RESEARCH REVIEW JOURNALS
Volume-04
Issue-03
March -2019

ISSN: 2455-3085 (Online)
RESEARCH REVIEW International Journal of Multidisciplinary
www.rrjournals.com [UGC Listed Journal]

Hiding Techniques for Data Publishing by Conserving Item hiding and Privacy Conserving

Shipra Varshney

Research Scholar, Department of Information Technology, GGISP University/ Dr. Akhilesh Das Gupta Institute of Technology & Management (Formerly Northern India Engineering College) NEW DELHI-110053, India

ARTICLE DETAILS

Article History
Published Online: 13 March 2019

Keywords
Perturbation, OSA-SIH, SIPRP, RSA-DSO, Privacy Preserving Data Mining, Sensitive Item Hiding

ABSTRACT

We lives in a world are live in a world where vast amounts of data are collected daily. Analyzing such data is an important need. Data mining is a current technology for extracting knowledge from a large amount of data. In real world business applications Data mining plays a major role by providing disparate techniques and algorithms. Privacy preserving is an important research area, which allows parties to cooperate in the knowledge extraction without revealing the extracted knowledge with any individual parties. Once the information is pooled among dissimilar nodes such that centralized or distributed, then data mining out-comes and results should guarantee the secret sharing of information. This paper presents and proposes performance analysis for the proposed OSA-SIH, SIPRP AND RSA-DSO methods and expands on the original data set. In the paper the techniques based on the evaluated dataset, augmented and optimal hiding of sensitive item is appeared and even for large item sets, ensuring time for optimal hiding. The findings of the techniques discussed in this paper demonstrated that the proposed techniques out performed than the already alive state of the art works in understanding of privacy preservation accuracy, time for optimal data hiding and degree of side effects on modified dataset as related to the state of the art works.

1. Introduction

Now a day's information is played major role in decision making in an organization. We are in the world of information processing society. Data is the major valuable resource of any business or organization. There is a huge amount of sensitive data produced by various business operational applications. Sharing information between various sources through authorized channel is an important task. Data Mining is kind knowledge discovery system, through this data can extract from different sources. While sharing information through different channels or extracting information from different external sources, the key factor is protecting data from unauthorized accesses. Data Mining is the process of extracting the data or knowledge from different data resources like data base and Data warehouse. Data Mining is one of the emerging research area which deals with privacy and security of data. There are many Privacy Preserving Data mining (PPDM) and Privacy Preserving Distributed Data mining (PPDDM) techniques are used to solve privacy issues in Centralized Data mining environment and Distributed data mining Environment respectively. In real world applications and business Data mining has a vital role not only by providing different techniques but also by giving different algorithms. The data is extracted from varying sources for business processing, for business purposes and it is not only the need of the hour but also crucial to secure the data of individuals or groups.

When the information shared among different nodes such that centralized or distributed, then data mining results should ensure the secret sharing of information. Multilevel Trust in Privacy Preserving Data Mining (MT-PPDM) [1] is based on the concept of aggregated data without the probability of accessing the information by the third parties, safeguarding trust level. Reducing Side Effects in Privacy Preserving Data Mining (RSE-PPDM) [2] used hiding missing artificial utility algorithm to minimize the number of deleted transactions and number of side effects.

2. Privacy Conserving Data Mining

One method among privacy preserving techniques which is PPDM loads with loading valid data mining results without leaking the essential sensitive data values. The need for privacy is sometimes due to individual respect or it can be motivated by business interests [4]. Privacy- preserving data mining (PPDM) is a new area of research in data mining. Its ultimate goal is to develop efficient algorithms that allow to extracting relevant knowledge from a large amount of data, while preventing sensitive information from disclosure or inference[3].In paper [5] PPDM is defined as "getting valid data mining results without learning the underlying data values". PPDM goal is to produce valid data mining results through privacy requirements [6][13].

PPDM research usually follows one of the three approaches: (1) data hiding, raw sensitive data can be modified, blocked, trimmed out from the original database, in order for the users of the data not to be able to compromise another person's privacy; (2) rule hiding, in which sensitive data can be extracted from the data mining process, it excluded for use, because confidential information may be derived from the released knowledge or data. This problem is commonly known as the "database inference problem;" and (3) Secure Multiparty Computation (SMC), where distributed data are encrypted before released or shared for computations; thus, no party knows anything except its own inputs and the results. Privacy Preserving Data Mining which is also known as PPDM is a fast growing research area. Privacy-Preserving Data Mining (PPDM) is a data mining and statistical databases

innovative field where data mining algorithms are analyzed for side-effects in data privacy. Learning essential and sensitive data values without breaching and leaking the sensitive information is coined as privacy sensitive or privacy enhanced data mining dealing with the results without breaching the underlying data values. There are so many different methods and techniques which can be used in a PPDM context from a technical perspective [12]. The main objectives and goals of a PPDM algorithm are mentioned below:

i. Prevent the discovery of sensible information.
ii. Being uncompromised to access and to use the non sensitive data.
iii. Extent usable on large amounts of data.
iv. The exponential computational complexity must be less.

To achieve optimized results while preserving the privacy of the data subjects efficiently, five dimensions need to be considered and listed below:

1) The distribution of the basic data
2) The modification of the basic data
3) Mining method being used
4) If basic data or rules are to be hidden and
5) Additional methods for privacy preservation used.

3. Basic Taxanomy of Privacy

Altman privacy theory [7] views privacy management as a dialectic and dynamic boundary disclosure which are explained as follows:

I. **Privacy as colloquial activity:** This says that if privacy is measured as dialectic then it mostly depends upon experiences and expectations not only of the users but also the ones with whom they interact.
II. **Privacy as a robust boundary disclosure:** Privacy as a robust and vital boundary regulation and continuous process of a deciding parameter to decide a line or a boundary between public and private.

According to Gurses classification [8] privacy problems can be divided into three categories which are i) privacy as control ii) privacy as confidentiality and iii) privacy as practice

I. **Privacy as controller:** The organizations if by any means disclose the collected information to the unauthorized third parties or to a broader public then it routes to a privacy violation path. The policies which are defined by the users as privacy settings or organizations as access control can be articulated in Privacy. Access control, privacy settings, purpose based access control, auditing are some of the examples of privacy research in this epitome.
II. **Privacy as secrecy**: This mainly centers on the data disclosure problem. Privacy is leaked if the information crosses and goes beyond its visibility scope. Privacy as confidentiality makes a modest disclosure such that the information cannot be linked back to the individual. There are some example like private retrieval systems, anonymous authentication protocols and anonymous communication networks.
III. **Privacy as exercise:** Privacy now a day's not limit to an individual's matter but it is surely a matter of social concern. Now mostly users decide their privacy and its dimensions predicted and grounded on the community they live in. It is difficult to understand for the users to extrapolate how they should regulate their sensitive information and if this information gets disclosed what are the consequences they could made on it. For example, P3P and privacy mirror are the technologies adapting to privacy as a practice.

Discretion is the ability to build our own path without any external influence or impediment. Based on discretion Papacharissi et al. [9] outlooks privacy as self expression of social relationship and as useful commodity.

I. Privacy and the uniqueness: The individual's identity is unique but basically it is social in nature. Various social interactions and dealing with multiple audiences with collective and collaborative experiences gives a sense of who we truly are.
II. Privacy and the formulation of social relationships: Sharing personal information with the public makes it lose its meaning and inherent value. In an OSN platform the individuals form social relationships which encourage the loss of privacy. Hence, it is a challenging task to prevent privacy on such platforms which is explicitly designed for sharing.
III. Privacy as nicety stuff: Now days everybody access web and various web platforms. These platforms offer services which are social in nature. Making money out of the personal information is very common these days on the web therefore such web based platforms information privacy is one of the unmatched luxury commodities.

4. Privacy Concerns

Some ways in which privacy concerns raised by data mining are as follows (Oliveira et al 2004) [10][11]:

I. The implicit patterns involving information about persons that can be derived from data in the data mining process vs. the explicit nature of the personal data (in records) extracted in traditional database retrieval techniques.
II. Single database to extract the information of a person v/s multiples databases possibly data warehouses to retrieve information is used.
III. The use of "open-ended" queries to discover information on relationships and associations about individuals and groups of individuals vs. (traditional) specific queries to retrieve information about relationships and associations that are already known to exist.
IV. The unanticipated and non predicted feature of information about persons gained from data mining vs. the generally predictive aspect of information retrieved from traditional database techniques.
V. The public attributes of the information of the persons which is derived through the data mining process vs. the private or intimate nature of the information about persons retrieved and exchanged in traditional database-exchange techniques.
VI. The potentiality of creating the new groups or categories of persons based and planted on patterns of information extracted from data mining vs. the sheer extraction of information of the individuals

themselves from personal data available to traditional techniques of database retrieval.

5. Optimized Social Ant Based Sensitive Item Hiding

An efficient framework called Optimized Social Ant Based Sensitive Item Hiding (OSA-SIH) technique is established to maximize the privacy preservation accuracy for data publishing. Besides, the proposed OSA-SIH technique decreases the rate of side effects on the modified dataset at relatively lesser time interval. This is obtained based on user operational conditions-based sensitive items, social ant-based relative item set distribution and Ant-based based Orthogonal Multiplicative and Transformational algorithm. Then, the proposed OSA-SIH technique gives the detailed description which is presented in the forthcoming sections. Tong Yi and Minyong Shi (2015) intended the privacy protection method utilizes the association rules among sensitive attributes and data publication for many sensitive attributes. Privacy protection method provides better protection for privacy and also preserves significant relationships for improving the system performance. Though, the hiding the available sensitive attributes from the database is the major issue. In order to overcome such limitations, the OSA-SIH technique is proposed to hide the sensitive items for the privacy protection using user operational conditions-based sensitive items when compared to the privacy protection technique[14][15].

6. User Operational Condition Based Sensitive Items

The proposed OSA-SIH technique is utilized based on the performance of designing the user operational conditions-based sensitive items with better efficient. User operational condition approach is implemented with the objective of hiding the sensitive item from global frequent items for distributed dataset is being shared during the privacy preservation. Then the block diagram of user operational conditions-based sensitive items using proposed OSA-SIH technique is represented in the figure 1 should be streamlined, and the users should be made more aware of their rights regarding the internet.

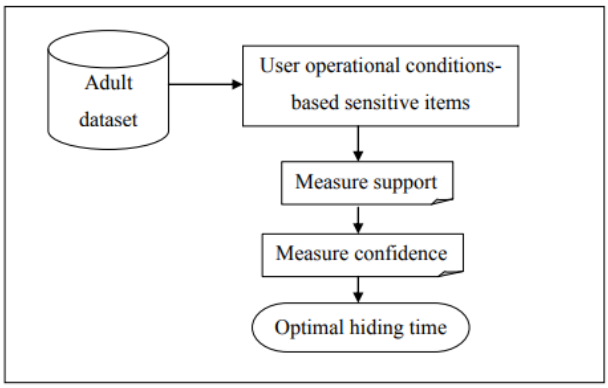

Figure 1. Block diagram of user operational conditions-based sensitive items

Let us consider a dataset '*D*' where '$I = I_1, I_2, \ldots I_n$' represents the items, consisting of '*n*' transaction comprises of the set of items in such a way that $T \subseteq I$. Then, the association rule is of the form

$$P \rightarrow Q, \text{where } P \in I \,\&\, Q \in I \qquad (1)$$

Where '*P*' and '*Q*' are said to be the antecedent and consequent of rule respectively. Relative strength of an item with respect to its strong or weak nature is evaluated using two factors namely, support and confidence of the item. The first factor to be measured for sensitive item hiding is the support and is mathematically formulated as given below.

$$S(P \rightarrow Q) = S(P \in Q) = ((P \in Q)/n) \qquad (2)$$

From (2), support 'S' measures the proportion of transactions that includes both 'P' and 'Q' respectively, with 'n' denoting the total number of transactions involved during sensitive item hiding. The second factor to be measured for sensitive item hiding is the confidence formulated as given below.

$$C(P \rightarrow Q) = ((P \cup Q)/P) = (S(P \cup Q)/S(P)) \qquad (3)$$

From (3), the percentage for a transaction is the **confidence 'C'** that contains **'P'** and also contains **'Q'.** A rule is significant if its support and confidence are higher than the user designated Support Threshold Value (STV) and Confidence Threshold Value (CTV). As a result, using the ant-based relative item set distribution algorithm not all the items are retrieved, but only a very small member that satisfies the 'STV' and 'CTV 'are retrieved. In this way, the time for optimal hiding is significantly reduced.

Experimental Setup:

Input: A dataset 'D' and Items $=I_1, I_2, \ldots In$, Support threshold value STV, Confidence threshold value CTV.

Output: augmented and optimal sensitive item hiding.

Steps are:

Step 1: Begin with each dataset 'D'

Step 2: Begin with each item set 'I' in 'D'

Step 3: Evaluate support 'S' for sensitive item hiding using a function ()

Step 4: Evaluate confidence 'C' for sensitive item hiding using a function ()

Step 5: Check two conditions:

5.1: If 'S<STV' and

5.2: If 'C< CTV' then move to Step 6:

Step 6: Evaluate optimal hiding of sensitive item using ()

Step 7: Evaluate orthogonal multiplication and transformational process using ()

Step 8: Else

Step 9: Go to step 2

Step 10: End of Step 5

Step 11: End of Step 2

Step 12: End of Step 1

Description:

The objective behind the use of orthogonal multiplicative and transformational data perturbation is to improve the privacy preservation accuracy during the data perturbation process. The orthogonal multiplicative and transformational data perturbation in proposed OSA-SIH technique uses orthogonal transformation.

Let us consider two datasets 'L' and 'M' of size 'i * n matrix' and 'j * n matrix' respectively with orthogonal matrix represented as 'O'. Now the mathematical formulation for the orthogonal multiplicative and transformational data perturbation for two datasets 'L' and 'M' is as given below:

$$A = LO; \qquad B = MO; \qquad (5)$$

$$AA^T = LL^T; \qquad BB^T = MM^T \qquad (6)$$

$$AB^T = L\,OO^T\,B^T = LM^T \quad (7)$$

From (5), (6) and (7), by applying an orthogonal matrix based on socially cohesive relational rate between sensitive and non sensitive item sets, all the pair distances and similarities from column vectors 'A and B' are preserved in an efficient manner in the perturbed data. At the same time, both the sensitive and non sensitive items and the transformation process are kept secret, whereas only the perturbed data is viewed by the third user. As a result, the privacy preservation accuracy is improved in a significant manner. Figure 3 shows the ant- based orthogonal multiplicative and transformational algorithm.

The Ant-based Orthogonal Multiplicative and Transformational (AOMT) algorithm includes four main steps. The first step measures the support for sensitive item hiding. The second step evaluates the confidence value for sensitive item hiding. Next, a comparison is made between the support threshold 'STV' and confidence threshold 'CTV' with the evaluated confidence 'C' and support value 'S'. Followed by this, optimal hiding of sensitive item and orthogonal multiplicative and transformational process is performed. If the values of support 'S' and confidence 'C' are less than the support threshold 'STV' and confidence threshold 'CTV' respectively, item hiding is performed, otherwise, the same operation is performed with other transactions. In this way, privacy preservation accuracy is ensured in an efficient manner.

7. Discussion

Experiments are actually conducted by using java language for the suggested OSA SIH, SIPRP, Existing techniques and RSA-DSO techniques specifically Multilevel Trust in Privacy Preserving Data Mining (MLT PPDM) developed by Yaping Li et al. (2012 Protocol and) for protected mining of association rule created by Tamir Tassa (2014). The enhancement of privacy preservation is actually obtained using adult data set. The adult data set is actually taken out of the Faculty of California Irvine data repository. It's additionally called as "Census Income" dataset. The income dataset consists of the fourteen attributes as well as the number of instances is actually 48842. The attributes are actually age, job class, education, marital status, occupation, native country, level, present employment style etc. Sticking to the performance metrics used for analyzing the suggested OSA SIH, RSA-DSO and SIPRP type with the present strategies with graph representations. The overall performance metrics consist of privacy preservation accuracy, time for optimal data hiding and selection of sensitive rules, processing time, number of hidden rules as well as rate of privacy rate.

1) Impact of privacy preservation accuracy:

Privacy preservation accuracy helps for measuring the amount of privacies preserved perturbed copies within PPDM with regard to the total number of perturbed copies thought for the data publishing. The privacy preservation accuracy is mathematically formulated as provided below in equation 1:

PPA (Number of privacy perturbed copies / Total number of perturbed copies)* 100 equation 1

From the above equation 1, privacy protection precision 'PPA' is estimated by rate (%). In the event that the privacy protection accuracy is high, at that point the technique is supposed to be increasingly proficient

Age (Number of perturbed copies)	Privacy Preservation Accuracy (%)				
	Existing MLT-PPDM	Existing protocol for secure mining of association rule	Proposed OSA-SIH	Proposed SIPRP	Proposed RSA-DSO
10	64.48	55.36	70.35	74.02	77.28
20	70.48	60.45	74.41	78.14	81.52
30	72.54	62.51	75.39	79.32	83.58
40	71.26	61.98	73.28	77.41	82.3
50	73.96	65.93	76.97	80.34	84.32
60	77.43	68.43	80.32	84.26	88.35
70	80.44	70.54	84.36	87.12	90.48
80	83.17	73.37	87.14	90.48	93.32
90	86.11	76.43	89.53	92.22	95.28
100	87.84	78.94	91.26	93.52	97.12

Table 1: Tabulation of privacy preservation accuracy

Table 1 shows the privacy safeguarding accuracy dependent on the quantity of bothered duplicates for data publishing utilizing the proposed OSA-SIH, SIPRP, RSA-DSO techniques and the current MLT-PPDM, Protocol for secure mining of affiliation rule strategies. The tabulation shown above tells the number of annoyed duplicates taken from the scope of 10 to 100 for leading examination. While expanding the quantity of irritated duplicates, privacy safeguarding accuracy is additionally expanded in all the strategies. Table 1 show that the proposed RSA-DSO model adequately accomplishes higher privacy protection accuracy when contrasted with the other proposed and existing strategies.

Figure 2 shows the proportion of privacy safeguarding accuracy utilizing the proposed OSA-SIH, SIPRP, RSA-DSO techniques and existing strategies, for example, MLT-PPDM created by Yaping Li et al. (2012) and Protocol for secure mining of affiliation rules created by Tamir Tassa (2014).

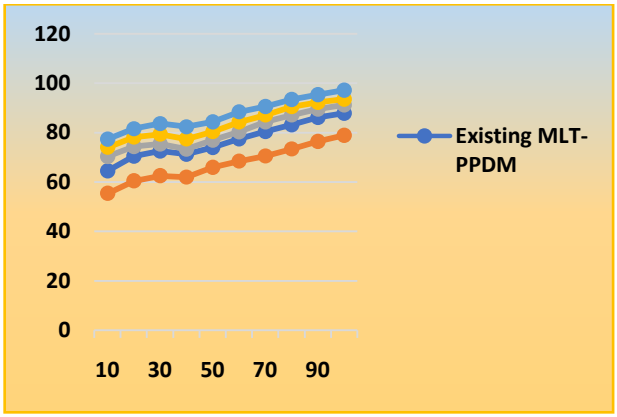

Figure 2: Measure of the privacy preservation accuracy

By the figure 2, it's apparent that the proposed RSA -DSO design effectively raises the privacy preservation accuracy when in comparison to the other group suggested as well as existing techniques. This particular privacy preservation enhancement is actually attained to the proposed RSA -DSO type with the assistance of discrete swarm optimization algorithm in the population census data publishing process. Additionally, the discrete swarm optimization algorithm uses the workout function for removing the sensitive rule in a major fashion which in turns raising the privacy preservation accuracy. Thus, the proposed RSA -DSO design raises the privacy preservation accuracy by nineteen % when compared to the current Protocol as well as MLT PPDM for the protected

mining of association rule strategies. Proposed SIPRP technique raises the privacy preservation accuracy by fifteen % as well as the suggested OSA SIH strategy raises the privacy preservation accuracy by eleven % when compared to the current Protocol and MLT PPDM for protected mining of association rule methods.

2) Impact of time for optimal hiding:

Time for optimal hiding is actually calculated by the quantity of time taken for one-time transaction to the total amount of transactions. Time for optimal hiding is mathematically estimated by equation 2:

TOH = Single transaction time * Total number of transaction equation 2

In equation 2, Time for Optimal Hiding 'TOH' actually calculated in phrases of milliseconds (ms).Lower period for optimal hiding guarantees much better functionality of the method.

Total number of transactions	Time for Optimal Hiding (ms)				
	Existing MLT-PPDM	Existing protocol for secure mining of association rule	Proposed OSA-SIH	Proposed SIPRP	Proposed RSA-DSO
5	6.62	5.43	4.22	3.32	2.22
10	7.56	6.66	5.85	4.85	3.85
15	9.32	8.42	7.55	6.46	5.55
20	10.52	9.67	9.11	7.81	6.67
25	10.79	9.59	8.62	7.52	6.89
30	13.32	12.42	11.46	10.02	8.9
35	15.31	14.01	13.14	12.01	10.8
40	17.36	16.34	15.33	14.22	12.95
45	18.53	17.13	16.34	15.13	14.32
50	19.3	18.31	17.56	16.34	15.3

Table 2: Tabulation of time for the optimal hiding

Table 2 displays the time for optimal hiding that primarily is dependent upon the total amount of transactions of the network utilizing the suggested as well as existing techniques. Total amount of transactions is taken out of the assortment of five to fifty for experimental objective. By the table 2 it's apparent that, the time for optimal hiding is actually enhanced for the corresponding rise in the total amount of transactions using all of the techniques. Nevertheless, the proposed RSA - DSO design creates little time for hiding the sensitive items in comparison with various other methods.

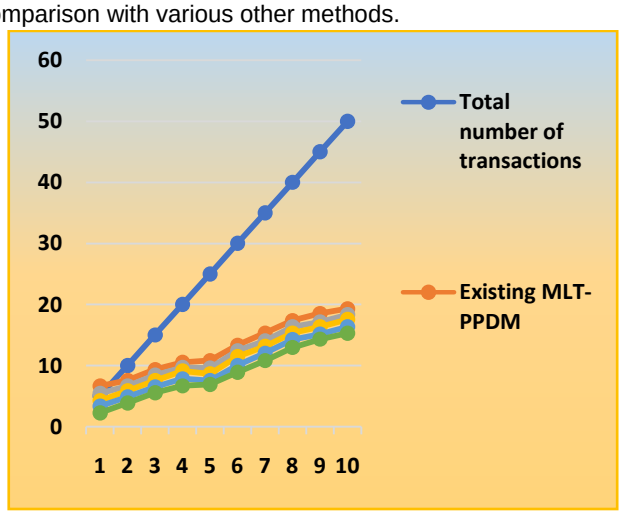

Figure 3 Measure of the optimal hiding

Figure 3 shows the degree of time for optimal hiding making use of the suggested OSA- SIH, SIPRP, Existing techniques and RSA-DSO techniques including MLT - PPDM created by Yaping Li et al. (2012 Protocol and) for protected mining of association rule created by Tamir Tassa (2014). By the figure 3, it's apparent that time for optimal hiding for the complete amount of transactions are comparatively decreased in the proposed RSA- DSO design when in comparison to the other group suggested as well as existing techniques. This particular reduction of time for optimal hiding is actually attained to the proposed RSA -DSO type with the application of reinforcement type for hiding the data item. Hiding the data item is primarily depending on the probability to take place the sensitive data item is actually evaluated only one time. Additionally, RSA -DSO design sort the sensitive data items exactly where the optimal typical items are actually reinforced which in turns lowering the time for optimal hiding. Thus, the proposed RSA -DSO design minimizes the time for optimal hiding by thirty-nine % when compared to the current Protocol as well as MLT - PPDM for protected mining of association rule strategies. Proposed SIPRP technique decreases the time for optimal hiding by twenty-eight % and suggested OSA SIH method decreases the time for optimal hiding by nineteen % when compared to the current Protocol and MLT - PPDM for protected mining of association rule methods.

3) Impact of the number of hidden rules

Hidden rules are very helpful in comparing the usefulness of the suggested strategies. The number of hidden rules is actually calculated depending on the number of sensitive rules hidden by having no unwanted side effects to the total number of sensitive rules in PPDM. It's mathematically formulated in equation 3 below:

NHR = (number of sensitive rules hidden / total no. of sensitive rules) *100 equation 3

From the situation above, the number of hidden rules NHR' are actually assessed in term of percent (%). Greater number of sensitive rules hidden represents the technique is believed to be a lot more secure while data publishing.

Number of Sensitive Rules	Number of Hidden Rules (%)				
	Existing MLT-PPDM	Existing protocol for secure mining of association rule	Proposed OSA-SIH	Proposed SIPRP	Proposed RSA-DSO
10	44	52	60	71	65
20	46	54	62	73	67
30	48	56	64	75	69
40	50	58	66	77	71
50	52	60	68	79	73
60	54	62	70	81	75
70	56	64	72	83	77
80	59	66	75	87	80
90	63	71	78	90	83
100	69	75	82	95	88

Table 3: Tabulation of the number of hidden rules

Table 3 shows the number of hidden rules for the data publishing making use of the suggested as well as existing techniques. Selection of sensitive rules is actually taken out of the assortment of ten to hundred for the experimental objective. By the table 3 it's apparent that, the number of hidden rules is

actually enhanced for the corresponding rise in the number of sensitive rules using all of the techniques. Nevertheless, the proposed SIPRP technique creates maximum number of hiding rules when as compared to the other techniques.

Figure 4 illustrates the number of hidden rules for the population census data publishing procedure with the suggested OSA SIH, SIPRP, RSA -DSO techniques as well as the present techniques including MLT PPDM created by Yaping Li et al. (2012) and the Protocol for protected mining of association rule created by Tamir Tassa (2014). As shown in the figure 4, the proposed SIPRP strategy greatly improves the number of hidden rules when as compared to the other methods.

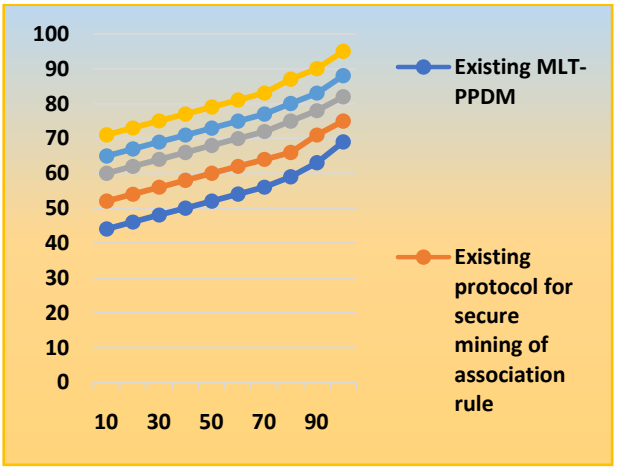

Figure 4: Measure of the number of hidden rules

This handy enhancement of the number of hidden rules is actually done at the proposed SIPRP process with the assistance of particle swarm optimization mechanism. This particular mechanism computes the placement as well as velocity of every particle by means of hiding the sensitive rules coming from the public by utilizing the suggested SIPRP technique. Particle swarm optimization mechanism is actually applied for preserving the attached information in the population census data publishing process. Thus, the amount of the hidden rules of proposed SIPRP strategy is enhanced by forty-two % when compared to the current Protocol as well as MLT PPDM for the protected mining of association rule strategies. Likewise, the proposed RSA -DSO design advances the number of hidden rules by twenty-three % as well as the suggested OSA SIH strategy improves the number of hidden rules by seventeen % when compared to the current Protocol and MLT PPDM for the protected mining of association rule methods.

8. Conclusion

OSA-SIH which is Optimized Social Ant Based Sensitive Item Hiding is a technique with outlook of equality and quality privacy preservation for distributed data mining with ideal and superlative side effects on original dataset that has been designed. The objective of providing such a design is to ensure high quality privacy preservation of the data items of corresponding user's privileges for distributed data and to decrease the time for optimal hiding for various user requested item set distribution. A user operational conditions-based sensitive item are designed as a measure for identifying the support and confidence value and proposed a proposed a system to measure the global frequent item sets for distributed data item being shared based on user query. The proposed social ant-based relative item set distribution provides privacy preservation accuracy for large item sets through multiplicative and transformational data perturbation technique. Added to this Ant-based Orthogonal Multiplicative and Transformational algorithm with probability function supports the privacy preservation accuracy. Experimental evaluation is conducted with the Adult Data Set extracted from UCI repository to provide high quality privacy preservation of data items and measured the performance in terms of privacy preservation accuracy, optimal time hiding and rate of side effects on answering user query requests. Performances results reveal that the proposed OSA-SIH technique provides higher level of privacy preservation accuracy efficiency and also strengthen the optimal time hiding on high dimensional dataset. The proposed OSA-SIH technique provides 12.85% high rate of privacy preservation accuracy and minimizes the time for optimal hiding by 17.47% when compared to state of the art works.

References

1. Li Y, Chen M, Li Q, Zhang W. Enabling multilevel trust in privacy preserving data mining. IEEE Transactions on Knowledge and Data Engineering. 2012 Sep; 24(9):1598–612.
2. Lin CW, Hong TP, Hsu HC. Reducing side effects of hiding sensitive item sets in privacy preserving data mining. The Scientific World Journal. 2014; 2014:12.
3. W. Stallings, "Network and Internetwork Security Principles and Practice", IEEE Press, New Jersey, 462 p., May 1995
4. K. SrinivasaRao" An Insight in to Privacy Preserving Data Mining Methods" The SIJ Transactions on Computer Science Engineering & its Applications (CSEA), Vol. 1, No. 3, JulyAugust 2013.
5 Yehuda Lindell, BennyPinkas" Privacy Preserving Data Mining"
6. C. Clifton, M. Kantarcioglu, and J. Vaidya. Defining Privacy For Data Mining. In Proc. of the National Science Foundation Workshop on Next G eneration Data Mining, pages 126-133, Baltimore, MD, USA.
7. Irwin Altman. Privacy: A conceptual analysis. Environment and behavior, ERIC, volume, 8(1):7–29, 1976.
8. Claudia Diaz and Seda Gurses. Understanding the landscape of privacy technolo- ¨ gies1, 2012.
9. Zizi Papacharissi and Paige L Gibson. Fifteen minutes of privacy: Privacy, sociality, and publicity on social network sites. In Privacy Online, pages 75–89. Springer, 2011.
10. Paulet R, Md Kaosar G, Yi X, Bertino E. Privacy-preserving and content-protecting location based queries. IEEE Transactions on Knowledge and Data Engineering. 2014 May; 26(5):1200–10.
11. Pervaiz Z, Walid G, Ghafoor A, Prabhu N. Accuracy-constrained privacy-preserving access control mechanism for relational data. IEEE Transactions on Knowledge and Data Engineering. 2014 Apr; 26(4):795–807.
12. Li T, Li N, Zhang J, Molloy I. Slicing: A new approach to privacy preserving data publishing. IEEE Transactions on Knowledge and Data Engineering. 2012 Mar; 24(3):561–74.

13. Goryczka S, Xiong L, Fung BCM. m-Privacy for Collaborative Data Publishing. 7th International Conference on Collaborative Computing: Networking, Applications and Work sharing (Collaborate.Com). 2011 Oct 15-18. p. 1–10.
14. Nabeel M, BertiIssue E. Privacy preserving delegated access control in public clouds. IEEE Transactions on Knowledge and Data Engineering. 2014 Sep; 26(9):2268–80.
15. Nabeel M, Bertino E. Privacy-preserving fine-grained access control in public clouds. IEEE Computer Society Technical Committee on Data Engineering. 2012 Dec; 35(4):1–10.
16. Glu MK, Clifton C. Privacy-preserving Distributed Mining of Association Rules on Horizontally Partitioned Data. 12th International Conference on Hybrid Intelligent Systems (HIS). 2012. p. 2–13.

