International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 16 (2017) pp. 5637-5642

Fuzzy-Rough set Approach to Attribute Reduction

Pallab kumar Dey [1] and Sripati Mukhopadhyay[2]

[1]*Department of Computer Science, Kalna College, Kalna-713409, W.B., India.*

Orcid: 0000-0003-2229-1526

[2]*Department of Computer Science, The University of Burdwan, Bardhaman-713104, W.B., India.*

Abstract

Attribute Reduction has a significant role in different branches of artificial intelligence like machine learning, pattern recognition, data mining from databases etc. This paper deals with reduction of unimportant attribute(s) for classification and decision making, using Fuzzy-Rough set. A survey of Fuzzy-Rough set based methods for attribute reduction is presented here.

Keywords: Attribute Reduction; Rough set; Fuzzy set; fuzzy-rough set.

INTRODUCTION

In real world application, databases flourish with objects and extraneous attributes. Irrelevant attributes deteriorate performance of machine learning and pattern recognition. Irrelevant attributes may hamper data mining task by misguiding information and increasing time. Noisy and misleading attributes may be removed to improve machine learning classification. Fewer numbers of attributes can assist for data realizing by visualization also. It is also expensive to store unwanted data. Retaining same classification ability, redundant attribute may be removed.

Attribute reduction is important application of Rough Set proposed by pawlak [1]. Rough set deal with equivalence relation, so it is applicable to limited discrete data only. We have to apply discretization for application on real life data which may induce information loss. To solve these problem researchers proposed many wing of Rough sets. This paper will focus on merging concept of rough and fuzzy set i.e, fuzzy-Rough Sets [2-9]. It can deal with fuzziness of objects. The organisation of the paper is as follows: Section II is devoted to Rough-Fuzzy concepts which are essential for this problem analysis. Section III is devoted to survey of existing works in the field with analysis of works.

OVERVIEW OF FUZZY-ROUGH SETS

To analyze inexact, uncertain and vague knowledge pawlak's rough set theory [1] is the most prominent tool. The rough sets theory provides a technique to deal with vague and imprecise data. Objects with the same information are indiscernible considering available information. Information system may be represented as four tuple (U,A,V,f) where U is a non empty finite set of objects, A is a non empty finite set of attributes, $\forall a\in A$: V_a is the domain of attribute a, $V=\cup V_a$ is the domain of A, f is a mapping $f: U\times A\rightarrow V$, $f(x, a)\in V_a$ is the value that x holds on a. Any subset B of A determines a binary relation $I(B)$ called indiscernibility relation defined as:

$$I(B)=\{(x_i, x_j)\in U\times U \mid \forall a\in B,\ a(x_i)=a(x_j)\} \qquad (1)$$

If $(x_i, x_j)\in I(B)$, then x_i and x_j are indiscernible by attributes from B. The equivalence classes of indiscernibility relation B are denoted by $[x_i]_B$ and computed as:

$$[x_i]_B=\{x_j\in U \mid (x_i, x_j)\in I(B)\} \qquad (2)$$

Let $X\subseteq U$, B-lower and B-upper approximation of a set can be defined as:

$$\underline{B}X=\{x_i \mid [x_i]_B\subseteq X\} \qquad (3)$$

$$\overline{B}X=\{x_i \mid [x_i]_B\cap X\neq \emptyset\} \qquad (4)$$

Order pair $(\underline{B}X, \overline{B}X)$ is called rough set of X. If P and Q be equivalence relations over U, then positive region can be defined as:

$$POS_P(Q)=\cup_{X\in U/Q}\underline{P}X \qquad (5)$$

Positive region contain all objects that can be classified to classes of U/Q using the information of P. If $P, Q\subseteq A$, , then Q depends on P in a degree k $(0\leq k\leq 1)$, determined by

$$k=\gamma_P(Q)=\frac{|POS_P(Q)|}{|U|} \qquad (6)$$

$k=1$ signify, total dependency of Q on P. $0<k<1$ signify, partial dependency of Q on P. $k=0$ signify, no dependency of Q on P. If an attribute is removed from conditional attributes then by using change in dependency degree, significance of the attribute can be obtained. If an attribute $a \in P$, significance of the attribute a on Q is determined by

$$\sigma_P(Q,a)= \gamma_P(Q) - \gamma_{P-a}(Q) \qquad (7)$$

Reduction of attributes is a important application of Rough Set. If removing of some attributes does not affect classification of the dataset then these attributes are redundant. If attribute $a \in B$ and $I(B)=I(B-\{a\})$ then *'a'* is dispensable, otherwise *'a'* is indispensable. Minimal subset of attributes that maintain same partition as whole set of attributes is called Reduct. Set of all indispensable attributes of the Universe is called Core; also it may be defined as intersection of all reduct. In decision system, attributes set *'A'* consist of conditional attribute set *'C'* and decision attribute set *'D'* with no intersection.

If $C' \subseteq C$ is a D-reduct of C then C' is minimal subset of C such that $\gamma(C,D)=\gamma(C',D)$.

Minimum subset with $\gamma_{c'}(D)=1$ is the minimal reduct. By computing dependencies of all possible subset of C minimal reduct may be determined but not applicable for large data set.

Rough sets can be used to reduce the number of attributes as a form of reduct, when data set have discrete attribute values. Rough sets cannot deal with real valued attribute data sets. Due to noise, same value may differ slightly but it will be treated as different value in rough set. This problem may be solved by discretization but information will be lost.

The concept of fuzzy rough set was first proposed by Dubois and Prade [2]. Fuzzy rough set theory has been analyzed further [3-9] in terms of property and axioms. Fuzzy rough set is emerging tool frequently used in pattern recognition and machine learning. Significant application of fuzzy rough set is attribute reduction for crisp and real valued attribute data sets. Indiscernibility of Rough set and vagueness of fuzzy set has been merged in Fuzzy-Rough sets to deal with uncertainty more precisely.

Crisp equivalence classes are key concept of rough set and fuzzy equivalence classes are key concept of fuzzy-rough set [2]. A fuzzy binary relation of a fuzzy set R on a non empty set U, is called fuzzy equivalence relation if it satisfies

1) Reflexivity: $\mu_R(x,x)=1 \ \forall x \in U$

2) Symmetry: $\mu_R(x,y)=\mu_R(y,x) \ \forall x,y \in U$

3) Transitivity:

$\mu_R(x,z) \geq sup_y \min\{\mu_R(x,y), \mu_R(y,z)\} \ \forall x,y,z \in U$

From fuzzy equivalence relation, fuzzy equivalence class $\mu_{[X]_R}$ for objects close to x may be defined as:

$$\mu_{[X]_R}(y) = \mu_R(\mathrm{x}, \mathrm{y}),\ y \in U \qquad (8)$$

The fuzzy P-lower and P upper approximation are defined as:

$$\mu_{\underline{P}X}(F_i) = \inf_{\mathrm{X}} \max\{1-\mu_{F_i}(\mathrm{x}), \mu_X(\mathrm{x})\}\ \forall \mathrm{i} \qquad (9)$$

$$\mu_{\overline{P}X}(F_i) = \sup_{\mathrm{X}} \min\{\mu_{F_i}(\mathrm{x}), \mu_X(\mathrm{x})\}\ \forall \mathrm{i} \qquad (10)$$

where F_i denotes a fuzzy equivalence class belonging to U/P. crisp upper and lower approximations are same as above definitions, as the memberships of individual objects to the approximations are not explicitly available. So fuzzy lower and upper approximations are herein redefined [10] as:

$$\mu_{\underline{P}X}(X) = \sup_{\mathrm{F} \in \mathrm{U}/P} \min\{\mu_F(\mathrm{x}), \inf_{\mathrm{y} \in \mathrm{U}} \max\{1-\mu_F(\mathrm{y}), \mu_X(\mathrm{y})\}\} \qquad (11)$$

$$\mu_{\overline{P}X}(X) = \sup_{\mathrm{F} \in \mathrm{U}/P} \min\{\mu_F(\mathrm{x}), \sup_{\mathrm{y} \in \mathrm{U}} \min\{\mu_F(\mathrm{y}), \mu_X(\mathrm{y})\}\} \qquad (12)$$

The tuple $\langle \underline{P}X, \overline{P}X \rangle$ is called a fuzzy-rough set. If all equivalence classes are crisp then these definition degenerate to traditional rough set.

ATTRIBUTE REDUCTION MODELLING

Jensen and Shen first proposed fuzzy-rough dependency function and fuzzy-rough quick reduct algorithm to compute reduct [10-12]. Using fuzzy positive region of a fuzzy set (where A= C∪D), membership of an object X∈ U may be defined as:

$$\mu_{POS_C(D)}(X) = \sup_{\mathrm{X} \in \mathrm{U}/D} \{\mu_{\underline{C}X}(\mathrm{x})\} \qquad (13)$$

and dependency function may be defined as:

$$\gamma'_p(Q) = \frac{\left|\mu_{POS_C(D)}(X)\right|}{|U|} = \frac{\sum_{\mathrm{X} \in \mathrm{U}} \mu_{POS_C(D)}(X)}{|U|} \qquad (14)$$

In crisp case, U/P contains set of objects grouped together that are indiscernible. In traditional rough set, if $R(\subseteq A)$ is a reduct of attribute set A then $\gamma(A)$ and $\gamma(R)$ are identical and equal to 1 if data set is consistent. But in fuzzy-rough set it may not be as object may belongs to many fuzzy equivalence classes, so reduce total dependency. With this issue fuzzy-rough quick reduct algorithm [10] has been proposed as given in Algo.1.

Algo1: fuzzy-rough quick reduct algorithm

Input: Incomplete information System S,
S={ Aj,Vij : j=1,2,...,k; i=1,2,...,n where Vij may be missing}
k=number of Attributes, n=number of Objects

Output: Complete Information System
S'={ Aj ,Vij :j=1,2,...,k; i=1,2,...,n where Vij not null}

Step 1. $R \leftarrow \{\}$; $\gamma'_{best}=0$; $\gamma'_{prev}=0$;

Step 2. do

Step 3. $T \leftarrow R$

Step 4. $\gamma'_{prev} = \gamma'_{best}$;

Step 5. $\forall x \in (C-R)$

Step 6. If $\gamma'_{R\cup X}(D) > \gamma'_T(D)$

Step 7. $T \leftarrow R \cup \{x\}$

Step 8. $\gamma'_{best} = \gamma'_T(D)$

Step 9. $R \leftarrow T$

Step 10. Until $\gamma'_{best} == \gamma'_{prev}$;

Step 11. Return R

It starts off with an empty set and add those attributes, one at a time, that result in the maximum increase in fuzzy-rough dependency function γ'. Algorithm was tested and shown better performance with some real life data sets like web categorization. Maximum afford on attribute reduction using fuzzy rough set try to improve efficiency of dependency function or this algorithm.

Fuzzy-rough quick reduct algorithm is not always convergent [13, 14]. As measure of fuzzy-rough degree of dependency is non monotonic, so search may terminates after reaching local optimum though global optimum may lie elsewhere in the search space. With increasing number of input variables, computational complexity of the algorithm increases exponentially.

This algorithm may return a reduct with superfluous attribute due to the non optimality of the search heuristic [15]. Computational efficiency of this algorithm has been improved by concept of computational domain [14]. For large attribute sets complexity of calculating the Cartesian product of fuzzy equivalence classes becomes high. For some data set fuzzy lower approximation might not be a subset of the fuzzy upper approximation. It is meaningless as it indicates; there is more certainty in the upper than the lower.

Same Fuzzy-rough quick reduct algorithm using alternative definitions of fuzzy-rough lower and upper approximations [16], has been proposed in [15]. Here monotonicity of fuzzy rough dependency function has been shown. Complexity of the algorithm remains same but explosive growth of the number of fuzzy equivalence classes is avoided. Only one fuzzy similarity relation is used to compute the fuzzy lower approximation for one subset.

Most of the fuzzy-rough attribute reduction method use lower approximation ignoring upper approximation. But upper approximation contains information regarding the degree of uncertainty of objects. Similar quick reduct algorithm using Fuzzy-rough based uncertainty degree has been proposed [15]. Minimizing uncertainty degree fuzzy-rough reduct may be computed. Computation of fuzzy equivalence classes may be avoided but due to computation of fuzzy lower and upper approximations, fuzzy boundary region is more costly than that of the fuzzy lower approximation only. Axiomatic approach and constructive approaches are different way to develop knowledge representation of fuzzy rough sets. Fuzzy approximation operators and axiomatic study of fuzzy rough set are building block of axiomatic approach. Fuzzy binary relation, fuzzy T-similarity relation and other fuzzy relations substitute the equivalence relation in constructive approach. Attribute reduction based on Constructive approaches. Some theorems to describe the impacts of fuzzy approximation operator on attribute reduction are presented [17] and review of constructive approach has been presented. A variable precision fuzzy-rough set was proposed [18, 19]. Here fuzzy membership of a sample to lower and upper approximations with fuzzy inclusion was computed. A new model was proposed [18, 19] to handle noise of misclassification using variable precision fuzzy-rough sets. Shannon's Information entropy has been extended [20, 21] to measure the dependency between conditional attributes and decisions in fuzzy rough sets. It proposed unsupervised and supervised reduction algorithms for hybrid data based on this measure. Proposed measure has been utilized to calculate the uncertainty in the fuzzy-rough approximation spaces and reduce hybrid data.

Application of Granular computing in many research area attract researcher. Fuzzy information granulation and granular computing has a great significance in fuzzy set and rough set based approach [22, 23]. Fuzzy-rough set based approach appears to be an efficient tool for granular computing. Hybrid data reduction model has been proposed based on Fuzzy-rough set model for granular computing [24].

Noise in data cannot be tolerated in basic fuzzy-rough model. Variable precision fuzzy-rough model may take rational decision in the conditions of imprecision. So for efficient implementation of granular computing variable precision fuzzy-rough model is more effective. Here positive region of the decision by fuzzy inclusion and variable precision fuzzy inclusion has been computed and four attribute significance measures are defined. Forward attribute reduction based on variable precision fuzzy-rough model (FAR-VPFRS) has been proposed.

Most of the fuzzy-rough heuristic feature selection algorithms are computationally time-consuming. So an accelerator, called forward approximation, which combines sample reduction and dimensionality reduction together, has been proposed [25]. Heuristic process of fuzzy-rough feature selection may be speed up using this accelerator. An improved algorithm is proposed using the accelerator. It has been shown by experiments, that modified algorithms are much faster than their original counterparts.

Classical fuzzy-rough model cannot describe sample classification well because it merely maintain dependency function maximum, so a new fuzzy rough set model was introduced in [26]. Fuzzy decision of a sample by using the concept of fuzzy neighbourhood has been defined. To characterize fuzzy information granules, a parameterized fuzzy relation is introduced and a new fuzzy dependency function is proposed. Then significance measure of a candidate attribute has been define and a greedy forward algorithm (Heuristic algorithm based on fitting fuzzy rough sets) for attribute reduction has been proposed. It is shown that algorithm is more feasible and effective for those data sets for which different categories exhibit a large degree of overlap.

A fuzzy-rough based dimensionality reduction method is proposed [27] that simultaneously selects attributes and extracts features using the concept of feature significance. Insignificance features are discarded and significant features are used to select or extract a feature in next iteration. New attribute set may contain some original features and/or some extracted new features. Effectiveness of the proposed algorithm with comparison of other methods is shown with real life data.

Retrieval of proper reduct can not be possible by these algorithms due to their stopping criteria but an over reduct or sub reduct is also acceptable to save running time. But it may not be acceptable in many application areas like medical field as we may lose some information or unnecessary attribute may diverge. Discernibility matrix based approach is another important way for attribute reduction by which, proper reduct may be retrieved. Discernibility matrix of Pawlak's rough sets has been extended and fuzzy discernibility matrix has been proposed for attribute reduction using fuzzy rough sets [28, 29, 15] .

Every element of fuzzy discernibility matrix is a fuzzy set with each attribute belongs to a certain degree. If attribute *'a'* belongs to the fuzzy clause $\mathbf{C_{ij}}$ then measure of fuzzy discernibility measure is determined by [15]:

$$\mu_{C_{ij}}(a) = N(\mu_{R_a}(i,j)) \tag{15}$$

Where N denotes fuzzy negation and $\mu_{R_a}(i,j)$ is the fuzzy similarity of object *i* and *j*. Every element of fuzzy indiscernible matrix is a set of attributes and their corresponding membership

$$C_{ij} = \{a_X \mid a \in C, X = \mu_{C_{ij}}(a)\} \tag{16}$$

and fuzzy discernibility function may be defined as

$$f_D(a_1^*,...,a_m^*) = \wedge\{\vee C_{ij}^* \mid 1 \le j < i \le |U|\} \tag{17}$$

where $C_{ij}^* = \{a_x^* \mid a_X \in C_{ij}\}$ and the function returns values in [0,1]. To find reducts from the fuzzy discernibility function we have to fetch the minimal assignment of the value 1 to the variables such that formula is maximally satisfied.

For decision system only different decision values are included in crisp discernibility matrix. For fuzzy rough set for decision feature q it may be defined as

$$f_D(a_1^*,...,a_m^*) = \wedge\{\vee C_{ij}^* \leftarrow q_N(\mu_{R_q}(i,j)) \mid 1 \le j < i \le |U|\} \tag{18}$$

It has been proved [15] that considering individual satisfaction of each clause for a given set of features may produce better result. Degree of satisfaction of a clause $\mathbf{C_{ij}}$ for a subset of feature P is defined as

$$SAT_P\{C_{ij}\} = \cup\{\mu_{C_{ij}}(a)\} \tag{19}$$

Taking decision feature q it may be defined as

$$SAT_{P,q}\{C_{ij}\} = SAT_P\{C_{ij}\} \leftarrow \mu_{C_{ij}}(q) \tag{20}$$

Total satisfiability of all clauses for a subset P may be defined as

$$SAT(P) = \frac{\sum_{i,j \in u, i \ne j} SAT_{p,q}\{C_{ij}\}}{\sum_{i,j \in u, i \ne j} SAT_{C,q}\{C_{ij}\}} \tag{21}$$

where *C* is the set of all conditional attributes. If this value reaches 1 for a subset *P,* then subset is a fuzzy rough reduct. The algorithm start with considering reduct candidate *P* as empty set and each conditional feature of discernibility function is measured according to heuristic used.

Fuzzy discernibility matrix was constructed and a algorithm was developed to find proper reduct [28]. Due to compute and search every element in discernibility matrix, its computational complexity is high. Only minimal elements in the discernibility matrix are necessary to find reducts. Based on this idea algorithms were developed to find reducts using minimal elements in discernible matrix [29] . First algorithm has been used to find all minimal elements in discernibility matrix. From set of minimal elements we get discernibility function by which all reduct may be found. But in most cases only one reduct is sufficient, for that another algorithm has been proposed for finding one reduct. By experimental comparison it has been shown that proposed algorithm can improve efficiency by reducing computational complexity to find reduct.

Discernibility matrix and Discernibility function are main concept of this method. To find proper reduct we have to consider every element in discernibility matrix with high computational complexity.

Attribute reduction methods with fuzzy decision attributes have been proposed [30]. Inconsistent fuzzy decision system has been defined. Discernibility matrix is used to compute all reduct. Here two algorithms presented for reduct computation, one is to find all reduct which is time consuming and other is close approximation to minimal reduct which is fast. There is no previous work for retrieving all reduction for continuous decisions in fuzzy rough sets. To prove its effectiveness, several experiments are presented with comparison of other methods.

Multi-modality attributes may be addressed by Multi-kernel learning by using different kernels to extract information coming from different attributes. But fuzziness in fuzzy classification cannot be covered. To handle fuzzy and uncertain attribute reduction the model of multi-kernel fuzzy rough sets was developed by integrating multi-kernel learning with fuzzy rough sets [31]. An attribute reduction algorithm has been proposed for large scale multi-modality fuzzy classification based on this model. Effectiveness has been proved by experimental results.

For dynamic data, most of the algorithms for attribute reduction with fuzzy rough set have to re-compute a reduct with sample arriving where one sample or multiple samples arrive successively. To find a reduct from such datasets, fuzzy rough set based methods have been proposed [32]. At the arrival of one sample or multiple samples, the relative discernibility relation is updated for each attribute. Depending upon updated relation, algorithm for incremental attribute reduction (IARS) that updates reducts when a new sample arrives and another algorithm for incremental attribute reduction (IARM) that updates reducts with multiple samples have been proposed. Experimental result shows its efficiency with short time.

CONCLUSION

To select important attributes from a given real-valued dataset, attribute reduction using fuzzy rough set is most effective technique. There are different ways for attribute reduction using fuzzy rough set. But most of them can be categorized into two broad ways, dependency function based and discernible matrix based approach. By using dependency function based approach, we can get approximate reduct (may not be exact) in short time. By using discernible matrix based approach we may fetch exact reduct but time complexity is high. Also same method is not applicable for all type of data. So according to requirement and nature of data, we have to choose a method for attribute reduction.

REFERENCES

[1] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Vol.9, Kluwer Academic Publishers, Dordrecht, 1991.

[2] D. Dubois, H. Prade, “Rough fuzzy sets and fuzzy rough sets”, International Journal of General Systems, Vol.17, No.1, pp. 191–209, 1990.

[3] N.N. Morsi, M. M. Yakout, “Axiomatics for fuzzy-rough sets”, *Fuzzy Sets and Systems,* Vol.100, pp. 327–342,1998.

[4] W. Wu, J. Mi, W. Zhang, “Generalized fuzzy-rough sets”, *Information Sciences*, Vol.151, pp. 263–282, 2004.

[5] W. Wu, W. Zhang, “Constructive and axiomatic approaches of fuzzy approximation operators”, *Information Sciences*, Vol.159, pp. 233–254, 2004.

[6] M. D. Cock, C. Cornelis, E.E. kerre, “Fuzzy rough sets: the forgotten step”, *IEEE Transactions on Fuzzy Systems*, Vol.15, No.1, pp. 121–130, 2007.

[7] T.P. Hong, Y.L. Liou and S.L. Wang, “Fuzzy rough sets with hierarchical quantitative attributes”, *Int. J. Expert Systems with Applications,* Vol.36, No.3, pp. 6790–6799, 2009.

[8] D.S. Yeung, D.G. Chen, E.C.C. Tsang, J.W.T. Lee and X.Z. Wang, “On the generalization of fuzzy rough sets”, *IEEE Trans. Fuzzy Systems*, Vol.13, No.3, pp. 343–361, 2005.

[9] “Q. Hu, L. Zhang, S. An, D. Zhang and D. Yu, “On robust fuzzy rough set models”, *IEEE Transactions on Fuzzy Systems,* Vol.20, No.4, pp. 636–651, 2012.

[10] R. Jensen, Q. Shen, “Fuzzy-rough attribute reduction with application to web categorization”, *Fuzzy Sets and Systems, Elsevier,* Vol.141, pp. 469–485, 2004.

[11] Q. Shen, R. Jensen, “On robust fuzzy rough set models”, *Pattern Recognition,* Vol. 37(7), pp. 1351–1363, 2004.

[12] R. Jensen, Q. Shen, “Semantics-preserving

dimensionality reduction: Rough and fuzzy-rough-based approaches", *IEEE Transactions on Knowledge and Data Engineering,* Vol.16, pp. 1457-1471, 2004.

[13] R. B. Bhatt, M. Gopal, "On fuzzy-rough sets approach to feature selection", *Pattern Recognition Letter*, Vol.26, No.7, pp. 965–975, 2005.

[14] R.B. Bhatt, M. Gopal, "On the compact computational domain of fuzzy rough sets", *Pattern Recognition Letter*, Vol.26, No.11, pp. 1632–1640, 2005.

[15] R. Jensen, Q. Shen, "New approaches to fuzzy-rough feature selection", *IEEE Transactions on Fuzzy Systems,* Vol.17, No.4, pp. 824-838, 2009.

[16] A. M. Radzikowska and E. E. Kerre, "On fuzzy-rough sets approach to feature selection", *Fuzzy Sets Systems,* Vol.126, No.2, pp. 137–155, 2002.

[17] S. Zhao and E. C. C. Tsang, "on fuzzy approximation operators in attribute reduction with fuzzy rough sets", *Information Sciences,* Vol.178, pp. 3163–3176, 2008.

[18] A. Mieszkowicz-Rolka and L. Rolka, "Variable precision fuzzy rough sets", *Transactions on Rough sets, Springer,* Vol. LNCS-3100, pp. 144-160,2004.

[19] S. Zhao, E. C. C. Tsang, and D. Chen, "The model of fuzzy variable precision rough sets", *IEEE Transactions on Fuzzy Systems,* Vol.17, No2, pp. 451–467, 2009.

[20] Q. H. Hu, D. R. Yu, Z. X. Xie, and J. F. Liu, "Fuzzy probabilistic approximation spaces and their information measures", *IEEE Transactions on Fuzzy Systems,* Vol.14, No.2, pp. 191–201, 2006.

[21] Q. H. Hu, D. R. Yu and Z. X. Xie , "Information-preserving hybrid data reduction based on fuzzy-rough techniques", *Pattern Recognition Letters,* Vol.27, No.5, pp. 414–423, 2006.

[22] L. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic", *Fuzzy Sets and Systems,* Vol.19, pp. 111–127, 1997.

[23] Y.Y. Yao, "Information granulation and rough set approximation", *Int. J. Intelligence Systems,* Vol.16, No.1, pp. 87–104, 2001.

[24] Q. Hu., Z. Xie and Daren Yu, "Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation", *Pattern Recognition,* Vol.40, pp. 3509 – 3521, 2007.

[25] Y. Qiana, Q. Wanga, H. Chenga, J. Lianga and C. Dangb, "Fuzzy-rough feature selection accelerator", *Fuzzy Sets and Systems,* Vol.258, pp. 61–78, 2015.

[26] C. Wang, Y. Qi, M. Shao, Q.Hu, D. Chen, Y. Qian and Y. Lin, "A Fitting Model for Feature Selection with Fuzzy Rough Sets", *IEEE Transactions on Fuzzy Systems,* Vol.PP, No.PP, pp. PP, 2016.

[27] P. Maji and P. Garai, "Fuzzy Rough simultaneous Attribute selection and Feature Extraction Algorithm", *IEEE Transactions on Cybernetics,* Vol.43, No.4, pp. 1166–1177, 2013.

[28] E. C. C. Tsang, D. G. Chen, D. S. Yeung and X. Z. Wang, "Fuzzy probabilistic approximation spaces and their information measures", *IEEE Transactions on Fuzzy Systems,* Vol.16, No.5, pp. 1130–1141, 2008.

[29] D. Chen, L. Zhang, S. Zhao, Q. Hu and P. Zhu, "A noval algorithm for finding reducts with fuzzy rough sets", *IEEE Transactions on Fuzzy Systems*, Vol.20, No.2, pp. 385–389, 2012.

[30] Q. He, C. Wu, D. Chen, S. Zhao, "Fuzzy rough set based attribute reduction for information Systems with fuzzy decisions", *Knowledge Based Systems,* Vol.24, pp. 689–696, 2011.

[31] Q. Hu, L. Zhang, Y. Zhou, and W. Pedrycz, "Large-Scale Multi-Modality Attribute Reduction with Multi-Kernel Fuzzy Rough Sets", *IEEE Transactions on Fuzzy Systems,* Vol. PP, No. PP, pp. PP, 2017.

[32] Y. Yang, D. Chen, H. Wang, E. Tsang and D. Zhang, "Fuzzy rough set based incremental attribute reduction from dynamic data with sample arriving", *Fuzzy Sets and Systems,* Vol.312, No.2, pp 66–86, 2017.