A Novel Technique To Prevent The Sensitive Data Using Non-Cryptographic Methods

[1]Mr.V Dilip Kumar, [2]K Rajini [3]J Suja Pushpa [4]M Monica
[1]Assistant Professor [2, 3, 4] Final B Tech Students [1,2,3,4] Department of Computer Science Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada.

Abstract: *Current approach in information, means, data mining, and security automation have gave rise to a new era of exploration, known as Privacy Preserving Data Mining (PPDM). Privacy conserve data mining has become progressively popular because it allows sharing of privacy conscious data for analysis aspiration. Several data mining algorithms, consolidate privacy preserving structure, have been matured that allow one to extract applicable knowledge from large amount of data, while disguise responsive data or information from acknowledgment or inference. Finally, we share assignment for future analysis. We provide a review of the state-of-the-art design for privacy and consider the model technique for isolation preserving data mining.*

Keywords*-Privacy-preserving data mining (PPDM), data mining, research challenges, privacy preserving techniques, slicing.*

I. INTRODUCTION:

Technology on the rise has always appeal ease of use, easy access, scalability and most particularly retreat of user data .Privacy has been a involvement since the invent of internet. Every single development from booking tickets to making international economic transactions data is stored in electronic form. Computerized data is as accessible to as data in its physical form. Various algorithms past few years have been on the front to provide maximum conservation to private data and to overcome the conditions of existing algorithms etc. are some of the PPDM approach that have proved to be adequate to prevent data when it is being reserve. However independently the methods have few deficiencies, thus a different access for PPDM can be made. Connecting the PPDM techniques administer more robust, durable and secure algorithm. Thus in our research we have mainly attract on connecting two techniques of PPDM i.e. Slicing and Cryptography. Both techniques have demonstrated to provide best separation to user data, thus this excite us to combine these techniques to achieves a new algorithm that administer a simple yet an efficient way of derive data without risking the confidentiality of data. The paper is coordinated as follows. In category I, we give the basic concept of PPDM and its capability. In Section II, we characterize the related work done in the range of PPDM. In section III, we have Specified the research objection in PPDM. Section IV consist of the proposed procedure that tries to affected the flaws of PPDM. And finally we complete in Section V.

II. RELATED WORK

Yehuda Lindell, Benny Pinkas [3], presented establishment to secure multiparty computing and its application to privacy-preserving data mining. The accepted errors that are traditional in the conserve data mining is achieve with secure multiparty calculation techniques and the issues complex in the efficiency are consider and also determine the difficulties in compose highly efficient covenant.

Sweety R. Lodha, S. Dhande [4], explained encryption algorithm resolve at three different levels in the paper Web table Security Algorithms in this paper Encryption is divided into three contrasting levels i.e. Storage-level encryption, Database-level encryption, Application-level encryption. Storage-level encryption encrypts the data in the stored in subsystem and hence assure the static data stored. From a database point of view, storage-level encryption is transpired thus any changes to current applications are escape easily. Database-level encryption provides guarantee when data is being inserted or recapture from database. Application-level encryption observe the encryption and reading process at application level where the data is generates Within the application that begin the data into the system encryption is achieve; the data is encrypted and then sent, thus consistently the

data is reserved stand encrypted data is retrieve which is finally decode again within the utilization.

Yuan-Hung Kao, Tung-Shou Chen and Jeanne Chen [5], expected a novel hybrid conservation scheme that conserve the privacy of information and the gather knowledge in data mining. The prospective scheme accommodate the privacy – preserving data mining technique with that of knowledge-protect anti-data mining capability. The given scheme allows user to adapt the amount of preservation on personal level.

Hanumantha Rao Jalla and P N Girija [6] , proposed an algorithm that direction the problem of respective customers related to their privacy concern. Authors expected a conversion technique. This basis of this technique has been attribute from Walsh- Hadamard conversion (WHT) and one of its fundamentals i.e. Rotation in their paper. An orthogonal matrix is achieve by the WHT, it transfers entire data into new sphere and also maintain the length between the data evidences. Techniques which are analytical based can be used to modernize the records, so by administer Rotation transformation this complication is resolved. Inverse matrix is one this capability

Sativa Lohiya and Lata Ragha [7], proposed a hybrid capability in which randomization and abstraction is used in their paper. In this access the data is first randomized and then abstraction is performed on the modified or randomized data. This technique assures private data and regenerate original data with better efficiency and with no clues loss.

Tiancheng Li, Ninghui Li,Ian Molloy and Jian Zhang [1], presented a new access called slicing. Slicing is used to conserve privacy of micro data. The constraint of observation and took decision are overcome by this method. Advantage is preserved well in Insider Threats while conserve against threats associated to privacy of data. Analysis show that data utility is much better conserve by allowance than generalization and its ability is more than bucketization in case of assignment that involve the sensitive characteristic.

III. RESEARCH CHALLENGES

Now- a-days, Data Mining is used in much utilization. There are convinced areas where data mining if used without privacy may cause genuine affects .These areas are the main exploration objection and are mentioned below.

Internal and independent attacks, Cyber impendence. One of the major intimidation people face today is Cyber Crime [9]. Since most of our report is stored on electronic publishing and a lot of data is also accessible on internet or networks. Aggression on such areas might be critical and destructive for an individual. For example, contemplate the Banking system. If operator attack a bank's instruction system and empty the explanation, the bank could lose gathering of dollars. Therefore guarantee of data is a analytical issue. There are two types of risk

–Outsider or assistant. An attack on Information System from someone outside the grouping is called foreigner threat, such as hackers, hacking Bank's analog systems and causing havocs. A more critical complication is the insider threat. Accomplice threat can be due to an squatter present in the grouping. Members of an system have studied their policies and business proceeding and know every bit of the intelligence so it can affect the organization's data assets.

B. Fraud in Credit Cards and Individual's continuation Theft Another area which desire consideration is disclose frauds and thefts. Frauds may be credit card pressure. These can be disclosing by analyze acquisition made of enormous amounts [9]. A similar and a more conscious theft is identity theft. Here one estimates to be an status of another person by connection that person's personal report and carrying out all types of concession under the other person's name. By the time, the owner finds out it is often far too late-the victims may already have lost millions of dollars due to presence theft**.**

C Flaws unparticular techniques

PPDM has a huge list of capability with different approach and thought. However every individual technique in its own has some flaws which development the challenge for scheming a better algorithm, the respective flaws are stated below.

• Anonymization: Since Anonymization generates translate data, its accuracy of applications on the data is reduced [8]. Available or unavailable characteristic in external table are difficult to resolve in k- anonymity model.

• Cryptographic Technique: For huge directory this algorithm doesn't proves to be a strong technique as this technique fails to protect the output of estimation [8]. Thus as a result mining the result may break the privacy of individual's record [9].

• Data Perturbation: conserve the original data becomes difficult in some anxiety access. Data mining technique is to be elected based on the design using which noise has been received in data [9].

• Randomization: Each records are treated individually heedless of their local density [8].

• Generalization: A appreciable amount of information is lost for high structural data in observation [1].

• Bucketization: enrollment confession is not stopped in this method and clear departure between sensitive characteristic and quasi-identifying attributes is a must for this method additional the method is inapplicable [1].

IV.PROPOSED METHODOLOGY

Our concept of merging contrasting techniques aims on linking cryptographic technique and slicing. Cryptography has different way to provide privacy. Authentication, Encryption, key transaction, etc is some of the basic techniques which when converted provide a high level of preservation thus making it nearly demanding to break into an individual's privacy. Cryptography has been one of the most used privacy hindrance technique in multiparty data estimation. This method prevents flow of computations. portion was one of the techniques imported that overcame the defect of generalization and bucketization. Membership exposure and preserving better data utility are the convenience of slicing. Slicing as the name says separation the data set or attributes diagonally as well horizontally.

Since cryptography aims at assure leakage of private computing result and slicing aims at preserving better data service each method holds some drawback. Thus our concept combine different level evidence and database level slicing. Connecting these two path ensures user level privacy and index level privacy. A robust algorithm is thus introduced in this paper

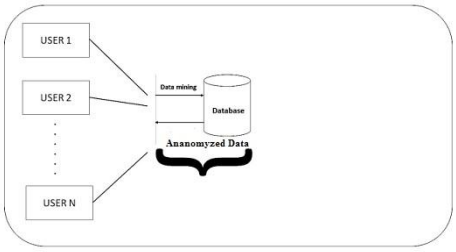

Fig. 1. Basic framework of the proposed approach

Data posted in database may have unknown capacity. Thus in order to handle all dimensions of data exclusively high –dimensional data slicing acts as a major backing factor. Since slicing slices the dataset on the bias and vertically, it aims at breaking cooperative across the files but at same time preserving the federation within each column. Slicing assure database level surveillance as sliced data may have least association with other records thus compressing the risk of leaking added private data which is not desired. As shown in figure 1,a request from user is prepared and thus the data highly correspond with the data requested are organize together using the slicing algorithm. Thus cryptography provide user level privacy whereas slicing ensures database matched privacy. Cryptography and Slicing form a robust hybrid technique for privacy conserve in data mining.

V. CONCLUSION

This paper has introduced a robust, stable and effective method for preserving data privacy at different platform. Since data online are the most vulnerable one this hybrid technique can be used over internet. Implementation of the algorithm guarantees security to a higher extent. However further research may make this technique much more unpredictable and difficult to break. The two level authentication proves to be an impact factor as a fresh approach of key exchange and authentication are used at the same time. Further research may reduce the overhead on the two level authentication algorithm. Slicing at the basic level supports cryptography in the given approach thus plays an important role at database level. Hybrid techniques have always proved to be a better approach for PPDM, thus overcoming different flaws and providing a better mean for preserving data privacy.

REFERENCES

[1] Tiancheng Li, Ninghui Li, Jian Zhang, and Ian Molloy, ―Slicing: A New Approach for Privacy Preserving Data Publishing‖, in proceedings of IEEE Transactions on Knowledge and Data Engineering, Vol. 24, No. 3,pp. 561-574, Mar. 2012.

[2]Anand Sharma and Vibha Ojha, ―Implementation of Cryptography for Privacy Preserving Data Mining‖, in International Journal of Database Management Systems (IJDMS), Vol.2, No.3, Aug. 2010.

[3]Yehuda Lindell and Benny Pinkas, ―Secure Multiparty Computation for Privacy-Preserving Data Mining‖ ,The Journal of Privacy and Confidentiality, Number 1, pp. 59-98,2009.

[4] Sweety R. Lodha and S. Dhande, ―Web Database Security Algorithms‖, in International Journal of Advance Research in Computer Science and Management Studies (ijarcsms), Volume 2, Issue 3, pp.293-299, Mar 2014

[5]Tung-Shou Chen, Jeanne Chen and Yuan-Hung Kao, ‖A Novel Hybrid Protection Technique of Privacy Preserving Data Mining and Anti-Data Mining‖, in Information Technology Journal, Volume 9, Issue 3,pp. 500-505,2010.

[6] Hanumantha Rao Jalla and P N Girija,―An Efficient Algorithm For Privacy Preserving Data Mining Using Hybrid Transformation‖, in International Journal of Data Mining & Knowledge Management Process (IJDKP) Volume 4, Number 4, July 2014.

[7]Savita Lohiya and Lata Ragha,―Performance Analysis of Hybrid Approach for Privacy Preserving in Data Mining‖, in proceedings of Int. J. on Recent Trends in Engineering and Technology, Volume 8, Number. 1, Jan. 2013.

[8] Dharmendra Thakur, Prof. Hitesh Gupta, ―An Exemplary Study of Privacy Preserving Association Rule Mining Techniques‖, in proceedings of International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE) , Volume 3, Issue11, pp. 893-900, Nov. 2013.