ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE
International Journal of Advanced Research in Computer and Communication Engineering

ICACTRP 2017
International Conference on Advances in Computational Techniques and Research Practices
Noida Institute of Engineering & Technology, Greater Noida

Vol. 6, Special Issue 2, February 2017

Concept of Rough Set Theory and its Applications in Decision Making Processes

Pooja Chaturvedi[1], A K Daniel[2], Kumari Khusboo[3]

Research Scholar, Dept of Computer Science & Engg, M. M. M. University of Technology, Gorakhpur, India[1]

Associate Professor, Dept of Computer Science &Engg, M. M. M. University of Technology, Gorakhpur, India[2]

Assistant Professor, Dept of Computer Science & Engg, Noida Institute of Engineering & Technology, Noida, India[3]

Abstract: The application which deals with the uncertain or vague information, the rough set theory is a useful and major method in the decision making. Rough set theory provides a feasible approach for decision rule extraction. This paper discusses the basic concepts of rough set theory as data representation, lower approximation, and upper approximation, attribute selection and rule formation. We have also discussed some application areas of rough set theory in the field of data mining and sensor networks.

Keywords: rough set theory, lower approximation, upper approximation, data mining, rule extraction.

I. INTRODUCTION

The rough set theory is a mathematical approach proposed by Pawlak in 1982 to address the decision-making problem where the information is not represented in exact values. The rough set theory finds its application in various problems such as image processing, data mining, medical informatics, pattern recognition, extracting interesting patterns in databases, in conjunction with other uncertain reasoning methods such as fuzzy logic, neural network and genetic algorithm. The aim of this paper is to provide the basic concepts related to the rough set theory and description of its applications in various fields.
The organization of the paper is as follows: basic concepts of rough set theory in section 2, applications of rough set theory in section 3, and section 4 concludes the paper.

II. ROUGH SET THEORY

Rough set theory is defined as an extension of set theory and is based on approximations. It is useful in the applications where the objects are represented using the attribute value pairs. The assumptions in rough set theory are as follows:

- Objects are represented as the pair of attribute and its values.
- Objects with the same attribute value pairs are indiscernible.

Basic concepts
The basic concepts related to the rough set theory are as follows:

a. Approximation space
An approximation space is defined as a pair U= (A, X) where A is a nonempty finite set of elements called as universe and X is an equivalence relation defined in A.

b. Information system
An information system is a pair S = (A, B), where
A is a nonempty finite set of elements called as universe and B is a nonempty finite set of properties, i.e., a: A→V_b for $b \in B$, where V_b is the domain of B.

c. Decision/Data table
A decision table is represented as $D = (A, B = F \cup \{d\}$, where attributes in F are called condition attributes and d is a designated attribute called the decision attribute.

d. Approximation of sets
Let U = (A, P) be an approximation space and P be asubset of A.

- The lower approximation of P by X in S is defined as:

$$\underline{P}\,X = \{e \in U \mid [e] \subseteq P\}$$

- The upper approximation of X by R in S is defined as:

$$\overline{P}(X) = \{e \in U \mid [e] \cap P \neq \varphi\}$$

where[e] represents the equivalence class containing e and [e] is called elementary set.

- A subset P of U is said to be P-definable in Uif $\underline{P}\,X = \overline{P}\,X$
- The boundary set $B_P(S)$ is defined as:

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

ICACTRP 2017

International Conference on Advances in Computational Techniques and Research Practices

Noida Institute of Engineering & Technology, Greater Noida

Vol. 6, Special Issue 2, February 2017

$$\overline{P}\,X = \underline{P}\,X\ .$$

- A set P is rough in U if its boundary set is nonempty.

e. Accuracy of Approximation

The accuracy of the approximation is defined as the ratio of lower and upper approximation of the rough set as:

$$a_B(X) = \frac{\left|\underline{P}(X)\right|}{\left|\overline{P}(X)\right|}$$

Where U= (A, X), $B \subseteq X$ and X $\subseteq$ A

$|X|$ represents the cardinality of X.

If $\alpha_B(X) = 1$, then X is crisp with respect to B.

If $\alpha_B(X) < 1$, then X is rough with respect to B.

f. Dependency of attributes

Considering M and N are two subsets of P then N depends on M to degree $k\ (0 \le k \le 1)$ is represented as $M ->_k N$ if

$$k = \gamma(M, N) = \sum_{X \in U/N} \frac{\left|\underline{M}(X)\right|}{|U|}$$

$$k = \gamma(M, N) = \frac{\left|POS_M(N)\right|}{|U|}$$

Where $POS_M(N) = \bigcup_{X \in U/N} M(X)$ is called the M-positive region of N.

If k= 1, then N depends totally on M.

If k<1, N depends partially (in degree k) on M.

g. Dispensable and Indispensable Attributes

Let $S = (U, A = M \cup N)$ be a decision table. Let m be an attribute in M then attribute m is dispensable in M if $POS_M(N) = POS_{(M - |m|)}(N)$ otherwise m is indispensable.

A table is said to be independent if all the attributes in the decision table are indispensable.

h. Reduct and Core

Let $S = (U, A = M \cup N)$ be a decision table. A subset Y of M is Core of M if $POS_X(N) = POS_M(N)$ and the subset $S' = (U, X \cup N)$ is independent if all the attributes in X are indispensable in S'.

Core of M is the set of attributes shared by all reducts of M.

$CORE(M) = \cap RED(M)$ whereRED(M) is the set of all reducts of M.

i. Rough membership function

Let S= (U, A), $B \subseteq A$ and X $\subseteq$ U then the rough membership function μ_X^B for X is a mapping from U to [0, 1] as:

$$\mu_X^B : U -> [0,1]$$

For all e in U, the degree of X belongs to X in view of the set of attributes B is defined as:

$$\mu_X^B(e) = \left|B(e) \cap X\right| / \left|B(e)\right|$$

Where B(e) denotes the block containing e.

j. Properties of rough membership function

The rough set membership function has several properties as follows:

$P1: \mu_X^B(e) = 1$ if e in $B_*(X)$

$P2: \mu_X^B(e) = 0$ if e in $U - B_*(X)$

$P3: 0 < \mu_X^B(e) < 1$ if e in $BN_B(X)$

$P4: \mu_{U-X}^B(e) = 1 - \mu_X^B(e)$ for any e in U

$P5: \mu_{X \cup Y}^B(e) \ge \max(\mu_X^B(e), \mu_Y^B(e))$ for any e in U

$P6: \mu_{X \cap Y}^B(e) \le \min(\mu_X^B(e), \mu_Y^B(e))$ for any e in U

Where $B_*(X)$ is the lower approximation of X in B and $B^*(X)$ is the upper approximation of X in B.

k. Rule extraction

The rough set based algorithms can be classified into three types on the basis of decision rules obtained as:

1. Algorithm which generates the minimum set of rules.
2. Algorithms which generates the exhaustive set of rules.
3. Algorithms which generates the satisfactory set of rules.

The decision rules are mainly generated for classifying the objects of a system into different groups and for performing the data extraction and knowledge discovery in the field of data mining. The algorithms used for learning the rules are shown in Table1:

TABLE 1 DECISION RULE GENERATION MECHANISM IN ROUGH SET THEORY

Algorithm	Approach
LEM2	-In this, the rules are learned from lower and upper approximations

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

ICACTRP 2017

International Conference on Advances in Computational Techniques and Research Practices

Noida Institute of Engineering & Technology, Greater Noida

Vol. 6, Special Issue 2, February 2017

	– It works in increment/non-increment modes.
RLEM2	– In this learning mechanism, the rules are learned from the lower and upper approximations and using multi set decision tables. -in this, the extension of basic SQL operators are used.
BLEM2	– In this approach, the rules are learned from the lower and upper approximation using the partition of a boundary set and the rules are dependent on Bayes' theorem and rough sets.

l. Tools based on rough set

The various tools used for the applications of rough set theory are shown in Table 2.

TABLE 2. TOOLS USED IN ROUGH SET THEORY

Tool	Usage
ROSE/ROSE2	It is used for Rough set data exploration in the knowledge discovery process in data mining.
4eMka	It is Dominance-based Rough Set Approach used to perform the Multi-criteria based classification.
JAMM	It is relatively a new decision support tool for analysis and is used for multi criteria based classification problems

III. APPLICATIONS OF ROUGH SET THEORY

The concept of the rough set theory can be applied in various fields such as:

Data Mining

Rough set theory can be applied to the field of data mining in different phases of knowledge discovery such as selection of core attribute, attribute extraction, reducing the datasets, generate the decision rule and pattern recognition.

Medical

A most common and critical challenge in the medical field is the abdominal pain in the children, which can be due to various reasons and it's a challenging task to correctly diagnosis the reason. Rough set theory based approaches can help the doctors in diagnosis by discharge observation/ further investigation and consultation.

Acoustical analysis

A rough set theory based prototype system is used to determine the generalised rules that describe the relationship between acoustical parameters of concert halls and sound processing approaches.

Power system security analysis

In this, the rough set decision rules help the knowledge engineers in the extraction of the facts of a set of examples for power system operation problems.

Spatial and meteorological pattern classification

The hierarchical rough set based learning method can ease the process of sunspot classification.

In the area of meteorology, the volumetric radar is used to detect the storm events which are responsible for severe weather conditions.

Intelligent control system

The rough set theory can be used in the intelligent control systems in combination with the fuzzy theory.

Measure the quality of a single subset

A heuristic based on the hybrid approach combining the ant colony system based algorithm and rough set theory is applied to measure the quality of a subset.

Wireless sensor networks

In the field of wireless sensor network the rough set theory can be used in various applications:

1. Fault detection in nodes using the rough set and neural network
2. Clustering by identifying the core attributes for the selection of cluster head.

IV. CONCLUSION

Rough set theory is a relatively new mathematical method to aid in the decision making processes involving uncertainty and imprecision. In this paper, we have discussed the basic concepts related to the rough set in terms of lower approximation, upper approximation and decision rule extraction. We have also discussed the application areas of the rough set in various fields ranging from knowledge discovery to communication networks. To develop an approach for improving the network efficiency by using the rough set theory concepts in the area of wireless networks is our future work.

REFERENCES

[1] http://idss.cs.put.poznan.pl/site/software.html
[2] C. C. Chan, "Fundamentals of Rough set theory and its applications.
[3] Z. Pawlak, "Rough Sets", International Journal of Computer and Information Sciences, Vol.11, 341-356 (1982).
[4] Z. Pawlak, Rough Sets - Theoretical Aspect of Reasoning about Data, Kluwer Academic Pubilishers (1991).
[5] S. Rissino, G. L. Toress, Rough set theory-Fundamental theory, concepts, principals, data extractions and applications,Data Mining and Knowledge Discovery in Real Life Applications, pp. 438, 2009.
[6] F. Coaquira, E. Acuna, "Applications of Rough Set Theory in Data Preprocessing for Knowledge Discovery, Proceedings of World Congress on Engineering and Computer Science, 2007.
[7] F. Liang, L. Zhang, P. Sun, "Study on the Rough-set-based Clustering Algorithm for Sensor Networks, Bulletin of Electrical Engineering and Informatics, Vol.3, No.2, June, 2014, pp 77-90.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

ICACTRP 2017

International Conference on Advances in Computational Techniques and Research Practices

Noida Institute of Engineering & Technology, Greater Noida

Vol. 6, Special Issue 2, February 2017

[8] H. Xu, R. Zhang, C. Lin, Y. Ma, "Novel Approach of Fault Diagnosis in Wireless Sensor Networks Node Based On Rough Set and Neural Network Model", International Journal of Future Generation Communication and Networking Vol. 9, No. 4 (2016), pp. 1-16.

[9] Pawlak, Z., "Rough sets: basic notion," Int. J. of Computer and Information Science 11, 344-56, (1982).

[10] Grzymala-Busse, J.W., "Learning from examples based on rough multisets," Proc. of the 2nd Int. Symposium on Methodologies for Intelligent Systems, Charlotte, North Carolina, October 14-17, 325-332, (1987).

[11] Chan, C.-C., "Distributed incremental data mining from very large databases: a rough multiset approach," Proc. The 5th World Multi-Conference on Systemics, Cybernetics and Informatics, SCI 2001, Orlando, Florida, July 22-25,(2001), 517-522.

[12] Chan, C.-C., "Learning rules from very large databases using rough multisets," LNCS Transactions on Rough Sets, (J.

[13] Peters, A. Skowron, J.W. Grzymala-Busse, B. Kostek, R.W. Swiniarski, M. Szczuka Eds.), Vol. 1, pp. 59 – 77,Springer –Verlag Berlin Heidelberg, 2004.

[14] Chan, C.-C. and S. Santhosh, "BLEM2: Learning Bayes' rules from examples using rough sets," Proc. NAFIPS 2003, 22nd Int. Conf. of the North American Fuzzy Information Processing Society, July 24 – 26, 2003, Chicago, Illinois, pp. 187-190.

[15] Greco, S. Matarazzo, B., Slowinski, R., "A new rough set approach to evaluation of bankruptcy risk," in Zopounidis,C. (Ed.), Operational Tools in the Management of Financial Risks. Kluwer Academic Publishers, Dordrecht, pp. 121–136 (1998).

[16] Greco, S., Matarazzo, B., Slowinski, R., "The use of rough sets and fuzzy sets in MCDM," in Gal, T., Stewart, T.,Hanne, T. (Eds.), Advances in Multiple Criteria Decisions Making. Kluwer Academic Publishers, Dordrecht, pp. 14.1–14.59 (Chapter 14) (1999).

[17] Greco, S., Matarazzo, B., Slowinski, R., "Rough sets theory for multicriteria decision analysis," European Journal of Operational Research 129(1), 1 – 47 (2001).

[18] Greco, S., Matarazzo, B., Slowinski, R., "Rough sets methodology for sorting problems in presence of multiple attributes and criteria," European Journal of Operational Research 138(2), 247 – 259 (2002).

[19] Greco, S., Matarazzo, B., Slowinski, R., Stefanowski, J., "An algorithm for induction of decision rules consistent with the dominance principle," in Ziarko, W. and Yao, Y. (Eds.), RSCTC 2000, LNAI 2005, Springer-Verlag Berlin Heidelberg, pp. 304 – 313 (2001).

[20] Greco, S., Matarazzo, B., Slowinski, R., Stefanowski, J., Zurawski, M., "Incremental versus non-incremental rule induction for multicriteria classification," Peters, J.F. et al. (Eds.): Transactions on Rough Sets II, LNCS 3135, Springer-Verlag Berlin Heidelberg, pp. 33 – 53 (2004).

[21] Z. Xi, "The study on the Routing Algorithm for Wireless Sensor Networks Based on Energy Balance", IJACT, vol. 5, no. 2, (2013), pp. 8-16.

[22] Arpan Kumar Dubey and Adwitiya Sinha, Congestion Control for Self Similar Traffic in Wireless Sensor Network",Contemporary Computing (IC3), 2015 Eighth International Conference on 20-22 Aug. , 2015, IEEE.