International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-3 Issue-4, April 2014

Data Hiding and Secured Data Storage with Access Control towards Multiparty Protocols

T.S.Keerthiga, S.Sarika

Abstract—Secure multiparty protocols is used as third party protocols in the data hiding and security. The major problem is, there is no Security Scheme operated for Data Storage Services between Multi Party protocols. To overcome this look-ahead approach, specifically for secure multiparty protocols to achieve distributed k-anonymity, which helps parties to decide if the utility benefit from the protocol is within an acceptable range before initiating the protocol. The look-ahead operation is highly localized and its accuracy depends on the amount of information. The system deals with Generalization approach, with a common Identification. Suppression approach, used for Hiding User Identity. In the secure random key algorithm, an Authentication Key is generated before a user change/update the data for Verification. Entire Data is encrypted to ensure Security.

Index Terms—Security, Privacy, k-anonymity, Multi Party protocols.

I. INTRODUCTION

In the data mining,data will access the cluster of data into user data to get the information.A Look-ahead approach is to provide security.The main purpose in this project is to hide the person details where hacker cannot retrieve any information.Eventhough third party can see it but entire details they cannot view it.privacy preserving technique is used to prevent the security in the form of sensitive and non-sensitive data.SMC protocols used for privacy preserving technique where the communication and computation cost is high. The protocol can calculate the output only within their function.Linking attacks are performed by adversaries to form the attributes of individuals in the dataset.k-anonymity is based on privacy preserving technique to prevent attacks on shared databases.Data might be horizontally or vertically over multiple parties,the main purpose is participation using larger dataset to create a better utilized k-anonymization. Optimal distributed k-anonymization is done over mainfunctionality to form the fall back function.To minimize the overfitting, instead of using the exact statistics histograms to form the number of tuples. In the descendant preserving distributed k-anonymization is to form decision predicate. The histogram contributes inorder to form the amount of information bounded by their local anonymizations.Decision predicate used in SMC protocol where benefit cost is high. A look ahead on vertically partitioned data involves a comparative quantification of utility over different projection to form cost metrics.In look–ahead SMC the parties have main functionality,to form some information.

Manuscript published on 30 April 2014.
***** Correspondence Author (s)
T.S. Keerthiga, Computer Science and Engineering, Sathyabama University, Chennai, India.
S.Sarika, Computer Science and Engineering, Sathyabama University, Chennai, India.

Here, they used credit card transaction where in the employee details they can view only their information they cannot see others information.In that by preventing security sensitive information can hide the details.

II. RELATED WORK

M.Kantarclu and C.Clifton(2004) developed a paper for secure mining of association rules over horizontally partitioned data. Here cryptographic technique is used.Association rule is used as a limitation. The paper presents a solution that preserves such secrets – the parties learn almost nothing beyond the global results. The solution is efficient: The additional cost relative to previous non-secure techniques is encryptions, and a constant increase in the number of messages. Here, the problem of privately computing association rules in vertically partitioned distributed data has also been addressed. The vertically partitioned problem occurs when each transaction is split across multiple sites, with each site having a different set of attributes for the entire set of transactions.

R.C.-W. Wong, J. Li, A.W.-C. Fu, and K. Wang(2006) developed a paper for protecting the individual identification. Privacy preserving technique is also used in data mining applications. Here the limitation used to minimize a distortion effect.In this k-anonymity property is done by quasi identifier. The *k*-anonymity model assumes a quasi-identifier,which is a set of attributes that may serve as an identifier in the data set. Hence, propose an (α, k)-anonymity model to protect both identifications and relationships to sensitive information in data. In global recoding method for the (α, k)-anonymity problem. Next they propose a local-recoding algorithm which is more scalable and result in less data distortion. In global recoding, all values of an attribute come from the same domain level in the hierarchy.In local recoding,values may be generalized to different levels in the domain.One advantage is that an anonymous view has uniform domains but it may lose more information.

B.-C. Chen, K. LeFevre, and R. Ramakrishnan (2007) developed a paper for measuring disclosure and sanitizing data in the presence of external knowledge.To improve the computational efficiency over the techniques.It is based on the special forms of knowledge,where it is efficient and scalable. Hence, propose a novel multidimensional approach to quantifying an adversary's external knowledge. This approach allows the publishing organization to investigate privacy threats and enforce privacy requirements in the presence of various types and amounts of external knowledge.

S.R. Ganta, S.P. Kasiviswanathan, and A. Smith (2008) developed a paper for designing schemes with good composition Properties in order toAttacks the information. Here the limitation used is auxilary information. It indicates that several currently proposed partition-based anonymization schemes, including k-anonymity and its variants, are vulnerable to composition attacks.

Hence, propose the reason about privacy in the presence of independent anonymized releases of overlapping population. Our experimental study indicates that several currently proposed partition-based anonymization schemes, including k-anonymity and its variants, are vulnerable to composition attacks.

G.Ghinita, P.Karras, P.Kalnis, and N.Mamoulis(2007) developed a paper for the microdata without revealing the sensitive information leading privacy preserving paradigms of k-anonymity and l-diversity.k-anonymity protects against the identification of individual record.l-diversity against the association of individual with specific sensitive information. The anonymization process is inefficient in terms of computation and I/O cost.k-anonymity is commonly achieved by generalization or suppression lead to information loss. Here, presenting a framework for solving efficiently the k-anonymity and l-diversity problems,by mapping the multidimensional quasi-identifiers.It is more prominent when applied to l-diversity,and it is difficult inorder to cover the variety of partitioning to the overlapping ranges. Optimal partitionings consisting of only consecutive ranges with respect to each individual value of the sensitive attribute. The inefficiency arises from the fact that the resulting partitioning may contain overlapping groups therefore, numerous possible combinations must be examined. A framework for solving the k-anonymity and l-diversity problems, by mapping the multidimensional quasi dentifiers.Both problems are NP-hard in the multidimensional space.

Hence ,proposing a framework for privacy preservation techniques.To protect privacy is not sufficient due to the existence of quasi inentifiers in the microdata. M.E.Nergiz and C.Clifton,developed a paper for increasing the need for anonymized data and risks of anonymization.

This possess a trade off value from the knowledge to the shared information. One solution to this problem is anonymity ensuring the disclosure data cannot be linked to the individual where the data is to share the information. k-anonymity and related techniques have received considerable attention,datasets are anonymized to distort the data loss than k-anonymization to comparable private levels. Therefore, it enables a meaningful bridge between human understandable policy and mathematically sound standards for anonymity.It correlates the risk and cost of private violation.sensitive data in disclosed dataset in l-diversity to form some technique.Dataset is not available but statistical properties of public dataset is known.It is also possible to use randomization instead of generalization on the private dataset,more advanced bayesian and statistical techniques are required.

III. EXISTING SYSTEM

In existing system, many SMC protocols for privacy preserving data mining suffer from high computation and communication costs. The high overhead of SMC Protocol raise the question of whether the information gain after the protocol execution is more than the cost. This is the valid question for protocols working on horizontally or vertically portioned data.

While publishing person specific sensitive data, simply removing uniquely identifying information (SSN, name) from data is not sufficient to prevent identification because partially identifying information, quasi identifiers, (age, sex, nation,etc.) can still be mapped to individuals (and possibly to their sensitive information such as salary) by using an external knowledge.

The goal of privacy protection based on k-anonymity is to limit the linking of a record from a set of released records to a specific individual even when adversaries can link individuals Previous work does not accurately achieve the goal of privacy data preserving method. Drawback in the existing system are The computational and communication cost is high, and does not worth for the implemented method.Possibilities of linking attacks are high by the adversary who knows the quasi identifiers values without generalization.

IV. PROPOSED SYSTEM

In proposed system, focus on the SMC protocol for distributed k- anonymity. k-Anonymity is a well-known privacy preservation technique proposed to prevent linking attacks on shared databases. One way to enable effective data mining while preserving privacy is to anonymize the data set that includes private information about subjects before being released for data mining.

The way to anonymize data set is to manipulate its content so that the records are to form k-anonymity. Two common manipulation techniques used to achieve k-anonymity of a data set are generalization and suppression.

Generalization refers to replacing a value with a less specific but semantically consistent value, while suppression refers to not releasing a value at all. the hacker has entered the query to retrieve a details of the persons who are working in Java, so that the server will display the details hiding the department and salary.

Generalization contains the details of the persons who works in software field. When hiding the name of the country, the server will display its continent and Employee id instead of their name By hiding all these sensitive information, the hacker will not able to view the exact employee's details.

The user wants to update the information, where an authentication key will be generated and send to the legitimate user's mobile for an authentication process. The user have to enter that key for verification.Entire details hide using RC4 algorithm.At the same time third party can see but entire details they cannot view it.Data accessibility and filtering is to provide security.Secure random key algorithm ,the authentication key is generated by the user and receive the alert message.

Here the methodology is highly localized in computation, it is fast and requires little communication cost.Linking attacks are minimized using suppression and generalization methods.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

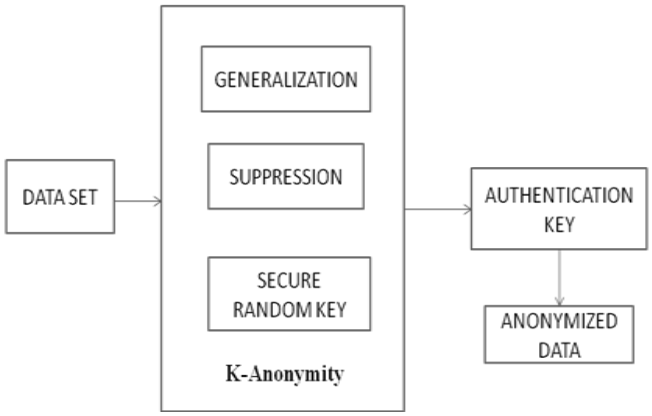


Fig.4.1 Block Diagram of Secure data storage.

In the secure data storage all data to be stored in the table to view some information.The process is done through secure look-ahead in the form of k-anonymity ,where k-anonymity is to identity the information.

Generalization hide the information like department,salary.It is to view all the information.Suppression hide the information like country,nationality.It is to view the specific information.Authentication key is generated by the user and validate the key.

A. Algorithm Implementation

In generalization algorithm,the hacker has entered the query to retrieve a details of the persons who are working in Java, so that the server will display the details hiding the department and salary. It also contains the details of the persons who works in software field.

In suppression algorithm, When hiding the name of the country, the server will display its continent and Employee id instead of their name .

By hiding all these sensitive information, the hacker will not able to view the exact employee's details. Though the user wants to update the information, an authentication key will be generated and send to the legitimate user's mobile for an authentication process. The user have to enter that key for verification.

B. Modules

K-Anonymity look ahead approach:

Here it describes a fast look ahead of k-anonymization of horizontally partitioned data. The look ahead returns an upper bound on the probability that k-anonymity will be achieved at a certain utility. Utility is quantified by commonly used metrics from the anonymization literature.

For each and every user, the server will set the access privileges. So that if the user will not able to cross beyond their limits. Some user will only be allowed to view the content. For some user, they can view and not to modify the data and update the data. So by setting the access privileges, they can provide more security for the users.

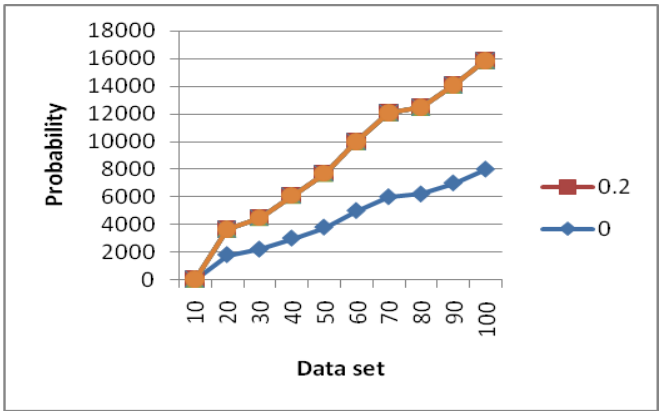
Distributed K-Anonymity approach:

Here they use a larger data set to create a better utilized k-anonymization. As increasing data size, fewer tuples need to be suppressed or generalized to satisfy k-anonymity, in other words k-anonymization can be satisfied with lower level mappings.

Also it is possible for the hacker to hack the data from the server. To prevent the data from this issue, we can encrypt the entire data in the server using RC4 Algorithm. Therefore if the hacker hacks the data, they aren't able to view the data. Only the data will decrypted in the legitimate users ends.

Decision predicate:

The amount of benefit parties get from the SMC is sufficiently high. Each party involved in an SMC expects to gain from SMC either locally or globally depending on the application. In this work, they assume that parties require the local benefit to exceed some threshold before they proceed with the SMC protocol. However, in most applications, without total knowledge of all private tables, parties can only have some confidence that SMC will meet their expectations.



V. CONCLUSION

Secure multiparty protocols is used as third party protocols in the data hiding and security is given through this protocol is done by access rights.

To overcome this look-ahead approach, specifically for secure multiparty protocols to achieve distributed k-anonymity, which helps parties to decide if the utility benefit from the protocol is within an acceptable range before initiating the protocol.

k-Anonymity is a well-known privacy preservation technique proposed to prevent linking attacks on shared databases. One way to enable effective data mining while preserving privacy is to anonymize the data set that includes private information about subjects before being released for data mining.

Generalization refers to replacing a value with a less specific but semantically consistent value, while suppression refers to not releasing a value at all. the hacker has entered the query to retrieve a information. An authentication key will be generated and send to the legitimate user's mobile for an authentication process. Entire details hide using RC4 algorithm.At the same time third party can see but entire details they cannot view it.

REFERENCES

[1] Mehmet Ercan Nergiz,Abdullah Ercument Cicek,Thomas B.Pedersen,and Yucel Saygin,"A Look-Ahead Approach to Secure Multiparty Protocols",IEEE Transactions on knowledge and data engineering,VOL.24,NO.7,JULY 2012.

[2] M.Kantarclu and C.Clifton, "Privacy-Preserving Distributed Mining of association Rules on Horizontally Partitioned Data,"IEEE Trans. Knowledge and Data Eng., vol. 16, no. 9,pp. 1026-1037,Sept. 2004.

[3] R.C.-W. Wong, J. Li, A.W.-C. Fu, and K. Wang, "(α, k)-Anonymity:An Enhanced K-Anonymity Model for Privacy Preserving Data Publishing," Proc. 12th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD '06), pp. 754-759, 2006.

[4] B.-C. Chen, K. LeFevre, and R. Ramakrishnan,"Privacy Skyline:Privacy with Multidimensional Adversarial Knowledge," Proc.33rd Int'l Conf. Very Large Data Bases (VLDB '07), pp. 770-781, 2007.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

[5] S.R. Ganta, S.P. Kasiviswanathan, and A. Smith, "CompositionAttacks and Auxiliary Information in Data Privacy," Proc. 14thACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining(KDD '08), pp. 265-273, http://doi.acm.org/10.1145/1401890.1401926, 2008.

[6] G.Ghinita,P.Karras,P.Kalnis,and N.Mamoulis,"Fast Data Anonymization with Low Information Loss",Proc.33rd Int'l conf.Very Large Data Bases(VLDB'07),PP.758-769,2007.

[7]M.E.Nergiz,M.Atzori,andC.Clifton,"Hiding the presence of individuals in Shared Databases",Proc.ACM SIGMOD Int'l Conf.Management of Data(SIGMOD'07),June-2007