Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 154 (2019) 194–198

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Congress of Information and Communication Technology, ICICT 2019

Research on Feature Selection/Attribute Reduction Method Based on Rough Set Theory

Shi Qiang Wang[1], Cai Yun Gao[2], Chang Luo[3,*1], Gui Mei Zheng[1], Yan Nian Zhou[1]

[1] Air and Missile-Defence College, Air Force Engineering University, Xi'an 710051, China.
[2] Basic Department, Air Force Engineering University, Xi'an 710051, China.
[3] Troops of 78092, ChengDu, 610000, China

Abstract:

Radar emitter signal is interfered by various noises during the propagation process, and the signal-to-noise ratio varies widely. Therefore, the features that play a key role in sorting and classifying or identification signals are often difficult to find. In addition, the extracted features are usually subjective and speculative, so it is necessary to select the features that can characterize the maximum difference mode information between the modulated signal categories and the changes in the signal-to-noise ratio. That is, the selected features also have good separability at low SNR. In this paper, based on rough set theory the feature selection method is studied, which lays a foundation for the feature selection of radiation source signals by rough set theory.

Keywords: radar emitter signal, rough set theory, attribute set, feature selection, attribute reduction.

1 Introduction

As the mathematical tool, Rough set theory is proposed by Pawlak. Rough set theory(RST) is used to deal with inaccurate and incomplete information [1,2]. At present, in many fields the theory has been successfully used, such as learning of machine, knowledge discovery, mining of data, and analysis of decision. Among them, the problem of attribute reduction without reducing the ability of distinguishing different objects in information systems has always

* Corresponding author. Tel.: +(86) 13402975961
[1] *E-mail: 1048768838@qq.com*

10.1016/j.procs.2019.06.029

been one of the core issues among the research of RST [3]. Rough set theory mainly includes related concepts such as equivalence relation, definition of rough set and dependence.

2 Rough Set Equivalence Relation

Assume that the $S=\{U,A\}$ is an system of information, in the set the U represents a non-empty finite set of objects, which is called, a domain, A representing a the same set of attributes, defining a mapping $a:U\rightarrow V_a$ from U to V_a for each attribute $a\in A$, where V_a represents a set of attribute a, called domain of attribute a. Set $P\subseteq A, x,y\in U$, then $IND(P)=\{(x,y)\in U\times U \mid \forall a\in P, a(x)=a(y)\}$ is called the equivalent relationship about P. If $(x,y)\in IND(P)$, it is said that x and y are equivalent about the set P of attribute in the system of information *S*, that is, x and y cannot be distinguished by the attribute in the attribute set P. $IND(P)$ is called P -indistinguishable relationship, if $x\in U$,then the class $[x]_{IND(P)}$ of equivalence on the attribute set P could be defined as [1]:

$$[x]_{IND(P)}=\{y \mid y\in U,\ y\ IND(P)\ x\} \tag{1}$$

It represents a collection of elements that are equivalent to element x under equivalent relationship $IND(P)$.

A division $U/IND(P)$ of U is decided by the equivalent relationship $IND(P)$, or abbreviated as U/P (here P is an equivalence relationship $IND(P)$, hereinafter P referred to $IND(P)$ as ambiguity). Assuming the attribute set P contain m attribute values $\{a_1,a_2,...,a_m\}$, U partition contains n subsets $\{X_1,X_2,...,X_n\}$ determined by the $IND(P)$, and for the set B and C, the operator $\otimes$ is defined as:

$$B\otimes C=\{X\cap Y:\forall X\in B,\forall Y\in C, X\cap Y\neq \phi\} \tag{2}$$

Then $U/IND(P)$ can be calculated by:

$$U/IND(P)=\otimes\{a_i\in P:U/IND(a_i), i=1,2,...,m\}=\{X_1,X_2,...,X_n\} \tag{3}$$

To assume a set $X\subseteq U$, then the lower approximation P about X is defined as: $\underline{P}X=\{x\in U:[x]_P\subseteq X\}$, the upper approximation P about X is defined as: $\overline{P}X=\{x\in U:[x]_P \cap X\neq\phi\}$.

The concept of the upper and lower approximation of set X is shown in figure 1.

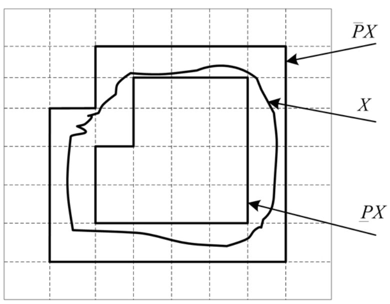

Figure 1 Schematic diagram of the upper and lower approximations of the set X

3 Definition of Rough Sets

Pawlak defines the set $[x]_{IND(P)}$ as the of equivalence classes, which is determined by equivalence relations to form a P1-Rough Set (PRS1). Obviously, the P1-rough set is a subset set, namely:

$$\mathrm{PRS1} = \left\{ [x]_{IND(P)} \mid X \subseteq 2^{U} \right\} \tag{4}$$

By means of the description of the upper and lower approximations, another definition of the rough set equivalent to PRS1 can also be given, called P2-Rough Set (PRS2). It is designed by:

$$\mathrm{PRS2} = \left\{ \left\langle X_1, X_2 \right\rangle \right\} = \left\{ \left\langle \underline{P}X, \overline{P}X \right\rangle \right\} \tag{5}$$

PRS1 and PRS2 are collectively referred to as the Pawlak rough set [4].

Assume P and Q to be a subset of attribute set A, then:

(1)The P-positive region of Q is:

$$POS_P(Q) = \bigcup_{X_j \in U/Q} \underline{P}X_j \tag{6}$$

(2) The P-boundary region of Q is:

$$BND_P(Q) = \bigcup_{X_j \in U/Q} \underline{P}X_j - \bigcup_{X_j \in U/Q} \overline{P}X_j \tag{7}$$

(3) The P-negative region of Q is:

$$NEG_P(Q) = U - \bigcup_{X_j \in U/Q} \overline{P}X_j \tag{8}$$

In the formula, X_j denotes the jth of the n divisions of U determined by $IND(Q)$. $POS_P(Q)$ denotes the sets of objects that according to P which could be correctly classified into U/Q in U; $BND_P(Q)$ denotes a collection of objects that may be the equivalence class of U/Q; $NEG_P(Q)$ denotes those sets of objects that cannot be classified correctly into the class, that is equivalent to U/Q.

4 Reduction of Attribute

In [1], Pawlak analyzes the dependence and independence of attributes and sets, and gives related definitions. Most of the early algorithms based on rough set-based attribute reduction evolved from this definition.

Definition 1 Set $a \in P$, if $POS_P(Q) = POS_{(P-\{a\})}(Q)$, then the attribute a is Q-can be saved in P, otherwise it is called Q-cannot be saved; if each attribute a in P is Q-cannot be saved, it is said P is independent to Q, otherwise it is called dependent.

The form of the dependency is defined as follows:

(1) If and only if $IND(P) \subseteq IND(Q)$, Q relies on P, namely $P \Rightarrow Q$;

(2) If and only if $P \Rightarrow Q$ and $Q \Rightarrow P$, namely $IND(P) = IND(Q)$, P is equivalent to Q, namely $P = Q$;

(3) If and only if $P \Rightarrow Q$ or $Q \Rightarrow P$ are not true, P is independent to Q, namely $P \neq Q$;

(4) If and only if $k = \gamma_P(Q) = |POS_P(Q)|/|U|$, Q dependent to P on the degree of dependence k $(0 \leq k \leq 1)$, denoted as $P \Rightarrow_k Q$:

①If k=1, it is said Q is completely dependent on P, namely $P \Rightarrow_1 Q$,and is also denoted as $P \Rightarrow Q$;

② If $0 < k < 1$, then it is said Q is depend on P partly;

③ If $k = 0$, then it is said Q is completely independent of P.

In the formula, $|U|$ represents the cardinality of the set $U \neq \phi$, that is, the number of objects contained in the set.

Definition 2 If $R \subseteq P$ is independent and R and P is equivalent, it is said R is a reduction of the relational family set P , denoted RED(P). The set of all unsolvable relationships in the family set P is called the core of P , denoted as CORE(P), and CORE(P)=∩RED(P). Attributes in the kernel are important attributes that affect classification, and are attributes that must be included in any reduction [5].

Commonly, the classification is decided by one or several attributes. And it is meaning that the classification is not decided by the a little differences for all attributes in the database of relational. The attribute reduction concept and kernel provides a powerful mathematical tool for extracting important attributes in the system, but considering the non-uniqueness of reduction, that is, the family set P contains multiple reductions, and the minimum reduction of the decision table is NP-hard problem [6], the main cause of NP-hard is the combination explosion of attributes, so far it is failed to find efficient and complete solutions, so many scholars have carried out systematic research and provided many effective new methods, such as data analysis method, attribute reduction algorithm based on information entropy, dynamic reduction algorithm, incremental algorithm, discernible matrix algorithm, reduction method based on genetic algorithm, etc. [7~12].

However, the Pawlak rough set is defined on the basis of the classical equivalence relation and the equivalence class, and is only suitable for dealing with nominal variables. However, the numerical data that is widely used in practical applications cannot be directly processed. In the fields of finance, medical, scientific research and engineering applications, numerical variables are ubiquitous, such as spectral signals in vibration analysis, temperature, current, and voltage signals in transformer state analysis. Researchers are introducing machine learning methods such as rough sets to deal with them. In this kind of data, the discretization algorithm is often used to transform the numerical attribute into a symbolic attribute. The commonly used method is the cloud transformation discretization method [13], the supervised local discretization method [14], based on the attribute importance discretization method [15] and so on. However, the discrete processing cannot well preserve the difference in the real value of the attribute value, so that the integrity of the information content is difficult to guarantee, resulting in different degrees of information loss. The result of the calculation process depends largely on the discretization effect. Therefore, the use of rough set theory to deal with such problems has certain limitations.

5 Conclusion

In this paper, the equivalence relations, definitions, and reduction of attribute based on RST, that is, feature selection, are studied. Feature selection (also known as attribute reduction) refers to the process of selecting some of the most effective features from a set of features to reduce the dimensionality of the feature space, that is, selecting a feature subset that optimizes an evaluation criterion from the original feature set. The basic principle is to select category-related features and eliminate redundant features. However, as analyzed in the paper, the Pawlak rough set is defined on the basis of the classical equivalence relation and equivalence class, and is only suitable for dealing with nominal variables. However, the numerical data that exists widely in practical applications cannot be directly processed. Therefore, the next step is to study the two aspects of rough set and fuzzy set to solve the attribute reduction problem in the fuzzy attribute information table.

6 Acknowledgments

This paper was supported in part by the National Natural Science Foundation of China under the Grant no. 61601499, 61701527 and 61601503.

References

1. Pawlak Z. Rough set[J]. International Journal of Computer and Information Sciences, 1982, 11(5): 341-356.
2. Pawlak Z, Skowron A. Rudiments of rough sets [J]. Information Sciences, 2007, 177(1):3-27.
3. Cornelis C, Jensen R, Hurtado G, et al. Attribute selection with fuzzy decision reducts[J]. Information Sciences, 2010, 180(2): 209-224.
4. Wang Guoyong, Yao Yi Yu, Yu Hong. A Review of Rough Set Theory and Applied Research[J]. Chinese Journal of Computers, 2009, 32(7): 1229-1246.
5. Chang Chunguang, Wang Dingwei, Hu Yuyuan et al. Case attribute reduction technique based on rough set[J]. Control Theory and Applications, 2006, 23(6): 867-872.
6. He Ming, Ma Guoliang, Sun Lifeng. Attribute Reduction Algorithm Based on Ant Colony Algorithm and Rough Set Theory[J]. Journal of Beijing University of Technology, 2010, 36(9): 1292-1296.

7. Wang Guoyong, Yu Hong, Yang Dachun. Decision Table Reduction Based on Conditional Information Entropy[J]. Chinese Journal of Computers, 2002, 25(7): 759-766.
8. Wang Jue, Wang Ren, Miao Duanqian et al. "Data Concentration" Based on Rough Set Theory[J]. Chinese Journal of Computers, 1998, 21(5): 393-400.
9. Yu H, Wang G Y, Yang D C, et al. Knowledge reduction algorithms based on rough set and conditional information entropy[C]. proceedings of SPIE: data mining and knowledge discovery: theory, tool, and technology IV, Orlando, USA, 2002, 4730: 422-431.
10. Wei L, Li H R, Zhang W X. Knowledge reduction based on the equivalence relations defined on attribute set and its power set[J]. Information Sciences, 2007, (177): 3178-3185.
11. Hu Q H, Zhao H, Xie Z X, et al. Consistency based attribute reduction[C]. Lecture Notes in Computer Science, 2007, 4426: 96-107.
12. Tao Zhi, Xu Baodong, Wang Dingwei et al. Rough Set Knowledge Reduction Method Based on Genetic Algorithm[J]. Systems Engineering, 2003, 21(4): 116-122.
13. Song Jing, Yan Jianhuai. Decision Tree Construction Based on Rough Sets Based on Cloud Transform and Characteristic Relationship[J]. Journal of SouthWest JiaoTong University, 2010, 45(2): 312-316.
14. Guan Xin, Yi Xiao, He You. A New Rough Set Attribute Reduction Method and Its Application[J]. Control and Decision, 2009, 24(3): 464-467.
15. Zhang Jianjun, Zhang Jingbo. A new decision table discretization method based on rough set theory[J]. Journal of Xidian University, 2004, 31(3): 369-472.