Association Rule Hiding Techniques for Privacy Preserving Data Mining: A Study

Gayathiri P
Research Scholar,
Department of Computer Science,
Bharathiar University, Coimbatore-641 046
TamilNadu, India

Dr. B Poorna
Principal,
SSS Jain College for Women,
T.Nagar, Chennai,
TamilNadu, India

***Abstract*—Association rule mining is an efficient data mining technique that recognizes the frequent items and associative rule based on a market basket data analysis for large set of transactional databases. The probability of most frequent data item occurrence of the transactional data items are calculated to present the associative rule that represents the habits of buying products of the customers in demand. Identifying associative rules of a transactional database in data mining may expose the confidentiality and privacy of an organization and individual. Privacy Preserving Data Mining (PPDM) is a solution for privacy threats in data mining. This issue is solved using Association Rule Hiding (ARH) techniques in Privacy Preserving Data Mining (PPDM). This research work on Association Rule Hiding technique in data mining performs the generation of sensitive association rules by the way of hiding based on the transactional data items. The property of hiding rules not the data makes the sensitive rule hiding process is a minimal side effects and higher data utility technique.**

Keywords*—Association rule mining; transactional data; privacy preservation; Association Rule Hiding (ARH); Privacy Preserving Data Mining (PPDM)

I. Introduction

Data belongs to a person or an organization may have different sensitive levels. These data are made available only for authorized persons. So ensuring the protection of sensitive data by access restriction is not a complete method. This may affect the utility of the data mining result and with help of the knowledge the user may re-identify sensitive data items from non-sensitive data is known as Inference Problem. The privacy preserving data mining is to provide a solution for protecting sensitive information by developing a data mining techniques which could be applied on databases without affecting the accuracy of data mining result and without violating the privacy of individuals is the motivation for this research.

Data mining is the method of determining patterns in large data sets with artificial intelligence, machine learning, statistics and database systems. The aim of data mining process is to extract information from a huge volume of data set to have logical structural representation of the data item in the transactional database. It is utilized to mine significant and useful information or knowledge from large database. Protected or private information extracted by data mining methods leads to the risk of threats to privacy. Association rule mining is a technique in data mining to recognize the regularities created in large volume of data. The method is cooperated by allowing third party to recognize and disclose hidden private information for an individual or organization.

Privacy-preserving data mining with association rule denotes the area of data mining that looks to preserve sensitive information from unnecessary or unlawful disclosure. Privacy information comprises personal or confidential information in business like social security numbers, home address, credit card numbers, credit ratings, purchasing behavior, medical records and best-selling services. The privacy preservation data mining requires guarantee for hiding of sensitive information in efficient manner. The association rule hiding technique protects the sensitive data indirectly under the scanner. Also it fails to hide data items which are not sensitive. It affects the privacy of rules and the utility of the data mining results.

This paper is organized as follows: Section 2 discusses survey with existing techniques of Association Rule Hiding (ARH) for Privacy Preserving Data Mining (PPDM), Section 3 shows the Association Rule Hiding (ARH) for Privacy Preserving Data Mining (PPDM), Section 4 identifies the possible comparison between them, Section 5 discusses about the limitations of the existing techniques and Section 6 concludes the paper, key areas of research is given for improving the selection of sensitive rules for enhancing the business transactions. It also preserves the association rules for maintaining the privacy in database.

II. Literature Survey

Privacy-Preserving Data Mining of Association Rules from Outsourced Transaction Databases technique [3] is developed with an encryption scheme. Encryption/Decryption (E/D) model was used to change the client data before it is shipped to the server. But, the mined results are not intended for sharing and remain exposed. Attack able to identify the intricacies of the rule preservation and data item property supports are not true supports. To Secure Association Rules, Secure Multi-party Computation (SMC) algorithm [4] is introduced to hide the association rules in a horizontally distributed database. The combination of private item subsets is calculated using SMC algorithm. Though, secure protocol is not relying on the commutative encryption and transfer. The SMC algorithm fails to secure the transaction items.

A perturbation-based PPDM with Multilevel Trust (MLT-PPDM) [5] is developed to preserve the privacy of data and association rules at different levels. This method preserved multiple perturbed copies, data miner perform resistance to diversity attacks and reconstruct the original data more accurately. MLT-PPDM permits the data owners for designs perturbed copies of data for different trust levels. However, the data set does not re-anonymize after it is updated with insertions and deletions. Efficiently Hiding Sensitive Item set with Transaction Deletion Based on Genetic Algorithms [2] is planned to enhance the chosen transactions deleted, so minimizing the side effects in Privacy-preserving data mining (PPDM) technique. But, predefined item set and a missing item set are non-sensitive item set that affect the rules being disclosed.

A Hiding Sensitive Association Rules [1] with Limited Side Effects are designed for PPDM. Heuristic method is involved for raising the hidden sensitive rules quantity in hiding sensitive association rule. The side effects minimized are not taken for correlation among rules that wipes out the creativity of the association rule. Privacy preserving data mining attains data mining goals without showing the privacy information of the individuals to the public users. A novel Hiding-Missing-Artificial Utility (HMAU) algorithm [6] is designed for hiding the sensitive itemsets during the transaction deletion process. Privacy preserving data mining (PPDM) is presented to hide the sensitive information. HMAU algorithm reduces the side effects through transaction depletion and the transaction with minimal HMAU value removed from the database. But, the noise addition and data modification are the significant problems to hide the sensitive information in PPDM.

III. Association Rule Hiding Techniques for Privacy Preservation

Privacy Preserving Data Mining (PPDM) is used to extract relevant knowledge from large amount of data and protects the sensitive information from the data miners simultaneously. Privacy preserving data mining is a hot spot in data mining. Privacy Preserving Data Mining (PPDM) solves the issues of designing accurate models about combined data without access to exact information in individual data record. Association Rule Hiding is a PPDM technique use with Association Rule Mining method in transactional database.

An itemset is a set of products and transaction maintains simultaneously for a given set of items. The support of an itemset I in a transaction database is the percentage of transactions having I in the whole database. An itemset is frequent when the support is higher than a minimum support threshold (MST).

For two itemsets X and Y where $X \cap Y = \emptyset$ .The confidence of an association rule $X \rightarrow Y$is the probability that number of times Y occurs given that X occurs is equal to $Sup_{X \cup Y}$divided bySup_X. When $X \rightarrow Y$ holds in the database if $X \cup Y$ is frequent and its confidence is higher than a minimum confidence threshold (MCT). This rule is called the strong association rule. Association rule mining is used to discover all strong rules in the database.

A. Association Rule Mining using Selection Technique for Sensitive Rules

Privacy-preserving data mining (PPDM) is designed to minimize the privacy threats. Privacy threats are decreased by sensitive information hiding process from databases. These types of data having the confidential information result in the privacy threats when the data gets misused. Heuristic methods are used to choose the suitable data for sanitization to hide the sensitive information. In hiding the sensitive information process, side effects of missing cost and artificial cost are created. The beneficial method is used to choose the hidden sensitive information based on the NP-hard problem in sanitization process.

An extensive work on privacy-preserving data mining (PPDM) is carried out in various contexts. A general characteristic of frameworks is the patterns mined from the data that are planned to distribute with parties other than the data owner. The significant difference between the work and issues is: both the fundamental data and the mined results are not planned for sharing and stays private to the data owner. A conservative frequency-based attack model [3] is used where the server recognizes the correct set of items in the owner's data. Furthermore, it recognizes the exact support for all items in the original data.

Patterns with read/write Support Encrypted TDB B*

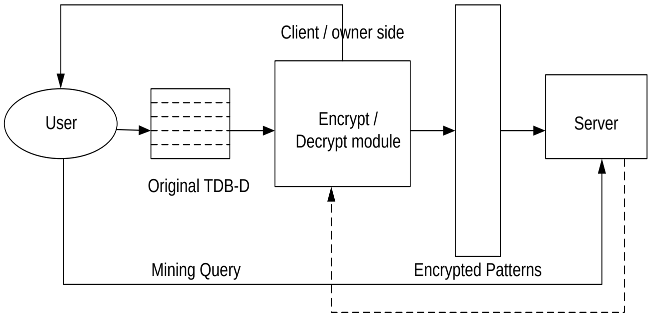

Fig. 1. Architecture of mining-as-service paradigm

The client/owner encrypts the data using encrypt/decrypt (E/D) module which is considered as a black box from its viewpoint. Encrypt/decrypt (E/D) module is used for converting the input data into an encrypted database. On the other hand, the server performs data mining operations and transmitted the patterns in the encrypted form to the owner of the data. The encryption scheme contains property where the revisited supports are not true supports. The E/D module regains the true individuality of the returned patterns and the true supports. An encryption scheme is named as RobFrugal. It is used to change the client data before it send to the server.

An alternative protocol is designed in [4] using simplicity, efficiency and privacy. Particularly, protocol fails in depending on commutative encryption and oblivious transfer. The solution is not completely secured. It gives large information to a small number of feasible combinations not same as protocol that discloses information.

B. Rule hiding for Privacy Preservation

The association rule hiding technique is to remove the sensitive rules from the transactional database during association rule mining. ARH technique protects sensitive data items by concealing the sensitive rules from miners and discloses all the non-sensitive rules to the miners. Data perturbation is used by Privacy Preserving Data Mining (PPDM) approach takes single-level trust on data miners. The technique establishes the ambiguity regarding individual values than the data released to the third parties for data mining purposes. In single trust level assumption, a data owner creates disturbed copy of its data with an amount of uncertainty. This assumption is restricted in many functions where a data owner trusts the data miners at various levels. An innovative dimension of Multi-Level Trust (MLT) [5] contains new demands for perturbation based PPDM. In contradiction to the single-level trust situation where only one perturbed copy is released and several perturbed copies of the similar data is presented for the data miners at various trusted levels.

The additional trust in data miner resulted in the less perturbed copy access. It also contains the access to the perturbed copies exist at lower trust levels. Additionally, data miners access multiple perturbed copies in forms. With diversity maintained across perturbed copies, the data miner on the other hand produced an exact reconstruction of the original data than permitted by the data owner. It is known as the diversity attack. It comprises the colluding attack situation where adversaries join their copies to increase an attack. It also incorporates the situation where an adversary uses public information to execute the attack by themselves. Preventing diversity attacks is the significant issue in solving the MLT-PPDM problem.

A compact prelarge GA-based (cpGA2DT) algorithm is designed in [2] to perform hiding operation of the sensitive itemsets while deleting transaction. The designed algorithm solves the issues of the evolutionary process by implementing both the compact GA-based (cGA) mechanism and the pre-large concept. A fitness function that was flexible in nature was structured using three adjustable weights to identify suitable transactions deleted to securitize the sensitive itemsets with minimal side effects of hiding failure, missing cost and artificial cost. A GA algorithm minimizes the memory needs by not taking the crossover and mutation operations but mimic the performances of traditional GAs.

C. Association Rule Hiding Techniques with Minimal Side Effects

The common technique of PPDM is to sanitize the database for hiding the information that is sensitive. A novel hiding-missing-artificial utility (HMAU) algorithm is designed in [6] to hide sensitive itemsets during transaction deletion. The transaction through the higher ratio of sensitive to non-sensitive one is chosen to delete.

In order to hide sensitive itemsets, three side effects were considered known as hiding failures, missing itemsets and artificial itemsets. Data sanitization is used to hide the sensitive knowledge from reveal in PPDM. To reduce the side effects, minimal distortion of the databases is required.

The transactions with any of the sensitive itemset are designed to locate the minimal HMAU values between transactions. The transaction with minimal HMAU value is directly taken away from the database. The process gets iterated till all sensitive itemsets are hidden. To avoid exposing hidden sensitive itemsets, the minimum count is modernized in the deletion process.

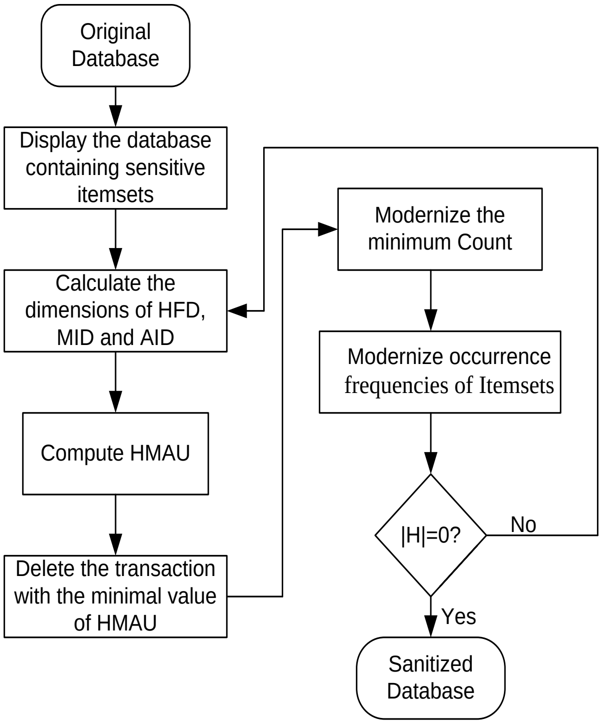
Original Database
Display the database containing sensitive itemsets
Calculate the dimensions of HFD, MID and AID
Compute HMAU
Delete the transaction with the minimal value of HMAU
Modernize the minimum Count
Modernize occurrence frequencies of Itemsets
|H|=0?
No
Yes
Sanitized Database

Fig. 2. HMAU Algorithm

A new heuristic method is designed in [1] that changes few transactions in the transaction database to reduce the supports or confidences of sensitive rules without any higher side effects. Connection between rules is not apparent to attain the goal. Heuristic methods are used for incrementing the hidden sensitive rules and minimize the number of modified entries. Rejected side effects are removed in the rule hiding process. The complete sensitive rules are hidden without unauthentic rules that are falsely created.

IV. Comparison of Association Rulehiding Techniques for Privacy Preservation & Suggestions

In order to evaluate the privacy perseveration using association rule hiding, number of data is taken to execute the experiment. Various parameters are used to calculate the privacy preserving in association rule hiding of the data mining techniques.

A. Privacy Preserving Level

Privacy preserving level is described as the level at which the data is privately transacted to the corresponding user without showing to the public users. It also increases the information delivery to the private users. It is measured in terms of percentage (%).

TABLE I. TABULATION FOR PRIVACY PRESERVING LEVEL OF ASSOCIATION RULE HIDING TECHNIQUES FOR PRIVACY PRESERVATION

Number of Data (Number)	Privacy Preserving Level (%)	
	PPMAR-OTD Techniques	SMC Algorithm
10	61	69
20	64	71
30	68	74
40	71	78
50	74	82
60	78	85
70	80	88

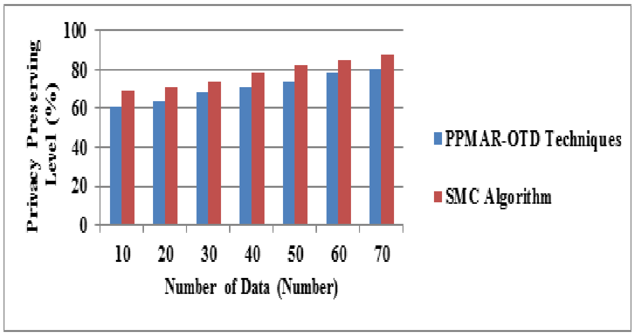

Fig. 3. Privacy Preserving Level of Association Rule Hiding Techniques for Privacy Preservation

Fig. 1 describes the privacy preserving level of association rule hiding techniques for privacy preservation. The privacy preserving level comparison takes place on existingPrivacy-Preserving Mining of Association Rules from Outsourced Transaction Databases (PPMAR-OTD)technique and Secure Multi-party Computation (SMC) algorithm. The experiment shows that SMC Algorithm has 9.37% higher privacy preserving levelthanPPMAR-OTD technique.

B. Data Utility Rate

Data utility rate is defined as the amount of data utilized for privacy preserving using association rule hiding techniques. It is measured in terms of percentage (%).

$$Data\ Utility\ Rate = \frac{Amount\ of\ data\ utilized\ for\ preserving\ privacy}{Total\ number\ of\ data}$$

TABLE II. TABULATION FOR DATA UTILITY RATE OF ASSOCIATION RULE HIDING TECHNIQUES FOR PRIVACY PRESERVATION

Number of Data (Number)	**Data Utility Rate (%)**	
	PPMAR-OTD Techniques	**SMC Algorithm**
10	68	51
20	72	54
30	75	56
40	78	58
50	80	61
60	81	64
70	84	68

The data utility rate comparison takes place on existing Privacy-Preserving Mining of Association Rules from Outsourced Transaction Databases (PPMAR-OTD) technique and Secure Multi-party Computation (SMC) algorithm.

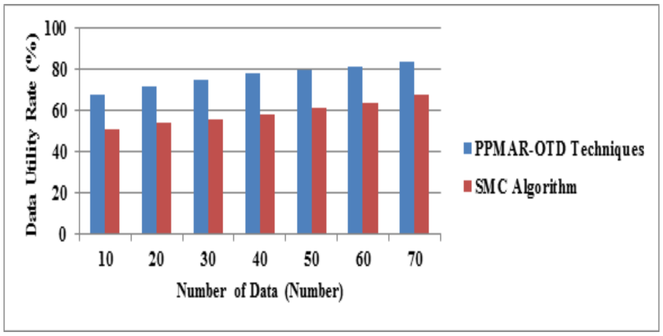

Fig. 4. Data Utility Rate of Association Rule Hiding Techniques for Privacy Preservation

Fig. 2 explains the data utility rate of association rule hiding techniques for privacy preservation. The experiment shows that PPMAR-OTD technique has 23.53% higher data utility rate than SMC Algorithm.

C. Efficiency (in terms of Side Effects)

Efficiency is defined as the number of data hided without any side effects to the total number of data given. It is measured in terms of percentage.

$$Efficiency\ (\%) = \frac{Data\ hiden\ without\ any\ side\ effects}{Total\ number\ of\ data\ given}$$

TABLE III. TABULATION FOR EFFICIENCY OF ASSOCIATION RULE HIDING TECHNIQUES FOR PRIVACY PRESERVATION

Number of Data (Number)	**Efficiency (%)**	
	MLT-PPDM	**cpGA2DT Algorithm**
10	75	65
20	78	68
30	81	71
40	83	74
50	85	77
60	87	81
70	89	83

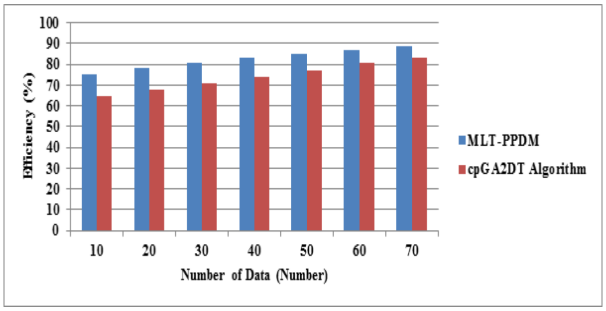

Fig. 5. Efficiency of Association Rule Hiding Technique for Privacy Preservation

Fig. 3 demonstrates the efficiency of association rule hiding techniques for privacy preservation. The efficiency comparison takes place on existing compact prelarge GA-based (cpGA2DT) Algorithm and Multi-Level Trust Privacy Preserving Data Mining (MLT-PPDM).The experiment shows that MLT-PPDM is 10.34% higher efficient than cpGA2DTAlgorithm.

D. *Execution Time*

Execution time is defined as the time taken to hide the data with minimum side effects. Execution time is measured in terms of milliseconds (ms).

Fig. 4 describes the execution time of association rule hiding techniques for privacy preservation. The execution time comparison takes place on existing compact prelarge GA-based (cpGA2DT) Algorithm and Multi-Level Trust Privacy Preserving Data Mining (MLT-PPDM). The experiment shows that cpGA2DT Algorithm consumes 33.86% lesser time for execution than MLT-PPDM.

TABLE IV. TABULATION FOR EXECUTION TIME OF ASSOCIATION RULE HIDING TECHNIQUES FOR PRIVACY PRESERVATION

Number of Data (Number)	Execution Time (ms)	
	MLT-PPDM	cpGA2DT Algorithm
10	21	15
20	25	18
30	28	20
40	31	23
50	34	26
60	37	29
70	40	32

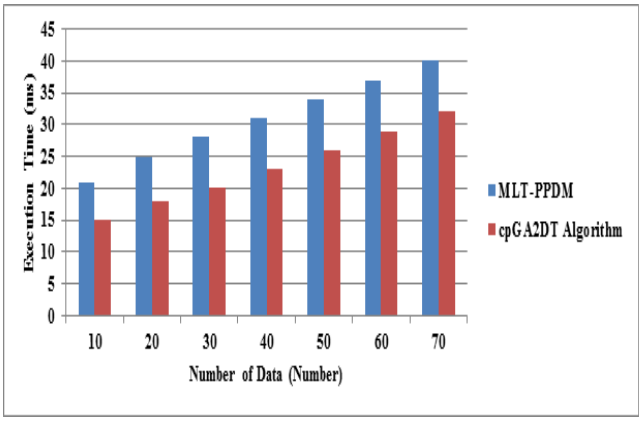

Fig. 6. Execution Time of Association Rule Hiding Technique for Privacy Preservation

E. *Memory Requirement*

Memory requirement is defined as the amount of memory space required forhiding the data using the association rule hiding techniques.It is measured in terms of mega bytes (MB).

TABLE V. TABULATION FOR MEMORY REQUIREMENT OF ASSOCIATION RULE HIDING TECHNIQUES FOR PRIVACY PRESERVATION

Number of Data (Number)	Memory Requirement (MB)	
	HMAU Algorithm	Heuristic Method
10	12	17
20	15	19
30	18	21
40	21	24
50	24	28
60	27	32
70	30	35

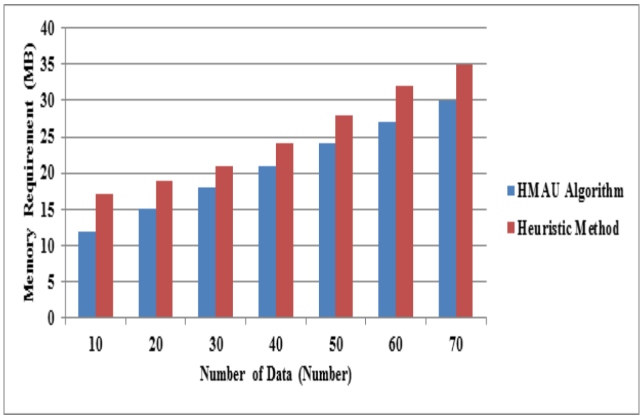

Fig. 7. Memory Requirement of Association Rule Hiding Technique for Privacy Preservation

Fig. 5 illustrates the memory requirement of association rule hiding techniques for privacy preservation. The memory requirement comparison takes place on existing Heuristic Method and Hiding-Missing-Artificial Utility (HMAU) algorithm. The experiment shows that HMAU Algorithm takes 21.59% lesser memory space than Heuristic Method.

F. *Hiding Failure Rate (in terms of Side Effects)*

Hiding failure rate is defined as the ratio of number of sensitive itemsets before sanitization to the number of sensitive itemsets after sanitization. It is measured in terms of percentage (%).

Fig. 6 shows the hiding failure rate of association rule hiding techniques for privacy preservation. The hiding failure rate comparison takes place on existing Heuristic Method and Hiding-Missing-Artificial Utility (HMAU) algorithm.

TABLE VI. TABULATION FOR HIDING FAILURE RATE OF ASSOCIATION RULE HIDING TECHNIQUES FOR PRIVACY PRESERVATION

Number of Data(Number)	Hiding Failure Rate (%)	
	HMAU Algorithm	Heuristic Method
10	25	18
20	28	21
30	31	24
40	35	27
50	37	30
60	39	33
70	41	36

The experiment shows that Heuristic Method has 26.63% lesser hiding failure rate than HMAU Algorithm.

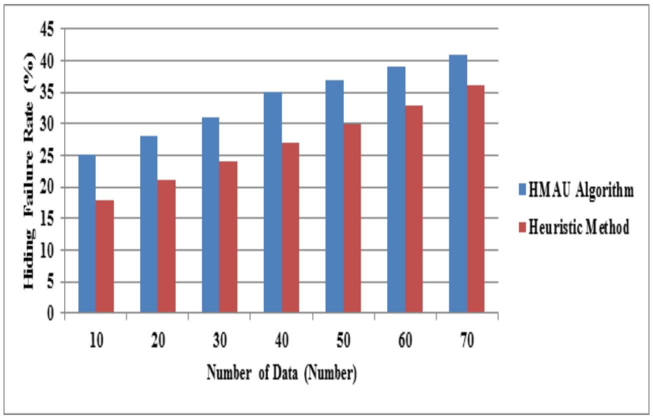

Fig. 8. Hiding Failure Rate of Association Rule Hiding Technique for Privacy Preservation

V. DISCUSSION ON LIMITATION OF PRIVACY PRESERVATION IN TRANSACTIONAL DATABASE

An encryption scheme designed with encryption/decryption module was used to transform client data before it send to the server. An attack model was designed to expose secret information and knowledge for privacy preserving mining. Usually, association rule mining task is performed in a shared privacy-preserving framework. In privacy preserving mining, the mined result is not aimed at sharing and stays as private. Attack model failed to know the details for encryption algorithms. Encryption scheme preserve the support item count values, attacker module can only work based on false support item counts.

Heuristic methods are used to improve the level of hidden sensitive rules quantity. The modified database is used to hide sensitive rules with limited side effects. Efficient mechanisms are needed to increase the speed of the rule hiding process for large databases. The association rules generated from the modified database have item sets does not appear in original transaction database. The side effect minimization fails to retain the correlation among rules on the modified transactional database. Secure Multi-party Computation (SMC) algorithm computes the union of private subsets. Secure mining of association rules is located in distributed databases in horizontal manner. The leakage information delivers, the protocol of item sets exposed, were insecure. Secure protocol is not based on the commutative encryption and transfer.

MLT-PPDM introduces the flexibility dimension that permits the data owners to make perturbed copies of data for various trust levels. In MLT-PPDM, data miners have an ability to approach several perturbed copies. Multiple perturbed copies and data miners achieves diversity attacks to modernize the original data more correctly. The department fails to have more power in reconstructing the original data with many copies. The data set does not re-anonymized after it is modified with insertions and deletions. Less perturbed copies are not used by data miners at lower trust levels.

A. Related Works

Direct and Indirect Discrimination Prevention method [9] was designed to evaluate the discriminatory frequent item sets between original and modified transactional database during data mining process. Discrimination-free data models are produced from transformed data set without damaging data quality and mining based on single measure. However, discrimination fails to include any measure to remove redundant information. To provide with a minimum extension to the original database, Border based approach with hiding algorithm [11] was designed to present the sensitive knowledge hiding. Border approach provides globally optimal solution for sensitive frequent item set hiding. However, the border approach fails to change the original data set properly, lead to information leakage and redundant frequent item sets. The regenerated frequent patterns were not present in the initial data set.

Locality-Sensitive Hashing (LSH) based Blocking Approach with a Homomorphic Matching Technique [10] is designed for recognizing the candidate record pairs. The matching of pairs is designed using a basic protocol performing simple distance computations. Matching Technique is used for Privacy-Preserving Record Linkage. Though, it fails to create exact results because of the used anonymization format. Because of improper encoding, the initial distances fails to preserve.In order to obtain higher computational overhead, Local NN-search and Global data reorganization technique [8] is implemented for Sensitive Transactional Data. But, anonymization of personal data is not enough in various applications and the approach is not suitable because of the high dimensionality of the data.

An improved Gaussian Function based Perturbation Technique [7] was designed for preserving privacy of association rules and private data of individuals in an outsourced business transaction database. Gaussian Function based perturbation technique [7] preserved the privacy of association rules generated from the dataset and the sensitive frequent item sets. However, it is highly complex for distributed high volume dataset in cloud environment. A group incremental feature selection algorithm [12] was developed to locate the new feature subset in a short interval of time, when multiple objects are added to a decision table. Incremental feature selection algorithm is derived from information entropy and it manages an effectual as well as well-organized mechanism. Though, the time complexity does not include the computational time of entropies.

B. Future Direction

The future direction of the privacy preservation using association rule hiding techniques needs to handle the confidentiality of sensitive rules in terms of better data utility and optimal side effects on the modified transactional databases. As each user may have different concern over privacy, user-oriented privacy preserving techniques can be developed. Parallel algorithms could be developed to prevent revealing of sensitive association between items and to improve the performance of the algorithm for large and dynamic datasets. Most of the proposed research works are concentrating on side effects and numbers of sensitive rules are hidden from sanitized database. Those are not clearly stated about number of rules are hidden in each iteration, number of levels in multi-level sensitive rule hiding, number of scan needed for the database, computational efficiency in terms of memory and CPU time. In future, these objectives are also being considered and new techniques are to be proposed for hiding the sensitive association rules in privacy preserving data mining.

VI. CONCLUSION

Based on the obtained nature of the survey, existing privacy preservation techniques in data mining using association rule hiding techniques has less privacy preserving level and also involves higher amount of side effects. At the same time, the utility of the data is also very low. As well, it takes higher execution time and so the efficiency gets decreased. The survey shows that while sending the data to the destination, the public user access the data and so the privacy is not maintained when it reaches to the destination. These types of issues decrease the effectiveness of the existing systems. The wide range of experiments on existing techniques calculates the relative performance of several privacy preserving techniques and its limitations. For this reason the new privacy preservation technique using association rules hiding techniques are planned to design. Finally from the result, the research work can be carried out in privacy preservation using association rule hiding techniques to attain minimal side effects with higher data utility.

REFERENCES

[1] Yi-Hung Wu, Chia-Ming Chiang, and Arbee L.P. Chen, Senior Member, IEEE Computer Society, “Hiding Sensitive Association Rules with Limited Side Effects”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 19, NO. 1, JANUARY 2007.

[2] Chun-Wei Lin, Binbin Zhang, Kuo-Tung Yang and Tzung-Pei Hong, “Efficiently Hiding Sensitive Itemsets with Transaction Deletion Based on Genetic Algorithms”, Hindawi Publishing Corporation, the Scientific World Journal, Volume 2014, Article ID 398269 September 2014.

[3] FoscaGiannotti, Laks V. S. Lakshmanan, Anna Monreale, Dino Pedreschi, and Hui (Wendy) Wang, “Privacy-Preserving Mining of Association Rules From Outsourced Transaction Databases”, IEEE SYSTEMS JOURNAL, VOL. 7, NO. 3, SEPTEMBER 2013.

[4] TamirTassa, “Secure Mining of Association Rules in Horizontally Distributed Databases”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 26, NO. 4, APRIL 2014.

[5] Yaping Li, Minghua Chen, Qiwei Li, and Wei Zhang, “Enabling Multilevel Trust in Privacy Preserving Data Mining”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 24, NO. 9, SEPTEMBER 2012.

[6] Chun-Wei Lin, Tzung-Pei Hong and Hung-Chuan Hsu, “ Reducing Side Effects of Hiding Sensitive Itemsets in Privacy Preserving Data Mining”, Hindawi Publishing Corporation, the Scientific World Journal, Volume 2014, Article ID 235837 April 2014.

[7] VineetRichhariya., and PrateekChourey., “A Robust Technique for Privacy Preservation of Outsourced Transaction Database” International Journal of Research in Engineering & Technology (IJRET), Vol. 2, Issue 6, Jun 2014, 51-58.

[8] Gabriel Ghinita, Member, IEEE, PanosKalnis, and Yufei Tao, “Anonymous Publication of Sensitive Transactional Data”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 2, FEBRUARY 2011.

[9] Sara Hajian and Josep Domingo-Ferrer, Fellow, IEEE, “A Methodology for Direct and Indirect Discrimination Prevention in Data Mining”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 7, JULY 2013.

[10] DimitriosKarapiperis and Vassilios S. Verykios, Member, IEEE, “An LSH-Based Blocking Approach with a Homomorphic Matching Technique for Privacy-Preserving Record Linkage”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 27, NO. 4, APRIL 2015.

[11] ArisGkoulalas-Divanis, Member, IEEE, and Vassilios S. Verykios, Member, IEEE, “Exact Knowledge Hiding through Database Extension”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 21, NO. 5, MAY 2009.