ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 5, May 2014

Mining Indirect Positive and Negative Association Rules: A Novel Approach

B.Ramasubbareddy, A.Govardhan, A.Ramamohanreddy

Professor, S.V College of Engineering, Tirupathi, India

Professor & Principal, JNTUH college of Engineering, Karimnagar, India

Professor, S.V.U College of Engineering, S.V University, Tirupati, India

ABSTRACT: Indirect association is a new kind of infrequent pattern, which provides a new way for interpreting the value of infrequent patterns and can effectively reduce the number of uninteresting infrequent patterns. The concept of indirect association is to "indirectly" connect two rarely co-occurred items via a frequent itemset called mediator, and if appropriately utilized it can help to identify real interesting "infrequent itempairs" from databases. Indirect association rule is said to be positive (Negative) if mediator set contains presence (presence or absence) of items. Existing indirect association mining methods mine positive mediator sets. To the best of our knowledge, no research work has been conducted on mining indirect negative associations. In this paper, we propose an approach for mining indirect negative associations. The proposed method can discover all positive and negative indirect association between itemsets.

KEYWORDS: Data mining; positive and negative association rules; indirect association**.**

I. INTRODUCTION

Association rule mining is a data mining task that discovers associations among items in a transactional database. Association rules have been extensively studied in the literature for their usefulness in many application domains such as recommender systems, diagnosis decisions support, telecommunication, intrusion detection, etc. Efficient discovery of such rules has been a major focus in the data mining research. From the celebrated *Apriori* algorithm [1] there have been a remarkable number of variants and improvements of association rule mining algorithms [2]. A typical example of association rule mining application is the market basket analysis. In this example, the behavior of the customers is studied with reference to buying different products in a shopping store. The discovery of interesting patterns in this collection of data can lead to important marketing and management strategic decisions. For instance, if a customer buys bread, what are chances that customer buys milk as well?. Depending on some measure to represent the said chances of such an association, marketing personnel can develop better planning of the shelf space in the store or can base their discount strategies on such associations/correlations found in the data. All the traditional association rule mining algorithms were developed to find positive associations between items.

In [14], a new class of patterns called indirect associations has been proposed and its utilities have been examined in various application domains. Consider a pair of items X and Y that are rarely present together in the same transaction. If both items are highly dependent on the presence of another itemset M, then the pair (X, Y) is said to be indirectly associated via M. There are many advantages in mining indirect associations in large data sets. For example, an indirect association between a pair of words in text documents can be used to classify query results into categories [14]. For instance, the words *coal* and *data* can be indirectly associated via *mining*. If only the word *mining* is used in a query, documents in both *mining* domains are returned. Discovery of the indirect association between *coal* and *data* enables us to classify the retrieved documents into *coal mining* and *data mining*. There are also potential applications of indirect associations in many other real-world domains, such as competitive product analysis and stock market analysis [14].

This paper is structured as follows: the next section contains preliminaries about Indirect Association Rules, In Section3 existing strategies for mining indirect association rules are reviewed. The proposed algorithm is presented in Section 4 for finding all valid indirect association rules for pairs of multiple itemsets. Section 5 contains conclusions and future work.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 5, May 2014

II. BASIC CONCEPTS AND TERMINOLOGY

Let I = {i_1, i_2,. . . , i_m} be a set of m items. A subset X $\subseteq$ I is called an itemset. A k-itemset is an itemset that contains k items. Let D = {T_1, T_2,. . . , T_n } be a set of n transactions, called a transaction database, where each transaction $T_{j,}$ j = 1, 2, . . . , n, is a set of items such that $T_j \subseteq$ I. Each transaction is associated with a unique identifier, called its TID. A transaction T contains an itemset X if and only if X $\subseteq$ T. The support of an itemset X is the percentage of transactions in D containing X. An itemset X in a transaction database D is called "frequent itemset" if its support is at least a user-specified minimum support threshold viz., minsup. Accordingly, an infrequent itemset is an itemset that is not a frequent itemset.

A. NEGATIVE ASSOCIATION RULES

An association rule is an implication of the form X ⇒ Y, where X ⊂ I, Y ⊂ I, and X ∩ Y = ∅. Here, X is called the antecedent and Y is called the consequent of the rule. The confidence of an association rule X ⇒ Y is the conditional probability that a transaction contains Y, given that it contains X. The support of rule X ⇒ Y is defined as: sup(X ⇒ Y) = sup(X ∪ Y). Negative association was first pointed out by Brin et al. in [6]. Since then, many techniques for mining negative associations have been developed [10, 14, 18]. In the case of negative associations we are interested in finding itemsets that have a very low probability of occurring together. That is, a negative association between two itemsets X and Y, denoted as X ⇒ ˥Y or Y ⇒ ˥X, means that X and Y appear very rarely in the same transaction. Mining negative association rules is computational intractable with a naive approach because billions of negative associations may be found in a large database while almost all of them are extremely uninteresting. This problem was addressed in [10] by combining previously discovered positive associations with domain knowledge to constrain the search space such that fewer but more interesting negative rules are mined. A general framework for mining both positive and negative association rules of interest was presented in [18], in which no domain knowledge was required and the negative association rules were given in more concrete expressions to indicate actual relationships between different itemsets. However, although the sets of the positive and negative itemsets of interest in the database were minimized in this framework, the search space for negative itemsets of interest was still huge. Another problem was that it tended to produce too many negative association rules, thus the practical application of this framework remained uncertain. An innovative approach has proposed in [34]. In this generating positive and negative association rules consists of four steps: (1) Generate all positive frequent itemsets L (P1) (ii) for all itemsets I in L(P1), generate negative frequent itemsets of the form ┑ (I1 I2) (iii) Generate all negative frequent itemsets ┑ I1 ┑ I2 (iv) Generate all negative frequent itemsets I1 ┑ I2 and (v) Generate all valid positive and negative association rules . Authors generated negative rules without adding additional interesting measure(s) to Support-Confidence frame work.

B. INDIRECT ASSOCIATION

Indirect association is a new kind of infrequent pattern, which provides a new way for interpreting the value of infrequent patterns and can effectively reduce the number of uninteresting infrequent patterns. The concept of indirect association is to "indirectly" connect two rarely co-occurred items via a frequent itemset called mediator, and if appropriately utilized it can help to identify real interesting "infrequent itempairs" from databases. Indirect association is closely related to negative association, they are both dealing with itemsets that do not have sufficiently high support. Indirect associations provide an effective way to detect interesting negative associations by discovering only "infrequent itempairs that are highly expected to be frequent" without using negative items or domain knowledge.

Definition (Indirect Association). A pair of itemsets X and Y is indirectly associated via a mediator M, if the following conditions hold:

1. sup(X, Y) < t_s (Itepair Support Condition)

2. There exists a non-empty set M such that

(a) sup(X ∪ M) ≥ t_f , sup(Y ∪ M) ≥ t_f ; (Mediator Support Condition)

(b) dep(X , M) ≥ t_d, dep(Y, M) ≥ t_d, where dep(P, Q) is a measure of the dependence between itemsets P and Q. (Mediator Dependence Condition)

The thresholds above are called itemset pair support threshold (t_s), mediator support threshold (t_f), and mediator dependence threshold (t_d), respectively. In practice, it is reasonably to set $t_f \geq t_s$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 5, May 2014

Condition 1 is needed because an indirect relationship between two items is significant only if both items rarely occur together in the same transaction. Otherwise, it makes more sense to characterize the pair in terms of their direct association.

Condition 2(a) can be used to guarantee that the statistical significance of the mediator setIn particular, for market basket data, the support of an itemset affects the amount of revenue generated and justifies the feasibility of a marketing decision. Moreover, support has a nice downward closure property which allows us to prune the combinatorial search space of the problem.Condition 2(b) ensures that only items that are highly dependent on the presence of x and y will be used to form the mediator set.

Over the years, many measures have been proposed to quantify the degree of dependence between attributes of a dataset. From statistics, the Chi-Square test is often used for this purpose. However, the drawback of this approach is that it does not measure the strength of dependencies between items [29]. Furthermore, the Chi-Square statistic depends on the number of transactions in the database. As a result, other statistical measures of association are often used, including Pearson's Φ coefficient, Goodman and Krushkal's λ, Yule's Q and Y coefficients, etc [26].

Interest factor is another measure that has been used quite extensively to quantify the strength of dependency among items [21, 22, 23].

Definition: Given a pair of itemsets, say X and Y, its' IS measure can be computed using the following equation:

$$IS(X,Y) = \frac{P(X,Y)}{P(X).P(Y)} \quad (1)$$

Where P denotes the probability that the given itemset appears in a transaction

C. INDIRECT NEGATIVE ASSOCIATION

Indirect association rule is said to be Indirect Negative Association Rule if mediator set used to generate it, contains both presence and absence of items. Thus, for a given indirect itemset pair X, Y and mediator itemset M ($=X^1Y^1$) where X^1 and Y^1 may be positive and/or negative itemsets. If both X^1 and Y^1 are positive then (X, Y/M) is said to be an indirect positive association rule otherwise it is said to be an indirect negative association rule.

III. RELATED WORK IN INDIRECT ASSOCIATION RULE MINING

It is observed that automated document translation systems tend to produce lexicon translation tables that are full of indirectly-associated words [25]. A lexicon translation table encodes the probability that two words from different languages being semantically equivalent to another. The presence of indirect association can pollute the resulting tables, thereby reducing the overall precision of the system. An iterative strategy was proposed in [25] to clean up existing translation tables by finding only the most probable translations for a given word.

The notion of internal and external measures of similarity between attributes of a database relation was introduced in [24]. Internal similarity between two attributes x and y is a measure whose value depends only on the values of x and y columns. Conversely, external measure takes into account data from other columns (called the probe attributes). Their notion of probe attributes is similar to mediators for indirect association in [24]. However, their sole purpose of using probe attributes is to perform attribute clustering.

Indirect association is closely related to the notion of negative association rules [27]. In both cases, we are dealing with itemsets that do not have sufficiently high support. A negative association rule discovers the set of items a customer will not likely to buy given that he/she bought a certain set of other items. Typically, the number of negative association rules can be prohibitively large and the majority of them are not interesting to a data analyst. The use of domain knowledge, in the form of item taxonomy, was proposed in [27] to decide what constitutes an interesting negative association rule. The intuition here is that items belonging to the same parent node in taxonomy are expected to have similar types of associations with other items. If the observed support is significantly smaller than its expected value, then there is a negative association exists between the items. Again, unlike indirect association, these types of regularities do not specifically look for mediating elements.

Another related area is the study of functional dependencies in relational databases. Functional dependencies are relationships that exist between attributes of a relation. However, the emphasis of functional dependencies is to find dependent and independent attributes for applications such as semantic query optimization [28] and reverse engineering [28].

In [31], authors proposed an efficient algorithm, called HI-mine, based on a new data structure, called HI-struct, for mining the complete set of indirect associations between items. Experimental results show that HI-mine's performance

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 5, May 2014

is significantly better than that of the previously developed algorithm for mining indirect associations on both synthetic and real world data sets over practical ranges of support specifications.

In [32], IAM algorithm proceeds in four phases: an initialization phase, a pruning phase, a bridge itemset calculation phase, and a ranking phase. The purpose of the initialization phase is to allocate the memory needed. The second phase is a process of pruning for the purpose of minimizing the search space of problem. The threshold value of pruning is min-sup(s). The third phase, the Bridge Itemset Calculation Phase, is the most important for this algorithm. The last phase, a ranking phase, is mainly to finish the ranking operation according to' the closeness value in the linked vector C for the purpose of providing decision makers the most useful indirect association rules

IV. ALGORITHM

In literature, a little work was done on generating indirect positive associations between pair of items only. In this paper, we propose a new method which generates indirect positive and negative associations between pair of itemsets. This method contains three algorithms. Algorithm1 finds set of all frequent itemsets and set of all Valid Candidates (VC). An itemset V is said to be Valid candidate if sup (V) $\leq t_s$ and all subsets of V are frequent. Algorithm 2 finds set of all indirect positive association rules between pairs of itemsets. Though Algorithms 1 and 2 were presented in [33], they are stated in this paper for self-containment. In this paper, proposed Algorithm 3 finds set of all indirect negative associations between pair of itemsets in which mediator set is of the form $\neg X^1 \neg Y^1$. In addition, Algorithm 4 finds set of all indirect negative associations between pair of itemsets in which mediator set is of the form $(\neg X^1)\ Y^1$.

Algorithm1 : Finding Positive Frequent(P), and ValidCandidates (VC)
Input: TDB- Transactional Database, m_s, t_s
Output: P- Positive Frequent itemsets, VC- ValidCandidates,
Method:
Find P_1, the set all frequent 1-itemsets
for(K=2;P_{k-1} != Φ ; K++)
{ $C_K = P_{K-1} \bowtie P_{K-1}$
// Pruning infrequent itemsets
for each c $\in C_K$ {
if any sub-set of c is not a member of P_{K-1} then $C_K = C_K - \{c\}$
}
// find positive frequent itemsets P_k and Valid Candidates (VC) in C_K
for each c in C_K {
if support(c)$\geq m_s$ then $P_k = P_k \cup \{ c \}$
if support(c)$\leq t_s$ then VC= VC U { c }
}
P= P U P_K
}
return P,VC

Algorithm2: Mining Indirect Positive Association
Input: P, VC, t_f, t_d, IAR= Ø
Output: Indirect Positive Association Rules
Method:
for each l (= X U Y) ∈ VC {
for each I ∈ P {
if (support(X U I) $\geq t_f$ && support (Y U I)$\geq t_f$)
if (dependency(X U I) $\geq t_d$ && dependency (Y U I)$\geq t_d$)
IAR= IAR U (X,Y/I)
} }
return IAR

Algorithm3: Mining Indirect Negative Association
Input: L_1- frequent 1-itemset, m_s-minsup, VC, t_f, t_d, IAR= Ø

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 5, May 2014

Output: Indirect Negative Association Rules
Method:

$C_2 = \{\rceil\{i_1\}\rceil\{i_2\} | i_1, i_2 \varepsilon L_1, i_1 \neq i_2\}$
for $(k = 2; C_k \neq \emptyset; k{++})$
{ for all $I = \rceil X \rceil Y \in C_k$
  { if $supp(I) \geq m_s$
    { **for** each $l\ (= X \cup Y) \in VC$
      { **if** ($support(X \cup I) \geq t_f$ && $support(Y \cup I) \geq t_f$)
        if ($dependency(X \cup I) \geq t_d$ && $dependency(Y \cup I) \geq t_d$
        IAR= IAR U (X,Y/I)
      }
    }
  else
  { for all $I \notin XY$ **do**
  Cand =check candidates(I, i) // i, is one the items of DB, is not a member of I
    $C_{k+1} = C_{k+1}$ U Cand
// S the set of positive itemsets whose supports are known
    if Cand $\neq \emptyset$, XY {i} $\notin$ S and $(\nexists I^1 \subseteq XY\{i\})(supp(I^1) = 0)$ **then**
    $S_{k+1} = S_{k+1}$ **U {** XY {i }}
  }
}
compute support of itemsets in S_{k+1}
$S = S \cup S_{k+1}$
}
return IAR;

Algorithm 4: Mining Indirect Negative Association
Input: L_1- frequent 1-itemset, m_s-minsup, VC- ValidCandidates, $t_{f\text{-}}$ mediator Support, t_d- mediator dependency, IAR= Ø
Output: Indirect Negative Association Rules
Method:

$C_{1,1} = \{\neg\{i_1\}\{i_2\} | i_1, i_2 \in L_1, i_1 \neq i_2\}$
for $\{k = 1; C_{k,1} \neq \emptyset; k{++}\}$
{ **for** $\{p = 1; C_{k,p} \neq \emptyset; p{++}\}$ // *$C_{k,p}$- k length negative and p length positive*
  { **for** all $I \in C_{k,p}$
    { **if** $supp(I) \geq m_s$
      { **for** each $l\ (= X \cup Y) \in VC$
        { **if** ($support(X \cup I) \geq t_f$ && $support(Y \cup I) \geq t_f$
          if ($dependency(X \cup I) \geq t_d$ && $dependency(Y \cup I) \geq t_d$
        IAR= IAR U (X,Y/I)
        }
      }
    }
  for all joinable $I_1, I_2 \in L_{k,p}$ do
  { $X = I_1$.negative, $Y = I_1$.positive$\cup I_2$.positive
    $I = \neg XY$
  if $(\nexists X^1 \subset X)(supp(\neg X^1 Y) \geq ms)$ and $(\nexists Y^1 \subset Y)(supp(\neg XY^1) < ms)$ then
    insert I into $C_{k,p+1}$
    if (XY $\notin$ S and $\nexists I^1 \subset XY$, $supp(I^1) = 0$ then $S_{k,p+1} = S_{k,p+1} \cup \{XY\}$
    }
  compute support of itemsets in $S_{k,p+1}$
  $S = S \cup S_{k,p+1}$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 5, May 2014

```
            }
            for all X ∈ L_{k+1}, i ∈ L_1 do
                if (∄X^1 ⊂ X)(¬X^1{i} ∈ L then  C_{k+1,1} = C_{k+1,1} ∪¬X{i}
        }
        return IAR;
```

V. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION

To evaluate the performance of proposed algorithm experiments are performed on two synthetic transactional databases containing 5400 and 12000 transactions each and implemented on java platform. We concentrate on mediator support(t_f) which is a support of itemset and mediator and mediator dependency(t_d) which is estimated by "Eq. (1)".

Data set consisting of 5400 transactions with mediator support as 0.2,0.25,0.3,0.35; mediator dependence as 0.4,0.45,0.5,0.55 and the total number of rules generated as 205,20,63,31 and 13 respectively. Figure 1 shows the graph showing the mediator support and mediator dependency vs. total number of rules

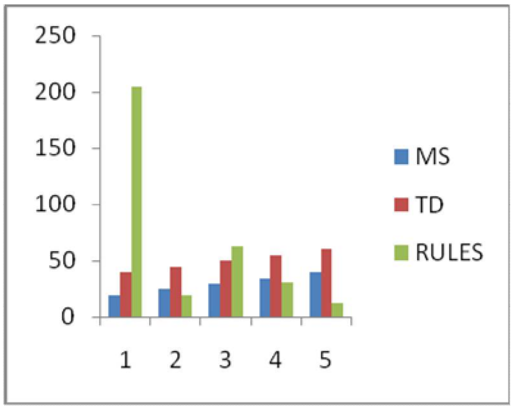

Figure 1: graph showing the mediator support and mediator dependency vs. total number of rules for 5400 transactions

Figure 2 is generated by considering 12000 transactions with mediator support as 0.2,0.25,0.3,0.35,0.4 mediator dependence as 0.4,0.45,0.5,0.55,0.6 and the total number of rules generated 35,31,7,7 and 6 respectively

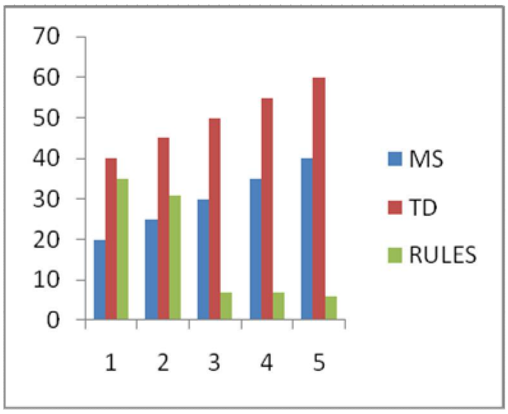

Figure 2: graph showing the mediator support and mediator dependency vs. total number of rules for 5400 transactions

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 5, May 2014

VI. CONCLUSION AND FUTURE WORK

Currently, available indirect association mining algorithms mine indirect positive associations between itempairs [14,16] and later we have extended to itemsets [33]. Further, in this paper, we propose an algorithm to discover all indirect positive and negative associations between itemsets. In indirect association mining for itempairs, algorithms require two join operations. To overcome this disadvantage we have proposed a new algorithm to mine indirect associations between itempairs and itemsets. This algorithm features, performing only one join operation and generating indirect positive and negative associations for pair of items and itemsets. In future we propose to elaborate this work by conducting experiments on large databases to test the scalability. Threshold selection is another issue that needs further investigation.

REFERENCES

1. R. Agarwal, C. Aggarwal, and V. V. V. Prasad. Depth first generation of long patterns. In Proceedings of ACM-SIGKDD International Conference on Knowledge Discovery and Data Mining, 2000.
2. R. Agarwal, C. Aggarwal, and V. V. V. Prasad. A tree projection algorithm for generation of frequent itemsets. In Journal of Parallel and Distributed Computing (Special Issue on High Performance Data Mining), 2000.
3. B.Ramasubbareddy, A.Govardhan, and A.Ramamohanreddy. Mining Positive and Negative Association Rules, IEEE ICSE 2010, Hefeai, China, August 2010
4. B.Ramasubbareddy, A.Govardhan, and A.Ramamohanreddy. Adaptive approaches in mining negative association rules. In intl. conference on ITFRWP-09, India Dec-2009.
5. B.Ramasubbareddy, A.Govardhan, A.Ramamohanreddy, An Approach for Mining Positive and Negative Association Rules, Second International Joint Journal Conference in Computer, Electronics and Electrical, CEE 2010
6. S. Brin, R. Motwani, J. Ullman, and S. Tsur. Dynamic itemset counting and implication rules for market basket data. In Proceedings of the International ACM SIGMOD Conference, pages 255–264, Tucson, Arizona, USA, May 1997.
7. P. Tan and V. Kumar. Interestingness measures for association patterns: A perspective. In KDD 2000 Workshop on Postprocessing in Machine Learning and Data Mining, Boston, MA, August 2000.
8. P. Tan and V. Kumar. Mining indirect associations in web data. In Proc of WebKDD2001: Mining Log Data Across All Customer TouchPoints, August 2001.
9. P. Tan, V. Kumar, and H. Kuno. Using sas for mining indirect associations in data. In Proc of the Western Users of SAS Software Conference, 2001.
10. A. Savasere, E. Omiecinski, and S. Navathe. Mining for strong negative associations in a large database of customer transactions. In Proceedings of the 14th International Conference on Data Engineering, pages 494–502, Orlando, Florida, February 1998.
11. P. Tan and V. Kumar. Interestingness measures for association patterns: A perspective. In KDD 2000 Workshop on Postprocessing in Machine Learning and Data Mining, Boston,MA, August 2000.
12. P. Tan and V. Kumar. Mining indirect associations in web data. In Proc of WebKDD2001:Mining Log Data Across All Customer TouchPoints, August 2001.
13. P. Tan, V. Kumar, and H. Kuno. Using sas for mining indirect associations in data. In Proc of the Western Users of SAS Software Conference, 2001.
14. P. Tan, V. Kumar, and J. Srivastava. Indirect association: mining higher order dependencies in data. In Proceedings of the 4th European Conference on Principles and Practice of Knowledge Discovery in Databases, pages 632–637, Lyon, France, 2000.
15. P. Tan, V. Kumar, and J. Srivastava. Selecting the right interestingness measure for association patterns. In Proceedings of the 8th International Conference on Knowledge Discovery and Data Mining, Edmonto, CA, July 2002.
16. Q. Wan and A. An. Efficient mining of indirect associations using hi-mine. In Proceedings of 16th Conference of the Canadian Society for Computational Studies of Intelligence, AI 2003, Halifax, Canada, June 2003.
17. Wong and C. J. Butz. Constructing the dependency structure of a multi-agent probability network. IEEE Transactions on Knowledge and Data Engineering, 13(3):395-415, 2001.
18. X. Wu, C. Zhang, and S. Zhang. Mining both positive and negative association rules. In Proceedings of the 19th International Conference on Machine Learning (ICML-2002), pages 658–665, Sydney, Australia, July 2002.
19. M. Zaki and M. Orihara. Theoretical foundations of association rules. In Proceedings of the 3rd ACM-SIGMOD Workshop on Research Issues in Data Mining and Kownledge Discovery, Seattle, WA, June 1998.
20. Z. Zheng, R. Kohavi, and L. Mason. Real world performance of association rule algorithms. In Proceedings of International Conference on Kownledge Discovery in Databases, August.
21. S.Brin,R.Motwani, and C.Silverstein. Beyond market baskets: Generalizing association rules to correlations. In Proc. ACM SIGMOD Intl.Conf.Management of Data, pages 265-276. Tuscon, AZ, 1997.
22. T.Brijs, G.Swinnen, K.Vanhoof, and G.Wets. Using association rules for product assortment decisions: A case study. In Proc.of the fifth ACM SIGKDD Conf on Knowledge Discovery and Data Mining, pages 254-260, San Diego, Calif, August 1999.
23. Robert Cooley, Chris Clifton.Topcat: Data mining for topic identification in a text corpus. In Procedings of the 3rd European Conference of Principles and Practice of Knowledge Discovery in Databases, 1999.
24. G.Das, H.Mannila, and P.Ronkainen. Similarity of attributes by external probes. In Proc. Of the Fourth ACM SIGKDD Intl Conf on Knowledge Discovery and Data Mining, pages 23-29, New york, NY, 1998.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 5, May 2014

25. D.Melamed. Automatic construction of clean broad-coverage translation lexicons. In 2nd Conference of the Association for Machine Translation in the Americas (ATMA 96) 1996.
26. H.T.Reynolds. The Analysis of Cross-Classifications. Macmillan Publishing Co., New York, 1997.
27. A. Savasere, E. Omiecinski, and S. Navathe. Mining for strong negative associations in a large database of customer transactions. In Proceedings of the 14th International Conference on Data Engineering, pages 494–502, Orlando, Florida, February 1998.
28. Z.Tari, o.Bukhres, J.Stokes, and S.Hammoudi. The reengineering of relational databases based on key and data correlations. In S.Sspaccapietra and F.Maryanski, editors, searching for semantics: Data Mining, Reverse Engineering, etc. Chapman and Hall, 1993.
29. R.Winkler and W.Hays. Statistics: Probabilty, Inference and Decision. Holt, Rinehart &Winston, New York, second edition, 1975.
30. G.E.Weddell. Reasoning about functional dependencies generalized for semantic data models. ACM Transactions on Database Systems, 17 (1): 32-64, March 1992.
31. Qian Wan, Aijun An, An Efficient Approach to Mining Indirect Associations, Kluwer Academic Publishers, Boston. Netherlands. Pages 1–26
32. Lei Li, Fanjiang Xu, Hongbing Wang, Chundong She and Zhihua Fan, IAM: An Algorithm of Indirect Association Mining, Proceedings of the 2004 International Conference on Intelligent Mechatronics and Automation Chengdu.China August 2004
33. B.Ramasubbareddy,A.Govardhan,A.Ramamohanreddy, Mining Indirect Association between Itemsets, proceedings of Intl conference on Advances in Information Technology and Mobile Communication-AIM-2011 published by Springer LNCS , April 21-22, 2011, Nagapur, Maharastra, India
34. Chris Cornelis, peng Yan, Xing Zhang, Guoqing Chen: Mining Positive and Negative Association Rules from Large Databases , IEEE conference 2006.