International Journal of Soft Computing 9 (2): 109-116, 2014
ISSN: 1816-9503

Enhanced Applicability of Privacy Preservation for Perturbed Data in Multi-Partitioned Data Set

V.S. Prakash and A. Shanmugam
Department of Computer Applications, Bannari Amman Institute of Technology,
638401 Sathyamangalam, Tamil Nadu, India

Abstract: The perturbation technique has been widely considered for privacy preserving in data mining for different datasets. Generally, multi-partitioned datasets comprises of both vertical and horizontal data sets which is being a current demand of e-Business and e-Commerce data mining environment. In perturbation process, arbitrary noise from a recognized distribution is processed as privacy susceptible data, prior the data is thrown to the data miner. Consequently, the data miner rebuilds estimation to the unique data distribution from the perturbed data and exercises the renovated delivery for data mining principles. Owing to the count of noise, loss of information versus conservation of privacy is a constant transaction in the perturbation based techniques. The question is to what level the users are disposed to cooperate with their privacy? This is a preference that amends from individual to individual. To assess a tradeoff among data privacy and simplicity of individual's data, the first research is to describe the data perturbation technique with validation and authentication. Diverse individuals may have diverse approaches towards confidentiality, based on traditions and cultures. Unfortunately, the earlier perturbation based privacy preserving data mining techniques do not permit the individuals to decide their preferred privacy levels. This is a negative aspect as privacy is an individual choice. In this study, researchers propose an individually adaptable perturbation model which enables the individuals to choose their own privacy levels. The effectiveness of the proposed model lies is the enhancement of the Applicability of Privacy Preservation for Perturbed Data in Multi-partitioned datasets (APPDM) demonstrated by diverse experiments conducted on both synthetic and real-world data sets. Based on the experimental evaluation, researchers propose a simple, valuable and resourceful method to construct data mining models from perturbed data and enhance the process of privacy preservation.

Key words: Data mining, privacy, security, multi-partitioned dataset, data perturbation

INTRODUCTION

Data mining study compacts with the extraction of potentially helpful information from huge collections of data with a diversity of relevant areas such as customer association organization, market basket examination and bioinformatics. The mined information can be in the type of models, clusters or categorization models. Association rules in a superstore for instance could explain the association between items bought collectively. Customers could be collected in segments for enhanced customer association organization. Classification demonstrations could be erected on client profiles and shopping presentation to do beleaguered marketing. The authority of data mining tools to remove concealed information from huge collections of data; guide to the augmented data compilation efforts by companies and government agencies. Obviously this elevated privacy disquiets pertaining were to composed data. In reply to that, data mining, researchers happened to deal with privacy concerns by mounting individual data mining techniques, under the structure of "privacy preserving data mining". It is different when compared to usual data mining techniques. Privacy preserving data mining can be useful to databases without breaching the privacy of persons. A data perturbation process can be basically explained as follows. Before the data possessors distribute their data they need to alter the data in definite way which will be concealing outfitting for the responsive information while protecting the scrupulous data assets which is dangerous is building significant data mining models. Perturbation techniques have to touch the inherent tradeoff among preserving data privacy and preserving data efficacy as perturbing data frequently decreases data efficacy. The data perturbation procedure is shown in Fig. 1.

Achieving privacy conservation for perturbed data clustering is a demanding problem. To lecture this problem, data proprietors must not only congregate

Corresponding Author: V.S. Prakash, Department of Computer Applications, Bannari Amman Institute of Technology, 638401 Sathyamangalam, Tamil Nadu, India

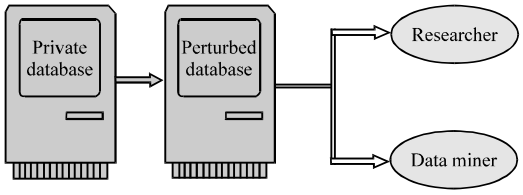

Fig. 1: Data perturbation

privacy necessities but also pledge suitable clustering results. The essential question processed in this study is: How can associations defend personal data subjected to clustering, gathering their requirements to hold decision making or endorsing communal benefits? Obviously, sharing data for clustering facades novel confronts for narrative uses of data mining knowledge. In this study, researchers propose an individually adaptable perturbation model, solemnly which enables the individuals to choose their own privacy levels.

Literature review: In e-Business Data Mining Models, privacy becomes a key issue in preserving individual's data on product/service transactions. Nevertheless, the precision and revelation of the product/service augment the amount of operation to more novel and accessible clients. To assess a tradeoff among data privacy and precision of individual's data, many data perturbation techniques are presented with substantiation of several researchers.

The rising participatory sensing of requests rely on individuals' efforts which might be extremely connected with individuals' responsive information or individual data. Therefore, privacy fortification is vital to hearten individual contribution for participatory intellection applications to produce dependable and high excellence data. An extensively used method for privacy conservation is data perturbation which appends noises to unique data at the customer side to defend individuals' isolation and permits the information server to rebuild the figures of the original data. Zhang *et al.* (2012) discover a serious susceptibility of existing data perturbation algorithms in which an opponent might develop for reinstating other users' private information (mean, variance and the sharing of unique data) from the agitated data, since all the participants distribute the similar noise allocation. Perturbation technique is a very significant method in privacy preserving data mining. In (Liu *et al.*, 2008) defeat of information versus conservation of privacy is recognized as a trade off among the users in it.

The exercise of clustering as a data examination tool has elevated disquiets regarding the contravention of individual privacy. Data objects that have been spitted into clusters employing k-means clustering are troubled by processing geometric alterations on the clusters in such a way that the object association of every cluster and direction of objects inside a cluster stay behind the same (Dhiraj *et al.*, 2009). In addition, the effectiveness of the dataset must be measured as a conversion that takes place. Such data conversion crisis privacy customary must be assembled and the efficacy must be optimized is which called as NP-hard crisis. Natwichai (2010) proposed an estimate algorithm for the data conversion crisis. The focused data processing addressed in this study is categorization of employing connection rule or associative categorization.

Privacy-Preserving Data Mining (PPDM) is one of the current inclinations in privacy and security investigation. Current advances in data compilation, data distribution and associated technologies have instated a novel period of research offering Data Mining algorithms which should be reassessed from a diverse point of view that enhances confidentiality protection. Banu and Nagaveni (2009) discovered all the features of privacy concerns in data mining, particularly connected with clustering and gives a method for privacy preserving grouping with a theoretical banking situation. Here, the researcher proposed a representation for grouping horizontally partitioned or central data sets using an easy PCA based alteration approach. The Naive Bayes categorization has been employed as one of its applicability in case of evaluation dataset (Keshavamurthy *et al.*, 2010).

Along with the existing privacy preserving techniques, data anonymization presents a effortless and efficient way to defend the responsive data. Nevertheless, in most of the connected algorithms, data particulars are misplaced and the outcome dataset is far less instructive than the unique one (Yang, 2008). In topical years of data mining requests, an efficient method to protect privacy is to anonymize the dataset that comprise private information before being unconfined for mining (Deivanai *et al.*, 2011). Liu *et al.* (2006) discovers the outlook of using multiplicative subjective projection matrices for secrecy conserving distributed data mining. Kamakshi and Babu (2010) proposed a structure that permits universal alteration of unique data using randomized data perturbation technique and the modified data is then offered as a conclusion of client's query through cryptographic technique. A protocol (Sang *et al.*, 2012) can be employed to sustain such investigates in a privacy-sensitive manner. In this research, researchers plan to provide privacy preservation scheme for perturbed data in multi-partitioned datasets. Privacy-Preserving Data

Mining (PPDM) distress the predicament of realizing data mining tasks, Vaidya and Clifton (2009) devoid any straight admittance to the exclusive data sets as the providers preserve isolation on their data.

MATERIALS AND METHODS

Proposed enhanced security mechanism for perturbed data in multi-partitioned dataset: The proposed research is efficiently designed to enhance the process of privacy preservation for perturbed data present in the multi-partitioned datasets. The proposed enhanced security mechanism for perturbed data in multi-partitioned is processed under three different phases. The first phase describes the process of sharing of multi-partitioned data with the users. The second phase describes the process of clustering the multi-partitioned data with the divisive k-neighbor clustering procedure. The third phase describes the process of enhancing the privacy preservation mechanism for perturbed data in multi-partitioned dataset. The architecture diagram of the proposed enhanced security mechanism for perturbed data in multi-partitioned datasets is shown in Fig. 2.

Multi-partitioned data set consists of data which have been divided from any logical database. The data has been partitioned into two types: horizontally data partition and vertically data partition. Horizontal partitioning engages setting diverse rows into diverse tables. A general outline of vertical partitioning is to divide dynamic data from static data in a table where the dynamic data is not used as often as the static. Generating a view for the two tables, re-establishes the unique table with a performance penalty while increasing the performance by accessing the static data, e.g., for statistical analysis.

From the Fig. 2, it is being observed that the privacy preservation is achieved by adapting the adaptable perturbation model. Before adaptation of privacy preservation scheme, divisive k-means clustering is applied to the multi-partitioned dataset which begin with individual inclusive cluster and at every step, divides a cluster until only singleton clusters of entity points stay behinds. In this case, researchers call for to choose at every step which cluster to be divided and how to carry out the split. The divisive k-means cluster is efficiently used for privacy preservation mechanism which overcomes the issue of data perturbation technique that raises the ambiguity among the clustered datasets results in unreliability. The divisive k-Means is supported by the idea that an axis point can symbolize a cluster. After clearing the ambiguity of the partitioned dataset, the adaptable perturbation model is used to enhance the privacy preservation scheme among the partitioned datasets.

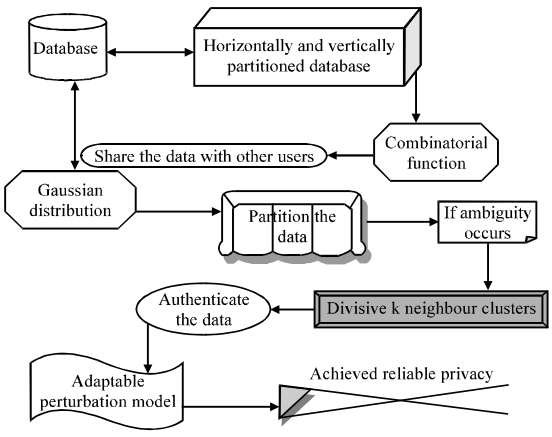

Fig. 2: Architecture diagram of the proposed APPDM

Sharing the multi-partitioned dataset using combinatorial function: The critical design of data perturbation is to modify the data so that valid individual data values cannot be improved while preserving the value of the data for statistical summaries. The data doesn't reacts the real values of private data even if a data item is linked to individual the individual's privacy is not violated. If user A and B wants to partition the data in the database, out first work present an appropriate technique in a horizontal and vertical manner. To partition the dataset horizontally, gradient descent model is adapted here. For vertical partition of data in the dataset, kth element vector technique is employed. After partitioning the data either horizontally or vertically from the database they have some sets of data to share it with other users in a safe and secure manner. That is, no other third party should involve in this sharing data between the users. Then, users can partition the data from the database in a vertical manner. They receive some sets of vertically partitioned data sets.

To share the data sets with different users along with privacy preservation, the first research presented a technique named data perturbation technique which is used to share the data from different users and unite all those data to get one complete true data sets. The combinatorial function is used to preserve the data sets which are to be shared among the users and it does not allow the third party members to seize the data. The combinatorial function will allow the users to share both the horizontal and vertical partitioned data sets to share with different users.

Clustering the partitioned dataset with divisive k-neighbor clusters: Here in this research, researchers are going to present a divisive k-means clustering to

remove the ambiguity occurred in multi-partitioned dataset. The alternative of hierarchical clustering is called top-down clustering or divisive clustering. Researchers establish all datasets in one cluster. The cluster is divided using a Flat Clustering algorithm. This practice is functioned recursively until every dataset is in its individual singleton cluster.

k-Means algorithms are admired and extensively used for Clustering Methods. They divide the data to diminish the principle:

$$E=\sum_{j=1}^{k}\sum_{x_i\in S_j} d^2\left(x_i, s_j\right) \quad (1)$$

Where:

K = The amount of clusters

$s_j=\sum_{x\in S_j} x_i/|S_j|$ = The axis of cluster of S_j

d(a, b) = The Euclidean distance

Divisive k-Means Clustering algorithms regularly occupy an arbitrary primary division or centers and continually recompute the centers supported on division and then re-evaluating the division based on the centers. Such, process can be established to congregate a minimum restriction whereas the crisis of recognizing the inclusive minimum is NP-hard. Researchers suggest a hierarchical divisive k-Means algorithm that reduces the same principle as the standard k-means with clustering process planned as a hierarchical process.

For a given set of N items to be grouped and an N×N distance (or resemblance) matrix, the procedure of divisive hierarchical clustering is this:

1) Establish by conveying every item to a cluster so that if you enclose N datasets, you now enclose N clusters each comprising just one item. Let the distances (resemblances) among the clusters will be similar as the distances (similarities) among the items they include
2) Discover the contiguous (most similar) pair of clusters and combine them into a particular cluster so that now you contain one cluster less
3) Calculate resemblances between the new group and each of the old groups
4) Reiterate steps 2 and 3 until all items are grouped into a distinct cluster of size N(x)

Divisive clustering initiates from the top indulging the entire dataset as a cluster. It constantly divides a present cluster (a leaf node in a binary tree) until the amount of clusters achieves a pre-defined value K or some other ending measures are met.

Adaptable perturbation model for privacy preservation scheme: The perturbation technique is processed on establishing noise with not considerably altering the allocation of the unique data. Consequently, diverse geometric techniques are functioned to the perturbed data to modernize the original allocation. Information hammering in opposition to privacy protection is forever a trade off in this method. The amount to which researchers agitate the unique data can noticeably concern the data mining consequences and consequently donate to the possible hazard of privacy revelation. At the similar time, deciding the suitable stage of perturbation is not insignificant. Consider the subsequent noise condition system, Let $a_1, a_2, ..., a_n$ be the unique values of a one-dimensional allocation as recognition of n autonomous identically dispersed (iid) arbitrary variables, all has the similar allocation as the arbitrary variable A. Let $b_1, b_2, ..., b_n$ be the arbitrary values exercised to deform the unique data, b_i is the recognition of n independent identically dispersed (iid) arbitrary variables, all has the similar allocation as the arbitrary variable B. Either consistent or Gaussian allocation with the subsequent properties is used to produce arbitrary variable B:

- Uniform distribution: the arbitrary variable has an identical allocation over a time. The mean of the arbitrary variable is 0
- Gaussian distribution: the arbitrary variable has a standard allocation with $\mu = 0$ and standard deviation σ
- Given $a_1+b_1, a_2+b_2, ..., a_n+b_n$ (perturbed data W) and collective possibility distribution F_B (noise), determine possibility distribution F_A (of unique data)

Researchers can observe that the noise calculation process explained above is only one step that is researchers adjoin the noise to the unique data and then concern the renovation algorithm to approximate the unique distribution. This is termed as an one phase distribution model.

In two phase perturbation model as shown in Fig. 3, researchers first split the province of the W into determined periods. After producing the $w_i = a_i+b_i$, researchers compute the fixed interval $[l_k, l_{k+1}]$ which w_i falls. As a substitute of employing w_i through the renovation phase, researchers use a w'_i that is produced consistently from $[l_k, l_{k+1}]$. Clearly w' is formed as iid. If the periods used for case are selected minute sufficient, the second phase does not result the collective distribution function (c.d.f) of W. To see the fact that W and W' have

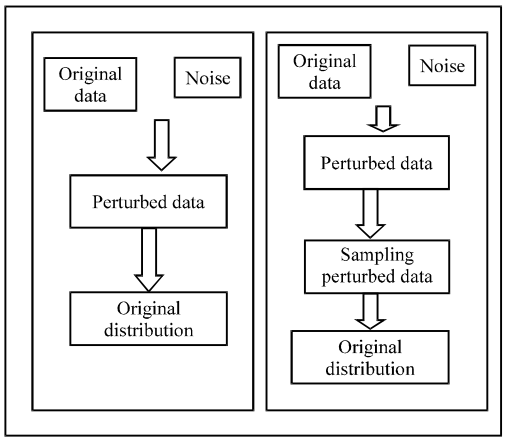

Fig. 3: Two phase perturbation model

similar cumulative allocation functions for minute intervals; check the reflection on the association among the c.d.f of W and W'. In Fig. 3, W'-few communicates to the case where total three periods are used to build the W' data set, likewise W'-med is formed utilizing six periods; W'-many is formed using 20 periods.

Researchers can observe that when the period is escalating, the c.d.f of W' is getting nearer to the c.d.f of W. In observation, researchers put the number of periods such that c.d.f of W' is secure to W.

The two-phase perturbation model could be simply customized to integrate diverse individual privacy inclinations. Note that for every w_i, researchers are sampling a arbitrary position from the period where $w_i \in [l_k, l_{k+1}]$. Obviously this sample could be completed using diverse intervals for diverse users. For instance, a user who is more careful could employ the period $[l_{k-1}, l_{k+2}]$ for creating w'_i (i.e., twice the original size).

Using this reality, researchers can portray independently flexible perturbation technique as follows:

- The system first adds a random noise $B = (b_1, b_2, ..., b_i, ..., b_n)$ to the unique data values $A = (a_1, a_2, ..., a_i, ..., a_n)$ to get $W = (w_1, w_2, ... ,w_i, ... ,w_n)$
- User i decide his/her privacy level with diverse privacy levels shown in Fig. 4
- Based on the user's confidentiality level selection, the system pertain a time length $[l_i, l_j]$ that proceeds to the selected privacy level. Later on, w'_i is formed by sampling consistently from the time $[l_i, l_j]$
- w'_i value is sent to the data miner

The association among diverse privacy inclination and perturbation point is shown in Fig. 4. Every privacy level described in Fig. 4 will have an analogous period

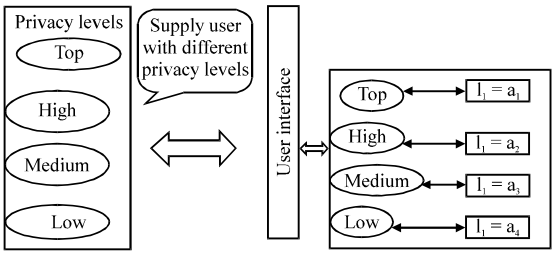

Fig. 4: Privacy levels vs. perturbation levels

length and this diverse interval duration for diverse individuals will facilitate the individual flexibility. Once the data miner obtains the w'_i values, Bayesian based rebuilding approach could be exercised to rebuild the unique data allocation. Obviously, if the duration of the selected intervals is huge, the specified data will appear more like a arbitrary model. If the duration of the selected interval is small w'_i, then it will be much stopped to the w_i. At the similar time if the number of users who prefer to contain an outsized interval to example from is little, cumulative allocation function F_W and $F_{W'}$ will not be too much diverse. In the proposed APPDM Model, customers can decide diverse time lengths to adapt their privacy levels.

Experimental evaluation: The proposed system enhancing the Applicability of Privacy Preservation for Perturbed Data in Multi-partitioned datasets (APPDM) has been implemented in Java. The experiments run on an Intel P-IV machine with 2 GB memory and 3 GHz dual processor CPU. Researchers are going to compare the proposed system which enhances the Applicability of Privacy Preservation for Perturbed Data in Multi-Partitioned Datasets (APPDM) with the previous researches and an existing technique or single partitioned datasets. Using combinatorial function, the datasets are partitioned effectively as such horizontally or vertically. So, the scalability of the products/services became less. After that the strength of the data set to be shared remains identical from the beginning of the dividing process employing divisive k-Means Clustering algorithm. After successfully removing the ambiguity occurred over dataset in this research, researchers efficiently present an individually adaptable privacy preservation scheme to enhance the privacy of perturbed data in multi-partitioned datasets. The performance of the proposed enhancing the applicability of privacy preservation for perturbed data in multi-partitioned datasets is measured in terms of:

- Privacy level
- User density

- Perturbed data objects
- Adversary attack rate
- Scalability

RESULTS AND DISCUSSION

The proposed system enhances the applicability of privacy preservation for perturbed data in multi-partitioned datasets is reliably made for enhancing the privacy preservation in multi-partitioned dataset. The proposed APPDM allowed the users to share their file with other users by achieving a safe and secure communication. An experimental evaluation has also been carried out with the benchmark dataset to estimate the performance of the proposed APPDM. Figure 5 describes the performance of the proposed APPDM and compare the results with an existing techniques for single partitioned datasets and also with the previous researches.

Figure 5 describes the privacy level of the users in the environment with their shared data in it. The privacy level of the proposed system enhances the applicability of privacy preservation for perturbed data in multi-partitioned datasets is compared with an existing techniques for single partitioned datasets and with the previous researches CPPDP (Cluster Based Privacy Preserving Data Perturbation Technique) and CF (Combinatorial Function).

Normally, multi-partitioned data contains both vertical and horizontal data sets which are present is command of e-Business and e-Commerce data mining background. In e-Business data mining representations, privacy turns into an issue in preserving individual's data on product/service transactions. In the proposed APPDM, the privacy level of the individuals' data is efficiently preserved and processed over the adaptable data format. In the previous researches, the dataset in the database are efficiently partitioned over both horizontally and vertically and shared with the users' involved in the transaction.

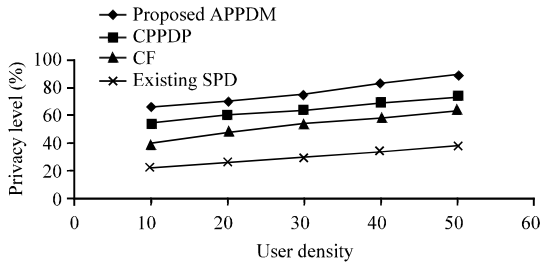

Fig. 5: User density vs. privacy level

While partitioning the dataset there is a great chance for the ambiguity raised in the partitioned dataset. To remove the repeatability of data, the k-Divisive algorithm is presented. Both the previous research does not concentrates more on the privacy preservation. So, compared with existing techniques for single partitioned datasets and with the previous researches CPPDP (Cluster Based Privacy Preserving Data Perturbation Technique) and CF (Combinatorial Function), the proposed APPDM provide a high privacy level for individuals' data on partitioned datasets. The variance in privacy levels is 40-50% high in the proposed APPDM.

Figure 6 describes the adversary attack rate of accessing unauthenticated data and the efficiency of a privacy preservation of data in the network. The adversary attack rate of the proposed system enhances the applicability of privacy preservation for perturbed data in multi-partitioned datasets is compared with an existing techniques for single partitioned datasets and with the previous researches CPPDP (Cluster Based Privacy Preserving Data Perturbation Technique) and CF (Combinatorial Function).

Figure 6 describes the adversary attack rate of accessing unauthenticated data of the users involved in the transaction of message among different users based on the number of perturbed objects present. In the proposed APPDM for privacy preservation, researchers have implemented successfully individually adaptable perturbation techniques which followed the preservation of privacy in multi-partitioned dataset. Since, each objects/data in the dataset are clustered until each item in the dataset is clustered with a singleton cluster size there is a less chance of multi-partitioned data set to be hacked by the adversaries and also the privacy level of individuals' data is determined based on the two phase perturbation model by sampling the perturbed data objects. So, in the proposed APPDM, the adversary rate of accessing the multi-partitioned data in the dataset is very less compared to an existing techniques for single partitioned datasets and with the previous researches CPPDP (Cluster Based Privacy Preserving Data

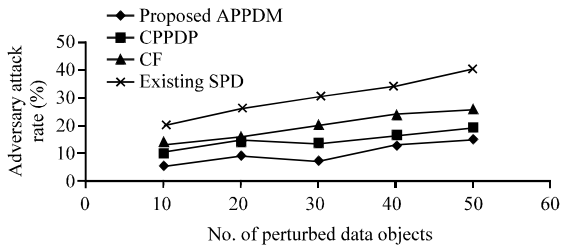

Fig. 6: No. of perturbed data objects vs. adversary attack rate

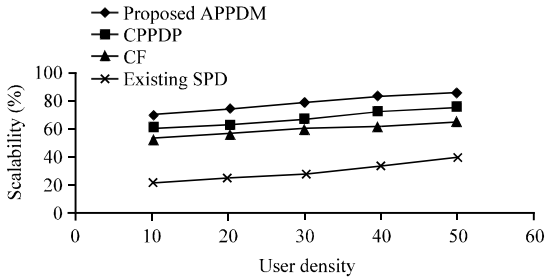

Fig. 7: User density vs. scalability

Perturbation Technique) and CF (Combinatorial Function). The variance in the adversary attack rate in the proposed APPDM is 50-60% low.

Figure 7 describes the scalability of the privacy of the users in the environment with their shared data in it. The privacy level of the proposed system enhances the applicability of privacy preservation for perturbed data in multi-partitioned datasets is compared with an existing techniques for single partitioned datasets and with the previous researches CPPDP (Cluster Based Privacy Preserving Data Perturbation Technique) and CF (Combinatorial Function).

Since, the research work efficiently provides a secure perturbed data transaction among different users, the scalability of privacy preservation in the proposed APPDM is high when compared to the existing simple technique for single partitioned dataset.

Finally, it is being observed that the research work efficiently achieved the cluster based privacy preserving data perturbation technique to mine multi-partitioned data sets. The privacy preservation is done by adapting the step by step process and at last it achieves better results in terms of privacy level, adversary rate and scalability compared to the other researches.

CONCLUSION

In this research, researchers have efficiently achieves the cluster based privacy preserving data perturbation technique to mine multi-partitioned data sets. Generally, multi-partitioned datasets combined both horizontal and vertical partitioned datasets. The proposed APPDM work efficiently deliberates the privacy preservation mechanism by sampling the perturbed data in a subsequent manner. To share the data with other users, researchers first partition the datasets effectively in both horizontal and vertical manner. Then, clustering process is done efficiently which enhances the privacy preservation scheme by adapting the two phase perturbation model.

Compared to an existing Combinatorial Function (CF) for multi-partitioned dataset, the proposed system enhances the applicability of privacy preservation for perturbed data in multi-partitioned datasets performs well and an experimental evaluation has been carried over with bench data sets obtained from popular e-Business/e-Commerce sites (Amazon, e-Bay etc.) estimates the performance of the proposed system enhances the applicability of privacy preservation for perturbed data in multi-partitioned datasets in terms of privacy level, adversary attack rate and scalability.

REFERENCES

Banu, R.V. and N. Nagaveni, 2009. Preservation of data privacy using PCA based transformation. Proceedings of the International Conference on Advances in Recent Technologies in Communication and Computing, October 27-28, 2009, Kottayam, Kerala, India, pp: 439-443.

Deivanai, P., J.J.V. Nayahi and V. Kavitha, 2011. A hybrid data anonymization integrated with suppression for preserving privacy in mining multi party data. Proceedings of the International Conference on Recent Trends in Information Technology, June 3-5, 2011, Chennai, Tamil Nadu, China, pp: 732-736.

Dhiraj, S.S.S., A. Khan, W. Khan and A. Challagalla, 2009. Privacy preservation in k-Means clustering by cluster rotation. Proceedings of the IEEE Region 10 Conference, January 23-26, 2009, Singapore, pp: 1-7.

Kamakshi, P. and A.V. Babu, 2010. Preserving privacy and sharing the data in distributed environment using cryptographic technique on perturbed data. J. Comput., 2: 115-119.

Keshavamurthy, B.N., M. Sharma and D. Toshniwal, 2010. Privacy-preserving Naive Bayes classification using trusted third party and different offset computation over distributed databases. Proceedings of the 1st International Conference on Parallel Distributed and Grid Computing, October 28-30, 2010, Solan, India, pp: 362-365.

Liu, K., H. Kargupta and J. Ryan, 2006. Random projection-based multiplicative data perturbation for privacy preserving distributed data mining. IEEE Trans. Knowl. Data Eng., 18: 92-106.

Liu, L., M. Kantarcioglu and B. Thuraisingham, 2008. The applicability of the perturbation based privacy preserving data mining for real-world data. Data Knowl. Eng., 65: 5-21.

Natwichai, J., 2010. An approximation algorithm for privacy preservation of associative classification. Proceedings of the International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology, May 19-21, 2010, Chiang Mai, Thailand, pp: 127-131.

Sang, Y.P., H. Shen and H. Tian, 2012. Effective reconstruction of data perturbed by random projections. IEEE Trans. Comput., 61: 101-117.

Vaidya, J. and C.W. Clifton, 2009. Privacy-preserving Kth element score over vertically partitioned data. IEEE Trans. Knowl. Data Eng., 21: 253-258.

Yang, W., 2008. Knowledge reserving in privacy preserving data mining. Proceedings of the 2nd International Symposium on Intelligent Information Technology Application, Volume 2, December 20-22, 2008, Shanghai, China, pp: 855-859.

Zhang, F., L. He, W. He and X. Liu, 2012. Data perturbation with state-dependent noise for participatory sensing. Proceedings of the IEEE 31st Annual IEEE International Conference on Computer Communications, March 25-30, 2012, Orlando, FL., USA., pp: 2246-2254.