A Threshold Free Implication Rule Mining

Alex Tze Hiang Sim, Maria Indrawan, Bala Srinivasan

***Abstract*—Typically, before association rules are mined, a user needs to determine a support threshold in order to obtain only the frequent item sets. Having users to determine a support threshold attracts a number of issues. We propose an association rule mining framework that does not require a pre-set support threshold. The framework is developed based on implication of propositional logic. The experiments show that our approach is able to identify meaningful association rules within an acceptable execution time.**

***Index Terms*—association rule mining, propositional logic, implication, threshold free.**

I. Introduction

Association Rule Mining (ARM) is a unique technique. It has the advantage to discover knowledge without the need to undergo a training process. It discovers rules from a dataset, and each rule discovered has its importance measured against interesting measures such as *support* and *confidence*.

Although ARM technique does not involve model selection, it necessitates a cut-off support threshold to be predefined to separate frequent patterns from the infrequent ones. Two item sets are said to be *associated* if they co-occurred together frequently, and only the frequent ones are reported. There are major disadvantages to having a predefined threshold. Firstly, some rules are inevitably loss if the support threshold is set inaccurately. In addition, it is usually not possible to remove the support threshold in order to find infrequent items because ARM relies on a downward closure property of *support*, which necessitates a threshold to search for frequent item sets. That is, if an item set passes a minimum support requirement then all its sub sets also passes this requirement. If this threshold is waived then there will be no pruning opportunity, which results in an exponential search space. As a result, search could not be completed within feasible time. In summary in the traditional association rule mining, a minimum support threshold is needed, and should be determined accurately in order to produce useful rules for users.

Based on the above limitations, we investigate the possibility of developing a new association rule mining framework that work without having to determine a support threshold. We based our framework on the notion of implication of propositional logic. We explain in details our proposed model in section 3 after discussion of previous work is presented in section 2. Several experiments and discussion on the results are presented in section 4. Finally, conclusion is made in section 5.

Manuscript received December 27, 2007. Sim Author is supported by Monash PhD scholarship. All authors are with the Monash University, Melbourne, Australia. (e-mail: {Tze.Sim,Maria.Indrawan, Bala.Srinivasan}@infotech.monash.edu.au).

II. Previous Work

More recently, it is accepted that infrequent rules are also important because it represents knowledge not found in frequent rules, and these infrequent rules are often interesting [1],[2],[3],[4],[5]. In addition to missing infrequent item rules, the traditional algorithm such as apropri [6] also does not report the existence of negative associations.

Association among infrequent items and negative associations have been relatively ignored by association mining algorithm mainly due to the problem of large search space and the explosion of total number of association rules reported [1],[2],[3],[4],[5],[7],[8]. Some of these rules may in fact are noise in the data. There are some attempts to find infrequent association such as that of [9]. This work proposed a generalise association using correlation. Correlation is measured by chi-square. However, at small expected values, the measure of chi-square has limitation of measuring the association accurately and, hence, results may be inaccurate. In addition, the authors' algorithm relies on a modified support hence, is not really suitable to find infrequent rules except the ones that are above a threshold. [10] finds independent rules measured by interest (leverage) and below a minimum support threshold. Authors in [10] also use [11] measure, which is derived from correlation, and necessitates a minimum confidence threshold. Mining below a minimum support threshold is similar to having a maximum support threshold. In addition, measure used in [11] will inherit the drawbacks of a correlation measure in [9],[12]. [13] filters uninteresting rules using leverage as a measure. [14],[15] finds rules using measure such as leverage or lift; these can be performed without other thresholds in place. Since rules found are independent from a minimum support threshold, theoretically it could find all infrequent rules. Rules found using leverage however measures co-occurrence but not the real implication [23].

There is relatively little research to find infrequent and interesting association rules. Two fundamental constraints are (i) selection of measure used and (ii) use of this measure to search for infrequent and interesting rule directly without post-processing the found rules. The measure should justify the search time in discovering rules. This new measure must contain properties that can be used to search for infrequent association rules directly. Otherwise it is meaningless to have a measure but not the rules.

III. Association Rule Pairs

We study the frequency of occurrences between two item sets and proposed our approach that association rules can be mined without a minimum support threshold.

In our study of the definition on *association*, we found that

ISBN: 978-988-98671-8-8

association is defined in many ways of which can be referred to a number and different types of relationships among item sets. A typical definition of association is *co-occurrence* [17]. Association can also be generalized into *correlation* [9] or *dependence* rule [12]. Each has their merits. For the purpose of our model, we define association using implication of propositional logic in that an implication must be supported by its *contrapositive*. Such association rules mined has implications stronger than the typical associations based on co-occurrences.

To illustrate our proposed framework, consider table 1 that contains relations between a rule antecedent (LHS), X and a rule consequence (RHS), Y as an association rule. The rule antecedent X consists of a combination of items, called an antecedent item set *A*. An antecedent item set *A* may exist, represented by *A*, or absence, represented by ¬*A*. Similarly, the rule consequence Y may contain existence or absence of consequence item set *C*. They are represented as *C* and ¬*C*. The frequency of occurrence of *A* and *C* is represented by Q1, *A* and ¬*C* by Q2, ¬*A* and *C* by Q3, finally, ¬*A* and ¬*C* by Q4. The total of occurrence of *C* is represented by covC, the occurrence of ¬*C* is given by m-covC. The same representations applied to *A* and ¬*A*.

Table 1: Frequency of occurrences among antecedent and consequence item set

		A rule consequence (RHS), Y		
		C	¬*C*	Total
A rule antecedent (LHS), X	*A*	Q1	Q2	covA
	¬*A*	Q3	Q4	m-covA
	Total	covC	m-covC	m

We now present the principle behind the pairing of the rules used in our mining algorithm. We consider an implicational association exists when the contrapositive of a positive rule does not contradict the positive rule. This definition of a strong association is based on the logic definition in [19] that states, if a statement is true then its contrapositive is always true (and vice versa). If a statement is false then its contrapositive is always false (and vice versa). Hence we should only report on positive association when the statistics of its contrapositive support the positive association. It follows, a contrapositive is the *inverse* and *converse* of an association, which has the same frequency of occurrence with only the *inverse* of an association. To show the distinct differences between an *implication*, its *inverse*, and its *contrapositive*, we list their forms and names according to propositional logic as follows.

i) $A \Rightarrow C$; an implication (a association rule)
ii) $\neg A \Rightarrow \neg C$; inverse of an implication
iii) $\neg C \Rightarrow \neg A$; contrapositive of an implication

With referring to table 1, the statistics of the frequency of occurrence of an implication is shown in Q1, which is the total occurrences of *A* and *C* together. Similarly, the frequency of occurrence of an inverse or a contrapositive is shown in Q4; both have the same value.

This pairing opportunity arises when Q3 and Q2 are smaller than Q1 and Q4. That is, frequency of occurrences from other forms of relationships does not overpower (i) and (iii). For example, if we found that Q3 has a higher frequency than Q1 and Q4 then the rule $A \Rightarrow C$ should not be reported because the absence of A has a stronger association to *C* than existence of *A*. As a result, the frequency of occurrence of associations between item sets in rules (i) and (iii) are lower than other negative associations, such as "$\neg A \Rightarrow C$" and "$A \Rightarrow \neg C$", which weaken the associations between a positive rule and its contrapositive.

Since we choose to report $A \Rightarrow C$ only if Q3 and Q2 are smaller than Q1 and Q4, we write these two conditions being (Q1,Q4) > Q2 and (Q1,Q4) > Q3. Based on these enforcements, our model will only report rules that have the highest statistics value in Q1 and Q4. There are many advantages in associated with these enforcements. Among these, the model enforces that a positive rule will always be supported by its contrapositive statistics, and will not be contradicted by other forms of negative association. That is positive rules found are mathematically true. In addition, positive rules are also supported by its inverses, which according to proportional error reduction [16], [24] are associations that may further improve prediction error on a consequence. Besides, such a rule also meets both the necessary and sufficient conditions of another rule in logic.

Hence, following propositional logic and proportional error reduction, we could mine rules, which is strong in implication. While an implication is true, the strength of its ability to improve prediction on a targeted item (a consequence item set) is given by interestingness measure, *H* denoted as follows,

$$H = \begin{cases} lambda, if\ Q1 > (Q2, Q3), Q4 > (Q2, Q3) \\ 0, otherwise \end{cases} \quad (2)$$

where, lambda = (min(covC, m – covC) – min(Q1, Q2) – min(Q3, Q4)) / min(covC, m – covC) (3)

We call these association rule pairs, *coherent* rules, and its measure, a *coherent* rule measure.

It is interesting to highlight here that although we have only shown the work based on positive association as a starting point, the model will be able to report negative association such as the association "¬*A* ⟹ ¬*C*", and consider the positive association "*A* ⟹ *C*" as its 'inverse'. This means, negative association rules may be found, and is supported by positive rules.

IV. Search for Coherent Rules

Our introduction has mentioned that removing support threshold and finding negative associations requires searching in large data space. In order to reduce the search space, we introduce two possible pruning opportunities in the search algorithm.

A. First Pruning Opportunity

We use the statistical conditions (Q1 > Q3) within property of coherent rule measure *H*, to prune away super item sets, which does not meet the condition. It follows, this statistical condition has a downward closure property. When use in a depth manner of search, any item set, which does not meet the condition will have its super sets also not meeting the condition.

Theoretically, a depth search will result in a reduce value in Q1, and an increase value in Q3. As a result, search will stop at some depth level following a parent item set. For example, we label attributes from 1 to 100, if a combination of [1, 2, 10, 20] has Q1 ≤ Q3 then further search on [1, 2, 10, 20, *] can be pruned.

B. *Second Pruning Opportunity*

We also use a moving window and forecast to prune item sets. It follows if an item set does not have forecast strength at least a moving window value, its super item set will also not.

A moving window (*mw*) of size *w*% is calculated using a momentary maximum strength (*tp*) of rule pairs searched, where $mw = tp - tp \times wo\%$. For example, if maximum strength searched, *tp* is at 1, and an opening, w% allowed is at 5%, then any item set with strength value between 0.95 to 1.00 is included. Item set with strength value outside this range are discarded following a downward closure property of the forecast measure of *H*. This concept is practical following 20-80 rule of thumb that a smaller percentage of knowledge or rules such as 20% may closely represent a larger 80% of knowledge.

We use this approach to find all strongest rule pairs within a top percentage because we do not trade expensive computational cost for weaker rule pairs, which fall outside this opening window. We give another example that if a dataset contains a rule pair that have the strongest measure *H* = 0.6, and we just want to have all other rules that have strength close to this strength value for example within 5%. Using this forecast-to-prune technique, rule pairs with strength within a vicinity of $0.6 \pm 0.6 \times$**5%** (i.e. 0.58 to 0.60) will be printed but others pruned. It is important to state here, we do not know how many rules will be reported because it is determined by algorithm accordingly. This is in contrast to technique such as [15] where a user need to specify the number of rules needed, and it is possible that some rules with the same strength will be lost [18]. A smaller opening window will produce a smaller number of rules, hence require a shorter time to search for it comprehensively.

To forecast the strength of a rule pairs rule, we calculate the highest possible measure value by simulating the lowest value on min(Q1,Q2). As a result, *H'* gives us the highest and potential strength value that may be attended by an item set and its super item set. It follows if maximum of forecast measure value on item set [1, 2, 10, 20] is lower than the lowest value of a moving window then this combination together with all its child item set can be pruned. These prunings reduce the exponential search space. Our search and forecasting procedures is further explained in a following algorithm in pseudo-language format.

Algorithm CRS(⟨L,C,D,wo%⟩,⟨covC,szD⟩,A,mw,tp,CR)

Inputs:
1. *L*: super item set
2. *C*: a target; a consequence item set, *C*∈*L*
3. D: a database of transaction records, tr⊆L
4. wo%: an opening of a moving window, 0≤wo≤100

Optional Recursive Inputs*:
5. covC: a coverage value on *C*
6. szD: total number of transactions
7. *A*: an antecedent item sets from *L*, *A*⊂*L*
8. mw: a dynamic lowest opening window value, 0≤mw≤1
9. tp: a momentary maximum strength
10. CR: a dynamic database of rule pairs rules with its coherent measure value, CR⊆*L*, *H*

* These recursive inputs may be optional if declared as 'global'.

Output: CR

```
//function CRVal(Q1, Q2, covC, szD)
//1.  Q3 := covC - Q1
//2.  Q4 := H – covC
//3.  if Q3 < 0 //occur only in forecast
//     3.1  H := 1
//4.  else if (Q1>Q3) && (Q4>Q2) &&
//              (Q1>Q2) && (Q4>Q3), then
//     4.1 cov := min(covC, H - covC)
//     4.2 H := (cov - min(Q1, Q2) - min(Q3, Q4)) / cov
//4.  end
//5.  return Q3, H

1. A ⟵ get_a_combination(L) // any depth search algorithm or its variations //
2. Q1 := covA := coverage(A, C, D)
//1st forecast below //
3. (Q3, H) := CRVal(Q1, 0, covC, szD)
4. if (H≥mw) and (Q1>Q3), then
   4.1 Q1 := support(A, C, D)
   //2nd  forecast below //
   4.2 (Q3, H) := CRVal(Q1, 0, covC, szD)
   4.3 if (H≥mw) and (Q1>Q3), then
      4.3.1 Q2 = covA - Q1
        //Get real measure value //
      4.3.2 (Q3, H) := CRVal(Q1, Q2, covC, szD)
      4.3.3 if (Q1>Q3), then
         4.3.3.1 if tp < H, then
            4.3.3.1.1 tp := H
         4.3.3.2 end
         4.3.3.3 mw := tp × (100 - wo)%
         4.3.3.4 CR := CR ∪ (A, C, H)
         4.3.3.5 remove cr ∈ CR, where H ∈ cr < mw
         4.3.3.6 CRS(⟨L,C,D,wo%⟩,A,mw,tp,CR)
      4.3.4 end
   4.4 end
5. end
6. return CR
```

V. Experiments and Discussions

We have conducted a number of experiments. In this paper, we report the results of three main categories of experiment. In the first category, we want to show that our association rule mining framework can find infrequent association that may be difficult to find in traditional association rule mining. The zoo data set is used in this experiment. The second experiment shows that our proposed framework requires less post-processing in generating the rule compared to the traditional association mining algorithm. That is, instead of finding too many rules, our algorithm finds smaller number of rules. The experiment for this purpose is conducted in the mushroom data set. Lastly, we measure the performance of our framework by testing its scalability. For this performance test, we created three sparse artificial datasets, and another three dense artificial datasets. In both zoo and mushroom dataset, we use the *classes* as the consequences in order to find association rules directly from data. On artificially generated datasets we use the last items as consequences.

A. *Zoo dataset*

Zoo dataset [20] is a collection of animal characteristics and their categories in a zoo. This dataset is chosen because animal characteristics in each category are very well known. As a result, it is easier to verify the correctness and interestingness of rules mined. Zoo dataset contains seven categories of animals including *mammalia* and *amphibian*. While mammalia type of animal such as elephants, buffalos, and goats are frequently observed in this zoo, amphibian type of animal such as frog and toad are relatively rare. We run our search algorithm without setting a minimum support threshold to obtain all rules within a window of a top 5%, and each rule contains not more than five items. We report the results as follows,

A total of 16 rules are found on mammalia type of animals. All rules have strength of 1.0 out of 1.0. We verify the correctness of these rules based on known knowledge on this category of animal. For example, all mammalia such as goat has no feather but has milk and backbone therefore feather(0), milk(1), and backbone(1) are reported associated with mammalia(1). We list all rules contains not more than four items (due to length of paper) in table 2.

Table 2: Rules describe *mammalia*

Antecedent Item Set		Conseq. Item Set
milk(1)	⇒	*mam.*(1)
feathers(0),milk(1)	⇒	*mam.*(1)
milk(1),backbone(1)	⇒	*mam.*(1)
feathers(0),milk(1),backbone(1)	⇒	*mam.*(1)
milk(1),breathes(1)	⇒	*mam.*(1)
feathers(0),milk(1),breathes(1)	⇒	*mam.*(1)
milk(1),backbone(1),breathes(1)	⇒	*mam.*(1)
milk(1),venomous(0)	⇒	*mam.*(1)
feathers(0),milk(1),venomous(0)	⇒	*mam.*(1)
milk(1),backbone(1),venomous(0)	⇒	*mam.*(1)
milk(1),breathes(1),venomous(0)	⇒	*mam.*(1)

We found these rules describe mammalia correctly. In fact, the first and the shortest rule $milk \Rightarrow mammalia$ describe a fundamental characteristic of a mammalia explicitly. From literature review, the second rule may be deemed redundant in comparison with the first rule because inclusion of an additional item set, feather(0), which cannot further increase the strength of rule. The strength of the first rule is already at its maximum at 1.0; any further inclusion of items may be redundant. Such a consideration however is application dependent. We could use both items, feathers(0) and milk(1) to describe mammalia more comprehensively at the same strength of 1.0. That is, an animal of mammalia does not have feather but milk. If we discard feather(0), we loss this item as a descriptive.

We run the search for amphibian, and found a total of 136 rules. Again, we could not find any incorrect rules. These rules have strength 1.0. While studying at these rules, we are surprised by the fact that amphibian like frog is toothed! We confirm this via answer.com, and this is indeed correct. That is, frog in this zoo is toothed.

Comparing the two experiments, there is a large difference in their total number of occurrence in the overall transaction records. 41% of transaction records contain mammalia, in comparison, only 4% of transaction records contains amphibian. That is, search for amphibian is a search for infrequent association rules, which is often missed by most association rule mining technique that demands a minimum support threshold. If we set minimum support threshold to be higher than 4% and use a typical association rule mining technique, we loss rules describing amphibian. In comparison, our technique does not necessitate a minimum support threshold, it finds all necessary rules.

On execution time wise, each running time takes less than 3 seconds on a notebook computer Pentium Centrino 1GHz with 1.5G of main memory and running Windows XP Home Edition. Zoo dataset contains 101 transactions and 43 item sets. The search space on a target is $2^{2(n-1)}$ - ($2^{(n-1)}$ - 1) where $2^{2(n-1)}$ is the total number of both positive and negative rules, and ($2^{(n-1)}$ - 1) is the total number of positive rules using a single consequence item set as a target. In this case, zoo dataset contains 2E+25 combinations of item sets. We use an optimistic assumption to grasp the size of the search space; we assume only one computation cycle time (1 / 1GHz) is needed to form and to validate a combination of item set in a single transaction. Based on this optimistic assumption, it follows that a search without pruning would require at least 6E+10 years to complete. In comparison, our search time is feasible. From these two experiments, we conclude that association rule pairs are useful to discover knowledge (both frequent and infrequent) from dataset.

B. *Mushroom dataset*

In our next experiment, we run our search algorithm on mushroom dataset [21] which contains 8124 transactions and 119 items. To grasp the search space, if one computation cycle time is needed to form a combination, it takes at least 3E+58 years to complete. Our search for both poisonous and edible mushrooms is completed within 17 seconds with 6 rules found. We list these rules in Table 3(a) and Table 3(b).

Table 3(a): Rules describe edible mushroom

Antecedent Item Set		Conseq. Item Set
odor.almond	→	*Edible*
odor.almond, stalk-color-below-ring.orange	→	*Edible*

Table 3(b): Rules describe poisonous mushroom

Antecedent Item Set		Conseq. Item Set
cap-color.green, odor.spicy, gill-attachment.free	→	*Poisonous*
cap-color.green, odor.spicy, gill-attachment.free, stalk-color-below-ring.orange	→	*Poisonous*
cap-color.green, gill-attachment.free, stalk-color-below-ring.cinnamon	→	*Poisonous*
cap-color.green, gill-attachment.free, stalk-color-below-ring.orange, stalk-color-below-ring.cinnamon	→	*Poisonous*

We leave the correctness of these results to domain experts since we are no expert. The strengths of these rules are around 0.77 out of 1.0, this suggests that there may exist some exceptional cases besides these strongest rules.

In comparison, a typical association rule mining technique such as *apriori* reports more than 100 thousands of rules with confidence value at 100%. Some of these rules are not interesting, and one way to filter these are to select high confidence rules with positive leverage values. Rules with positive leverage are rules that are dependent to each other. However, after filtering high confident rules with positive leverage, it still left us more than 100 thousands of rules for this dataset. Among these rules, it contains our six rules. We conclude from these observations that our approach produces rules that are concise and easier to apply.

C. *Artificial datasets*

We follow to generate a following three dense artificial datasets with an increase in complexity using the IBM synthetic data generator [22]. The symbols used in representing a dataset are explained below,

D: number of transactions in 000s
T: average items per transaction
N: number of items
L: number of patterns
I: average length of maxima pattern

The dense datasets have an average length of maxima pattern (I) close to average items per transaction (T), besides having a low number of patterns (L). These dense datasets have an increase number of items as follows,

i) D100T10N100L50I9,
ii) D100T10N500L50I9,
iii) D100T10N1000L50I9

We generate also sparse dataset with an increase in its number of items hence complexity,

i) D100T10N100L10000I4,
ii) D100T10N500L10000I4,
iii) D100T10N1000L10000I4

The results from experiments suggest that our search for association rule pairs is feasible within a linear or polynomial search time over an increase of complexity or items.

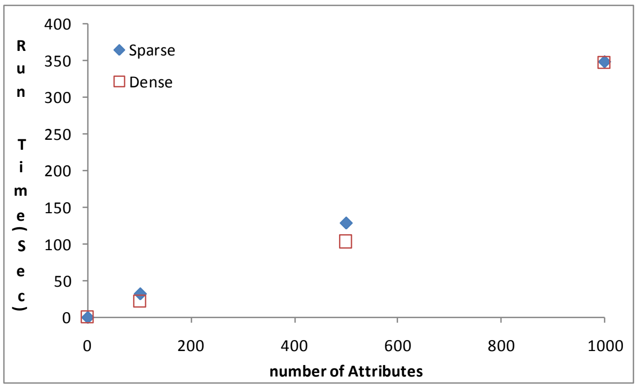

Fig 1: Search time on an increase complexity on dense and sparse dataset

VI. Conclusion

We conclude from our design of a threshold free association rule mining technique that a minimum support threshold may be waived. We may trade a heuristic minimum support threshold by employing other definition on association to avoid using a cut-off support threshold. Implication of propositional logic is a good alternative on the definition on association. Rules based on this definition may be searched and discovered within feasible time.

References

[1] Cohen, E.; Datar, M.; Fujiwara, S.; Gionis, A.; Indyk, P.; Motwani, R.; Ullman, J. D. & Yang, C., "Mining association rules with multiple minimum supports", *ACM*, 1999, pp. 337-341.
[2] Cohen, E.; Datar, M.; Fujiwara, S.; Gionis, A.; Indyk, P.; Motwani, R.; Ullman, J. D. & Yang, C., "Finding Interesting Associations without Support Pruning", *ICDE '00: Proceedings of the 16th International Conference on Data Engineering*, IEEE Computer Society, 2000, 489.
[3] Huang, Y.; Xiong, H.; Shekhar, S. & Pei, J., "Mining confident co-location rules without a support threshold", *SAC '03: Proceedings of the 2003 ACM symposium on Applied computing*, ACM, 2003, 497-501.
[4] Koh, Yun Sing and Rountree, Nathan., "Finding Sporadic Rules Using Apriori-Inverse", *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 2005, Vol. 3518, , pp. 97-106.
[5] Koh, Y. S.; Rountree, N. & O'Keefe, R., "Mining Interesting Imperfectly Sporadic Rules", *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 2006, 3918, 473-482.
[6] Srikant, R. & Agrawal, R., "Mining quantitative association rules in large relational tables", *SIGMOD Rec.*, ACM, 1996, 25, 1-12.
[7] Castelo, R.; Feelders, A. & Siebes, A., "MAMBO: Discovering Association Rules Based on Conditional Independencies", *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 2001, 2189, 289-298.
[8] Lin, W.; Tseng, M. & Su, J., "A Confidence-Lift Support Specification for Interesting Associations Mining", *PAKDD '02: Proceedings of the 6th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining*, Springer-Verlag, 2002, 148-158.
[9] Brin, S.; Motwani, R. & Silverstein, C., "Beyond market baskets: generalizing association rules to correlations", *SIGMOD Rec.*, ACM, 1997, 26, 265-276.
[10] Zhou, L. & Yau, S., "Efficient association rule mining among both frequent and infrequent items", *Comput. Math. Appl.*, Pergamon Press, Inc., 2007, 54, 737-749.
[11] Wu, X.; Zhang, C. & Zhang, S., "Mining Both Positive and Negative Association Rules", *ICML'02*, 2002, pp. 658-665.

[12] Silverstein, C.; Brin, S. & Motwani, R., "Beyond Market Baskets: Generalizing Association Rules to Dependence Rules", *Data Min. Knowl. Discov.*, Kluwer Academic Publishers, 1998, 2, 39–68.
[13] Piatetsky-Shapiro, G., "Discovery,analysis,and presentation of strong rules", *Knowledge Discovery in Databases*, 1991, pp. 229-248.
[14] Zhang, S. & Webb, G. I., "Further Pruning for Efficient Association Rule Discovery", *Lecture Notes in Computer Science*, Vol.2256, 2001, pp. 205-216.
[15] Webb, G. I. & Zhang, S., "k-Optimal Rule Discovery", *Journal Data Mining and Knowledge Discovery*, Springer Netherlands, 2005, 10(1), 39-79.
[16] Kviz F. J., "Interpretating Proportional Reduction in Error Measures as Percentage of Variation Explained", *The Sociological Quarterly*, Vol.22, 1981, pp. 413-420.
[17] Agrawal, R.; Imielinski, T. & Swami, A., "Mining association rules between sets of items in large databases", *SIGMOD Rec.*, ACM Press, 1993, 22, 207–216.
[18] Li, J., "On Optimal Rule Discovery", IEEE Transactions on Knowledge and Data Engineering, *IEEE Educational Activities Department*, 2006, 18, 460-471.
[19] Contraposition. [online]. Available: http://en.wikipedia.org/wiki/Contrapositive
[20] Zoo dataset. [online]. Available: http://magix.fri.uni-lj.si/orange/doc/datasets/zoo.htm.
[21] Frequent Item set Mining Dataset Repository. [online]. Available: http://fimi.cs.helsinki.fi/data/.
[22] IBM Quest Data Generator. [online]. Available: www.almaden.ibm.com/cs/quest/syndata.html
[23] Cornelis, C., Yan, P., Zhang, X., & Chen, G., "Mining Positive and Negative Association Rules from Large Databases", *IEEE Conference on Cybernetics and Intelligent Systems*, 2006, 1-6.
[24] Sim, A.T.H, Indrawan, M., Srinivasan, B., "Mining Infrequent and Interesting Rules from Transaction Records", AIKED, 2008, to be appeared.