

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 530483, 9 pages
<http://dx.doi.org/10.1155/2014/530483>



Research Article

A New Approach for Resolving Conflicts in Actionable Behavioral Rules

Peng Su,¹ Dan Zhu,² and Daniel Zeng³

¹ School of Mathematics and Computer Science, Dali University, Dali 671003, China

² College of Business, Iowa State University, Ames, IA 50011, USA

³ State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

Correspondence should be addressed to Peng Su; peng.su.casia@gmail.com

Received 7 April 2014; Accepted 29 April 2014; Published 5 August 2014

Academic Editor: Zheng Xu

Copyright © 2014 Peng Su et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Knowledge is considered actionable if users can take direct actions based on such knowledge to their advantage. Among the most important and distinctive actionable knowledge are actionable behavioral rules that can directly and explicitly suggest specific actions to take to influence (restrain or encourage) the behavior in the users' best interest. However, in mining such rules, it often occurs that different rules may suggest the same actions with different expected utilities, which we call conflicting rules. To resolve the conflicts, a previous valid method was proposed. However, inconsistency of the measure for rule evaluating may hinder its performance. To overcome this problem, we develop a new method that utilizes rule ranking procedure as the basis for selecting the rule with the highest utility prediction accuracy. More specifically, we propose an integrative measure, which combines the measures of the support and antecedent length, to evaluate the utility prediction accuracies of conflicting rules. We also introduce a tunable weight parameter to allow the flexibility of integration. We conduct several experiments to test our proposed approach and evaluate the sensitivity of the weight parameter. Empirical results indicate that our approach outperforms those from previous research.

1. Introduction

Data mining is the process of discovering patterns in large data sets and then processing these obtained patterns into understandable form for further use [1, 2]. The primary motive behind data mining process is to provide support in decision making process by detecting useful or actionable patterns from a large volume of data [3]. Extensive research has been conducted in data mining to discover hidden patterns from underlying data. However, the knowledge obtained through these data mining techniques is not always useful to the user, and it generally requires significant amount of expertise in postprocessing.

Tradition data mining studies concentrated primarily on predictive mining, where the cause and effect scenario is being described. But this information alone is not sufficient as it does not provide much benefit to the user. What becomes even more interesting and critical to business or

organizations is to make the mined patterns or knowledge actionable [4]. Knowledge is considered actionable if users can take direct actions based on such knowledge to their advantage. Actionability is a very important criterion measuring the interestingness of mined patterns. Among the most important and distinctive actionable knowledge are actionable behavioral rules that can directly and explicitly suggest specific actions to take to influence (restrain or encourage) the behavior in the users' best interest [5]. Mining such rules is also an active research area of social computing [6].

In mining such rules, all possible action combinations are considered in turn as a rule's condition. Therefore, candidate actionable behavioral rules may share actions in their conditions. This may lead to several conflicting rules with the same actions, yet different consequences and expected utilities. A rule's expected utility is the utility prediction when the actions suggested by the rule are actually taken. Hence, conflicting

rules give different utility predictions for the same actions. Naturally, the rule with the most accurate utility prediction should be reserved for the user and the others should be pruned. Therefore, the key problem is how to accurately evaluate the utility prediction accuracies of conflicting rules.

To resolve conflicting rules, a pruning method was introduced in [5]. Firstly, each general rule whose antecedent is part of another rule is pruned. Subsequently, the remaining rules sharing the same actions are consolidated into a new rule according to their supports. A new rule's expected utility is in fact the weighted average of the remaining rules' expected utilities with their supports as the weights. A major disadvantage of this method lies in the fact that the measure for pruning, namely, antecedent length, is not used consistently. That is to say, antecedent length is only used for some of the conflicting rules.

In this paper, to overcome the shortcomings of the previous method proposed in [5], a new rule ranking method is developed to resolve the conflicts in actionable behavioral rules. More specifically, an integrative measure, which linearly combines the measures of the support and antecedent length, is proposed to evaluate the utility prediction accuracies of conflicting rules. Furthermore, a tunable weight parameter will be introduced in the proposed measure to allow the flexibility of integration. We conduct an experiment to validate our proposed method and evaluate the performance of the proposed method against the previous method and random selection. The experimental results strongly suggest the validity and the superiority of our approach.

Related research on rule pruning in associative classification (AC) inspired our work. Most associative classification algorithms favor rules with large support and confidence values. Another important measure used to determine the precedence of the rules is rule antecedent length. Some AC algorithms, such as those in [7, 8], tend to prefer general rules (those with shorter antecedents) and consequently suffer from poor classification accuracy. On the contrary, other algorithms, such as those in [9, 10], tend to prefer specific rules, reducing the chance of misclassification. However, these methods may lead to more specific rules (i.e., those with longer antecedents). Thabtah argued that ranking of rules should not be limited to support and confidence parameters and proposed a rule ranking procedure that takes into account the class distribution frequency of each rule after considering confidence, support, and rule antecedent length [11].

Mining big data for complex decision making has been a major challenge for next generation data mining. The resolution of the conflicting rule may have significant implications. Actionable knowledge discovery (AKD) not only provides an important tool to decision makers to take appropriate actions but also delivers reliable and actionable outcomes to homeland security and to businesses, to name a few. The rest of the paper is organized as follows. Section 2 introduces the formal definition of actionable behavioral rules mining. Section 3 uses an example to illustrate the problem. Section 4 describes our proposed method. Section 5 presents our experimental study demonstrating the validity and the superiority of the proposed method. Section 6 reviews some of the related

works. Finally, Section 7 summarizes our contributions and the directions for future research.

2. Problem Definitions and Elaborations

To understand the domain in actionable behavioral rule mining, we first introduce some formal definitions listed in Table 1.

Within a behavioral information system, A_{be} is a set of behavior attributes describing the behaviors of the entity, and A_{en} is a set of environment attributes characterizing the environment in which the entity situates and having some causal influence upon behavior attributes. An action set S , also called a $|S|$ -action set, is defined as a finite nonempty set of actions such that $t_1 \cdot a \neq t_2 \cdot a$ for any $t_1, t_2 \in S$. We say that action set S holds, if every $t \in S$ holds. We say that observation o supports S , if $\rho(o, t \cdot a) = t \cdot v_t$ for every $t \in S$. The support of S is defined as

$$\text{sup}(S) = |\{o \in O \mid o \text{ supports } S\}|. \quad (1)$$

We call S a frequent action set, or frequent $|S|$ -action set, with regard to a user-specified threshold referred to as minsup , if $\text{sup}(S) \geq \text{minsup}$.

In effect, if the value of a changes from v_f to v_t , $e = (a, v_f, v_t)$ takes place. The observation o supports (S, e) , where S is an action set and e is an effect, if o supports S and $\rho(o, e \cdot a) = e \cdot v_t$. An effect-probability is defined as a pair $ep = (e, p)$, where e is an effect and $p \in [0, 1]$. We say that an effect-probability $ep = (e, p)$ takes place, if e takes place with probability p .

We call an actionable behavioral rule r an interesting actionable behavioral rule with regard to a user-specified threshold minutil , if $\text{util}(r) \geq \text{minutil}$.

Please note in the definition, except for the value of o^* , that the rest of the observation pertaining to environment or behavior attributes comes from a time period of a certain interval. Based on the recent observation, the forthcoming projection is predicted. Based on the value of o^* , the user decides whether or not to take a particular action and how it is going to influence the entity's behavior. Consider a medical scenario where a patient is suffering from a rare disease and the right way to treat this disease is not currently available to the doctors. It is very likely that in a period of time (e.g., a year) the condition of the patient is going to remain unchanged or worsen if proper medication and treatment are not provided. This signifies that a proper action should be taken within a specified time to get a better output. The objective of actionable behavior rule mining is to identify the useful actions that may be applied to improve the projected next observation o^* . In the above example the doctors would like to know the actions that can be taken to improve the patients' health condition.

In this paper, we assume that all the attributes (both environmental and behavioral) are categorical, while numerical attributes have been discretized in advance. The behavior attribute values are not restricted to be binary indicating whether a certain behavior occurs or not. The values can also describe how frequent the behavior occurs, the extent of the

TABLE 1: Concepts and definitions.

Concept	Definition	Explanations
Behavioral information system (BIS)	$I = (O, o^*, A, D, \rho)$	O : a finite nonempty set of observations $o^* \in O$: the projected next observation A : a finite nonempty set of attributes $A = A_{en} \cup A_{be}$ $D = \bigcup_{a \in A} D_a$ (D_a : value domain of attribute a) $\rho: O \times A \rightarrow D$: a function associating each observation with a set of attribute values
Action	$t = (a, v_f, v_t)$	$a \in A_{en}$ $v_f = \rho(o^*, a)$ $v_t \in D_a$ If $v_f \neq v_t$, t = standard action; else, t = nonstandard action. If the value of a is changed from v_f to v_t , $t = (a, v_f, v_t)$ holds.
Effect	$e = (a, v_f, v_t)$	$a \in A_{be}$ $v_f = \rho(o^*, a)$ $v_t \in D_a$
Actionable behavioral rule (ABR)	$r = (S, C)$	S : an action set C : a finite nonempty set of effect-probabilities $ C = \sum_{a \in A_{be}} D_a $ $r = (S, C)$: if S holds, each effect-probability in C will take place
Expected utility (EU)	$util(r) = \sum_{t \in S} util(t) + \sum_{ep \in C} util(ep \cdot e) \cdot ep \cdot p$	$util(t)$: utilities of action t $util(ep \cdot e)$: utilities of effect $ep \cdot e$

behavior, and so on. In the next section, we use an example to illustrate some of these concepts and the motivations. For detailed information, please refer to [5] for more explanations and examples.

3. An Illustrative Example

Consider an example of a Palestinian terrorist organization Hamas, an Arabic acronym for Harakat al-Muqawma al-Islamiyya. It has carried out sophisticated attacks on Israeli targets and has caused serious troubles for the Israeli government. The Israeli government wants to take some action that will either stop or improve the current situation. For instance, the following prediction rule has been proposed to the government against Hamas:

$$\left(\{(e_1, 1, 2), (e_2, 1, 3)\}, \left\{ \left((b_1, 3, 2), \frac{1}{9} \right), \left((b_1, 3, 1), \frac{3}{9} \right), \left((b_1, 3, 3), \frac{5}{9} \right) \right\}, 0.84 \right). \quad (2)$$

The rule says “if the Israeli government changes the degree of using lethal violence against Hamas from level 1 (not using lethal violence) to level 2 (using periodic lethal violence) and the degree of being in agreement with Hamas from level 1 (negotiation) to level 3 (major concession), the degree of terrorist attacks aiming at domestic targets launched by Hamas will change from level 3 to 2 with 1/9 confidence or 1 with 3/9 confidence or remain unchanged with 5/9 confidence, and the Israeli government will get a utility of 0.84.”

TABLE 2: Function ρ in behavioral information system.

	e_1	e_2	b_1	b_2
o^*	1	1	2	2
o_1	1	1	2	2
o_2	1	0	2	1
o_3	1	0	1	2
o_4	1	0	1	2
o_5	0	1	1	1
o_6	0	1	2	1
o_7	0	1	2	1
o_8	0	0	0	0
o_9	0	0	0	1
o_{10}	0	0	1	0

Consider a hypothetical behavioral information system for Hamas organization $I = (O, o^*, A, D, \rho)$, where $O = \{o^*, o_1, o_2, \dots, o_{10}\}$, $A_{en} = \{e_1, e_2\}$, $A_{be} = \{b_1, b_2\}$, $D_{e_1} = D_{e_2} = \{0, 1\}$, $D_{b_1} = D_{b_2} = \{0, 1, 2\}$, and ρ is presented in Table 2 (e.g., $\rho(o_1, e_1) = 1, \rho(o_3, b_2) = 2$). The values and the corresponding meanings of the attributes are shown in Table 3.

The problem of mining actionable behavioral rules is to mine all reliable and interesting actionable behavioral rules from a behavioral information system. The threshold minsup is used to assure that the rules are not found by chance. The threshold minutil is used to assure that the rules are sufficiently beneficiary to warrant deliberation by the user.

These actionable rules are important to the user in making decisions. However, rules identified sometimes tend

TABLE 3: Attributes and meanings.

Code	Meaning	Values	Label
e_1	Does the state use violence against the organization?	1	Not using lethal violence
		2	Using periodic lethal violence
		3	Using consistent lethal violence against the organization
e_2	To what degree has the organization succeeded in obtaining government agreement?	1	Negotiation
		2	Some concession
		3	Major concession
b_1	To what degree does the organization use violence domestically as a strategy?	1	Not using violence as a strategy
		2	Using violence as occasional strategy
		3	Using violence regularly as a strategy
b_2	To what degree does the organization use violence to target transnational entities as a strategy?	1	Not using violence as a strategy
		2	Using violence as occasional strategy
		3	Using violence regularly as a strategy

to be conflicting with each other. For example, the following four rules are conflicting, that is,

$$\begin{aligned}
 r_1: & \{ \{(e_1, 1, 0)\}, \{((b_1, 2, 0), 1/2), ((b_1, 2, 1), 1/2), ((b_1, 2, 2), 0)\}, 0.10), \\
 r_2: & \{ \{(e_1, 1, 0), (e_2, 1, 1)\}, \{((b_1, 2, 0), 0), ((b_1, 2, 1), 1/3), ((b_1, 2, 2), 2/3)\}, 0.13), \\
 r_3: & \{ \{(e_1, 1, 0), (e_2, 1, 1), (e_3, 0, 0)\}, \{((b_1, 2, 0), 2/3), ((b_1, 2, 1), 1/3), ((b_1, 2, 2), 0)\}, 0.21), \\
 r_4: & \{ \{(e_1, 1, 0), (e_4, 0, 0)\}, \{((b_1, 2, 0), 1/3), ((b_1, 2, 1), 1/3), ((b_1, 2, 2), 1/3)\}, 0.17).
 \end{aligned}$$

These rules are conflicting as they suggest the same standard-action $(e_1, 1, 0)$ with different expected utilities, 0.10, 0.13, 0.21, and 0.17. This means that if the user takes the action $(e_1, 1, 0)$, we cannot provide him/her with a definite utility estimation. Note that $(e_2, 1, 1)$, $(e_3, 0, 0)$, and $(e_4, 0, 0)$ are nonstandard actions and they cost the government nothing.

Earlier research has dealt with the issues involving conflicting rules. For example, following with the previous example, r_1 and r_2 were pruned first, as they are general rules with regard to r_3 . Subsequently, r_3 and r_4 were consolidated into the new rule $\{ \{(e_1, 1, 0)\}, \{((b_1, 2, 0), 1/2), ((b_1, 2, 1), 1/3), ((b_1, 2, 2), 1/6)\}, 0.19)$ whose expected utility is the weighted average of the expected utilities of r_1 and r_2 according to their supports. r_3 or r_4 is not pruned as neither is a general rule with regard to the other one.

Previous method prefers the rule with longer rule antecedent length while pruning the general rules. However, the same measure is not used for the rules which do not have general-specific relationship. The nonconsistent use of antecedent length may hurt the performance of the previous method. In addition, there may be another rule selection measure than rule antecedent length, that is, rule support.

4. Proposed Approach

The approach we propose to resolve conflicting rules is ranking rules according to their utility prediction accuracies.

The highest-order rule should be reserved for the user and the others should be pruned. The key strategy is to find appropriate measure for evaluating conflicting rules.

4.1. Candidate Measures for Evaluating Conflicting Rules.

There are two main measures in terms of which we can evaluate the utility prediction accuracies of conflicting rules: support and rule antecedent length (i.e., the number of actions in its antecedent). When all other conditions are the same, the higher is a rule's support, the more accurate is its utility prediction. This is due to the fact that higher support can reduce a rule's contingency. When all other conditions are the same, the longer is a rule's antecedent, the more accurate is its utility prediction. The main reason for giving a rule with longer antecedent length more accurate utility prediction than a rule with shorter antecedent length is that the former provides more evidence to make predictions.

However, it is not uncommon that these two measures are conflicting or mutually exclusive. That is to say, it is very likely that a rule may have a longer or shorter antecedent length and lower or higher support than another rule at the same time. Moreover, the precedence of these two parameters is unknown. Finally, the precedence of these two measures will be different for different datasets.

4.2. *The Rule Ranking Strategy.* Due to the conflict between the two measures, we linearly combine them using a tunable weight parameter. In particular, we proposed an integrative measure, score, which combines support and antecedent length to evaluate the utility prediction accuracies of the conflicting rules. This is presented as follows:

$$\text{score} = \alpha \cdot \frac{\sup(r \cdot S)}{(|O| - 1)} + (1 - \alpha) \cdot \frac{|r \cdot S|}{|D|}, \quad 0 \leq \alpha \leq 1, \quad (3)$$

where $r \cdot S$: antecedents of rule r , O : set of behavioral observations, D : set of conditional attributes, and α : a weight parameter tunable from 0 to 1.

When $\alpha = 0$, the measure depends ultimately on rule antecedent length. When $\alpha = 1$, the measure depends

Input: behavioral information system $I = (O, o^*, A, D, \rho)$, actionable behavioral rule $r = (S, C)$.

Output: *general* of r

```

(1) general  $\leftarrow 0$ 
(2) for each  $o \in O$ 
(3)   for each  $ep \in C$ 
(4)     if  $\rho(o, ep \cdot e \cdot a) = ep \cdot e \cdot v_i$ 
(5)        $ep \cdot \text{sup} \leftarrow ep \cdot \text{sup} + 1$ 
(6)   for each  $ep \in C$ 
(7)      $\text{general} \leftarrow \text{general} + |ep \cdot p - ep \cdot \text{sup}| / |O|$ 
(8) return general
    
```

ALGORITHM 1

ultimately on support. When $0 < \alpha < 1$, the measure depends on the combination of these two parameters. The larger α is, the more support contributes to the measure. The smaller α is, the more rule antecedent length contributes to the measure.

When several rules have the identical highest score, we can choose one of the rules randomly, which possibly in some cases may degrade performance. When the possibility of two or more rules with the same score is relatively high, there should be other parameters to consider in favoring one rule over another in order to limit rule random selection. Thus, we propose a rule ranking procedure that takes into account the class distribution frequency of each conflicting rule after considering score.

This rule ranking procedure adds upon previous rule ranking approaches by looking at the class distribution frequencies in the behavioral information systems and prefers rules whose effects are associated with labels that occur more frequently in the behavioral information systems. The computing procedure of this measure (*general*) is shown in Algorithm 1.

5. Experiments and Results

In this section, we empirically validate the proposed method for resolving conflicting rules. We aim to answer the following five questions. (1) Is the method valid? (2) For which values of α does the proposed method outperform the previous method? (3) Does the score-distribution method outperform the score method? (4) For which values of α does the proposed method outperform random selection? (5) When α increases, how will the performance change?

5.1. Experimental Design. To test the validity of our method, we conduct an experiment with the benchmark MAROB datasets (The Minorities at Risk Organizational Behavior Datasets: <http://www.cidcm.umd.edu/mar/>). The MAROB datasets cover several ethnopolitical organizations in the Middle East and North Africa. The datasets keep track of several attributes on a yearly basis from 1980 to 2004. We extracted three subdatasets about the Hezbollah Organization in Lebanon, the Kurdistan Democratic Party of Iran, and the Iraqi Communist Party, respectively. Utility values of the possible actions and effects from the viewpoints of

the corresponding governments were elicited from human analysts and normalized into the range of $[-1, 1]$.

Parameter α is set to each value in $\{0, 0.1, 0.2, \dots, 1\}$ which is believed to be a good representative of the set of permissible values of α ($[0, 1]$).

5.2. Evaluation Criterion. We use the mean absolute error (MAE), a standard measure to assess the closeness between predictions and eventual outcomes, as the criterion to evaluate the performance of different methods. The MAE is given by

$$\frac{1}{n} \sum_{i=1}^n |eu_i - au_i|, \tag{4}$$

where eu_i is the expected utility our method or another method estimates and au_i is the actual utility. Typically, the MAE threshold is set by the domain experts in security informatics to a reasonable value around 0.07 for field evaluation.

5.3. Experimental Results. Tables 4 and 5 show the experimental results on the three sub-MAROB datasets with 25 actual action sets that are antecedent of some candidate actionable behavioral rules. Note that there is no other actual action set that is antecedent of any candidate rule. The table entries present the actual utilities, the absolute errors of the proposed score method (Table 4), the proposed score-distribution method (Table 5), the previous method, and random selection with *minsup* set to 7. For the absolute errors, the means are presented.

Note that although the utilities of actionable behavioral rules-induced actions and effects were assigned by the experts in a subjective fashion, what is actually compared in our experiment is the absolute difference between the actual utility and the expected utility of the action set induced by a rule. For the same action set, different methods will yield different estimated probability distributions over the effects or outcomes. Generally speaking, the closer the estimated distribution of a rule's effects is to the actual realization, the closer the estimated utility is to the actual utility.

From Tables 4 and 5, we can see that the MAE of our approach is much smaller than the validity threshold

TABLE 4: Comparison of the results by score method and other methods.

	Actual utility	Absolute error										
		The previous method	Random selection	The proposed score method								
		$\alpha = 0$	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$	$\alpha = 1$
1	-0.1200	0.0500	0.0489	0.0400	0.0926	0.0926	0.0926	0.0926	0.0926	0.0880	0.0880	0.0880
2	-0.1200	0.0471	0.0457	0.0400	0.0926	0.0926	0.0926	0.0926	0.0926	0.0880	0.0880	0.0880
3	-0.0300	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933
4	-0.0400	0.1244	0.1244	0.1244	0.1244	0.1244	0.1244	0.1244	0.1244	0.1120	0.1120	0.1120
5	0.0638	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6	-0.0400	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0356	0.0356	0.0240
7	-0.1620	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
8	-0.06200	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533
9	-0.0300	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
10	0.1538	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467
11	-0.1720	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0533	0.0533	0.0640
12	-0.1620	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933
13	-0.0400	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0533	0.0533	0.0640
14	-0.8560	0.0000	0.0000	0.0000	0.0000	0.0000	0.0275	0.0275	0.0275	0.0259	0.0289	0.0289
15	-0.0560	0.0000	0.0000	0.0000	0.0000	0.0000	0.0314	0.0314	0.0259	0.0259	0.0289	0.0289
16	0.0768	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0300	0.0300
17	-0.8440	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0057	0.0057	0.0057	0.0057
18	0.0768	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0343	0.0300	0.0300
19	0.0712	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0171	0.0171
20	-0.0440	0.0933	0.0933	0.0914	0.0914	0.0914	0.0800	0.0800	0.0800	0.0800	0.0800	0.0800
21	-0.1000	0.0462	0.0462	0.0457	0.0457	0.0457	0.0440	0.0440	0.0440	0.0440	0.0440	0.0440
22	0.1168	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0300	0.0300
23	-0.0600	0.0629	0.0629	0.0629	0.0629	0.0629	0.1057	0.1057	0.1057	0.0933	0.0933	0.0933
24	-0.6660	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0086	0.0086	0.0086
25	0.1940	0.1371	0.1371	0.1371	0.1371	0.1371	0.1371	0.1057	0.1057	0.0933	0.0933	0.0933
26	-0.0760	0.0790	0.0790	0.0790	0.0790	0.0790	0.0854	0.0854	0.0854	0.0950	0.0950	0.0950
Mean		0.0465	0.0464	0.0457	0.0498	0.0520	0.0532	0.0534	0.0534	0.0531	0.0543	0.0543

TABLE 5: Comparison of the results by score-distribution method and other methods.

Actual utility	Absolute error														
	The previous method			Random selection			The proposed score-distribution method								
	$\alpha = 0$	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$	$\alpha = 1$				
1	-0.1200	0.0500	0.0651	0.0400	0.0400	0.0926	0.0926	0.0926	0.0926	0.0926	0.0926	0.0880	0.0880	0.0880	0.0880
2	-0.1200	0.0471	0.0619	0.0400	0.0400	0.0926	0.0926	0.0926	0.0926	0.0926	0.0926	0.0880	0.0880	0.0880	0.0880
3	-0.0300	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933
4	-0.0400	0.1244	0.1182	0.1244	0.1244	0.1244	0.1244	0.1244	0.1244	0.1244	0.1244	0.1120	0.1120	0.1120	0.1120
5	0.0638	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6	-0.0400	0.0500	0.0412	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0356	0.0356	0.0240	0.0240
7	-0.1620	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
8	-0.06200	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533	0.0533
9	-0.0300	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
10	0.1538	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467	0.1467
11	-0.1720	0.0267	0.0386	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0533	0.0533	0.0640	0.0640
12	-0.1620	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933	0.0933
13	-0.0400	0.0267	0.0386	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0533	0.0533	0.0640	0.0640
14	-0.08560	0.0000	0.0134	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0275	0.0275	0.0289	0.0289
15	-0.0560	0.0000	0.0144	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0314	0.0314	0.0289	0.0289
16	0.0768	0.0200	0.0234	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0300	0.0300
17	-0.08440	0.0000	0.0016	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0057	0.0057	0.0057	0.0057
18	0.0768	0.0200	0.0273	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0343	0.0343	0.0300	0.0300
19	0.0712	0.0200	0.0198	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0171	0.0171
20	-0.0440	0.0933	0.0873	0.0933	0.0914	0.0914	0.0914	0.0914	0.0914	0.0800	0.0800	0.0800	0.0800	0.0800	0.0800
21	-0.1000	0.0462	0.0455	0.0467	0.0457	0.0457	0.0457	0.0457	0.0457	0.0440	0.0440	0.0440	0.0440	0.0440	0.0440
22	0.1168	0.0171	0.0207	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0300	0.0300
23	-0.0600	0.0629	0.0809	0.0629	0.0629	0.0629	0.0629	0.0629	0.0629	0.1057	0.1057	0.1057	0.1057	0.0933	0.0933
24	-0.06660	0.0006	0.0026	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0086	0.0086
25	0.1940	0.1371	0.1252	0.1371	0.1371	0.1371	0.1371	0.1371	0.1371	0.1057	0.1057	0.1057	0.1057	0.0933	0.0933
26	-0.0760	0.0790	0.0773	0.0790	0.0790	0.0790	0.0790	0.0790	0.0790	0.0854	0.0854	0.0854	0.0854	0.0950	0.0950
Mean		0.0465	0.0496	0.0458	0.0457	0.0498	0.0520	0.0532	0.0534	0.0532	0.0532	0.0534	0.0531	0.0543	0.0543

0.07. This shows the validity of the proposed method. Furthermore, the proposed method outperforms the previous approach and random selection, when α is set to 0 or 0.1. In particular, the proposed method reaches the best performance when α is set to 0.1. The score-distribution method outperforms the score method when α is set to 0. The two methods yield the same results when α is set to the other values, as for these α values only one rule has the identical highest score. In general, MAE increases monotonically when α becomes larger. This suggests that, in the proposed integrative measure, rule antecedent length has much effect than support. The main reason may be that the small number of instances (fewer than 30) in each dataset leads to the small supports of the rules, and hence the differences among different supports are trivial.

6. Related Work

Previous research has looked into actionable knowledge discovery. For example, in discovering actionable knowledge for customer relationship management (CRM), Liu et al. proposed an approach to suggest actions to reclassify a customer from an undesired status to a desired one while postprocessing decision trees to maximum expected net profit [12]. However, these methods could miss some actions with higher net profit. To handle this problem, multiple trees with different subsets of “hard” attributes need to be built. To get optimal actions, the number of trees could be very large, especially when there are many “hard” attributes.

In order to increase profit of an institute to help devise a direct-marketing plan, Yang et al. proposed a lazy approach to use “role models” to generate advice and plans [13]. However, this method failed to provide rules in advance and will incur high computational cost when generating action recommendations. To improve the profitability of a bank, Schrodtr proposed action rules constructed from certain pairs of classification rules [14].

Ras and Tzacheva defined interesting action rules as the rules of the lowest cost [15]. Raś et al. proposed a heuristic strategy to construct interesting action rules [16]. Wang et al. combined the action forest algorithm to extract action rules and proposed a heuristic strategy to generate interesting action rules [8]. Tzacheva and Tsay introduced the notion of cost and feasibility of an action rule and proposed a graph based method to search and construct feasible action rules at the lowest cost [17].

Previous studies on mining action rules all produced actionable rules based on a certain pair of classification rules or a single classification rule. A main drawback of these approaches is that some interesting actionable rules can be missed. To address this problem, He et al. proposed another strategy in a support-confidence-cost framework to discover action rules directly from a database [18]. Raś et al. proposed an approach to generate association-type action rules [19]. Subsequently, Raś and Dardzińska proposed a bottom-up strategy to discover action rules without using preexisting classification rules [20].

7. Conclusions

Actionable behavioral rules have wide applicability. Finding such rules and applying them are of great importance in the field of data mining [21–27].

In this paper, to resolve conflicts in actionable behavioral rules, we propose a new ranking method using an integrative measure for evaluating the utility prediction accuracies of conflicting rules. The new measure combines the support and antecedent length using a tunable weight parameter. The experimental results strongly suggest the validity and the superiority of our approach.

While we have conducted a preliminary experiment using domain datasets, more comprehensive experiments with many large datasets drawn from various domains can be conducted to validate the generalizability of our findings. For different datasets, the proposed method can tune the weight parameter for the best performance.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

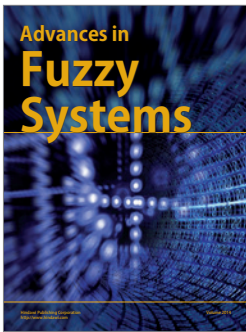
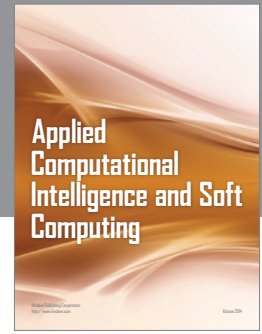
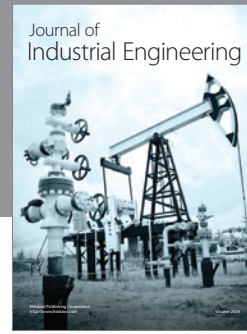
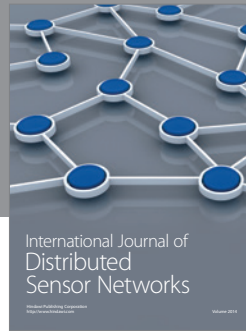
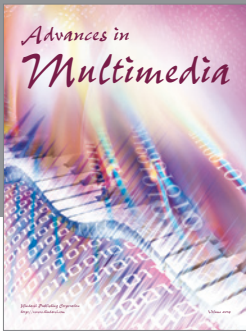
Acknowledgments

This work has been supported by NSFC (nos. 71025001, 91024030, and 70890084) and the Scientific Research Foundation of Yunnan Provincial Education Department (no. 2014Z137), and Doctoral Science Foundation of Dali University.

References

- [1] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, 1996.
- [2] D. Hand, H. Mannila, and P. Smyth, *Principles of Data Mining*, MIT Press, Cambridge, Mass, USA, 2001.
- [3] Silberschatz and A. Tuzhilin, “What makes patterns interesting in knowledge discovery systems,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 8, no. 6, pp. 970–974, 1996.
- [4] L. Cao, Y. Zhao, H. Zhang, D. Luo, C. Zhang, and E. K. Park, “Flexible frameworks for actionable knowledge discovery,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 9, pp. 1299–1312, 2010.
- [5] P. Su, W. Mao, D. Zeng, and H. Zhao, “Mining actionable behavioral rules,” *Decision Support Systems*, vol. 54, pp. 142–152, 2012.
- [6] D. Zeng, F.-Y. Wang, and K. M. Carley, “Social computing,” *IEEE Intelligent Systems*, vol. 22, no. 5, pp. 20–22, 2007.
- [7] W. Li, J. Han, and J. Pei, “CMAR: accurate and efficient classification based on multiple class-association rules,” in *Proceedings of the 1st IEEE International Conference on Data Mining (ICDM '01)*, pp. 369–376, December 2001.
- [8] K. Wang, S. Zhou, and Y. He, “Growing decision trees on support-less association rules,” in *Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '01)*, pp. 265–269, August 2000.

- [9] E. Baralis and P. Garza, "A lazy approach to pruning classification rules," in *Proceedings of the 2nd IEEE International Conference on Data Mining (ICDM '02)*, pp. 35–42, December 2002.
- [10] E. Baralis, S. Chiusano, and P. Garza, "On support thresholds in associative classification," in *Proceedings of the ACM Symposium on Applied Computing*, pp. 553–558, March 2004.
- [11] F. Thabtah, "Pruning techniques in associative classification: survey and comparison," *Journal of Digital Information Management*, vol. 4, no. 3, pp. 197–202, 2006.
- [12] B. Liu, W. Hsu, and S. Chen, "Using general impressions to analyze discovered classification rules," in *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining (KDD '97)*, pp. 31–36, 1997.
- [13] Q. Yang, J. Yin, C. X. Ling, and T. Chen, "Postprocessing decision trees to extract actionable knowledge," in *Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM '03)*, pp. 685–688, November 2003.
- [14] P. Schrodt, "Forecasting conflict in the Balkans using hidden Markov models," in *Programming for Peace*, T. Robert, Ed., pp. 161–184, Springer, 2006.
- [15] Z. Ras and A. Tzacheva, "In search for action rules of the lowest cost," in *Monitoring, Security, and Rescue Techniques in Multiagent Systems*, B. Dunin-Keplicz, A. Jankowski, A. Skowron, and M. Szczuka, Eds., pp. 261–272, Springer, 2005.
- [16] Z. W. Raś, L.-S. Tsay, A. A. Tzacheva, and O. Gürdal, "Mining for interesting action rules," in *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology (LAT '05)*, pp. 187–193, September 2005.
- [17] A. A. Tzacheva and L.-S. Tsay, "Tree-based construction of low-cost action rules," *Fundamenta Informaticae*, vol. 86, no. 1-2, pp. 213–225, 2008.
- [18] Z. He, X. Xu, S. Deng, and R. Ma, "Mining action rules from scratch," *Expert Systems with Applications*, vol. 29, no. 3, pp. 691–699, 2005.
- [19] Z. W. Raś, L.-S. Tsay, A. Dardzińska, and H. Wasyluk, "Association action rules," in *Proceedings of the IEEE International Conference on Data Mining Workshops*, pp. 283–290, December 2008.
- [20] Z. W. Raś and A. Dardzińska, "Action rules discovery without pre-existing classification rules," in *Rough Sets and Current Trends in Computing*, vol. 5306 of *Lecture Notes in Computer Science*, pp. 181–190, Springer, 2008.
- [21] X. Luo, Z. Xu, J. Yu, and X. Chen, "Building association link network for semantic link on web resources," *IEEE Transactions on Automation Science and Engineering*, vol. 8, no. 3, pp. 482–494, 2011.
- [22] Z. Xu, X. Luo, and L. Wang, "Incremental building association link network," *Computer Systems Science and Engineering*, vol. 26, no. 3, pp. 153–162, 2011.
- [23] Y. Liu, Y. Zhu, L. Ni, and G. Xue, "A reliability-oriented transmission service in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 12, pp. 2100–2107, 2011.
- [24] Y. Liu, Q. Zhang, and L. Ni, "Opportunity-based topology control in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 3, pp. 405–416, 2010.
- [25] C. Hu, Z. Xu, Y. Liu, L. Mei, L. Chen, and X. Luo, "Semantic link network based model for organizing multimedia big data," *IEEE Transactions on Emerging Topics in Computing*, 2014.
- [26] X. Liu, Y. Yang, D. Yuan, and J. Chen, "Do we need to handle every temporal violation in scientific workflow systems," *ACM Transactions on Software Engineering and Methodology*, vol. 23, no. 1, pp. 5:1–5:34, 2013.
- [27] L. Wang, J. Tao, R. Ranjan et al., "G-Hadoop: MapReduce across distributed data centers for data-intensive computing," *Future Generation Computer Systems*, vol. 29, no. 3, pp. 739–750, 2013.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

