FULL PAPER
International Journal of Recent Trends in Engineering, Vol 2, No. 5, November 2009

Rule Selection for Coronary Artery Disease Diagnosis Based on Rough Set

N.A. Setiawan[1,2], P.A. Venkatachalam[1], and Ahmad Fadzil M.H.[1]
[1] Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia
Email: noorwewe@ugm.ac.id
[2] Department of Electrical Engineering, Universitas Gadjah Mada, Indonesia

***Abstract*—The objective of this research is to develop rule selection method for filtering large number of extracted rules from Coronary Artery Disease (CAD) data set. Two stages rule selection is proposed. Selection based on support of individual rules is applied on the first stage. Rough Set based selection using attribute reduction concept is performed for the second stage. Selection based on unseen data set is carried out to increase the generalization of selected rules. Experiment on CAD data set show that the proposed method is able to select small number of rules while maintaining the quality of rule based classifier. The proposed method has better quality compared to previous rule selection methods.**

***Index Terms*—rule selection, rough set, coronary artery disease, diagnosis**

I. Introduction

The result of knowledge discovery process can be decision tree, association rules, decision rules, sequential pattern, etc. The most comprehensive and interpretable extracted knowledge is in the form of rules. Some rule induction algorithm such as rough set theory results in large number of rules. This large number makes interpretability of the knowledge becomes low. Lacking of interpretability will cut down the advantages of rule based systems which should be interpretable and easy to understand. The resulting large number of rules is because of noise, redundancy in input and/or training data sets. Rule pruning is the method to reduce the number of rules while maintaining the quality of the system.

Agotnes investigated genetic algorithm and the use of quality measure for individual rule to filter the large number of rules [1]. Predefined function is used to identify “good” rules. Two properties that are commonly used to define the function are accuracy and coverage [2]. Accuracy is used to measure the reliability or accuracy of single rule. Coverage is used to measure the importance or powerfulness of single rule. In general, a rule is said to be good if it is accurate and powerful. The performance is investigated using Receiver Operating Characteristic (ROC) and statistical hypothesis testing. Ten different formula of rule quality are used. Experiments using real data of acute appendicitis and coronary artery disease are conducted. For acute appendicitis, six to twelve rules filtered from 400 to 500 rules are obtained without significantly reducing the performance of the classifiers.

Corresponding author: N.A. Setiawan, Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia

Maddouri and Gammoudi conducted comparative study on semantic properties of rule interestingness measure [3]. More than forty measures of interestingness are investigated. These interestingness measures are then combined with semantic properties of rules to find which of these measures have good semantic properties. They found Zhang measure [3] has most of the semantic properties. There is no measure that can perform constantly better than others in all fields of applications.

Li and Cercone proposed another measure of rules which is called rule importance measure [4]. Based on medical cases, there are some routine exams that must be conducted by the doctor such as age, blood pressure, body temperature, sex, etc. There are some examinations or symptoms that are not always considered for patients by the doctor. Li and Cercone assumed that the most important examinations or symptoms should be included more frequently than the less important ones. Based on these and in term of rule generation, they defined that if a rule is generated more frequently across different rule sets, it can be said that this rule is more important than rules generated less frequently across those same rules. In term of rough set theory, rule importance measure equals the number of times a rule appears in all the generated rules from the reduct sets divided by number of reduct sets. In order to select important rules from large number of rules automatically, the concept of reduct is introduced. The concept of reduct is used by considering extracted rules as attributes of new reconstructed decision table. Because reduct represents important features of decision table, the attributes of reduct can be considered as important rules. If the new reconstructed decision table has core, then the attributes in the core are considered as the most important rules [5-7]. The applications of rule importance measure on different data sets is found in [5-8].

Rule importance measure based on reduct calculation is not feasible when the number of rules is very large. Finding the reduct will take high computation resource especially for large number of attributes with large number of objects. To overcome this problem, two stages of rule selection are proposed in this paper. The first stage is by using rule quality measure, which is support filtering. After getting acceptable number of rules with acceptable accuracy and coverage, rule importance measure is applied. It is worthless to apply rule importance measure on very low quality of rule sets. Li and Cercone applied rule importance measure on association rules and using training data to reconstruct new decision table. In this paper, modified rule importance measure is applied on decision rules and testing data is used instead of training data to increase the generalization.

II. Material and Methods

A Coronary Artery Disease Data Set

The CAD data sets contain 920 patients that are collected from Cleveland Clinic Foundation U.S. (303 patients); Hungarian Institute of Cardiology, Budapest, Hungary (294 patients); Veterans Administration Medical Center, Long Beach, California, U.S. (123 patients) and University Hospital, Zurich, Switzerland (200 patients). The results of CAD disease diagnosis are obtained by coronary angiography. There are 14 attributes of CAD data [9]. These attributes can be seen in Table I.

ANNRST is used to impute missing values in the CAD data sets [10-12]. After imputation, 661 objects is selected. The training data and testing data consist of 358 and 303 objects respectively.

B Rough Set Theory (RST)

Rough set theory (RST) is a relatively new mathematical and artificial intelligent technique developed bay Zdzislaw Pawlak, Warsaw University of Technology, in the early 1980 [13]. RST is especially useful to discover relationships in data. The discovering of relationship in data is called knowledge discovery or data mining. The result of knowledge discovery is understandable and meaningful knowledge extracted from data. RST method emerged as mathematical tool to manage uncertainties, ambiguity and vagueness from incomplete, inexact and noisy information [14].

For S=(U,A) and $B \subseteq A$, and $a \in B$, with $IND_S(B) = \{(x, x') \in U \times U \mid \forall a \in B, a(x) = a(x')\}$, it can be said that a is dispensable in B if $IND_S(B) = IND_S(B - \{a\})$, otherwise a is indispensable.

A set B is said to be independent if all its attributes are indispensable. Any subset B' of B is called a *reduct* of B if B' is independent and $IND_S(B') = IND_S(B)$.

Reduct can be defined as the minimal subset of attributes that have the same classification of elements as the universe which is the whole set of attributes. In other words, attributes that is not the element of reduct are redundant with respect to classification of elements of the universe. Relative reducts are based on certain objects. For example, reduct *i* is based on object *i*. To find the relative reduct based on object *i*. The other objects that belong to same decision is indiscernible simply because those objects are different from object *i*.

TABLE I
SUMMARY OF ATTRIBUTES (UCI HEART DISEASE DATABASE)

Attribute	Description	Value description
age	Age	Numerical
sex	Sex	1 if male; 0 if female
cp	Chest pain type	1 typical angina 2 atypical angina 3 non-anginal pain 4 asymptomatic
trestbps	Resting systolic blood pressure on admission to the hospital (mmHg)	Numerical
chol	Serum cholesterol (mg/dl)	Numerical
fbs	Fasting blood sugar over 120 mg/dl ?	1 if yes 0 if no
restecg	Resting electrocardiographic results :	0 normal 1 having ST-T wave abnormality 2 LV hypertrophy
thalach	Maximum heart rate achieved	Numerical
exang	Exercise induced angina?	1 if yes 0 if no
oldpeak	ST depression induced by exercise relative to rest	Numerical
slope	The slope of the peak exercise ST segment	1 upsloping 2 flat 3 downsloping
ca	Number of major vessels colored by fluoroscopy	Numerical
thal	Exercise thallium scintigraphic defects	3 normal 6 fixed defect 7 reversible defect
num	Diagnosis of heart disease (angiographic disease status / presence of coronary artery disease (CAD))	0 if less than 50% diameter narrowing in any major vessel (CAD no) 1 if more than 50% (CAD yes)

C Rule Extraction

When all relative reducts are determined, a set of decision rules can be generated from those reducts.

Consider $DS = (U, C \cup D)$ as decision table. $\forall x \in U$, series $c_1(x),..., c_k(x), d(x)$ can be defined, where $\{c_1,...,c_k\} = C$ and $\{d\} = D$. Hence, the decision rules can be generated in the form of $c_1(x),..., c_2(x) \to d(x)$. C can be the condition attributes of reduced form of decision table (reduct). As an example, reduct shown in Table II will have its rules as:

IF c_2 = 3 THEN d = 1. (1)
IF c_2 = 0 THEN d = 0. (2)
IF c_2 = 2 THEN d = 0. (3)
IF c_2 = 1 THEN d = 1 or d = 0. (4)

The set of rules (1-4) is said to be deterministic. Rule

(4) is indeterministic because its consequence is uncertain. The definition of rule support is described as how many objects match the corresponding rule. For example rule (3) has support = 2 because there are 2 objects that match the rule antecedent and consequent. Those objects are x_5 and x_7. Support can be used as rule filtering criterion when there are too many rules generated. Accuracy of certain rule can be calculated as its support divided by the number of objects that match its rule antecedent. Thus, rule (3) has accuracy $=\frac{2}{2}$. Coverage of a rule can also be calculated by dividing support by the number of objects that match its rule consequence. Hence, rule (3) has coverage $=\frac{2}{4}$.

D Rule Selection

RST based rule importance measure [7] is modified to select the rules by converting the rules to decision tables. Support filtering as one of rule quality based filtering is applied at the first stage of the selection. It is time consuming and high computation cost to apply only RST based rule selection because the number of attributes is very high. Filtering method based on rule support is applied to reduce the number of rules before applying rule importance measure to select the most importance rules. The modification is proposed by applying this method to decision system and converting rules to form decision tables based on testing data instead of training data for rule importance measurement.

Consider $R=\{Rule_1, Rule_2,..., Rule_j\}$ as a set of rules generated from training decision tables. If there are *i* objects on testing decision table, a new decision table $DS_{i\times(j+1)}$ can be formed. The value of $Rule_a$ attribute of object x_b is 1 if $Rule_a$ both its antecedent and consequent can be applied to x_b. The value is 0 if the rule cannot be applied. The value for column *j+1* equals decision value. With $a = 1,\ldots,j$ and $b = 1,\ldots,i$. The new decision table then can be reduced using RST reduct concept. The attribute of the shortest reduct are chosen as the selected rules based on their importance. Selected rule based classifier performance is calculated using classifier quality metrics such as accuracy and coverage. Consider Table III as the table of rules and their support. Using support filtering with stopping criteria based on accuracy and coverage, the number of rule will be reduced. If there are *l* number of removed rule, the number of rules become $j = k - l$. The *j* number of rules then applied to the testing data set which is the complete CAD data set to create new decision tables with rules as the attributes as shown in Table IV.

Reduct computation is applied to the Table IV using ROSETTA based on Johnson's algorithm [15,16]. The attributes of the reduct is then the selected rules. The selected rules then is used to classify the testing data. The Michalski, Torgo and Bradzil, which their formulae are empirical and Pearson and Cohen, which are based on statistics and the theory of contingency table is applied for comparison [1].

TABLE III
EXAMPLE OF K RULES WITH THEIR SUPPORT

Rules	Support
$Rule_1$	80
$Rule_2$	70
.	.
.	.
.	.
$Rule_{k-2}$	3
$Rule_{k-1}$	2
$Rule_k$	1

TABLE IV
DECISION TABLE WITH RULES AS ATTRIBUTES

$x \in U$	$Rule_1$	$Rule_2$	. . .	$Rule_{j-1}$	$Rule_j$	*d*
x_1	0	1	. . .	1	1	1
x_2	0	0	. . .	0	1	0
x_3	1	1	. . .	1	0	0
x_4	1	1	. . .	1	1	1
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
x_{i-2}	0	0	. . .	0	1	1
x_{i-1}	0	0	. . .	1	1	1
x_i	0	0	. . .	1	1	1

III. EXPERIMENT RESULTS

A. Rule Extraction

Selected Hungarian and Long Beach data sets with 358 objects are chosen to discover the knowledge. Discretization of numerical attributes is based on Boolean reasoning [15]. The results are in Table V.

RST is used in this study for extracting rules from data sets and extracts 3881 rules. "Support = *m*" means that rules, which have support below the value of *m* are filtered. For example when support = 19, then the rules that have support below 19 are filtered so the number of rules remains 206.

Rule filtering by removing the rules that have support below 20 results in 206 rules with their accuracy and coverage are 0.822 and 0.964. The rules cab be shown in Table VI. This set of rules from support based filtering is chosen and will be selected using RST rule importance measure.

TABLE V
DISCRETIZATION RESULT

Numerical Attributes	Discrete Value		
age	[*, 41)	[41, 53)	[53, *)
trestbps	[*, 129)		[129, *)
chol	[*, 241)		[241, *)
thalach	[*, 133)		[133, *)
oldpeak	[*, 0.3)		[0.3, *)
ca	[*, 1)		[1, *)

where [*, 41) means age < 41, [53,*) means age ≥ 53 and [41,53) means 41≤ age <53.

TABLE VI
EXTRACTED RULES

No	Decision Rules	Support
1	oldpeak([0.3, *)) AND slope(2) AND thal(7) => num(1)	81
2	cp(4) AND slope(2) AND thal(7) => num(1)	77
3	fbs(0) AND exang(1) AND oldpeak([0.3, *)) AND thal(7) => num(1)	73
4	cp(4) AND fbs(0) AND exang(1) AND thal(7) => num(1)	71
5	fbs(0) AND thalach([133, *)) AND oldpeak([*, 0.3)) AND ca ([*, 1)) AND thal(3) => num(0)	70
6	sex(1) AND cp(4) AND fbs(0) AND oldpeak([0.3, *)) AND thal(7) => num(1)	69
7	sex(1) AND exang(1) AND slope(2) AND thal(7) => num(1)	65
8	fbs(0) AND exang(1) AND slope(2) AND thal(7) => num(1)	64
.	.	.
.	.	.
.	.	.
201	trestbps([129, *)) AND chol([241, *)) AND oldpeak([*, 0.3)) AND slope(1) AND ca ([*, 1)) => num(0)	20
202	cp(4) AND trestbps([129, *)) AND chol([241, *)) AND restecg(0) AND thalach([*, 133)) AND slope(2) => num(1)	20
203	age([53, *)) AND trestbps([129, *)) AND restecg(0) AND exang(0) AND ca ([*, 1)) => num(0)	20
204	age([53, *)) AND restecg(1) AND slope(2) AND thal(7) => num(1)	20
205	age([41, 53)) AND trestbps([*, 129)) AND chol([*, 241)) AND thalach([133, *)) AND ca ([*, 1)) => num(0)	20
206	cp(4) AND trestbps([129, *)) AND chol([241, *)) AND fbs(0) AND restecg(0) AND thalach([*, 133)) => num(1)	20

B. Rule Selection

The new constructed decision table consists of 303 objects. The decision table has 206 conditional attributes with single decision, which is attribute *num*. RST reduct computation using ROSETTA Johnson's algorithm results in reduct with 27 conditions which represent 27 rules. The reduct attributes are {*Rule1*, *Rule10*, *Rule17*, *Rule19*, *Rule20*, *Rule23*, *Rule26*, *Rule30*, *Rule33*, *Rule49*, *Rule51*, *Rule66*, *Rule70*, *Rule73*, *Rule77*, *Rule81*, *Rule101*, *Rule114*, *Rule117*, *Rule118*, *Rule150*, *Rule154*, *Rule158*, *Rule160*, *Rule161*, *Rule180*, *Rule203*}. These are shown in Table VII.

TABLE VII
SELECTED RULES

No.	Rule	Rule No
1	oldpeak([0.3, *)) AND slope(2) AND thal(7) => num(1)	Rule1
2	fbs(0) AND thalach([133, *)) AND slope(1) AND ca ([*, 1)) AND thal(3) => num(0)	Rule10
3	fbs(0) AND ca ([1, *)) AND thal(7) => num(1)	Rule17
4	sex(1) AND fbs(0) AND thalach([133, *)) AND exang(0) AND ca ([*, 1)) AND thal(3) => num(0)	Rule19
5	sex(1) AND fbs(0) AND restecg(0) AND oldpeak([0.3, *)) AND thal(7) => num(1)	Rule20
6	chol([*, 241)) AND exang(0) AND thal(3) => num(0)	Rule23
7	trestbps([129, *)) AND restecg(0) AND ca ([*, 1)) AND thal(3) => num(0)	Rule26
8	cp(4) AND exang(1) AND slope(2) AND ca ([1, *)) => num(1)	Rule30
9	cp(2) AND chol([*, 241)) AND ca ([*, 1)) => num(0)	Rule33
10	sex(0) AND fbs(0) AND exang(0) AND oldpeak([*, 0.3)) => num(0)	Rule49
11	chol([241, *)) AND fbs(0) AND slope(2) AND thal(7) => num(1)	Rule51
12	trestbps([129, *)) AND slope(1) AND ca ([*, 1)) AND thal(3) => num(0)	Rule66
13	chol([241, *)) AND slope(2) AND ca ([1, *)) => num(1)	Rule70
14	exang(1) AND oldpeak([0.3, *)) AND ca ([*, 1)) AND thal(7) => num(1)	Rule73
15	cp(4) AND chol([241, *)) AND fbs(0) AND ca ([1, *)) => num(1)	Rule77
16	age([41, 53)) AND chol([*, 241)) AND thal(3) => num(0)	Rule81
17	sex(1) AND cp(4) AND trestbps([129, *)) AND chol([241, *)) AND thalach([*, 133)) AND slope(2) => num(1)	Rule101
18	age([41, 53)) AND sex(1) AND thalach([133, *)) AND exang(0) AND ca ([*, 1)) => num(0)	Rule114
19	thalach([133, *)) AND exang(1) AND thal(7) => num(1)	Rule117
20	trestbps([129, *)) AND chol([241, *)) AND fbs(0) AND exang(0) AND ca ([*, 1)) => num(0)	Rule118
21	age([53, *)) AND sex(1) AND trestbps([129, *)) AND fbs(1) AND slope(2) => num(1)	Rule150
22	sex(1) AND cp(4) AND thalach([133, *)) AND thal(7) => num(1)	Rule154
23	age([53, *)) AND cp(2) AND slope(1) => num(0)	Rule158
24	age([41, 53)) AND sex(0) AND thal(3) => num(0)	Rule160
25	trestbps([*, 129)) AND fbs(0) AND ca ([1, *)) => num(1)	Rule161
26	chol([*, 241)) AND restecg(0) AND oldpeak([0.3, *)) AND ca ([1, *)) => num(1)	Rule180
27	age([53, *)) AND trestbps([129, *)) AND restecg(0) AND exang(0) AND ca ([*, 1)) => num(0)	Rule203

These selected 27 rules have accuracy and coverage of 0.852 and 0.937 respectively on testing data set. Table VII shows the all of 27 selected rules based on RST reduct.

The comparison of between rule selection methods applied on training and testing data sets can be seen on Table VIII and Table IX. Two support based filtering are implemented. This selection method is based on support of rules on training data set and testing data set.

TABLE VIII
COMPARISON OF RULE SELECTION METHOD PERFORMANCE APPLIED ON TRAINING DATA SET

Selection Methods	Accuracy	Coverage	Number of rules
Proposed Method	1	0.816	27
Support Based (Training Data)	1	0.665	29
Support Based (Testing Data)	1	0.749	27
Michalski μ=0.5	1	0.682	27
Torgo	1	0.682	27
Brazdil	1	0.682	27
Pearson	1	0.682	27
Cohen	1	0.598	29

TABLE IX
COMPARISON OF RULE SELECTION METHOD PERFORMANCE APPLIED ON TESTING DATA SET

Selection Methods	Accuracy	Coverage	Number of rules
Proposed Method	0.852	0.937	27
Support Based (Training Data)	0.847	0.799	29
Support Based (Testing Data)	0.844	0.868	27
Michalski μ=0.5	0.845	0.785	27
Torgo	0.845	0.785	27
Brazdil	0.845	0.785	27
Pearson	0.845	0.785	27
Cohen	0.863	0.65	29

The formula of selection methods which are Michalski, Torgo, Brazdil, Pearson and Cohen are explained in [1].

Table VIII and Table IX show that the proposed method performs better than the others both testing and training data sets. The proposed method is also compared with other four classifiers such as decision tree, k-NN, LTF-C [17,18] and unfiltered RST rules on testing data. The results are shown in Table X. LTF-C and k-NN do not produce rules.

TABLE X
COMPARISON OF RULE SELECTION METHOD PERFORMANCE AND OTHER CLASSIFIERS APPLIED ON TESTING DATA SET

Methods	Accuracy	Coverage	Number of rules
Proposed Method	0.852	0.937	27
Unfiltered RST rules	0.815	1	4095
Decision Tree	0.856	0.482	83
LTF-C	0.812	1	-
k-NN	0.815	1	-

The proposed method is able to select high quality and important rules. Although the coverage decreases, the accuracy increases with only 27 rules. The coverage is still quite high and much better than decision tree.

IV. CONCLUSION

In this paper, RST based rule selection method is proposed. The proposed method uses hybrid approach of support filtering and RST rule importance selection. Support filtering is able to select high quality rules from large number of rules and RST rule importance is able to select the most important rules using reduct concept. By converting the rules into decision table based on testing data set, attribute reduction can be used to select the most important rules. The use of testing data instead of training data to convert rules into decision table gives better coverage when applied on the testing data. This is because testing data is involved in rule selection process. The experimental results show that the proposed method is able to select small number (27) of rules from large number (3881) of rule without significantly reducing the quality of classification. Small amount of data is easy to understand. The proposed method is comparable or even better than almost other methods.

In the future, more unseen data should be used as validation set to investigate the generalization improvement of the method by classifying unknown data, which is not involved in the selection process totally. The selected rule can be used as knowledge base for decision support system to diagnose coronary artery disease. In the future work, fuzzy reasoning can be considered for inference engine of that decision support system.

ACKNOWLEDGMENT

We gratefully acknowledge the financial and facility support of Universiti Teknologi PETRONAS to conduct this research project.

REFERENCES

[1] T. Agotnes, "Filtering large propositional rule sets while retaining classifier performance," in *Department of Computer and Information Science*: Norwegian Universtiy of Science and Technology, 1999, pp. 143.

[2] S. Tsumoto, "Accuracy and coverage in rough set rule induction," in *Rough Sets and Current Trends in Computing*, J. J. Alpigini, et. al., Ed. Berlin, Heidelberg: Springer-Verlag, 2002, pp. 373-380.

[3] M. Maddouri and J. Gammoudi, "On Semantic Properties of Interestingness Measures for Extracting Rules from Data," *Adaptive and Natural Computing Algorithms*, pp. 148, 2007.

[4] J. Li and N. Cercone, "Introducing a rule importance measure," in *Lecture Notes in Computer Science*: Springer, 2006, pp. 167.

[5] J. Li and N. Cercone, "Discovering and ranking important rules," presented at Granular Computing, 2005 IEEE International Conference on, 2005. pp. 506

[6] J. Li, P. Pattaraintakorn, and N. Cercone, "Rule evaluations, attributes, and rough sets: extension and a case study," in *Lecture Notes in Computer Science, Transactions on Rough Sets VI*, 2007, pp. 152.

[7] J. Li, "Rough set based rule evaluations and their applications." Waterloo, Ontario: University of Waterloo, 2007, pp. 191.

[8] J. Li and N. Cercone, "A rough set based model to rank the importance of association rules," in *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, 2005, pp. 109.

[9] D. J. Newman, S. Hettich, C. L. Blake, and C. J. Merz, "UCI Repository of machine learning databases," University California Irvine, Department of Information and Computer Science, 1998.

[10] N. A. Setiawan, P. A. Venkatachalam, and A. F. M. Hani, "Missing data estimation on heart disease using artificial neural network and rough set theory," presented at International Conference on Intelligent and Advanced Systems, Kuala Lumpur, Malaysia, 2007. pp. 129-133

[11] N. A. Setiawan, P. A. Venkatachalam, and A. F. M. Hani, "Missing attribute value prediction based on artificial neural network and rough set theory," presented at International Conference on Biomedical Engineering and Informatics, Sanya, Hainan, China, 2008. pp. 306-310

[12] N. A. Setiawan, P. A. Venkatachalam, and A. F. M. Hani, "A comparative study of imputation methods to predict missing attribute values in coronary heart disease data set," presented at 4th Kuala Lumpur International Conference on Biomedical Engineering, Kuala Lumpur, 2008. pp. 266-269

[13] Z. Pawlak, "Rough Sets," *International Journal of Computer and Information Sciences*, vol. 11, pp. 341-355, 1982.

[14] T. Munakata, *Fundamentals of the new artificial intelligence: beyond traditional paradigms*. New York: Springer, 1998.

[15] A. Ohrn, "Discernibility and rough sets in medicine: tools and applications," in *Department of computer and information science*. Trondheim: Norwegian University of Science and Technology, 1999, pp. 223.

[16] D. S. Johnson, "Approximation algorithms for combinatorial problems," *Journal of Computer and System Sciences*, vol. 9, pp. 256-278, 1974.

[17] J. Bazan and M. Szczuka, "RSES and RSESlib - A Collection of Tools for Rough Set Computations," in *Rough Sets and Current Trends in Computing*, 2001, pp. 106.

[18] J. G. Bazan, M. S. Szczuka, A. Wojna, and M. Wojnarski, "On the evolution of rough set exploration system," in *Rough Sets and Current Trends in Computing*, 2004, pp. 592.