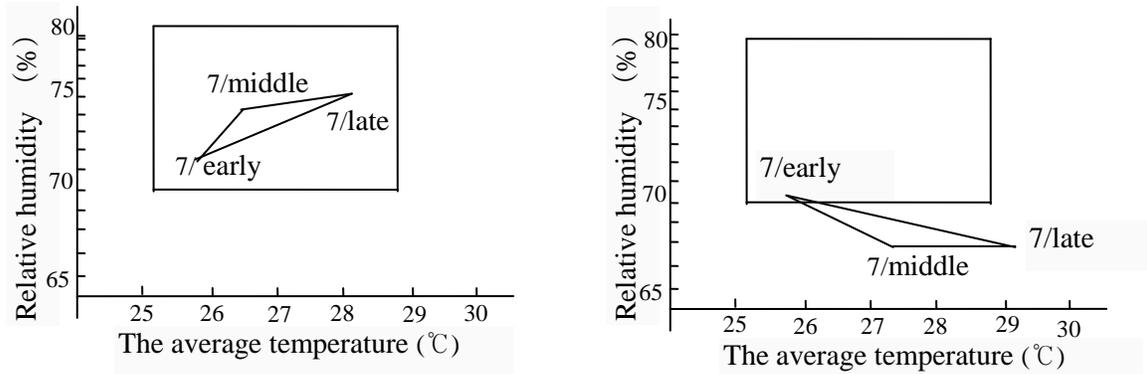
Figure 1. The Bioclimatic Char of Larvae of Dendrolimus Superans in July

4. Pests' Prediction Based on Rough Set Theory

The rough set theory is introduced to the forest pests forecasting, use rough set theory attribute reduction algorithm on forest pest forecasting the condition attribute set simplified to remove redundant information on forest pest forecasting after the rules were extracted and for about Jane, and calculate the prediction rule coverage and confidence to carry out the filtering rules, we propose a new theory based on rough sets forest pests forecasting model. As a method of data processing, analysis, rough set is kept in the same circumstances classification ability, through knowledge reduction, so as to arrive in decisions or classification rules.

For rough sets, the object of study is determined by temperature, humidity, rainfall and other multi-valued attribute collection description consisting of a collection of objects. For each object and its attribute has a symbolic value as its description. Therefore, objects, attributes, and descriptors constitute rough set data processing research space's three basic elements. This paper is tentatively applied rough set to the prediction of forest pests.

In general, based on rough set theory forest pest forecasting specific operation includes the following steps:

① Forest pest forecasting basic data preprocessing, including data collection, data completeness, data discredited.

② Removing redundant information on forest pest forecasting condition attribute set reduction.

③Draw pest forecasting rule sets on forest pest forecasting rules for extraction and reduction.

④ Calculation rules on confidence and coverage for rule by filtration to obtain the most valuable rules.
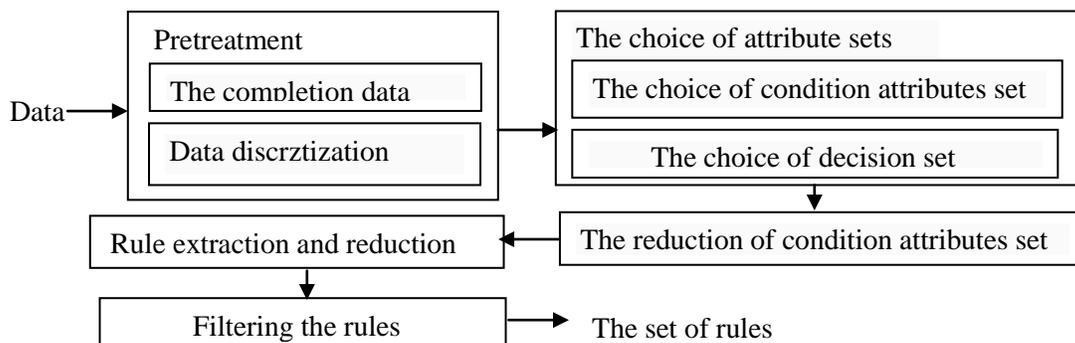
Figure 2 Forecasting Process of Diseases and Pests Based on Rough Set

5. Expert System Designs Based on Rough Set Theory

5.1. System Architecture Design

Forest pest forecasting based on actual demand, the thesis of the "forest pest forecasting expert systems" functionally relevant knowledge should consult with pests, forest pests and diseases and the outbreak of forecasting, forest pest forecast the degree of harm and damage evaluation , forest pest diagnosis, forest pest control, forest pest information query and so on. To ensure the adaptability of the system is running, the user ease of use and accuracy of the results predicted, "forest pest forecasting expert system" should also have pest information browsing, database / knowledge base maintenance system information and other auxiliary functions [9]. "Expert systems for predicting forest pests" system is mainly composed of knowledge base, inference engine, the interpreter, blackboards and several other components.

5.2. Design of Knowledge

The knowledge of forest pest forecasting including the rules knowledge base of forest pest forecasting, pest forecasting methods knowledge base and disease forecasting methods knowledge base. Including forest pest forecasting the occurrence of pests and forest pest forecast generating capacity forecast two kinds. In which the occurrence of forest pests forecasting including Interval method, effective temperature method, phrenology forecasting methods such as; generating capacity forecast of a valid credit method and mathematical statistics both methods. Forecasting disease occurs mainly on the year, the incidence of stage, disease severity and infection index four kinds of predictors for forecasting, different forecasting methods used indicators also vary, specific knowledge are stored in the knowledge base and pest prediction rules Knowledge disease among forecasting methods. As predicted target different, the prediction method is also different, so that, its principles, knowledge structure is different, the system mainly based on different mathematical models for different forecasting methods to organize, establish the appropriate knowledge base.

5.3. Design of Inference Engine

Forest pests forecasting expert system inference engine using backward reasoning design philosophy is: forest pest forecasting the selected target (the incidence of certain diseases or pests or the occurrence of the predicted forecast) will be credited to the fact that the selected target library and mixed with target-related knowledge in knowledge matches, get some sort of insect pests forecasting period or method of generating capacity, the user selects a prediction forecasting methods, will be credited to the fact that its library, call the relevant methods knowledge base, each condition to get the method, enter or select the value of each condition, the expert system based on different data models to be drawn on the occurrence of a pest or the occurrence of forecasting the amount of results, and the results provided to the users[10].

Backward reasoning based reasoning as follows:

① Give the target to solve.

②Check whether the target has been the fact that library. If, the target was established, reasoning successfully, go to (6). If the facts do not exist in the target database, then go to (3) step.

③Determine whether the target is evidence. If the goal is not to be confirmed by the user to the initial facts, go to (4). Otherwise, after asking the users continue the reasoning.

④Match the objectives and the knowledge in knowledge base, result applicable knowledge and form the applicable knowledge set, and then go to (5) step.

⑤Using the conflict resolution strategies selected a highest priority for further reasoning from the knowledge base knowledge and make the knowledge of the antecedent as a new target, and then go to (2) step.

⑥End.

6. System Implementation

6.1. Forest Pests Expert System Implementation

"Expert system for predicting forest pests" in accordance with versatility, low cost and easy maintenance requirements provide for the relevant departments, research institutions, agriculture, forestry and other related personnel with forest pests forecasting, damage assessment and control, such decision knowledge is as show in Figure.3. The system's main functions include pest forecasting and pest damage assessment, in addition, "forest pest forecasting expert system" also includes information model regions common tree species, forest pest forecasting related quiz, forest pest diagnosis, forest pest management, information inquiry demonstration areas and pests and other auxiliary functions. Systems having an information rich, vivid, friendly interface, which will help users understand and apply other characteristics.

Figure 3. Expert System Implementation

6.2. Analysis of Prediction Accuracy

This paper is based on historical data of pests and diseases occurred in the past 10 years in demonstration areas. Compare the forest pest expert system predicted forecast data with the actual occurrence forecasting, verify and analysis the expert system. Forecast results basically consistent with the actual situation, in line with rate reached 90%. To occur early fall larch disease, for example, the results of its forecast comparison with the actual situation is as show in Figure.4.

_____ Actual condition   _ _ _ _ Prediction result
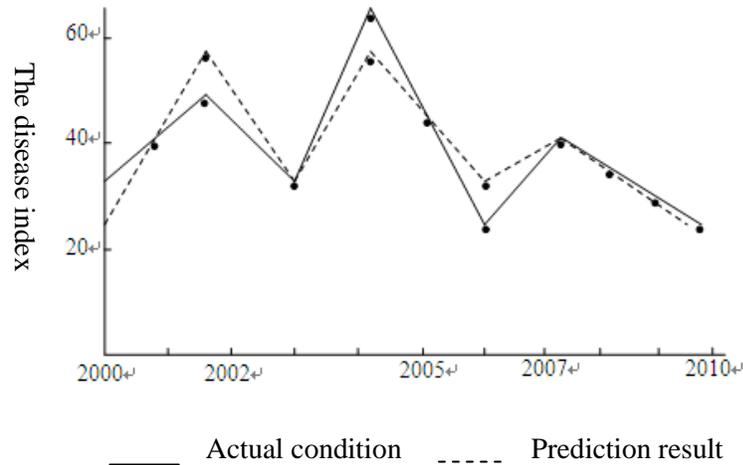
Figure 4. Prediction Results Compared with Actual Condition

7. Conclusion

In this paper, three-layer B / S architecture and standards-based J2EE development model establishment of forest pest forecasting expert system. This mode allows the performance of front-end application logic, data storage phase separation, through component development and deployment strategies, so that the whole system is clear and flexible framework to facilitate the deployment and expansion. Business processing use the way of Jsp + Java bean to achieve and through JDBC to access the data- layer data resources. Forest pest forecasting expert system mainly realizes common occurrence of forest pests and diseases and the occurrence of the amount of the forecast period in the main demonstration area. It can provide for forest pest control and quarantine stations and other security agencies with scientific control foundation and decision support in demonstration areas, which can minimize the occurrence of forest pests and diseases, caused direct economic losses and indirect economic losses.

Acknowledgements

References

[1] Y. H. Wei, K. F Hua and L. X. Zhen, "RS and GIS based forest Fire risk zone mapping in DA hingfan mountains", [J], Chinese geographical Science, vol. 3, **(2004)**, pp. 251-257.
[2] W. Suli, L. Quan-sheng, Q. Zhijie and L. Xiaojun, "Study of Setting up The Forest Resources Management Information System Based on WebGIS", [J], Chinese Geographical Science, vol. 1, **(2003)**, pp. 56-61.
[3] W. Hsiao and J. C. Trappey, "An integrated platform of collaborative project management and silicon intellectual property management for IC design industry", [J], Information Sciences, vol. 197, no. 15, **(2009)** , pp. 2576-2590.
[4] C. Kwok, O. Etzioni and D. S. Weld, "ScalingQuestion Answering tothe Web", [A], In:Proceedings of the 10th World WideWeb Conference,HongKong, **(2001)**, pp. 150-161.

[5] L. ZhiHai, Z. QingLiang and L. YuShan, "Research of equipment selection and matching expert system in fully mechanized caving face based on ontology", [J], Key Engineering Materials, **(2010)**, pp. 419-420, pp. 117-120.

[6] L. F. Zhang and J. Ai, "Method for Bridge Bearing Capacity Assessment Based on Analytic Hierarchy Process", [J], Transactions of Nanjing University of Aeronautics &Astronautics, vol. 26, no. 3, **(2009)**, pp. 236-241.

[7] R. E. Plant and N. D. Stone, "Knowledge-based systems in agriculture", [J], McGraw-Hill, **(1991),** New York, USA.

[8] G. L. Diaz., M. P. Jimenez. and F. Bastida, "Expert system for integrated plant protection in pepper", [J], Expert Systems with Applications, vol. 36, **(2009)**, pp. 8975-8979.

[9] G. Joseph and R. Gary, "Expert Systems Principles and Programming", [M], Beijing: Machine Press, **(2002).**

[10] S. Kaloudis, D. Anastopoulos, C. P. Yialouris and N. A. Lorentzos, "Pest identification expert system for forest protection", [J], Expert Systems with Applications, vol. 28, **(2005)**, pp. 445-452.

[11] Y. F. Su, S. Joan and M. Avi, "Distributing proprietary geographic data on the World Wide Web UCLA GIS database and map serve", [J], r. Computer Science, vol. 7, **(2000)**, pp. 741-749.

Authors

Honghan Zhang, Master supervisor, the committee member of economic management branch committee of Ministry of Education Liberal Arts Computer Basic Teaching, the committee member of Computer Basic Education Research Institution of China Higher Education, His main research fields include E-commerce and E-government, computer network and security.
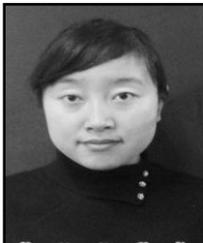
Yanrong Zhang, Doctor, Lecturer, teacher of School of the Computer and Information Engineering, Harbin University of Commerce. Her main research fields are E-commerce and Artificial Intelligence.