# Enhancing Rule Importance Measure Using Concept Hierarchy

Jiye Li, Nick Cercone, Serene W. H. Wong, and Lisa Jing Yan

Faculty of Computer Science and Engineering, York University
4700 Keele Street, Toronto, Ontario, Canada M3J 1P3
{jiye, swong, jingyan}@cse.yorku.ca, ncercone@yorku.ca

**Abstract.** A rule importance measure is used to evaluate how important are the rules which characterize a data set. This measure was designed based on association rules and it has been proven to be effective to enumerate the most important rules of all rules generated. However, since rule importance is an objective measure, its usage as a rule interestingness measure relies on the interpretation of domain experts. We propose to enhance the rule importance measure previously used by incorporating a weight biased attribute concept hierarchy. The new measure better reflects the importance of a rule by integrating with the domain knowledge. A geriatric care data set is used as our experimental data set. We show that this enhanced rule importance measure provides a knowledge oriented distinction of rules classified as important.

**Key words:** Association Rules, Rule Interestingness, Rule Importance Measure, Concept Hierarchy, Rough Sets

## 1 Introduction

Association rule algorithms are well known for discovering item-item associations among the transaction data set, and have been widely used in fields such as business data analysis, transaction management, and medical research. One of the challenging problems for association rule algorithms is that, given the characteristics of the application data set, there are usually enormous number of rules generated by the algorithms. How can one interpret and identify interesting rules among all those generated? One solution to this problem includes using interestingness measures [12] to evaluate and rank the generated rules. For example, given a grocery transaction data set, rules such as "80% of male customers who bought beer also bought diaper" may have a higher interestingness measure than "80% customers bought bread and milk together".

To evaluate the interestingness of the association rules, both subjective measures and objective measures are commonly used [4]. Subjective measures rely on the human (usually the domain experts) effort to evaluate rules manually. This approach is more accurate, though it is also more expensive and time-consuming to involve the domain experts for evaluation. The objective measures

P. Lenca and S. Lallich (Eds.): QIMIE/PAKDD 2009

include measures from statistics, machine learning and information theory fields, and can automate the evaluation process without the involvement of domain experts. Objective measures alone are not sufficient to provide solid evaluations because the data domain knowledge is not taken into consideration for rule evaluation. Therefore the optimal solution would be to integrate both the subjective and the objective measures together into the rule evaluations. A Rule Importance Measure (RIM) [7] was designed as an objective rule measure similar to the interestingness measures to evaluate how important the rules are. This measure is designed based on rough sets theory and association rules, and is illustrated as follows. ROSETTA [9] rough sets software was first used to generate multiple reducts. Apriori [1] association rule algorithm was then applied to generate rule sets for each data set based on each reduct. Some rules were generated more frequently than the others among the total rule sets. Such rules were considered as more important. The rule importance was defined as the occurrence of an association rule across all the rule sets. Experimental results show the RIM reduces the number of rules generated and at the same time provides a diverse measure of how important a rule is.

In this paper, we propose an enhanced measure for the rule importance measure using concept hierarchy. The motivation of this research is to design a rule measure that integrates the domain experts' opinions into the objective evaluations. Given a data set, we first develop a concept hierarchy based on its domain, and then weights are assigned to the attributes according to their corresponding hierarchy. The Rule Importance Measure generates rules measures with its importance. Then from the RIM rules, rules with higher weighted attributes and higher occurrence are considered as more important. We name this enhanced rule evaluation approach ERIM (Enhanced Rule Importance Measure). This weight biased rule measure integrates the domain knowledge together into the rule evaluation, therefore a knowledge oriented distinction of rules are suggested. Note that the rules evaluated by ERIM are to be used for classification or prediction purpose. The rules we are interested to evaluate all contain the decision attributes on the right hand sides of the rules, and the condition attributes on the left hand sides of the rules.

The contributions of our work are summarized as follows. We propose a novel rule evaluation approach based on concept hierarchy which integrates both the subjective measures and the objective measures; the proposed ERIM provides a knowledge oriented distinction of rules demonstrated by our case study.

The rest of the paper is organized as follows. We review the related work in Section 2. The proposed new measure with the usage of concept hierarchy to combine domain knowledge into the rule evaluations is discussed in Section 3. The data set and the case studies are discussed in Section 4 and 5. Section 6 provides the conclusions and future work.

## 2 Related Work

### 2.1 Association Rules

The association rule algorithm was first introduced in [1], and is commonly referred to as the apriori association rule algorithm. This algorithm is used to discover rules from transaction datasets. The algorithm first generates frequent itemsets, which are sets of items that have transaction support greater than the minimum support; then based on these itemsets, the association rules are generated which satisfy the minimum confidence. Association rule algorithms can be used to find associations among items from transactions. For example, in *market basket analysis*, by analyzing transaction records from the market, we could use association rule algorithms to discover different shopping behaviours such as, when customers buy bread, they will probably buy milk. This type of behaviour can be used in the market analysis to increase the amount of milk sold in the market. The association rule $\alpha \rightarrow \beta$ holds in the transaction set $D$ with *confidence* $c$ if $c\%$ of transactions in $D$ that contain $\alpha$ also contain $\beta$. The rule $\alpha \rightarrow \beta$ has *support* $s$ in the transaction set $D$ if $s\%$ of transactions in $D$ contain $\alpha \cup \beta$.

### 2.2 Rule Importance Measure

The Rule Importance Measure applies rough sets theory to association rules generation in order to evaluate association rules and thus improve their utilities. Rough sets theory [10] was proposed to classify imprecise and incomplete information. Reduct and core are the two important concepts in rough sets theory. A reduct is a subset of attributes that are sufficient to describe the decision attributes. Core represents the most important information of the original data set. The intersection of all the possible reducts is called the core. The rule importance measure (RIM) is defined as the percentage of the number of times a rule is generated among all the rule sets (represented as $RuleSets$) over the number of available rule sets. The rule importance measure is obtained by $RIM_i = \frac{|\{ruleset_j \in RuleSets | rule_i \in ruleset_j\}|}{n}$. The Rule Importance Measure is simple, quick, easy to compute; it provides a direct and objective view of how important a rule is.

### 2.3 Concept Hierarchy

Much research effort has been found on using concept hierarchy towards databases management, text categorizations, natural language processing and so on. Algorithms on discover associations between different items from levels of taxonomy (which is represented in hierarchies) was introduced in 1995 as the mining approach for generalized association rules [11]. As an example of recent applications, a keyword suggestion approach based on concept hierarchy has been proposed [3] to facilitate user's web search. A data mining system has been proposed to induce the classification rules using concept hierarchy [2]. Concept

Hierarchy can reflect the concepts and relationships of a given knowledge domain. Such hierarchies are useful towards generalization and specialization.

## 3 Enhanced RIM using Concept Hierarchy

Our motivation is to enhance the RIM by integrating the subjective measure into the rule evaluation. We use a concept hierarchy to embed a semantic relationship from the data domain into the knowledge evaluation. In this section, we discuss given a problem domain, how to build a concept hierarchy and combine such hierarchy to enhance the rule measure.

Let $T$ be a data set. $T = (U, C, D)$, where $U$ is the set of data records in the table, and $U \neq \phi$, $C$ is the set of the condition attributes and $D$ is the set of the decision attributes.

Let $s$ be the total number of concepts for a given data set. $c(k)$ $(1 \leq k \leq s)$ is the $k$th concept categorized from the concepts. $Attr_{c(k)}$ denotes all the attributes that belong to the concept $c(k)$. The weight of the concept $w_{c(k)}$ denotes the importance of the concept $c(k)$ from the domain expert's opinion. For a set of rules, the new measure $ERIM_i$ for $rule_i$ can be obtained by Eq 1.

$$ERIM_i = \sum_{p=1}^{l_i} w_{c(k),p} \tag{1}$$

, where $l_i$ is the number of attributes contained by $rule_i$ and $w_{c(k),p}$ is the weight of the $p$th attribute in this rule.

Since the weights $w_{c(k),p}$ are assigned by the domain experts, the greater the value of $ERIM_i$, the more interesting a rule becomes from the domain expert's opinion. Therefore, the $ERIM_i$ measure integrates subjective measures based on concept hierarchy into the rule evaluations.

The concept hierarchy and the weight of the concepts are pre-determined by the domain expert. Concept hierarchies for a given data domain may contain more than multiple levels of hierarchies. For example, given a grocery data domain, concept hierarchies may contain "meat", "seafood", "vegetables", and "soft drinks" as the second level concepts; under each category, there exists more hierarchies. "Meat" may contain "pork", "beef", "lamb" and so on as the sub-hierarchies. In this paper we illustrate the utilities of ERIM by both a six-level and a eight-level hierarchy from a given domain. Domains with more or less hierarchies may use ERIM approach similarly.

The enhanced RIM approach thus consists of two steps. The first step is to obtain the RIM for the given data set; and the second step is for each of the rules from RIM set, derive the ERIM using the Equation 1. Therefore, for each rule generated from a given data set, we have an objective measure to evaluate how important the rule is, and at the same time, we obtain the subjective measure to evaluate which rule is indeed important from the domain expert's opinion.

The procedure of the ERIM measurement is shown as follows:

```
1. Derive concept hierarchies for the given data domain;
2. Assign attributes to concept categories;
3. Assign weights to attributes that belong to each concept category;
4. Calculate the RIM to obtain rule sets ranked by the importance measure;
5. Calculate ERIM for each rule;
6. Combining both RIM and ERIM into rule evaluation.
```

## 4 Data Set

The geriatric care medical dataset used is from Canadian Study of Health And Aging (CSHA). It has 8547 instances of a population of 65 years old and up, of whom 1865 died during the 72 months of follow-up. 3458 of them are male, and 5089 of them are female. 44 self-report attributes were used. The 44 attributes include factors such as disabilities, sicknesses and stress situations. Disabilities refer to attributes such as whether patients could prepare their own meal, or use the telephone, or take medication, or go grocery shopping. Sicknesses refer to attributes such as whether they have a chest problem, or a heart problem, or a kidney problem. Stress situations refer to attributes such as whether they have trouble in life. The class attribute is a binary value indicating whether an individual has died during the 72 months of follow-up. Detailed description of the 44 attributes are available [6].

The sample reduct set of this data is {edulevel, eyesight, hearing, shopping, housewk, health, trouble, livealone, cough, sneeze, hbp, heart, arthriti, eyetroub, eartroub, dental, chest, kidney, diabetes, feet, nerves, skin, studyage, sex}. The reducts are used for the calculation of RIM. There are 14 core attributes generated for this data set. They are *eartroub, livealone, heart, hbp, eyetroub, hearing, sex, health, edulevel, chest, housewk, diabetes, dental, studyage*. All of these reducts contain the core attributes. After removing 12 inconsistent data entries in the medical data set, we obtain the data containing $8,535$ records[1].

## 5 Case Study

We illustrate in more detail how to use the ERIM measure in this section. The geriatric care data is used as our experimental data set.

### 5.1 ERIM - 6 levels

We derive the concept hierarchy by classifying the 44 attributes into 6 categories: sickness, minor sickness, disability, attitude, symptom and others. Sickness refers to significant sickness such as heart problem or chest problem. Minor sickness refers to problems which are common among a lot of older adults but are not

[1] Notice from our previous experiments that the core generation algorithm cannot return correct core attributes when the data set contains inconsistent data entries.

significant such as ear trouble. Disability refers to how well they satisfy their daily activities such as walking, cooking and dressing. Attitude refers to how happy they are and how they feel about themselves. Symptom refers to having some signs medically, but not a sickness yet. Others include attributes that are not strongly related to the sickness, i.e., education level, gender, study age, and age group. For a detailed description on how the attributes are categorized into 6 categories, refer to Table 1. The six level concept hierarchy is shown in Figure 1.

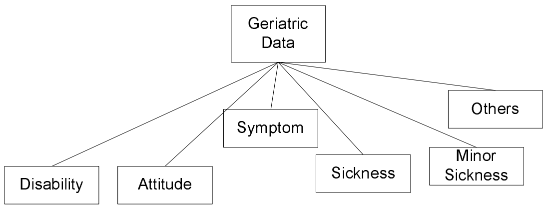


**Fig. 1.** 6-Level Concept Hierarchy

**Table 1.** Concept Hierarchy for the Geriatric Care Data

| Disability | Attitude | Symptom | Sickness | Minor Sickness | Others |
|---|---|---|---|---|---|
| dress, takecare | eyesight | eat | hbp | eyetroub | sex |
| walk, getbed | hearing | cough | heart | eartroub | studyage |
| shower, bathroom | health | tired | stroke | dental | age6 |
| phoneuse, walkout | trouble | sneeze | arthriti | stomach | edulevel |
| shopping, meal | livealon | | parkinson | bladder | |
| housewk, takemed | | | chest | bowels | |
| money | | | kidney | feet | |
| | | | diabetes | skin | |
| | | | nerves | fracture | |

**Table 2.** Weights for Concept Hierarchy of Table 1

| $c(i)$ | Disability | Attitude | Symptom | Sickness | Minor Sickness | Others |
|---|---|---|---|---|---|---|
| $w_{c(i)}$ | 6 | 2 | 1 | 30 | 1 | 1 |

We then assign weights for attributes of each concept category. Different weights are applied to different categories of attributes. The differences between weights are indications of different importance between attributes in terms of predicting the survival probability of an individual. For example, sickness is 30 times as important as symptoms, therefore the weight of sickness is assigned as 30 and the weight of symptoms is assigned as 1. Thus, different weights are applied to the sickness, minor sickness, disability, attitude, and symptom. The weights are assigned as follows: the sickness category has a weight of 30, the minor sickness category has a weight of 1, the disability category has a weight of 6, the attitude category has a weight of 2, and the symptom category and other attributes category also have a weight of 1 [13]. These weights are set after consultation with the domain expert, and is shown on Table 2.

**Calculating RIM and ERIM** The Rule Importance is calculated on this data set. For each reduct set, association rules are generated with $support = 30\%, confidence = 80\%$ [2]. We are interested in rules with survival status on the consequent part of the rules. Rule templates [5] are defined to ensure the desired form of rules are generated [7]. ERIM is also calculated for each of the rules generated by the concept hierarchy from Table 1 and Table 2 using Eq. 1.

As an example of calculating ERIM, suppose we have a rule as follows: *If a person lives alone, has diabetes and nerve problems, then this person has a higher chance of not surviving at the end of the observation period.* This rule contains three attributes, "livealone", "diabetes" and "nerves". The ERIM is calculated as

$$ERIM_i = \sum_{p=1}^{3} w_{c(k),p} = w_{c(livealone)} + w_{c(diabetes)} + w_{c(nerves)} = 2 + 30 + 30 = 62$$

We list all the rules generated ranked by their RIM and ERIM in Table 3. In this table, the first column indicates the original rule number ranked by the RIM approach [7]. The lower the rule number, the more important this rule is. We keep this original number for comparison purpose. The second column contains the generated rules; the third column indicates the ERIM measure of this rule and the fourth column indicates the RIM measure of the same rule in this row. (Note that for comparison purpose, we use the percentage of ERIM divided by the largest ERIM value from all the generated rules. The percentage value of ERIM is also applied on Table 6.)

**Observations and Discussions** The rule importance is an indication of how significant a rule is in term of its classification ability for the decision attribute. The ERIM indicated in the third column is listed to specify the interestingness considered by the domain experts. We compare the two measures and show the differences between the ERIM and RIM. We have the following observations from the experimental results.

- Same important rules are not always considered as interesting by the domain expert. As noted, rule No.3 has the same ranking of the RIM as rule

[2] Note that the values of support and confidence can be adjusted to generate as many or as few rules as required.

**Table 3.** Sample Rules Generated from the Geriatric Care Data Set - 6-Level Hierarchy

| No. | Selected Rules | ERIM-6level | RIM |
|---|---|---|---|
| 159 | hbp, stroke, kidney, nerve problem → negative survival | 100% | 32.56% |
| 109 | hbp, dental problem, kidney problem, nerve problem, fractures → negative survival | 76.67% | 43.02% |
| 22 | oftensneeze, hbp, diabetes, nerve problem → negative survival | 75.83% | 81.40% |
| 44 | oftencough, hbp, kidney, nerve problem → negative survival | 75.83% | 66.28% |
| 53 | oftensneeze, hbp, kidney, nerve problem → negative survival | 75.83% | 61.63% |
| 66 | hbp, diabetes, nerve problem, anyfractures → negative survival | 75.83% | 58.14% |
| 158 | stroke, dental, kidney, nerve problem → negative survival | 75.83% | 32.56% |
| 89 | hbp, stroke, diabetes → negative survival | 75.00% | 48.84% |
| 93 | stroke, arthritis, diabetes → negative survival | 75.00% | 46.51% |
| 100 | stroke, diabetes, nerve problem → negative survival | 75.00% | 45.35% |
| 150 | stroke, arthritis, kidney problem → negative survival | 75.00% | 33.72% |
| 7 | livealone, diabetes, hbp → negative survival | 51.67% | 100% |
| 11 | livealone, diabetes, nerve problem → negative survival | 51.67% | 95.35% |
| 127 | hearing problem, phoneuse, nerve problem → negative survival | 31.67% | 39.53% |
| 216 | oftensneeze, dental, kidney, skin→ negative survival | 27.50% | 1.16% |
| 3 | hearing, diabetes → negative survival | 25.67% | 100% |
| 6 | heart → negative survival | 25% | 100% |
| 2 | chest → negative survival | 25% | 100% |
| 128 | hearing problem, phoneuse, dental problem → negative survival | 7.50% | 39.53% |
| 8 | housework problem → negative survival | 5.00% | 100% |
| 24 | troublewithlife → negative survival | 1.67% | 81.40% |
| 4 | ear trouble → negative survival | 0.83% | 100% |
| 5 | eye trouble → negative survival | 0.83% | 100% |
| 9 | feet → negative survival | 0.83% | 96.51% |
| | ... | | |

No.4, but the ERIM of rule No.3 is much higher than that of rule No.4. The attributes "hearing", "diabetes" and "ear" are all core attributes, therefore these two rules both have the RIM as 100%. However, from Table 2, $w_{diabetes}$ belongs to the sickness concept, and $w_{hearing}$ falls into the attitude concept. The sum of these two weights is greater than $w_{eartrouble}$, which is considered as minor sickness. The same observation goes to rule No.127 and No. 128. Rule No.127 and No.128 have the same RIM, but rule No.127 contains attributes with larger weights than those of rule No.128. Therefore rule No.127 is considered as more interesting by domain expert. *This demonstrates the domain knowledge is necessary to distinguish rules with the same classification ability.*

– Rules that are considered as interesting by the domain expert do not necessarily have the same RIM. Rule No.22 and rule No.44 have the same ERIM, which indicate they have the interestingness degree by the domain expert. However, the RIM for rule No.22 is greater than that of rule No. 44. Note that the only difference of these two rules is No.22 contain attribute "diabetes", and No.44 contains attribute "kidney". "Diabetes" is a core attribute, but not the "kidney". *This demonstrates that, what is considered less interesting by objective measures may be more interesting by the domain experts.*
– Rules having low RIM can be considered surprisingly interesting by the domain expert. Note that rule No. 159 has a low importance of 32.56%, however, it is the most interesting rule ranked by the ERIM measure from

Table 3. This is because attributes “hbp”, “stroke”, “kidney” and “nerve” all fall into the sickness concept with weight 30 and the ERIM is very high; however, among these attributes, only “hbp” is a core attribute from the RIM measure. Same observation applies to rule No. 216. This rule is considered less important because there are less core attribute contained in the rule, and it is generated less frequently across multiple reducts. However, the ERIM of this rule is 27.50%. *This demonstrate that the objective measures alone may ignore very interesting rules considered by the domain knowledge.*
– ERIM measure can be used together with the RIM for distinction of more knowledge oriented rules.

Although the Rule Importance is different from other objective measures and it provides a diverse ranking of how important the rules are [7], this measure can certainly be enhanced with ERIM for a more complete view of rules using the concept hierarchy. Concept Hierarchy is derived by the domain experts. According to the different purposes of the knowledge evaluation, there may exist more than one concept hierarchy for a data domain. For example, in our case study, a frailty index [8] may be considered for assigning the weighted concept categories for the geriatric care data, if the purpose of the study is to consider the proportion of the deficits instead of the nature of the deficits. Neither the objective measure (i.e., RIM) nor the subjective measure (i.e., ERIM) alone is sufficient for a thorough knowledge evaluation. Through the experiments, we observed an integration of both the objective and the subject measures is an optimal approach for knowledge evaluation.

### 5.2 ERIM - 8 levels

In this section, we study how the number of concept hierarchies affects the rule evaluations. We derive the concept hierarchy by classifying the 44 attributes from the geriatric care data into 8 categories: severe sickness, sickness, moderate sickness, minor sickness, disability, attitude, symptom and others. Severe sickness refers to severe sickness such as stroke and diabetes. Sickness refers to significant sickness such as heart problem or chest problem. Moderate sickness refers to moderate problems such as bladder or facture. Minor sickness refers to problems which are common among a lot of older adults but are not significant such as ear trouble. Disability refers to how well they satisfy their daily activities such as walking, cooking and dressing. Attitude refers to how happy they are and how they feel about themselves. Symptom refers to having some signs medically, but not a sickness yet. Other category includes age, gender and so on. For a detailed description on how the attributes are categorized into 8 categories, refer to Table 4. The concept hierarchy is shown in Figure 2.

We then assign weights for attributes of each concept category. Different weights are applied to different categories of attributes. The weights are as follows: the severe sickness category has a weight of 30, sickness category has a weight of 20, moderate sickness has a weight of 2, the minor sickness category has a weight of 1, the disability category has a weight of 6, the attitude category

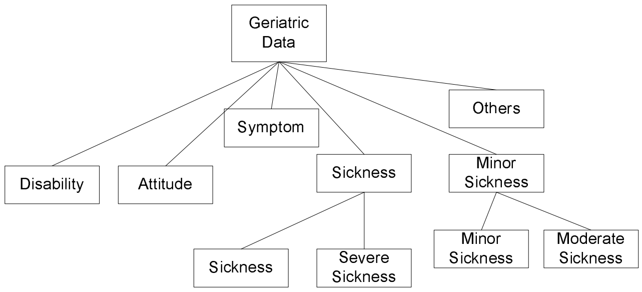


**Fig. 2.** 8-Level Concept Hierarchy

**Table 4.** Concept Hierarchy for the Geriatric Care Data

| Disability | Attitude | Symptom | Minor Sickness | Moderate Sickness | Sickness | Severe Sickness | Others |
|---|---|---|---|---|---|---|---|
| dress, takecare<br>walk, getbed<br>shower, bathroom<br>phoneuse, walkout<br>shopping, meal<br>housewk, takemed<br>money | eyesight<br>hearing<br>health<br>trouble<br>livealon | eat<br>cough<br>tired<br>sneeze | eyetroub<br>eartroub<br>dental<br>stomach<br>feet<br>skin | bladder<br>bowels<br>fracture | hbp<br>heart<br>arthriti<br>chest<br>kidney<br>nerves | stroke<br>parkinson<br>diabetes | Sex<br>studyage<br>age6<br>edulevel |

**Table 5.** Weights for Concept Hierarchy of Table 4

| $c(i)$ | Disability | Attitude | Symptom | Minor Sickness | Moderate Sickness | Sickness | Severe Sickness | Others |
|---|---|---|---|---|---|---|---|---|
| $w_{c(i)}$ | 6 | 2 | 1 | 1 | 2 | 20 | 30 | 1 |

has a weight of 2, and the symptom category and others each has a weight of 1 [13]. These weights are set after consultation with the domain expert, and is shown on Table 5.

We list all the rules generated ranked by their ERIM and RIM in Table 6 according to the same approach as shown in Section 5.2.

From Table 6 we observe that rules are ranked by ERIM in the similar order as in Table 3. For example, rule No.159 are ranked as the highest ERIM in both approaches. By using more detailed concept hierarchy, rules may be differentiated in a deeper level. For example, rule No.22 and No.66 have the same ERIM by using 6-level hierarchy. However, with 8-level hierarchy, the "Sickness" and "Minor Sickness" in Table 1 are further differentiated by more hierarchies with more weights in Table 4. The two different attributes comparing No.22 with No.66 are "oftensneeze" and "anyfractures". In the 8-level hierarchy, "anyfrac-

**Table 6.** Sample Rules Generated from the Geriatric Care Data Set Ranked by 8-level Hierarchy

| No. | Selected Rules | ERIM-8 | RIM |
|---|---|---|---|
| 159 | hbp, stroke, kidney, nerve problem → negative survival | 100% | 32.56% |
| 89 | hbp, stroke, diabetes → negative survival | 88.89% | 48.84% |
| 93 | stroke, athritis, diabetes → negative survival | 88.89% | 46.51% |
| 100 | stroke, diabetes, nerve problem → negative survival | 88.89% | 45.35% |
| 66 | hbp, diabetes, nerve problem, fractures → negative survival | 80.00% | 58.14% |
| 22 | often sneeze, hbp, diabetes, nerve problem → negative survival | 78.89% | 81.40% |
| 7 | live alone, hbp, diabetes → negative survival | 57.78% | 100.00% |
| 11 | live alone, diabetes, nerve problem → negative survival | 57.78% | 95.35% |
| | ... | | |
| 3 | hearing, diabetes → negative survival | 35.56% | 100.00% |
| 4 | ear trouble → negative survival | 1.11% | 100.00% |
| 5 | eye trouble → negative survival | 1.11% | 100.00% |
| 9 | feet problem → negative survival | 1.11% | 96.51% |
| | ... | | |

tures" is assigned with a higher weight. Therefore, rule No.66 is ranked higher than No.22 using the 8-level hierarchy in Table 6. This results indicate that more concept hierarchies represent finer-grained domain knowledge, therefore the interestingness of the rules are differentiated in a greater detail comparing to using less hierarchies.

## 6 Conclusion

In this paper we have proposed a novel approach for rule evaluation based on concept hierarchy. An enhanced Rule importance measure ERIM is shown to be effective on evaluating interesting rules from the domain expert's opinion. We demonstrate through a real world data set that the integration of both the objective and the subjective measures can provide a knowledge oriented distinction of rules. The advantages of ERIM are as follows: it combines both the subjective and the objective measures for rule evaluation; in the situation where the two rules have the same RIM, ERIM can be used to provide a knowledge oriented distinction. The concept hierarchy based weights are indications of interestingness reflecting domain knowledge.

In the future we plan to continue developing rule evaluation measures that combine both the objective measures and the subjective measures. As discussed in Section 3, concept hierarchy is limited by the purpose of knowledge evaluation and it is not automatable at this stage. The constructing of concept hierarchy as well as the assigning of attribute weights depend on the particular problem domain. These two components of our approach are time consuming and sometimes difficult to obtain from the problem domain expert. Domain experts and statistics information should play an important role. We are also interested in researching an automatic mechanism on developing the concept hierarchies to facilitate more efficient and more precise knowledge evaluations.

## Acknowledgements

We gratefully acknowledge the financial supports of the Natural Science and Engineering Research Council of Canada and Alpha Global-iT Inc.

## References

1. R. Agrawal and R. Srikant. Fast algorithms for mining association rules. In J. B. Bocca, M. Jarke, and C. Zaniolo, editors, *Proc. 20th Int. Conf. Very Large Data Bases, VLDB*, pages 487–499. Morgan Kaufmann, 12–15 1994.
2. M. E. M. D. Beneditto and L. N. de Barros. Using concept hierarchies in knowledge discovery. In A. L. C. Bazzan and S. Labidi, editors, *SBIA*, volume 3171 of *Lecture Notes in Computer Science*, pages 255–265. Springer, 2004.
3. Y. Chen, G.-R. Xue, and Y. Yu. Advertising keyword suggestion based on concept hierarchy. In *WSDM '08: Proceedings of the international conference on Web search and web data mining*, pages 251–260, New York, NY, USA, 2008. ACM.
4. L. Geng and H. J. Hamilton. Interestingness measures for data mining: A survey. *ACM Comput. Surv.*, 38(3):9, 2006.
5. M. Klemettinen, H. Mannila, P. Ronkainen, H. Toivonen, and A. I. Verkamo. Finding interesting rules from large sets of discovered association rules. In N. R. Adam, B. K. Bhargava, and Y. Yesha, editors, *Third International Conference on Information and Knowledge Management*, pages 401–407. ACM Press, 1994.
6. J. Li. *Rough Set Based Rule Evaluations and Their Applications.* PhD thesis, University of Waterloo, Waterloo, Canada, 2007.
7. J. Li and N. Cercone. Introducing a rule importance measure. In J. F. Peters, A. Skowron, D. Dubois, J. W. Grzymala-Busse, M. Inuiguchi, and L. Polkowski, editors, *T. Rough Sets*, volume 4100 of *Lecture Notes in Computer Science*, pages 167–189. Springer, 2006.
8. A. Mitnitski, X. Song, and K. Rockwood. The estimation of relative fitness and frailty in community-dwelling older adults using self-report data. *J Gerontol A Biol Sci Med Sci*, 59:M627–M632, 2004.
9. A. Øhrn. *Discernibility and Rough Sets in Medicine: Tools and Applications.* PhD thesis, Department of Computer and Information Science, Norwegian University of Science and Technology, Trondheim Norway, 1999.
10. Z. Pawlak. *Rough Sets: Theoretical Aspects of Reasoning about Data.* Kluwer Academic Publishers, Norwell, MA, USA, 1992.
11. R. Srikant and R. Agrawal. Mining generalized association rules. In *VLDB '95: Proceedings of the 21th International Conference on Very Large Data Bases*, pages 407–419, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc.
12. P.-N. Tan, V. Kumar, and J. Srivastava. Selecting the right interestingness measure for association patterns. In *KDD*, pages 32–41. ACM, 2002.
13. S. W. H. Wong and N. Cercone. Comparing classification methods for predicting survival probabilities in the elderly. *IEEE Conference on Bioinformatics and Biomedicine Workshop on Biomedical and Health Informatics*, 2008.