Feature Selection with Positive Region Constraint for Test-Cost-Sensitive Data

Jiabin Liu[1,2], Fan Min[2](✉), Hong Zhao[2], and William Zhu[2]

[1] Department of Computer Science, Sichuan University for Nationalities, Kangding 626001, China
liujb418@163.com

[2] Lab of Granular Computing, Zhangzhou Normal University, Zhangzhou 363000, China
minfanphd@163.com

Abstract. In many data mining and machine learning applications, data are not free, and there is a test cost for each data item. Due to economic, technological and legal reasons, it is neither possible nor necessary to obtain a classifier with 100 % accuracy. In this paper, we consider such a situation and propose a new constraint satisfaction problem to address it. With this in mind, one has to minimize the test cost to keep the accuracy of the classification under a budget. The constraint is expressed by the positive region, whereas the object is to minimizing the total test cost. The new problem is essentially a dual of the test cost constraint attribute reduction problem, which has been addressed recently. We propose a heuristic algorithm based on the information gain, the test cost, and a user specified parameter λ to deal with the new problem. The algorithm is tested on four University of California - Irvine datasets with various test cost settings. Experimental results indicate that the algorithm finds optimal feature subset in most cases, the rational setting of λ is different among datasets, and the algorithm is especially stable when the test cost is subject to the Pareto distribution.

Keywords: Feature selection · Cost-sensitive learning · Positive region · Test cost

1 Introduction

When industrial products are manufactured, they must be inspected strictly before delivery. Testing equipments are needed to classify the product as qualified, unqualified, etc. Each equipment costs money, which will be averaged on each product. Generally, we should pay more to obtain higher classification accuracy. However, due to economic, technological and legal reasons, it is neither possible nor necessary to obtain a classifier with 100 % accuracy. There may be an industrial standard to indicate the accuracy of the classification, such as 95 %. Consequently, we are interested in a set of equipments with minimal cost meeting the standard. In this scenario, there are two issues: one is the equipment

J.F. Peters et al. (Eds.): Transactions on Rough Sets XVIII, LNCS 8449, pp. 23–33, 2014.
DOI: 10.1007/978-3-662-44680-5_2

cost, and the other is the product classification accuracy. They are called test cost and classification accuracy, respectively. Since the classification accuracy only needs to meet the industrial standard, we would like to choose some testing equipments to minimize the total cost.

In many real applications of data mining, machine learning, pattern recognition and signal processing, datasets often contain huge numbers of features. In such a case, feature selection will be necessary [15]. Feature selection [10,32,42] is the process of choosing a subset of features from the original set of features forming patterns in a given dataset. The subset should be necessary and sufficient to describe target concepts, retaining a suitably high accuracy in representing the original features. Feature selection serves as a pre-processing technique in machine learning and pattern recognition application [1]. Consequently, it has been defined by many authors by booking at it from various angles [2].

According to Pawlak [27], minimal reducts have the best generalization ability. Hence many feature selection reduction algorithms based on rough set have been proposed to find one of them (see, e.g., [29,31,36,38,46]). On the other side, however, data are not free, and there is a test cost for each data item [9]. Therefore the classifier should also exhibit low test cost [19]. With this in mind, the minimal test cost reduct problem has been defined [17] and with some algorithms proposed to deal with it [6,7,21,25,35,41].

In this paper, we formally define the feature selection with positive region constraint for test-cost-sensitive data (FSPRC) problem. The positive region is a widely used concept in rough set [26]. We use this concept instead of the classification accuracy to specify the industrial standard. The new problem is essentially a dual of the optimal sub-reduct with test cost constraint (OSRT) problem, which has been defined in [20] and studied in [14,22,23]. The OSRT problem considers the test cost constraint, while the new problem considers the positive region constraint. As will be discussed in the following text, the classical reduct problem can be viewed as a special case of the FSPRC problem. Since the classical reduct problem is NP-hard, the new problem is at least NP-hard.

We propose a heuristic algorithm to deal with the new problem. The algorithm follows a popular addition-deletion framework. Since we do not require that the positive region is unchanged after feature selection, there does not exist a core computing stage. The heuristic information function is based on both the information gain and the test cost. It is deliberately designed to obtain a tradeoff between the usefulness and the cost of each feature.

Four open datasets from the University of California-Irvine (UCI) library are employed to study the performance and effectiveness of our algorithms. Experiments are undertaken with open source software Cost-sensitive rough sets (Coser) [24]. Results indicate the algorithm can find the optimal feature subset in most cases, the rational setting of λ is different among datasets, and the algorithm is especially stable when the test cost is subject to the Pareto distribution.

The rest of this paper is organized as follows. Section 2 describes related concepts in the rough set theory and defines the FSPRC problem formally. In Sect. 3, a heuristic algorithm based on λ-weighted information gain is presented.

Section 4 illustrates some results on four UCI datasets with detailed analysis. Finally, Sect. 5 gives some conclusions and indicates possibilities for further work.

2 Preliminaries

In this section, we define the FSPRC problem. First, we revisit the data model on which the problem is defined. Then we review the concept of positive region. Finally we propose feature selection with positive region constraint problem.

2.1 Test-Cost-Independent Decision Systems

Decision systems are fundamental in machine learning and data mining. A decision system is often denoted as $S = (U, C, D, \{V_a|a \in C \cup D\}, \{I_a|a \in C \cup D\})$, where U is a finite set of objects called the universe, C is the set of conditional attributes, D is the set of decision attributes, V_a is the set of values for each $a \in C \cup D$, and $I_a : U \rightarrow V_a$ is an information function for each $a \in C \cup D$. We often denote $\{V_a|a \in C \cup D\}$ and $\{I_a|a \in C \cup D\}$ by V and I, respectively. Table 1 is a decision system, which conditional attributes are symbolic values. Here C = {Patient, Headache, Temperature, Lymphocyte, Leukocyte, Eosinophil, Heartbeat}, d = {Flu}, and $U = \{x_1, x_2, \dots, x_7\}$.

A *test-cost-independent decision system* (TCI-DS) [19] is a decision system with test cost information represented by a vector, as the one in Table 2. It is the simplest form of the test-cost-sensitive decision system and defined as follows.

Table 1. An exemplary decision table

Patient	Headache	Temperature	Lymphocyte	Leukocyte	Eosinophil	Heartbeat	Flu
x_1	Yes	High	High	High	High	Normal	Yes
x_2	Yes	High	Normal	High	High	Abnormal	Yes
x_3	Yes	High	High	High	Normal	Abnormal	Yes
x_4	No	High	Normal	Normal	Normal	Normal	No
x_5	Yes	Normal	Normal	Low	High	Abnormal	No
x_6	Yes	Normal	Low	High	Normal	Abnormal	No
x_7	Yes	Low	Low	High	Normal	Normal	Yes

Table 2. An exemplary cost vector

a	Headache	Temperature	Lymphocyte	Leukocyte	Eosinophil	Heartbeat
$c(a)$	\$12	\$5	\$15	\$20	\$15	\$10

Definition 1. *[19] A test-cost-independent decision system (TCI-DS) S is the 6-tuple:*

$$S = (U, C, D, V, I, c), \tag{1}$$

where U, C, D, V and I have the same meanings as in a decision system, and $c : C \rightarrow R^+ \cup \{0\}$ is the test cost function. Here the test cost function can easily be represented by a vector $c = [c(a_1), c(a_2), \ldots, c(a_{|C|})]$, which indicates that test costs are independent of one another, that is, $c(B) = \sum_{a \in B} c(a)$ for any $B \subset C$.

For example, if we select tests Temperature, Lymphocyte, Leukocyte and Heartbeat, the total test cost would be \$5 + \$15 + \$20 + \$10 = \$50. This is why we call this type of decision system "test-cost-independent". If any element in c is 0, a TCI-DS coincides with a DS. Therefore, free tests are not considered for the sake of simplicity. A TCI-DS is represented by a decision table and a test cost vector. One example is given by Tables 1 and 2 [19].

2.2 Positive Region

Rough set theory [27] is an approach to vagueness and uncertainty. Similarly to fuzzy set theory it is not an alternative to classical set theory but it is embedded in it. Positive region is an important concept in rough set theory. It is defined by through lower approximation. Let $S = (U, C, D, V, I)$ be a decision system. Any $\emptyset \neq B \subseteq C \cup D$ determines an indiscernibility relation on U. A partition determined by B is denoted by U/B. Let $\underline{B}(X)$ denote the B-lower approximation of X.

Definition 2. *[27] Let $S = (U, C, D, V, I)$ be a decision system, $\forall B \subset C$, the positive region of D with respect to B is defined as*

$$POS_B(D) = \bigcup_{X \in U/D} \underline{B}(X), \tag{2}$$

where U, C, D, V and I have the same meanings as in a decision system.

In other words, D is totally (partially) dependent on B, if all (some) elements of the universe U can be uniquely classified to blocks of the partition U/D, employing B [26].

2.3 Feature Selection with Positive Region Constraint Problem

Feature selection is the process of choosing an appropriate subset of attributes from the original dataset [34]. There are numerous reduct problems which have been defined on the classical [28], the neighborhood (see, e.g., [11,12]), the covering-based [16,40,43–46], the decision-theoretical [37], and the dominance-based [4] rough set models. Respective definitions of relative reducts also have been studied in [8,29].

Definition 3. *[27] Let $S = (U, C, D, V, I)$ be a decision system. Any $B \subseteq C$ is called a decision relative reduct (or a relative reduct for brevity) of S iff:*

(1) $POS_B(D) = POS_C(D)$.

(2) $\forall a \in B, POS_{B-\{a\}}(D) \neq POS_B(D)$.

Definition 3 implies two issues. One is that the reduct is jointly sufficient, the other is that the reduct is individually necessary for preserving a particular property (positive region in this context) of the decision systems [17]. The set of all relative reducts of S is denoted by $Red(S)$. The core of S is the intersection of these reducts, namely, $core(S) = \cap Red(S)$. Core attributes are of great importance to the decision system and should never be removed, except when information loss is allowed [37].

Throughout this paper, due to the positive region constraint, it is not necessary to construct a reduct. On the other side, we never want to select any redundant test. Therefore we propose the following concept.

Definition 4. *Let $S = (U, C, D, V, I)$ be a decision system. Any $B \subseteq C$ is a positive region sub-reduct of S iff $\forall a \in B$, $POS_{B-\{a\}}(D) \neq POS_B(D)$.*

According to the Definition 4, we observe the following:

(1) A reduct is also a sub-reduct.
(2) A core attribute may not be included in a sub-reduct.

Here we are interested those feature subsets satisfying the positive region constraint, and at the same time, with minimal possible test cost. To formalize the situation, we adopt the style of [18] and define the feature selection with positive region constraint problem, where the optimization objective is to minimize the test cost under the constraint.

Problem 1. Feature selection with positive region constraint (FSPRC) problem.
Input: $S = (U, C, d, V, I, c)$, the positive region lower bound pl;
Output: $B \subseteq C$;
Constraint: $|POS_B(D)|/|POS_C(D)| \geq pl$;
Optimization objective: $\min c(B)$.

In fact, the FSPRC problem is more general than the minimal test cost reduct problem, which is defined in [17]. In case where $pl = 1$, it coincides with the later. The minimal test cost reduct problem is in turn more general than the classical reduct problem, which is NP-hard. Therefore the FSPRC problem is at least NP-hard, and heuristic algorithms are needed to deal with it. Note that the FSPRC is different with the variable precision rough set model. The variable precision rough set model changes the lower approximation by varying the accuracy, but in our problem definition, it is unchanged.

3 The Algorithm

Similar to the heuristic algorithm to the OSRT problem [20], we also design a heuristic algorithm to deal with the new problem. We firstly analyze the heuristic function which is the key issue in the algorithm. Let $B \subset C$ and $a_i \in C - B$, the information gain of a_i with respect to B is

$$f_e(B, a_i) = H(\{d\}|B) - H(\{d\}|B \cup \{a_i\}), \tag{3}$$

where $d \in D$ is a decision attribute. At the same time, the λ-weighted function is defined as

$$f(B, a_i, c, \lambda) = f_e(B, a_i) c_i^{\lambda}, \tag{4}$$

where λ is a non-positive number.

Algorithm 1. A heuristic algorithm to the FSPRC problem

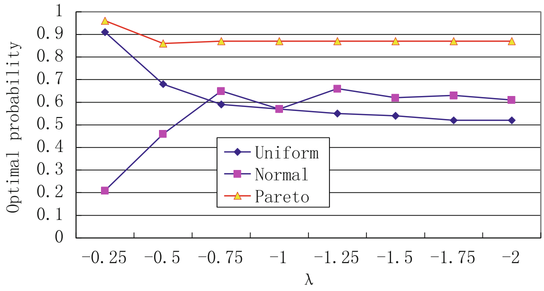
Input: $S = (U, C, D, V, I, c)$, p_{con}, λ
Output: A sub-reduct of S
Method: FSPRC

1: $B = \emptyset$; //the sub-reduct
2: $CA = C$; //the unprocessed attributes
3: **while** $(|POS_B(D)| < p_{con})$ **do**
4: For any $a \in CA$ compute $f(B, a, c, \lambda)$
//Addition
5: Select a' with maximal $f(B, a', c, \lambda)$;
6: $B = B \cup \{a'\}$;
7: $CA = CA - \{a'\}$;
8: **end while**
//Deletion, B must be a sub-reduct
9: $CD = B$; //sort attribute in CD according to respective test cost in a descending order
10: **while** $CD \neq \emptyset$ **do**
11: $CD = CD - \{a'\}$; //where a' is the first element in CD
12: **if** $(POS_{B-\{a'\}}(D) = POS_B(D))$ **then**
13: $B = B - \{a'\}$;
14: **end if**
15: **end while**
16: return B

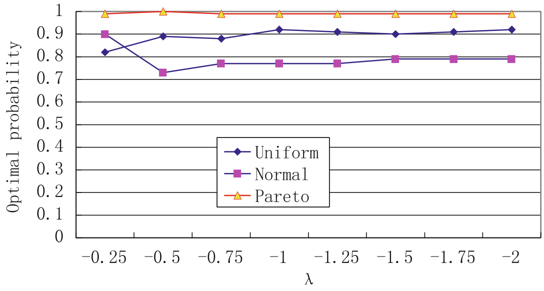
Our algorithm is listed in Algorithm 1. It contains two main steps. The first step contains lines 3 through 8. Attributes are added to B one by one according to the heuristic function indicated in Eq. (4). This step stops while the positive region reaches the lower bound. The second step contains lines 9 through 15. Redundant attributes are removed from B one by one until all redundant have been removed. As discussed in Sect. 2.3, our algorithm has not a stage of core computing.

4 Experiments

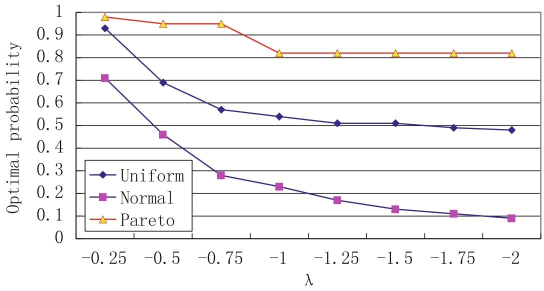
To study the effectiveness of the algorithm, we have undertaken experiments using our open source software Coser [24] on 4 different datasets, i.e., Zoo, Iris,



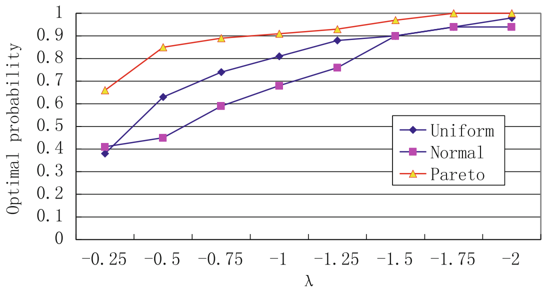
(a)



(b)



(c)



(d)

Fig. 1. Optimal probability: (a) zoo; (b) iris; (c) voting; (d) tic-tac-toc.

Voting, and Tic-tac-toe, downloaded from the UCI library [5]. To evaluate the performance of the algorithm, we need to study the quality of each sub-reduct which it computes. This experiment should be undertaken by comparing each sub-reduct to an optimal sub-reduct with the positive region constraint. Unfortunately, the computation of an optimal sub-reduct with test positive region constraint is more complex than that of a minimal reduct, or that of a minimal test cost reduct. In this paper, we only study the influence of λ to the quality of the result.

4.1 Experiments Settings

Because of lacking the predefined test costs in the four artificial datasets, we specify them as the same setting as that of [17] to produce test costs within [1, 100]. Three distributions, namely, Uniform, Normal, and bounded Pareto, are employed. In order to control the shape of the Normal distribution and the bounded Pareto distribution respectively, we must set the parameter α. In our experiment, for the Normal distribution, $\alpha = 8$, and test costs as high as 70 and as low as 30 are often generated. For the bounded Pareto distribution, $\alpha = 2$, and test costs higher than 50 are often generated. In addition, we intentionally set the constraint as $pl = 0.8$. This setting shows that we need a sub-reduct rather than a reduct.

4.2 Experiments Results

The experimental results of the 4 datasets are illustrated in Fig 1. By running our program in different λ values, the 3 different test cost distributions are compared. We can observe the following.

(1) The algorithm finds the optimal feature subset in most cases. With appropriate settings, it achieves more than 70 % optimal probability on these datasets.
(2) The result is influenced by the user-specified λ. The probability of obtained the best results is different with different λ values, where the "best" means the best one over the solutions we obtained, not the optimal one.
(3) The algorithm's performance is related with the test cost distribution. It is best on datasets with bounded Pareto distribution. At the same time, it is worst on datasets with Normal distribution. Consequently, if the real data has test cost subject to the Normal distribution, one may develop other heuristic algorithms to this problem.
(4) There is not a setting of λ such that the algorithm always obtain the best result. Therefore the settings might be learned instead of provided by the user.

5 Conclusions

In this paper, we firstly proposed the FSPRC problem. Then we designed a heuristic algorithm to deal with it. Experimental results indicate that the optimal solution is not easy to obtain. In the future, one can borrow ideas from [13,30,33,39] to develop an exhaustive algorithm to evaluate the performance of a heuristic algorithm. One also can develop more advanced heuristic algorithms to obtain better performance.

Acknowledgements. This work is partially supported by the Natural Science Foundation of Department of Education of Sichuan Province under Grant No. 13ZA0136, and National Science Foundation of China under Grant Nos. 61379089, 61379049.

References

1. Chen, X.: An improved branch and bound algorithm for feature selection. Pattern Recogn. Lett. **24**(12), 1925–1933 (2003)
2. Dash, M., Liu, H.: Feature selection for classification. Intell. Data Anal. **1**, 131–156 (1997)
3. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. AI Mag. **17**, 37–54 (1996)
4. Greco, S., Matarazzo, B., Slowinski, R., Stefanowski, J.: Variable consistency model of dominance-based rough sets approach. In: Ziarko, W.P., Yao, Y. (eds.) RSCTC 2000. LNCS (LNAI), vol. 2005, pp. 170–181. Springer, Heidelberg (2001)
5. Blake, C.L., Merz, C.J.: UCI repository of machine learning databases (1998). http://www.ics.uci.edu/~mlearn/mlrepository.html
6. He, H.P., Min, F.: Accumulated cost based test-cost-sensitive attribute reduction. In: Kuznetsov, S.O., Ślęzak, D., Hepting, D.H., Mirkin, B.G. (eds.) RSFDGrC 2011. LNCS, vol. 6743, pp. 244–247. Springer, Heidelberg (2011)
7. He, H.P., Min, F., Zhu, W.: Attribute reduction in test-cost-sensitive decision systems with common-test-costs. In: Proceedings of the 3rd International Conference on Machine Learning and Computing, vol. 1, pp. 432–436 (2011)
8. Hu, Q.H., Yu, D.R., Liu, J.F., Wu, C.: Neighborhood rough set based heterogeneous feature subset selection. Inf. Sci. **178**(18), 3577–3594 (2008)
9. Hunt, E.B., Marin, J., Stone, P.J. (eds.): Experiments in Induction. Academic Press, New York (1966)
10. Lanzi, P.: Fast feature selection with genetic algorithms: a filter approach. In: IEEE International Conference on Evolutionary Computation 1997. IEEE (1997)
11. Lin, T.Y.: Granular computing on binary relations - analysis of conflict and Chinese wall security policy. In: Alpigini, J.J., Peters, J.F., Skowron, A., Zhong, N. (eds.) RSCTC 2002. LNCS (LNAI), vol. 2475, pp. 296–299. Springer, Heidelberg (2002)
12. Lin, T.Y.: Granular computing - structures, representations, and applications. In: Wang, G., Liu, Q., Yao, Y., Skowron, A. (eds.) RSFDGrC 2003. LNCS (LNAI), vol. 2639, pp. 16–24. Springer, Heidelberg (2003)
13. Liu, Q.H., Li, F., Min, F., Ye, M., Yang, W.G.: An efficient reduction algorithm based on new conditional information entropy. Control Decis. (in Chinese) **20**(8), 878–882 (2005)
14. Liu, J.B., Min, F., Liao, S.J., Zhu, W.: A genetic algorithm to attribute reduction with test cost constraint. In: Proceedings of 6th International Conference on Computer Sciences and Convergence Information Technology, pp. 751–754 (2011)
15. Liu, H., Motoda, H.: Feature Selection for Knowledge Discovery and Data Mining. The Springer International Series in Engineering and Computer Science, vol. 454. Kluwer Academic Publishers, Boston (1998)

16. Ma, L.W.: On some types of neighborhood-related covering rough sets. Int. J. Approx. Reason. **53**(6), 901–911 (2012)
17. Min, F., He, H.P., Qian, Y.H., Zhu, W.: Test-cost-sensitive attribute reduction. Inf. Sci. **181**, 4928–4942 (2011)
18. Min, F., Hu, Q.H., Zhu, W.: Feature selection with test cost constraint. Int. J. Approximate Reasoning (2013, to appear). doi:10.1016/j.ijar.2013.04.003
19. Min, F., Liu, Q.H.: A hierarchical model for test-cost-sensitive decision systems. Inf. Sci. **179**, 2442–2452 (2009)
20. Min, F., Zhu, W.: Attribute reduction with test cost constraint. J. Electr. Sci. Technol. China **9**(2), 97–102 (2011)
21. Min, F., Zhu, W.: Minimal cost attribute reduction through backtracking. In: Kim, T., et al. (eds.) DTA/BSBT 2011. CCIS, vol. 258, pp. 100–107. Springer, Heidelberg (2011)
22. Min, F., Zhu, W.: Optimal sub-reducts in the dynamic environment. In: Proceedings of IEEE International Conference on Granular Computing, pp. 457–462 (2011)
23. Min, F., Zhu, W.: Optimal sub-reducts with test cost constraint. In: Yao, J.T., Ramanna, S., Wang, G., Suraj, Z. (eds.) RSKT 2011. LNCS, vol. 6954, pp. 57–62. Springer, Heidelberg (2011)
24. Min, F., Zhu, W., Zhao, H., Pan, G.Y., Liu, J.B., Xu, Z.L.: Coser: cost-sensitive rough sets (2012). http://grc.fjzs.edu.cn/~fmin/coser/
25. Pan, G.Y., Min, F., Zhu, W.: A genetic algorithm to the minimal test cost reduct problem. In: Proceedings of IEEE International Conference on Granular Computing. pp. 539–544 (2011)
26. Pawlak, Z.: Rough set approach to knowledge-based decision support. Eur. J. Oper. Res. **99**, 48–57 (1997)
27. Pawlak, Z.: Rough sets. Int. J. Comput. Inf. Sci. **11**, 341–356 (1982)
28. Pawlak, Z.: Rough sets and intelligent data analysis. Inf. Sci. **147**(12), 1–12 (2002)
29. Qian, Y.H., Liang, J.Y., Pedrycz, W., Dang, C.Y.: Positive approximation: an accelerator for attribute reduction in rough set theory. Artif. Intell. **174**(9–10), 597–618 (2010)
30. Skowron, A., Rauszer, C.: The discernibility matrices and functions in information systems. In: Intelligent Decision Support (1992)
31. Swiniarski, R.W., Skowron, A.: Rough set methods in feature selection and recognition. Pattern Recogn. Lett. **24**(6), 833–849 (2003)
32. Tseng, T.L.B., Huang, C.-C.: Rough set-based approach to feature selection in customer relationship management. Omega **35**(4), 365–383 (2007)
33. Wang, G.Y.: Attribute core of decision table. In: Alpigini, J.J., Peters, J.F., Skowron, A., Zhong, N. (eds.) RSCTC 2002. LNCS (LNAI), vol. 2475, pp. 213–217. Springer, Heidelberg (2002)
34. Wang, X., Yang, J., Teng, X., Xia, W., Jensen, R.: Feature selection based on rough sets and particle swarm optimization. Pattern Recogn. Lett. **28**(4), 459–471 (2007)
35. Xu, Z.L., Min, F., Liu, J.B., Zhu, W.: Ant colony optimization to minimal test cost reduction. In: Proceedings of the 2011 IEEE International Conference on Granular Computing. pp. 688–693 (2012)
36. Yao, J.T., Zhang, M.: Feature selection with adjustable criteria. In: Ślęzak, D., Wang, G., Szczuka, M.S., Düntsch, I., Yao, Y. (eds.) RSFDGrC 2005. LNCS (LNAI), vol. 3641, pp. 204–213. Springer, Heidelberg (2005)
37. Yao, Y.Y., Zhao, Y.: Attribute reduction in decision-theoretic rough set models. Inf. Sci. **178**(17), 3356–3373 (2008)

38. Zhang, W.X., Mi, J., Wu, W.: Knowledge reductions in inconsistent information systems. Chin. J. Comput. **26**(1), 12–18 (2003)
39. Zhao, H., Min, F., Zhu, W.: A backtracking approach to minimal cost feature selection of numerical data. J. Inf. Comput. Sci. **10**(13), 4105–4115 (2013)
40. Zhao, H., Min, F., Zhu, W.: Test-cost-sensitive attribute reduction based on neighborhood rough set. In: Proceedings of the 2011 IEEE International Conference on Granular Computing, pp. 802–806 (2011)
41. Zhao, H., Min, F., Zhu, W.: Test-cost-sensitive attribute reduction of data with normal distribution measurement errors. Math. Prob. Eng. **2013**, 12 pp (2013)
42. Zhong, N., Dong, Z.J., Ohsuga, S.: Using rough sets with heuristics to feature selection. J. Intell. Inf. Syst. **16**(3), 199–214 (2001)
43. Zhu, W.: Generalized rough sets based on relations. Inf. Sci. **177**(22), 4997–5011 (2007)
44. Zhu, W.: Topological approaches to covering rough sets. Inf. Sci. **177**(6), 1499–1508 (2007)
45. Zhu, W.: Relationship between generalized rough sets based on binary relation and covering. Inf. Sci. **179**(3), 210–225 (2009)
46. Zhu, W., Wang, F.: Reduction and axiomization of covering generalized rough sets. Inf. Sci. **152**(1), 217–230 (2003)

http://www.springer.com/978-3-662-44679-9

Transactions on Rough Sets XVIII
Peters, J.F.; Skowron, A.; Li, T.; Yang, Y.; Yao, J.; Nguyen, H.S. (Eds.)
2014, X, 131 p. 39 illus., 13 illus. in color., Softcover
ISBN: 978-3-662-44679-9