IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 3, January 2013
ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814
www.IJCSI.org

Neighborhood covering rough set model of fuzzy decision system

Bai-ting Zhao[1], Xiao-fen Jia[2]

[1] Electrical and Information Engineering College, Anhui University of Science and Technology, Huainan, China, 232001

[2] Electrical and Information Engineering College, Anhui University of Science and Technology, Huainan, China, 232001

Abstract

The neighborhood covering rough set model is established for the fuzziness of decision system. Firstly, the knowledge representation of fuzzy decision system is analyzed. Then fuzzy neighborhood relation is proposed to measure the fuzziness of the decision system. Finally prove that the classical indiscernibility relations and classical neighborhood relationship are special case of fuzzy neighborhood relations. Fuzzy neighborhood covering rough set model solved the problem of processing fuzzy attribute, and deal with the hybrid decision system which including both of fuzzy attributes and numerical attributes. The proved model expanded the applications of rough set.

Keywords: Fuzzy, Decision System, Neighborhood, Rough Set.

1. Introduction

The real world is diverse, complex and changeing, and the expressions of people to information are often imprecise, uncertain and fuzzy. The fuzziness is the basic characteristic of information uncertainty. The representation and processing of the fuzzy knowledge in fuzzy hybrid decision system has become one of the most important key issues in the research of artificial intelligence. Pawlak rough sets are based on the equivalence relations, to research rough approximation problems of distinct sets. The classical rough sets cannot deal with the fuzzy decision system which contains fuzzy attributes.

Fuzzy rough sets and rough fuzzy sets are important promotion of Pawlak rough sets model. Both rough set and fuzzy set can be used to deal with uncertainty and imprecision problems, therefore the organic combination of them is a good tool to process fuzzy information system. Dubois proposed a fuzzy rough sets model based on the promotion of rough fuzzy model to fuzzy approximation space. Mi improved axiomatic definition of the ambiguity, Feng proposed a new definition of the ambiguity. The different equivalence relations may have the same hierarchical structure are researched by Zhang ling. Zhang qinghua proposed information entropy sequence based on the general fuzzy relations, and the relations between fuzzy similarity relation, fuzzy equivalence relation, the hierarchical structure and the entropy sequence of the hierarchical structure.

2. KNOWLEDGE REPRESENTATION OF FUZZY DECISION SYSTEM

Knowledge is represented by knowledge system, the basic ingredient is the set of the studied object. The decision system can represent knowledge and describe the object by the attributes and attribute values.

Definition 1. Definite an knowledge system as $KRS=(U,A,V,f)$. $U=\{X_1,X_2,\cdots,X_n\}$ is a nonempty finite set, which be called the universe. $A=\{a_1,a_2,\cdots,a_N\}$ is a nonempty finite set of attributes. $V=\bigcup_{a\in A}V_a$, where V_a is a domain value of the attribute a, and $f:U\times A\to V$ is called the information function such that $f(x_i,a)\in V_a$ for each $a\in A, x_i\in U$.

Definition 2. Definite an decision system as $DT=(U,C\bigcup D,V,f)$. If $A=C\bigcup D$ and $C\bigcap D=\varnothing$, C is a finite set of condition attributes and D is a finite set of decision attributes, the information system is called decision system.

Definition 3. Definite fuzzy decision system as $FDS=(U,A,V,f,D)$. $U=\{x_1,x_2,\cdots,x_n\}$ is the set of objects, $A=\{a_1,a_2,\cdots,a_m\}$ is the set of attributes, $f=\{f_l:U\to V(l\le m)\}$ is the mapping of objects set to attributes set, and $D=\{D^{\sim}{}_j:U\to[0,1](j\le r)\}$ is fuzzy decision set.

Copyright (c) 2013 International Journal of Computer Science Issues. All Rights Reserved.

IJCSI
www.IJCSI.org

The the fuzzy decision system is the fuzzy knowledge representation of information systems, object attributes and attribute values in the system can be a good description of the object.

3. NEIGHBORHOOD ROUGH SET MODEL OF DECISION SYSTEM

Lin proposed neighborhood model in 1988. Hu build neighborhood rough set model and proposed hybrid reduce algorithm.

Definition 4. Let U is a nonempty finite set, if $\exists\Delta$ which is a real function corresponding to $\forall x \in U$ and satisfy the follows:

(1) $\Delta(x_i, x_j) \geq 0$,and $\Delta(x_i, x_j) = 0$ only when $x_i = x_j$;

(2) $\Delta(x_i, x_j) = \Delta(x_j, x_i)$;

(3) $\Delta(x_i, x_k) \leq \Delta(x_i, x_j) + \Delta(x_j, x_k)$.

$\triangle$is distance function of U, and $< U, \Delta >$ is distance space which also called metric space.

In Euclidean space of dimension N, to random two points of $x_i = (x_{1i}, x_{2i}, \cdots, x_{Ni})$ and $x_j = (x_{1j}, x_{2j}, \cdots, x_{Nj})$, the distance of them is:

$$\Delta(x_i, x_j) = \left(\sum_{i=1}^{N} (x_{li} - x_{lj})^2 \right)^{1/2}$$

Definition 5. Let $< U, \Delta >$ is a nonempty metric space, $x \in U$, $\delta \geqslant 0$, the points set $\delta(x) = \{y \mid \Delta(x, y) \leqslant \delta, y \in U\}$ is a closed sphere and is called δ neighborhood of x.

When the attributes both include numerical and characteristics values, let $B_1 \subseteq A$ is numerical attributes and $B_2 \subseteq A$ is characteristics attributes, then the neighborhood of sample x is:

(1) $\delta_{B_1}(x) = \{x_i \mid \Delta_{B_1}(x, x_i) \leq \delta, x_i \in U\}$

(2) $\delta_{B_2}(x) = \{x_i \mid \Delta_{B_2}(x, x_i) = 0, x_i \in U\}$

(3) $\delta_{B_1 \cup B_2}(x) = \{x_i \mid \Delta_{B_1}(x, x_i) \leq \delta \wedge \Delta_{B_2}(x, x_i) = 0, x_i \in U\}$

Let $NAS = (U, N)$ is a neighborhood approximate space and $X \subseteq U$, the lower and upper approximations of X in $NAS = (U, N)$ can be defined as follows,

$$\begin{cases} \underline{N}X = \{x_i \mid \delta(x_i) \subseteq X, x_i \in U\} \\ \overline{N}X = \{x_i \mid \delta(x_i) \cap X \neq \varnothing, x_i \in U\} \end{cases}$$

4. Neighborhood covering rough set model (NCRSM) for fuzzy decision system

Let U is a nonempty finite set, $C = \{C_1, C_2, \cdots, C_k\}$ is fuzzy covering of U , $\mathbf{I}$ is arbitrary contain, $D^{\mathbf{I}}$ is fuzzy inclusion degree of $\mathbf{I}$. A=$(F(\mathrm{U}), C, D^{\mathbf{I}})$ is fuzzy contains approximation space of $\mathbf{I}$. To $\forall X \in F(U)$ and $0 \leq l < u \leq 1$, the lower approximation $\underline{A}_u(X)$ and upper approximation $\overline{A}_l(X)$ of X about A can be defined as follows:

$$\underline{A}_u(X) = \bigcup\left\{C_i \in C \mid D^{\mathbf{I}}(C_i, X) \geq u\right\}$$
$$\overline{A}_l(X) = \bigcup\left\{C_i \in C \mid D^{\mathbf{I}}(C_i, X) > l\right\}$$

Definition 6. When the attributes both include clear and fuzzy values, numerical and characteristics values, let $B_1 \subseteq A$ is numerical attributes and $B_2 \subseteq A$ is characteristics attributes, let $B_3 \subseteq A$ is fuzzy attributes, then the neighborhood of sample x is:

1) $\delta_{B_1}(x) = \{x_i \mid \Delta_{B_1}(x, x_i) \leq \delta, x_i \in U\}$

2) $\delta_{B_2}(x) = \{x_i \mid \Delta_{B_2}(x, x_i) = 0, x_i \in U\}$

3) $\delta_{B_3}(x) = \{x_i \mid \Delta_{B_3}(f_l(x), f_l(x_i)) = 0, x_i \in U\}$

4) $\delta_{B_1 \cup B_2 \cup B_3}(x) = \{x_i \mid \Delta_{B_1}(x, x_i) \leq \delta \wedge \Delta_{B_2}(x, x_i) = 0 \wedge \Delta_{B_3}(f_l(x), f_l(x_i)) = 0, x_i \in U\}$

Further we can get the properties as follows:

1) $\delta(x_i) \neq \varnothing$,for $x_i \in \delta(x_i)$

Copyright (c) 2013 International Journal of Computer Science Issues. All Rights Reserved.

IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 3, January 2013
ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814
www.IJCSI.org

2) $\bigcup_{i=1}^{n} \delta(x_i) = U$

Neighborhood Particle tribe $\{\delta(x_i) \mid i = 1,2,\cdots,n\}$ can constitute a covering of U. However, due to $x_i \neq x_j$ can not ensure $x_j \notin \delta(x_i)$, so $\{\delta(x_i) \mid i = 1,2,\cdots,n\}$ generally does not constitute a partition of U.

The neighborhood equivalence relation extended to fuzzy and clear attributes, characteristics and numeric attributes coexist fuzzy decision system, we can get fuzzy neighborhood relations:

$$\begin{aligned} R(X) = \{(x,y) \in U^2 : \forall a \in X \\ \bigcap a(x) \neq ? \bigcap f_l(x) = f_l(y), \\ a(x) \in \delta(y,a) \bigcup a(y) \in \delta(x,a)\} \end{aligned}$$

From the basic nature of the neighborhood, the neighborhood relations satisfy reflexivity, symmetry, transitivity, so fuzzy neighborhood relations can be further simplified as:

$$\begin{aligned} R(X) = \{(x,y) \in U^2 : \forall a \in X \\ \bigcap a(x) \neq ? \bigcap f_l(x) = f_l(y), \\ a(x) \in \delta(y,a)\} \end{aligned}$$

The simplified fuzzy neighborhood relations meet reflexivity but not necessarily to meet symmetry and transitivity. The relaxation of the requirements to symmetry and transitivity can avoids overly conservative, and broaden the range of applications.

Theorem 1: The classic neighborhood relation is a special case of fuzzy neighborhood relation.

When the decision system is completely clear, fuzzy neighborhood relation degenerate to the classic neighborhood relation as follows:

$$\begin{aligned} R(X) = \{(x,y) \in U^2 : \forall a \in X \\ \bigcap a(x) \neq ?, a(x) \in \delta(y,a) \\ \bigcup a(x) = * \bigcup a(y) = *\} \end{aligned}$$

Theorem 2: The classic Pawlak indiscernibility relation is a special case of fuzzy neighborhood relation.

Let $FDS = <U, A, D>$ is FHDS, U is divided into N equivalence class by D as follows: $X_1, X_2, \cdots, X_N$, $B \subseteq A$ generate the neighborhood relation of U, so the upper approximation and lower approximation of decision D to B is :

$$\begin{cases} \underline{N_B} D = \{\underline{N_B} X_1, \underline{N_B} X_2, \cdots, \underline{N_B} X_N\} \\ N_B D = \{N_B X_1, N_B X_2, \cdots, N_B X_N\} \end{cases}$$

5. Simulation

The "Wpbc" and "Segmentation" of UCI machine learning database are used to verify the reasonable and validation of NCRSM model. Reduction and forecasting simulation compared with the Dubois fuzzy rough set model and fuzzy Radzikowska rough set model, and results as follows.

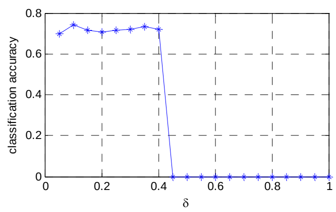
The data sample is divided into a training set and a test set, and then attributes are reduced by the algorithm based on attributes significant.

The SVM classifier is used as the evaluation function, in which used spline function. Raw data and reduction data are used to train the SVM separately, the prediction accuracy is used to evaluate the quality of the reduction. Shown form the results the fuzz neighborhood rough set model is better than Dubois and Radzikowska model.

Table 1 Prediction accuracy

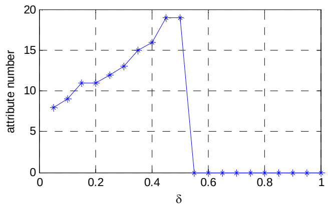
Data	*Raw*	*Dubois*	*Radzikowska*	*NCRSM*
Wpbc	84.6%	81.37 $\pm$ 6.02%	78.26 $\pm$ 3.89%	80.24 $\pm$ 3.93%
Segmentation	96.37%	94.26 $\pm$ 5.74%	95.41 $\pm$ 4.59%	94.74 $\pm$ 3.63%

In order to study the affection of the selection and the size of neighborhood operator to classification accuracy and attributes reduction number, the "Image" are used in reduction and prediction simulation. The results are shown in Fig.1.

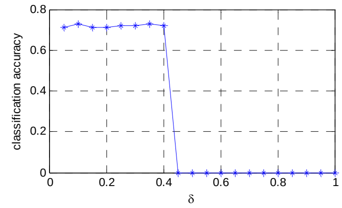


(a)

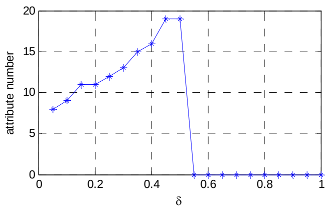
Copyright (c) 2013 International Journal of Computer Science Issues. All Rights Reserved.



(b)



(c)



(d)

Fig.1 Variation of accuracies and attribute numbers in reduction with δ of image

Figure.1(a) shows the classification accuracy changes with the different sizes δ which uses the 2-norm distance function. Figure.1(b) shows the reduction number of attributes changes with the different sizes δ which uses the 2-norm distance function. Figure.1(c) shows the classification accuracy changes with the different sizes δ which uses the infinite-norm distance function. Figure.1(d) shows the reduction number of attributes changes with the different sizes δ which uses the infinite-norm distance function.

Shown from Figure1.(a) and (c), when the value of δ is small, the particle size of classification is small. So fewer features can distinguish the decision attributes. When the value of δ is larger, the required characteristics to distinguish the decision attributes is also more. When the δ exceeds a certain degree, any characteristics will not distinguish any samples.

From Figure.1(b) and (c) we can see, the attributes number of reduction will increase with the increase of δ , but not monotonic changes. When the neighborhood operator exceeds a certain value, any feature is not sufficient to distinguish any samples and cannot obtain any reduction attributes. On the other hand, the increase of the attributes number of reduction may not improve classification performance. When the neighborhood operator is smaller, the increase of attributes can improve the accuracy of classification. Nevertheless, if attributes continue to increase, the accuracy of classification does not continue to be improved and even be declined slightly. Therefore the reasonable choice of neighborhood operator can obtain better classification effect.

Comparing Figure 1(a) (c) with Figure 1 (b) (d) we can see, using different distance functions has less affect to classification accuracy. The main factors affecting the classification accuracy is the size of the neighborhood. The value of δ depends on the specific classification problems. The general value is: [0.1 0.2].

6. Conclusion

Pawlak rough sets are based on the strict equivalence relation, which can only deal with characteristic attributes. But in practice, many information systems are hybrid systems which including characteristic attributes, numerical attributes and fuzzy attributes. In this paper a fuzzy neighborhood rough sets model is established for fuzzy hybrid decision system. The model uses the fuzzy neighborhood relation to measure the indiscernibility relation of the classical rough sets, and uses granulation to approximate the universe space, which can deal with fuzzy attributes directly. The neighborhood relation is a special case of the fuzzy neighborhood relation, and the fuzzy neighborhood relation can deal with the hybrid decision system include both numerical and fuzzy attributes.

The UCI machine learning databases are used in simulation experiments, and the simulation results proved the validity of the fuzzy neighborhood rough sets model which better than Dubois fuzzy rough sets model and Radzikowska rough sets model. Finally, we analyze the influence of the selection and size of neighborhood operator to the classification accuracy and the attributes number in reduction of fuzzy neighborhood rough set model. Fuzzy neighborhood rough set model is the promotion of Pawlak rough sets which provides effective solutions to the classification of hybrid decision system in the practical application.

Acknowledgement

The research is supported by the Youth Foundation of An hui University of science & technology of China under Gr ant No.12257, No.2012QNZ06, the Doctor Foundation of Anhui University of science & technology of China under Grant No.11223, and the Guidance Science and Technolog y Plan Projects of Huainan under Grant No.2011B31

Copyright (c) 2013 International Journal of Computer Science Issues. All Rights Reserved.

References

[1] WANG Guo-Yin, ZHANG Qing-Hua, MA Xi-Ao, YANG Qing-Shan. "Granular Computing Models for Knowledge Uncertainty", Journal of Software, Vol.22, No.4, 2011, pp. 676-694

[2] Dubois D, Prade H. "Rough fuzzy sets and fuzzy rough sets", Journal of General Systems, Vol.17, No.3,4, 1990, pp.191-208

[3] Dubois D, Prade H. "Putting rough sets and fuzzy sets together", In：Slowinski R, ed. Intelligent Decision Support：Handbook of Applications and Advances of the Rough Sets Theory. Boston: Kluwer Academic Publishers, 1992, 203-222.

[4] HU Jun, WANG Guo-Yin, ZHANG Qing-Hua. "Covering Based Generalized Rough Fuzzy Set Model", Journal of Software, Vol.21, No.5, 2010, pp. 968-977

[5] LUO Gong-zhi, YANG Xi-bei, YANG Xiao-jiang. "Limited dominance-based rough fuzzy set and knowledge reductions", Systems Engineering and Electronics, Vol.32, No.8, 2010, pp.1657-1661

[6] Huang bing, Wei da-kuan. "Distance-based rough set model in intuitionistic fuzzy information systems and its application", Systems engineering theory & practice, No.7, 2011, pp. 1356-1362

[7] Li Bai, Min Liu. "Fuzzy sets and similarity relations for semantic web service matching", Computers & Mathematics with Applications, Vol.61, No.8, 2011, pp. 2281-2286

[8] Degang Chen, Yongping Yang, Hui Wang. "Granular computing based on fuzzy similarity relations", Soft Computing, Vol.15, No.6, 2011, pp.1161-1172

[9] Hu Q. H, Liu J. F, Yu D. R. "Mixed Feature Selection Based on Granulation and Approximation", Knowledge Based systems, Vol.21, No.4,2008, pp.294~304

[10] Tripathy, B.K. Panda, G.K. " Approximate equalities on rough intuitionistic fuzzy sets and an analysis of approximate equalities ", International Journal of Computer Science Issues, Vol.9, No.2 2-3, 2012, pp.371-380.

[11] Li, Ju. Liu, Xiao-Ping. Du, Hui. " Research of image recognition based on rough set ", International Journal of Digital Content Technology and its Applications, Vol.6, No.9, 2012, pp.141-146,

Dr. Zhao received the Master. degree in control theory and control engineering from the Qingdao University of Science & Technology, in 2005. He received the Ph.D. degree in control science and engineering, from the Harbin Institute of Technology. Currently, he is a lectorate at Anhui University of Science & Technology, Electrical and Information Engineering College. His research interests include intelligent control and Rough sets.

Mrs. Jia received the Master. degree in control science and engineering, from the Harbin Institute of Technology. Currently, she is a lectorate at Anhui University of Science & Technology, Electrical and Information Engineering College. Her research interests include Rough sets and Image processing.

Copyright (c) 2013 International Journal of Computer Science Issues. All Rights Reserved.