# A NOVEL METHOD FOR FINDING PRIVACY PRESERVING ASSOCIATION RULE MINING

N V Muthu lakshmi[1]

[1] Research Scholar: Dept of Computer Science
S.P.M.V.V,
Tirupati, Andhra Pradesh, INDIA
nvmuthulakshmi@yahoo.co.in

Dr. K Sandhya Rani[2]

[2] Professor: Dept of Computer Science
S.P.M.V.V,
Tirupati, Andhra Pradesh, INDIA
sandhyaranikasireddy@yahoo.co.in

**Abstract**

Association rule mining is one of the significant research areas which explore the association between various item sets. The issue of privacy arises when several legitimate people share their data or knowledge for their mutual benefits. In case of centralized database, sensitive item sets are hidden by using association rule hiding approaches. Based on the execution time, the degree of optimality, the level of tolerance of side effects and guaranteed to get solution different association rule hiding approaches are exist. Among the commonly used approaches, heuristic approach is widely used since it guarantees to provide solution but causes some side effects. In this paper a heuristic based methodology is proposed to hide the sensitive item sets efficiently by adopting two criterions. This methodology protects private information by doing sanitization process but before participating in the sanitization process, the method analysis the side effects and select the most promising one to change so that side effects can be fully avoided or accepting few side effects which will not harm the informational accuracy.

***KEYWORDS: Association Rule Hiding, Heuristic Approach, Privacy Preservation, Centralized Database***

## 1. Introduction

Association rule mining is one of the important and widely used data mining techniques to explore hidden valuable information such as associations between item sets. In recent years, many organizations often share their information with legitimate parties to discover more useful information for decision making purposes and to enhance their competitive spirit. The issue of Privacy plays a vital role when several legitimate people share their resources in order to obtain mutual benefits but no one is interested to disclose their private data. In centralized database environment, various approaches are proposed for privacy preserving data mining and which can be categorized into data hiding and knowledge hiding approaches. In data hiding, the main concept is how the privacy of

data/information can be maintained before hiding process begins. In this approach, removal of private/sensitive information prior to its disclosure by adopting the techniques such as perturbation, sampling, generalization etc. to produce a sanitized database of the original database. In case of knowledge hiding approaches, instead of protecting the raw data, sensitive information from data mining results are protected by using distortion and blocking techniques.

Association rule hiding problem can be specified as follows:

Given a database D, and a set of sensitive rules SR, the aim of Association rule hiding is to prepare a sanitized database D0 from D such that when mining is performed on D0, all sensitive rules SR will be hidden and only non sensitive rules will be disclosed.

The goal of association rule hiding is to achieve the following

- No sensitive rule should be revealed when mining process is performed on sanitized database. In other words all sensitive rules must be hidden from the users.
- All non sensitive rules must be generated with sanitized database when mining is performed that is non sensitive rules should not be hidden from sanitized database.
- No new rules should be generated with sanitized database. In other words false rules should not be generated as a side effect of the sanitization process.

Association rule hiding approaches and earlier works are presented in the next section.

## 2. Association Rule Hiding

The three categories of Association rule hiding approaches are Heuristic, Border based and Exact approaches. Heuristic approaches have been getting focus of attention for majority of the researchers due to their efficiency, scalability and quick responses. However in some circumstances these heuristic based approaches suffer from undesirable side effects. So these side effects may degrade the performance of the hiding process of sensitive association rules. Heuristic approaches can be further subdivided into sensitive transaction identification methods, sensitive association clustering methods and sanitization matrix methods. The border based approaches utilizes the concept of borders to track the impact of altering transactions by greedy selecting those modifications while minimizing the side effects. These approaches focus on preserving the border of non sensitive frequent item sets rather than considering all non sensitive item sets during sanitization process. Third class of approach is non heuristic called exact, which conceive hiding process as constraint satisfaction problem. These problems are solved by integer programming. Compared to heuristic and border based, this guarantees quality for hiding sensitive information.
In this paper, a heuristic based method is proposed to hide sensitive association rules with limited side effects. The various works under heuristic approach are discussed as follows:

In [1], the authors presented various ideas to protect the privacy of individuals. An approach for hiding rules by replacing selected values by replacing selected value with unknowns instead of replacing with false values is proposed in [2]. A method to hide a rule by decreasing its support or confidence is proposed in [3 ]. In this method, the support or confidence is decreased one unit at a time, values of one transaction at a time are modified.
A sanitization technique is presented by the authors to block forward inference attack and backward inference attack to hide sensitive rules [4]. In [5], the authors presented a work which is the extension work of dassineni.et.al by improving and evaluating the association rule hiding algorithms which protects the sensitive knowledge by hiding the items whose support is maximum among the minimum length transactions. Multiple rule hiding approach is first proposed by the authors in [6]. These algorithms are efficient and require two scans of the database irrespective of

the number of sensitive items to hide. In [7], the author proposed three multiple association rule hiding heuristics that outperforms SWA by offering higher data utility and lower distortion. Data distortion approaches that operates on a sanitization matrix and then multiply with original database to obtain a sanitized database is presented in [8]. In this paper, the authors developed three sanitization algorithms namely hidden-first, non-hidden-first and hiding sensitive patterns completely with minimum side effects on non heuristic patterns. In [9], The authors proposed two new algorithms which rely on the maxmin criteria for the hiding of sensitive itemsets in an association rule hiding framework. These algorithms use the maxmin criteria to minimize the impact of the hiding process to the revised positive border [10]. By doing this it is possible to efficiently select items which must be hiding, while at the same time it ensures that non-border item sets are protected from hiding. In [10], the authors discussed the major categories of sensitive knowledge protecting methodologies such as Heuristic based, Border based, Constraints Satisfaction problem based and reconstruction based approaches. The authors also presented earlier works related to each category.

The following two criteria are introduced in the proposed novel method to hide the sensitive item sets efficiently. The Criteria1 specifies the efficient selection of victim item and Criteria2 helps to find the suitable supporting transaction for victim item in the sanitization process which minimizes the side effects.

**Criteria1:**
Suppose the item set $< A_i, A_j>$ is to be hidden, one can select either $A_i$ or $A_j$ as victim item which minimizes side effects. Victim item can be selected based on the following condition.
If number of times $<A_i>$ appears in non sensitive frequent item set is greater than number of times $<A_j>$ appears in non sensitive frequent item sets then $A_j$ be the victim item. If number of times $<A_i>$ appears in non sensitive frequent item set is less than number of times $<A_j>$ appears in non sensitive frequent item sets then $A_i$ be the victim item. If number of times $<A_i>$ appears in non sensitive frequent item set is equal to number of times $<A_j>$ appears in non sensitive frequent item sets then select Ai or Aj randomly as a victim item.
**Criteria2:**
After identifying the victim item to hide item set $<A_i, A_j>$, the minimum number of suitable transactions has to be selected from all supporting transactions for the item set $<A_i,A_j>$. The minimum number of transactions required to hide item set is based on the value of $<A_i,A_j>$.supp – MinTrans + 1. Once the minimum number of transactions is identified then one has to identify suitable transactions in such a way that least side number of side effects will occur over non sensitive frequent item sets. For each supporting transactions for item set $<A_i,A_j>$, weight is computed by using the following:
$W(T_g)$ = Number of dependant items with victim item - number of infrequent item sets associated with victim item.
Based on the weights of the transactions, the supporting transactions are sorted in ascending order and stored in MinT. The transactions in MinT are selected orderly for sanitization purpose to hide the sensitive item set $<A_i,A_j>$.

The proposed method utilizes the above Criterion1 and Criteria2 to hide sensitive item sets with minimum side effects is given in the next section.

## 3. Proposed Algorithm

To find privacy preserving association rule mining for a given database and the set of sensitive item sets, a methodology which adopts two criterions is proposed in this paper to obtain a distorted database which hides all sensitive item sets. In this methodology, split pattern is used which is taken from [11]. The authors in this paper suggested a procedure in which all the sensitive item sets whose length is greater than two are considered to find the pairs sub patterns. From this pair sub patterns only significant pair sub patterns are considered as sensitive to hide the sensitive patterns. This procedure is very important in a way that it avoids the problem of forward inference attack. In order to avoid forward inference attack problem, at least one sub pattern with length of two of the patterns

should be hidden. This split pattern procedure helps to speed up the hiding process. The terminology used in the proposed method is specified in the following table.

TABLE 1: Terminology used in proposed model

| S.NO | Terminology | Explanation |
|---|---|---|
| 1 | DB = $\{t_1,t_2,..t_N\}$ | A original database consisting of N number of transactions |
| 2 | I =$\{i_1,i_2,\ldots i_M\}$ | An item set of length M |
| 3 | $L_k$ | An item set of length k |
| 4 | $T_{nm}$ | The $n^{th}$ transaction of $m^{th}$ item |
| 5 | S =$\{ s_1, s_2, \ldots s_r\}$ | Set of sensitive item sets |
| 6 | MinSupport | User specified Minimum support threshold |
| 7 | Supp(J) | Number of transactions supporting item set J |
| 8 | MinTrans | Based on MinSupport, number of transactions required to support an item set to be frequent |
| 9 | MinConfidence | User specified Minimum confidence threshold |
| 10 | N | Size of original database, DB |
| 11 | $F_{DB}$ =$\{ L_1, L_2, L_3, \ldots L_k\}$ | A set consists of all frequent item sets |
| 12 | A → B | Association rule between item sets A and B |
| 13 | $F_S$ | The set consisting of sensitive item sets |
| 14 | $F_{NS}$ | The Set consisting of non sensitive frequent item sets |
| 15 | $F_{2S}$ | The set consisting of pairs determined by the procedure split pattern. |
| 16 | $<A_i,A_j>$ | The sensitive item set pair |
| 17 | $T_{AiAj}$ | Set of supporting transactions for item set $<A_i,A_j>$ |
| 18 | DB' | Distorted database which hides all sensitive item sets. |
| 19 | Victim item | An item which is selected from the sensitive item pair which produces least side effects or no side effects when modification is done over it. |
| 20 | Victim transactions | The transactions are selected to change. |
| 21 | MinT | A set consisting of suitable number transactions, which are to be modified to hide the sensitive item set |
| 22 | Count | Count gives number of times the victim item value has to be modified to hide sensitive item set pair. |
| 23 | $W(T_g)$ | Weight for transaction $T_g$ |

The algorithm for the proposed methodology is specified as follows:

Step 1. For a given database, DB and set of sensitive item sets, $F_s$ generate frequent item sets and store with their support values in $F_{DB}$.

Step 2. Let the sensitive item sets stored in $F_s$ then the non sensitive frequent item sets are obtained by subtracting $F_S$ from $F_{DB.}$

ie., $F_{NS} = F_{DB} - F_{S.}$

Step3. If any item sets in $F_S$ are having more than length of two, call the procedure split pattern to identify the

prominent pairs which are to be hidden in order to hide all the item sets whose length is greater than two.

Step4. After step 3 a vector $F_{2S}$ is prepared which consists of all two pair sensitive items.

Step5. The generated all pairs sensitive frequent item sets with their support values along with their supporting t transactions ID's are stored in a table $T_{S.}$

Step6. All the non sensitive frequent item sets that is F- $F_{2S}$ are stored along with their support values in a table $T_{NS.}$

Step7. For each item set in $F_{2S}$

If any non overlapping item set exists

go to step 12.

Else

The patterns $<A_i,A_j><A_j,A_k>$ are chosen

Consider $A_i$ or $A_j$ as victim item based on Criteria1

Step8. Find the intersection of supporting transactions for $A_iA_j$ and $A_jA_k$ as follows:

$T_{AiAjAk} = T_{AiAj} \cap T_{AjAk}$

Step9. Obtain the value for Count1 and Count2 as follows:

Count1 for $A_iA_j = <A_i,A_j>.Supp - MinTrans + 1$

Count2 for $A_jA_k = <A_j,A_k>.Supp - MinTrans + 1$

Step10., select minimum number of supporting transactions to be modified from a set MinT which is obtained by Criteria2.

From these counts, select smaller one and that many transactions selected from MinT and the victim item $A_j$ values are replaced with zero values. By performing this item set which has lower count will be hidden. To hide the item set which is having higher count value Count1 – Count2 number of transactions which are not yet processed will be selected from MinT for sanitization. To hide this item set the victim item set can be selected based on their dependencies with item sets in non sensitive item sets. Accordingly the victim item value will be replaced with zero in the selected transactions. With this the higher count item set is also hidden.

Step11. Modify $F_{2S}$ by removing the pairs $<A_i,A_j>$ and $<A_jA_k>$ from it. Go to step18.

Step12. For the sensitive item set pair $<A_i,A_j>$ in $F_{2S}$ find victim item by using criteria 1.

Step13. After identifying the victim item, find the supporting transactions for <Ai,Aj>.

Step14. Obtain the value for Count1 and Count2 as follows:

Count1 for $A_iA_j = <A_i,A_j>.Supp - MinTrans + 1$

Step15. Select Count1 number of transactions to be modified from a set MinT which is obtained by Criteria2.

Step16. The value of victim item in the selected transactions is replaced with zero value.

Step17. Update $F_{2S}$ by removing $<A_i,A_j>$ from it.

Step18. Repeat the above steps from step 7 until no more pair in the $F_{2S}$ to hide the remaining pairs of sensitive item sets in $F_{2S.}$

Step19. Finally distorted database, DB´ is obtained which hides all sensitive item sets in $F_{2S}$.

Step20. Stop the process.

4. Implementation of the Proposed Model

The proposed model is illustrated with sample database which consists of 5 attributes also called items for 8 transactions. Each transaction is represented by its TID value. The following table shows sample database.

Table 2: Sample Database

| TID\Item | $A_1$ | $A_2$ | $A_3$ | $A_4$ | $A_5$ |
|---|---|---|---|---|---|
| $T_1$ | 1 | 0 | 1 | 1 | 0 |
| $T_2$ | 0 | 1 | 0 | 1 | 1 |
| $T_3$ | 0 | 0 | 0 | 0 | 1 |
| $T_4$ | 1 | 1 | 0 | 0 | 0 |
| $T_5$ | 0 | 0 | 1 | 1 | 0 |
| $T_6$ | 0 | 0 | 1 | 1 | 1 |
| $T_7$ | 1 | 1 | 0 | 1 | 1 |
| $T_8$ | 0 | 0 | 1 | 0 | 0 |

Using apriori algorithm, frequent item sets are generated for the sample database based on user specified minimum support threshold value 40% and the results are given in table3.

Table 3: Frequent item sets and its support values

| Item set | Supp | Item set | Supp |
|---|---|---|---|
| $A_1$ | 3 | $<A_1,A_4>$ | 2 |
| $A_2$ | 3 | $<A_2,A_4>$ | 2 |
| $A_3$ | 4 | $<A_2,A_5>$ | 2 |
| $A_4$ | 5 | $<A_3,A_4>$ | 3 |
| $A_5$ | 4 | $<A_4,A_5>$ | 3 |
| $<A_1,A_2>$ | 2 | $<A_2,A_4,A_5>$ | 2 |

**Case I: Non Overlapping Pair Patterns**

Assuming the sensitive item sets$<A_1,A_4>$, $<A_2,A_5>$, which are to be hidden.

Let $F_S = \{<A_1,A_4>, <A_2,A_5>\}$

By invoking split procedure, we get $F_{2S} = \{ <A_1,A_4> , <A_2,A_5>\}$

Since the pairs in $F_{2S}$ are non overlapping patterns, each pair hast to be hidden individually.

Let us take the pair $<A_1,A_4>$ and find the victim item using Criteria1.

For item $<A_1>$, One time appeared in non sensitive frequent item set

For item $<A_4>$, In three non sensitive frequent item sets, $A_4$ is appeared.

So victim item is $A_1$. We have to replace 1 with zero for $A_1$.

Now we have to find suitable supporting transactions to change values of $A_1$ in order to hide $<A_1,A_4>$

Supporting transactions for $<A_1,A_4> = \{T_1,T_7\}$

Minimum number of supporting transactions of $<A_1,A_4>$ are to be determined to update $A_1$ values

To determine suitable and minimum number of transactions, count value is determined as

Count = $<A_1,A_4>$ . supp – MinTrans + 1

= 2 – 2 + 1 = 1

∴ Minimum number of transactions required to update $A_1$ is = 1

Now Criteria2 is applied to find which transaction is suitable to modify $A_1$ so that side effects are minimized and shown as follows:

For $T_1$, weight can be computed as

$W(T_1) = -1$ ∵ $<A_1,A_4>$ is a non sensitive frequent item set.

For $T_7$, weight can be computed as

$W(T_7) = +1-1=0$ ∵ $<A_1,A_2>$ is non sensitive frequent item set and $<A_1,A_5>$ is infrequent item set.

Sort the transactions in ascending order based on weight but both has same weight so we can select any transaction to modify.

Sored transactions is $\{T_1,T_7\}$.

$T_1$ is selected for modification to hide $<A_1,A_4>$

Now modify $A_1 = 0$ at $T_1$ to hide $<A_1,A_4>$

This modification changed item pair set $<A_1,A_3>$ support is decreased by one but still it is infrequent and no side effect occurred.

Hence $<A_1,A_4>$ pair is hidden and no side effects occurred.

Let us take the second pair $<A_2,A_5>$ By using Criteria1, victim item is selected.

For $A_2$, Two non sensitive item sets are associated with $A_2$.

For $A_5$, Only one non sensitive item set is associated with $A_5$.

∴ victim item is $A_5$.

Once victim item is selected, the next task is to find minimum number of transactions required to modify victim item so that $<A_2,A_5>$ is hidden with minimum side effects.

Supporting transactions for $<A_2,A_5> = \{T_2,T_7\}$

Minimum number of supporting transactions of $<A_2,A_5>$ are to be determined to modify $A_5$.

To determine suitable and minimum number of transactions, count value is determined as

Count = $<A_2,A_5>$ . supp – MinTrans + 1

= 2 – 2 + 1 = 1

∴ Minimum number of transactions required to update $A_5 = 1$

Now Criteria2 is applied to find which transaction is suitable to modify $A_5$ so that side effects are minimized and shown as follows:

For $T_2$, weight can be computed as

$W(T_2) = +1$ ∵ $<A_4,A_5>$ is a non sensitive frequent item set.

For $T_7$, weight can be computed as

$W(T_7) = +1-1 = 0$ ∵ $<A_4,A_5>$ is non sensitive frequent item set and $<A_1,A_5>$ is infrequent item set

Sort the transactions in ascending order based on weight

Sorted transactions is $\{T_7,T_2\}$. From this $T_7$ is selected for modification to hide $<A_2,A_5>$

Now modify $A_5 = 0$ at $T_7$ to hide $<A_2,A_5>$

This modification changed item pair set $<A_4,A_5>$ support is decreased by one but still it is frequent.

Hence $<A_2,A_5>$ pair is hidden and no side effects occurred.

From the pairs of sensitive frequent item sets in $F_{2S}$, no side effects occurred.

By doing the above process, distorted database is obtained and shown in table 4.

Table 4: Distorted Database, DB´

| TID\Item | $A_1$ | $A_2$ | $A_3$ | $A_4$ | $A_5$ |
|---|---|---|---|---|---|
| $T_1$ | 0 | 0 | 1 | 1 | 0 |
| $T_2$ | 0 | 1 | 0 | 1 | 1 |
| $T_3$ | 0 | 0 | 0 | 0 | 1 |
| $T_4$ | 1 | 1 | 0 | 0 | 0 |
| $T_5$ | 0 | 0 | 1 | 1 | 0 |
| $T_6$ | 0 | 0 | 1 | 1 | 1 |
| $T_7$ | 1 | 1 | 0 | 1 | 0 |
| $T_8$ | 0 | 0 | 1 | 0 | 0 |

**Case II: Overlapping Patterns**

Let the sensitive frequent item sets be $<A_2,A_4>$, $<A_4,A_5>$ and which are to be hidden.

Let $F_S = \{ <A_2,A_4>, <A_4,A_5>\}$ and find significant pairs of sensitive patterns by calling the split pattern procedure.

We get $F_{2S} = \{<A_2, A_4>, <A_4, A_5>\}$ and it is clear that $A_4$ appeared in both these pairs.

$\therefore$ victim item is $<A_4>$

Supporting transactions for $<A_2,A_4>$ is $\{T_2,T_7\}$

Supporting transactions for $<A_4,A_5>$ is $\{T_2, T_6,T_7\}$

Now we find common transactions by intersecting these two supporting transactions as

$= \{T_2,T_7\} \cap \{T_2, T_6,T_7\} = \{T_2,T_7\}$

Now we find minimum number of transactions required to hide these pairs by determining Count1 Count2 values and computations are shown below:

Count1 = $< A_2,A_4>$.Supp – MinTrans +1= 2 – 2 + 1 = 1

Count2 = $<A_4,A_5>$.Supp – MinTrans +1= 3 - 2 + 1 = 2

$\therefore$ Minimum number of transactions required to hide $< A_2,A_4>$ is only one.

$\therefore$ Minimum number of transactions required to hide $< A_4,A_5>$ is two.

Count2 is greater than Count1.

One suitable supporting transaction is required to modify $<A_4>$ value and this transaction can be computed as

For $T_2$, Weight can be computed as

$W(T_2)$ for $A_4$ is 0

For $T_7$, Weight can be computed as

$W(T_7)$ for $A_4$ is +1

Sort the transactions in ascending order based on weight

Sorted transactions is $\{T_2,T_7\}$. From this $T_2$ is selected for modification to hide $<A_2,A_4>$ and $<A_4,A_5>$.

Now modify $A_4 = 0$ at $T_2$.

This modification changed item pair sets $<A_2, A_4>$ and $<A_4, A_5>$ support is decreased by one and $<A_2,A_4>$ is hidden but $<A_4,A_5>$ is not hidden due to its support value is greater than MinSupport threshold.

This causes <$A_2$, $A_5$> pair is hidden and no side effects occurred.

Since Count2-Count1 is not zero, we have to find victim item using Criteria1 to hide item pair <$A_4$, $A_5$>.

Since $A_4$ > $A_5$ according to dependencies 2 > 1

∴ $A_5$ is the victim item.

Supporting transactions = {$T_2$,$T_6$,$T_7$}

Minimum number of suitable transactions can be determined by using Criteria2 is as

$W(T_2)$= +1 , $W(T_6)$= -1, $W(T_7)$= +1+1-1=1

Sorted transactions= {$T_6$, $T_2$, $T_7$}

From this set $T_6$ is selected to modify $A_5$ value to hide <$A_4$,$A_5$>

Now modify $A_5$=0 in transaction $T_6$ and causes no side effect.

By doing the above process, distorted database is obtained and shown in table 5.

Table 5. Distorted Database, DB´

| TID\Item | $A_1$ | $A_2$ | $A_3$ | $A_4$ | $A_5$ |
|---|---|---|---|---|---|
| $T_1$ | 1 | 0 | 1 | 1 | 0 |
| $T_2$ | 0 | 1 | 0 | 0 | 1 |
| $T_3$ | 0 | 0 | 0 | 0 | 1 |
| $T_4$ | 1 | 1 | 0 | 0 | 0 |
| $T_5$ | 0 | 0 | 1 | 1 | 0 |
| $T_6$ | 0 | 0 | 1 | 1 | 0 |
| $T_7$ | 1 | 1 | 0 | 1 | 0 |
| $T_8$ | 0 | 0 | 1 | 0 | 0 |

Hence the above overlapping patterns has no side effects and with minimum number of changes in the original database.

## 5. Conclusion

Privacy preserving association rule mining is a challenging task to researchers since many side effects occur when privacy is preserved in the database. Side effects cannot be avoided because correlation exists between item or item sets. In this paper a novel method is proposed related to heuristic approach to hide sensitive association rules specified by the users with minimum side effects. Two criterions are suggested in this paper to identify the victim item and selecting suitable supporting transactions efficiently for sanitization purposes. The functionality of the proposed method is illustrated with sample database by considering two cases related to existence of non overlapping and overlapping sensitive patterns. Especially in case of overlapping patterns, the Criteria1 and Criteria2 are useful to speed up the process of hiding sensitive item sets.

## References:

[1] C. Clifton and D. Marks. Security and privacy implications of data Mining (Feb'1996), In Workshop on Data Mining and Knowledge Discovery,pages 15–19.

[2] Y. Saygin, V. S. Verykios, and C. Clifton, Using unknowns to prevent discovery of association rules (Dec'2001), In ACM SIGMOD Record, volume 30(4), pages 45–54.

[3] E. Dasseni, V. S. Verykios, A. K. Elmagarmid, and E. Bertino. Hiding association rules by using confidence and support (April 2001). In I. S. Moskowitz,editor, Proceedings of the 4th Information Hiding Workshop, volume 2137, pages 369–383, Pittsburg, PA, USA, Springer Veralg, Lecture Notes in Computer Science.

[4] S. R. M. Oliveira, O. R. Zaiane, and Y. Saygin. Secure association rule sharing. In H. Dai, R. Srikant, and C. Zhang, editors, Proceedings of the 8th PAKDD Conference, volume 3056, pages 74–85, Sydney, Australia, May 2004. Springer Veralg Lecture Notes in Computer Science.

[5] V. S. Verykios, A. K. Elmagarmid, B. Elisa, Y. Saygin, and D. Elena. Association rule hiding. In IEEE Transactions on Knowledge and Data Engineering, volume 16(4), pages 434–447, Los Alamitos, CA, USA, April 2004. IEEE Computer Society.

[6] Oliveira SRM, Zaiane OR, Privacy Preserving Frequent Item set Mining (2002), In Proceeding of IEEE, International Conference on Privacy, Security and Data mining, Australia, PP 43-54. Amiri A, Data to Share: Protecting sensitive Knowledge with data sanitization, Decision Support Syst 2007,43 181-191.

[7] Guanling lee, Chien – Yu Chang, Arbee L.P Chen, Hiding Sensitive Patterns in Association Rules Mining (2004), Proceedings of the 28th Annual International Computer Software and Applications Conference (COMPSAC'04), IEEE.

[8] George V. Moustakides, Vassilios S. Verykios, A Max Min Approach for Hiding Frequent Item sets , Data & Knowledge Engineering, Science Direct, 65 (2008) PP 75-89.

[9] Guanling lee,Yi Chun Chen, Protecting Sensitive Knowledge in Association Patterns Mining, WIREs Data Mining Knowledge Discovery 2012,Vol 2, PP 60-68.

[10] En Tzu Wang, Gaunling Lee, Yu Tzu Lin, A Novel Method for Protecting Sensitive knowledge in Association Rules mining, from Ph.D., thesis , National Dong Hwa University Hualien, Taiwan.