ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

Privacy Preserving Data Sharing with Encrypted Anonymous ID Assignment

Akhila M, Nitha L Rozario

M.Tech Student, Dept. of CSE., Marian Engineering College, Kerala University, Trivandrum, Kerala, India

Asst. Professor, Dept. of CSE., Marian Engineering College, Kerala University, Trivandrum, Kerala, India

ABSTRACT: An algorithm for sharing of private data among N parties is developed. This work deals with efficient algorithms for assigning identifiers (IDs) to the nodes of a network in such a way that the IDs are anonymous using a distributed computation with no central authority. Given N nodes, this assignment is essentially a permutation of the integers {1….N} with each ID being known only to the node to which it is assigned. Resistance to collusion among other members is verified in an information theoretic sense when private communication channels are used. This assignments of IDs allows complex data to be shared and has applications in privacy preserving data mining, collision avoidance in communications and distributed database access. Existing and new algorithms for assigning IDs are examined with respect to trade-offs between communicational and computational requirements. New algorithms are built on top of secure sum data mining operations using Newtons identities.Markov chain representations are used to find statistics of number of iterations. . In this system, owner is assigned a randomly generated encrypted ID, with which data stored in the database is encrypted. This ensures confidentiality of the data. Another party can access data only if permission is granted by the owner. Also a comparison study based on different encryption methods are performed. The required computations are distributed without using a trusted central authority.

KEYWORDS: data mining; collision to resistance; Newtons identities

I. INTRODUCTION

It is today well understood that databases represent an important asset for many applications and thus their security is crucial. Data confidentiality is particularly relevant because of the value, often not only monetary, that data have. For example, medical data collected by following the history of patients over several years may represent an invaluable asset that needs to be adequately protected. Such a requirement has motivated a large variety of approaches aiming at better protecting data confidentiality and data ownership. Relevant approaches include query processing techniques for encrypted data and data watermarking techniques. Data confidentiality is not however the only requirement that needs to be addressed.

Today there is an increased concern for privacy. The availability of huge numbers of databases recording a large variety of information about individuals makes it possible to discover information about specific individuals by simply correlating all the available databases. Although confidentiality and privacy are often used as synonyms, they are different concepts: data confidentiality is about the difficulty (or impossibility) by an unauthorized user to learn anything about data stored in the database. Usually, confidentiality is achieved by enforcing an access policy,or possibly by using some cryptographic tools. Privacy relates to what data can be safely disclosed without leaking sensitive information regarding the legitimate owner [5]. Thus, if one asks whether confidentiality is still required once data have been anonymized, the reply is yes if the anonymous data have a business value for the party owning them or the unauthorized disclosure of such anonymous data may damage the party owning the data or other parties. Researchers have also investigated the relevance of anonymity and/or privacy in various application domains: patient medical records [4], electronic voting [5], e-mail [6], social networking [7], etc. Another form of anonymity, as used in secure multiparty computation, allows multiple parties on a network to jointly carry out a global computation that depends on data from each party while the data held by each party remains unknown to the other parties [8], [9]. A secure computation function widely used in the literature is secure sum that allows parties to compute the sum of their individual inputs without disclosing the inputs to one another. This function is popular in data mining multiparty computation [10], [11].

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

This work deals with efficient algorithms for assigning encrypted identifiers (IDs) to the nodes of a network in such a way that the IDs are anonymous using a distributed computation with no central authority. Given N nodes, this assignment is essentially a permutation of the integers{1,....,N} with each ID being known only to the node to which it is assigned. Our main algorithm is based on a method for anonymously sharing simple data and results in methods for efficient sharing of complex data. There are many applications that require dynamic unique IDs for network nodes [12]. Such IDs can be used as part of schemes for sharing/dividing communications bandwidth, data storage, and other resources anonymously and without conflict. The IDs are needed in sensor networks for security or for administrative tasks requiring reliability, such as configuration and monitoring of individual nodes, and download of binary code or data aggregation descriptions to these nodes. An application where IDs need to be anonymous is grid computing where one may seek services without divulging the identity of the service requestor.Wide acceptance of the grid technology has created pressure to add some features that were not part of its original design, such as security, privacy, and quality-of-service support.

The work further explores the connection between sharing secrets in an anonymous manner, distributed secure multiparty computation and anonymous ID assignment. Here the network is not anonymous and the participants are identifiable in that they are known to and can be addressed by the others.

II. TECHNIQUES USED

A. Matchmaking Protocol For Mobile Devices

In this several cryptographic protocols that seem suitable to solve the matchmaking problem. Matchmaking problem is a problem,where two parties ,each having a set of elements ,want to compute the intersection of their sets such that no party learn information other than the intersection elements.This protocol has three phases:initial phase,interest signing phase and matchmaking phase.

B. Review of Secure Sum Method

This method is used in case, if a group of hospitals with individual databases wish to compute and share only the average of a data item, such as the number of hospitals acquired infections, without revealing the value of data item for any member of the group

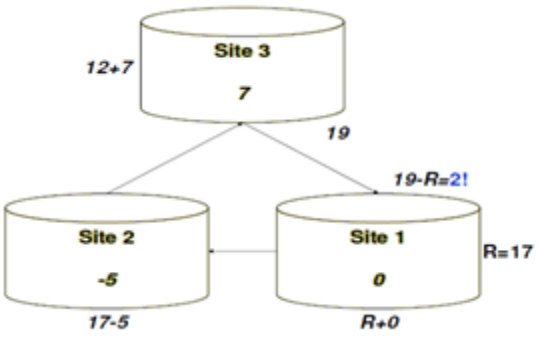

Fig. 1.Review of secure sum method

C.Transmitting Simple Data Using Power Sums

This method is used if a group of nodes wishes to share actual data values, rather than relying on only statistical information. However, here also data should be anonymous. A collusion resistance approach is developed, for this task of using secure sum as the communication mechanism. Data items are taken from a finite field, F, where F will be the field GF(P), where P is a prime number satisfying P>di for all i..

Power sums can be collected and shared using a single round of secure sum by sending them as an array and applying the method to the vectors transmitted and received.The power sums are symmetric functions, and thus no association is made between nodes and the value of data.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

D.Sharing Complex Data With AIDA

As the number of bits per data items and the number of nodes becomes larger, method of previous section becomes infeasible. Instead, to accomplish this sharing, indexing of nodes are performed. Here for this each node is provided with an unique identification(ID) or serial number. Each nodes ID is not known to other nodes in the network.This is then termed as anonymous ID assignment (AIDA).Here also data is shared using secure sum method.

E.Protocols for collision resistance

To prevent collision resistance from happening,a protocol against collusions is necessary for each party.Hence here an asymmetric protocol called Secure Product of Summations (SPOS) protocol is proposed, in which every party in the system is equal to any other.We can also prove that the protocol is full-private ((m-l)-private), where m is the number of parties.Here, a protocol is called *t*private if no coalition containing a most *t* parties is able to get additional information from its execution.From this protocol we can derive applicable protocols by which several problems in privacy-preserving data mining can be solved.

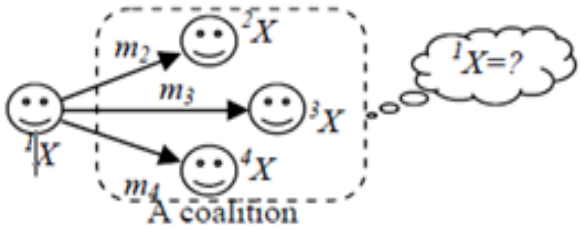

Fig. 2.Illustration of collusion

III. RELATED WORK

For mobile devices, a matchmaking protocol is used, which has three main phases namely: initial phase,interest signing phase and matchmaking phase. This protocol helps in finding common interests among people using mobile phones.In this, whenever matching interests are found people can come in contact with each other. But this method requires a trusted third party. Also in this, execution time increases with increase in number of interests.

In case of privacy preserving updates to anonymous and confidential databases, two secure protocols are used for privately checking whether a k-anonymous database retains its anonymity once a new tuple is being inserted to it. Since the proposed protocols ensure the updated database remains k-anonymous, the results returned from a user's (or a medical researcher's) query are also k-anonymous. But here there are problems like falling insertion,attack from malicious users due to the presence of untrusted third party,lack of efficiency in terms of number of messages exchanged and in terms of their sizes.

In case of privacy preserving data sharing for anonymous ID assignment,there is a lack of security and confidentiality.By knowing IDs, users can easily access other users databases, thereby hindering their privacy.

IV. EXISTING ALGORITHM

*Algorithm 1(Secure Sum):*Given nodes *n1,….,nN* each holding an data item *di* from a finitely representable abelian group, share the value *T=sum(di)* among nodes without revealing the values *di.*

1. Each node *ni* ,*i*=1,….,N chooses random values *ri,*1,……,*ri,*N such that
   *ri,*1+……+*ri,*N=*di*
2. Each random value *ri,j* is transmitted from node *ni* to node *nj.*The sum of all random numbers *ri,j* is the desired total *T.*
3. Each node *nj* then totals all random values received as:
   Sj=ri,j+….+rN,j
4. Now each node *ni* simply broadcasts *si* to all other nodes so that each node can compute:
   T=s1+…..+sN

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

*Algorithm 2(Finding AIDA):*Given nodes *n1,….nN*

1. Set the number of nodes A=0.
2. Each unassigned node *ni* chooses a random number *ri* in range 1 to *s.* A node assigned in the previous round chooses *ri=0.*
3. Random numbers are shared anonymously.
4. Let *q1,….,qk* denote a revised list of shared values with duplicate and zero values entirely removed.
5. Update the number of nodes assigned: *A=A+k.*
6. If A<N then return to step(2).

Two methods for AIDA

1 .Prime Modulus AIDA
2. Slot Selection AIDA

V. PROPOSED ALGORITHM

Here a comparison study of 4 different encryption methods :3DES,AES.Blowfish and RSA are performed and it is found that , among these 4 methods, RSA and then AES is found to have much better performance.Blowfish and 3DES is found to show almost equal performance.

After finding the best encryption method, created IDs are then encrypted using one of these best encryption methods.Once IDs are created, they are send to the corresponding users mobile phones.

By doing this, users can access other users database only after decrypting their IDs.This method provides much more security to the users private data.Thus data security and data confidentiality can be maintained.

VI. SIMULATION RESULTS

Here Fig.1. shows the working of secure sum method, consisting of 3 sites representing 3 nodes. Each of these nodes will obtain information only regarding the total value of each nodes individual data items, without even revealing their identities.Fig.2. shows how collision takes place between different nodes participating.Fig.3. shows a comparison study of 4 different encryption methods:3DES,AES,Blowfish and RSA based on time is performed.RSA being an asymmetric encryption, is found to take more time for encrypting the created ID. Similarly it will take more time for an attacker to decrypt the same and so is found to show better performance.

After RSA , next AES shows better performance. Blowfish and 3DES is found to show almost equal performance.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

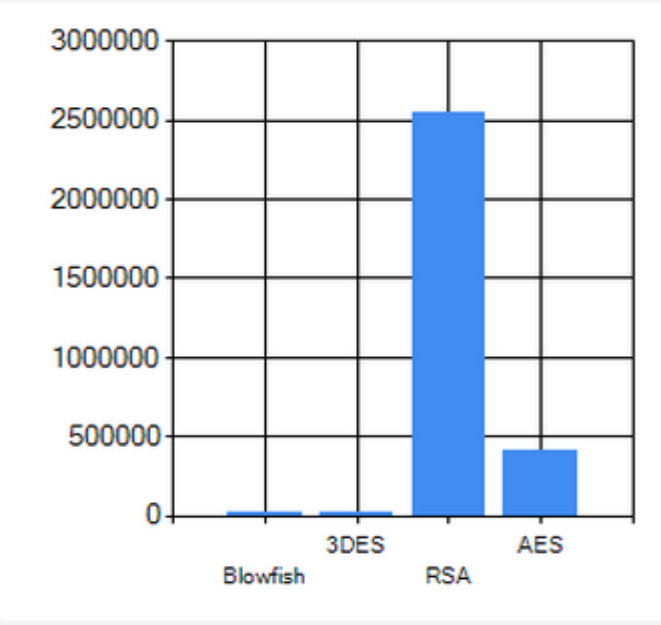

Fig .3.Comparison study based on time

V1. CONCLUSION AND FUTURE WORK

In the proposed work, IDs are provided in an encrypted form.These IDs are then send to the users mobile phones. Also here an alert is send to the users mobile, whenever an illegal user tries to login. This provides much more security. Here data confidentiality is also maintained.

As in this paper a comparison study about 4 different encryption methods are performed,we can find out one of the best encryption methods, and using this encryption method we can encrypt the generated IDs of the users database.

In future, we can try to perform a comparison study of much more encryption methods.Also some cryptographic approach can be used, that could guarantee finitely bound termination for AIDA.

REFERENCES

1. LARRY A.Dunning, "Privacy Preserving Data Sharing With Anonymous ID Assignment", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL.8,NO.2,FEBRURARY 2013.
2. Q.Xie and Hengarter, "Privacy preserving matchmaking for mobile social networking secure against malicious users,"in *Proc.23rd ann.IEEE conf.Privacy,Security* ,pp.252-259 *Trust*,Jul.2011.
3. A. Yao, "Protocols for secure computations,"in *Proc. 23rd Ann.IEEE Symp.Foundations of Computer Science,1*,IEEE Computer Society, *982, pp. 160-164.*
4. J. Wang,T. Fukasama,S. Urabe, and T. Takata, "A collusion-resistant approach to privacy-preserving distributed data mining,"*IEICE Trans.Inf.Syst (Inst.Electron.Inf.commun. Eng.),*vol E89-D, no. 11,pp. 2739-2747,2006.
5. J. Smith, Distributing identity[symmetry breaking distributed access protocols],"*IEEE Robot. Autom. Mag*., vol.6,no. 1, pp. 49-56, Mar.1999.
6. D. Jana, A. Chaudhuri and B. B.Bhaumik, "Privacy and anonymity protection in computational grid services,"*Int J.Comput.Sci. Applicat*,vol. 6,no. 1, pp.98-107,Jan.2009.
7. S. S. Separd,R.Dong,R. Kresman, and L. Dunning, "Anonymous id assignment and opt-out," in *Lecture Notes in Electrical Engineering*, S. Ao and L. Gleman,Eds. Newyork:Springer, , pp. 420-431, 2010.
8. J. A. Eidswick, "A proof of newton's power sum formulas," *Amer.Math.Monthly*, vol. 75, no. 4,pp.396-396,Apr.1968.
9. C. M. Grinstead and J. l. Snell, "Chapter 11:Markov chains," *introduction to prob*ability, 2nd ed. Providence, RI: Amer. Math. Society, pp. 510-510,1997.
10. A. Karr, Secure stastical analysis of distributed databases, emphasizing what we don't know," J. *Privacy Confidentiality, vol. 1, no. 2,pp.197-211,2009.*
11. Alberto Trombetta, Wei Jiang,Elisa Bertino and Lorenzo Bossi,"Privacy Preserving Updates to Anonymous and confidential Databases." D. G. Mead, Newton's identities,"*Amer.Math.Monthly*, vol. 99,no. 8,pp.749-749, Oct.1992.