Proceedings of International Conference on Advances in Computer Technology and Management (ICACTM)
In Association with Novateur Publications IJRPET-ISSN No: 2454-7875
ISBN No. 978-81-921768-9- 5
February, 23rd and 24th, 2018

AIDA: EFFICIENT ALGORITHM FOR ANONYMOUS SHARING OF PRIVATE DATA IN DISTRIBUTED NETWORKS

MRS. A.V.BANSOD
Lecturer, Department of Computer Engineering, Y.B.Patil Polytechnic Akurdi, Pune.
E-mail: archi.bansod@gmail

ABSTRACT
The network society places great demand on the dissemination and sharing of private data. As privacy concerns grow, anonymity of communications becomes important. This paper addresses the issue of anonymous ID assignment to nodes in a distributed network and how it can be integrated with secure mining algorithms to allow nodes that have privacy concerns, a capability to opt out of the mining computation. In this paper anonymous ID used for hiding the data sharing, also allows multiple partied 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. Technique is utilized iteratively to assign the nodes ID numbers ranging from 1…N,sanctions more complex data to be shared and has applications to other quandaries in collision avoidance in communications and distributed database access.We propose two algorithms for ID assignment and evaluate their performance. We use them in the design of a protocol that allows a node to opt out of data mining, and investigate the collusion resistance capability of the resulting protocol.

INDEX TERMS: Anonymizationand deanonymization; cloud and distributed computing systems; multiparty computation; privacy preserving data mining; privacy protection; security and trust in cooperative communications.

I.INTRODUCTION

The internet is very popular medium for communication for personal and business purpose as it supports anonymous connections. Enterprises also have valid causes to enlist in anonymous communication and avoid the penalties of persona revelation. For example, to permit dissemination of abstract facts and figures without disclosing the identity of t he entity the underlying facts and figures is affiliated with, or to protect whistle-blower's right to be anonymous and free from political or financial retributions. [2]The cloud base web services provide functionalities to server that it will capture users action.[10],[11]Anonymization deals with efficient algorithms for assigning identifiers (IDs) to the nodes of a network in such a way that the IDs are anonymous using the distributed computation with no central authority.[12]Given Nnodes, this assignment is essentially a permutation of the integers 1 to N with each ID being known only tothe node to which it is assigned. Our main algorithm is based on methods for anonymously sharing simple data with their results in methods for efficient sharing of complex data. There are many applications that require dynamic unique IDs for network nodes. [13]The IDs are needed in sensor networks for security or for administrative tasks requiring reliability, such as configuration and monitoring of individual's nodes, anddownload of binary code and data aggregation descriptions to these nodes. An application where the IDs need to be anonymous is grid computing, where one may seek services without divulging the identity of the service requestor. Existing and new algorithms for assigning anonymous IDs are examined and respect to tradeoffs between communication and computational requirements. Also, suppose that access to the database is strictly controlled, because data are used for certain experiments that need to be maintained confidential.[23] Clearly, allowing Alice to directly read the contents of the tuple breaks the privacy of Bob; on the other hand, the confidentiality of the database managed by Alice is violated once Bob has access to the contents of the database. Thus, the problem is to check whether the database inserted with the tuple is still k-anonymous, without letting Alice and Bob knows the contents of the tuple with the database respectively.

II. RELATED WORK

Many methods are presented for anonymous ID assignment; however every method is suffered from different kinds of limitations. In [4] A. Friedman, R. Wolff, and A. Schuster, "Providing k-anonymity in data mining, In this paper we present extended definitions of k-anonymity and use them to prove that a given data mining model does not violate the k-anonymity of the individuals represented in the learning examples. Our extension provides a tool that measures the amount of anonymity retained during data mining. We show that our model can be applied to various data mining problems, such as classification, association rule mining and clustering. We describe two data mining algorithms which exploit our extension to guarantee they will generate only k-anonymous output, and provide experimental results for one of them. Finally, we show that our method contributes new and efficient ways to anonymize data and preserve patterns during anonymization.

- In [7]Q. Xie and U. HengartnerThe success of online social networking and of mobile phone services has resulted in increased attention to mobile social networking. Matchmaking is a key component of mobile social networking. It notifies users of nearby people who fulfill some criteria, such as having shared interests, and who are therefore good candidates for being added to a user's social network. Unfortunately, the existing matchmaking approaches are troublesome from a privacy point of view. One approach has users' smart phones broadcast their owners' personal information to nearby devices. This approach reveals more personal information than necessary. The other approach

requires a trusted server that participates in each matchmaking operation. Namely, the server knows the interests and current location of each user and performs matchmaking based on this information. This approach allows the server to track users. This paper proposes a privacy-preserving matchmaking protocol for mobile social networking that lets a potentially malicious user learn only the interests (or some other traits) that he has in common with a nearby user, but no other interests. In addition, the protocol is distributed and does not require a trusted server that can track users or that needs to be involved in each matchmaking operation. We present an implementation and evaluation of our protocol on Nexus One smart phones and demonstrate that the protocol is practical.

• In [13]D. Jana, A. Chaudhuri, and B. B. Bhaumik,In computational grid computing, grid nodes spanning over several diverse computing resources belonging to heterogeneous administrative domains form the backbone of Virtual Enterprise [VE]. In order to offer service-on-demand, various service providers, requesters, brokers and administrators collaborate in request-response manner among each other in Service Oriented Virtual Enterprise through service registry, service discovery and service binding mechanisms. Security issues for integrated and collaborative sharing of computing resources across heterogeneous administrative domains are principal concern. At the same time, the privacy and anonymity are also of prime importance while communicating over publicly spanned network like web. The individual service providers or requesters may not reveal their true identity to one another for privacy needs. Also, computational grid services may be required to be availed anonymously within the grid framework to keep the personal sensitive information about the service requester protected. This paper focuses on the protection of privacy and anonymity of grid stakeholders in the service oriented computational grid framework. An extension of onion routing has been used with dynamic token exchange along with protection of privacy and anonymity of individual identity.

•In [17] A. Karr, over the past several years, the National Institute of Statistical Sciences (NISS) has developed methodology to perform statistical analyses that, in effect, integrate data in multiple, distributed databases, but without literally bringing the data together in one place. In this paper, we summarize that research, but focus on issues that are not understood. These include inability to perform exploratory analyses and visualizations, protections against dishonest participants, inequities between database owners and lack of measures of risk and utility.

• In[23]J. Castellà-Roca, V. Daza, J. Domingo-Ferrer, and F. Sebé, with the development of computer networks, situations where a set of players remotely play a game (e-gaming) have become usual. Often players play for money (e-gambling), which requires standards of security similar to those in physical gambling. Cryptographic tools have been commonly used so far to provide security to e-gambling. Holomorphic encryption is an example of such tools. In this paper we review the mental poker protocols, where players are assumed to remotely play poker. We focus on the key advantage of using cryptosystems with holomorphic properties (privacy homomorphisms) because they offer the possibility of manipulating cards in encrypted form.

III. IMPLEMENTATION DETAILS

3.1 Architecture

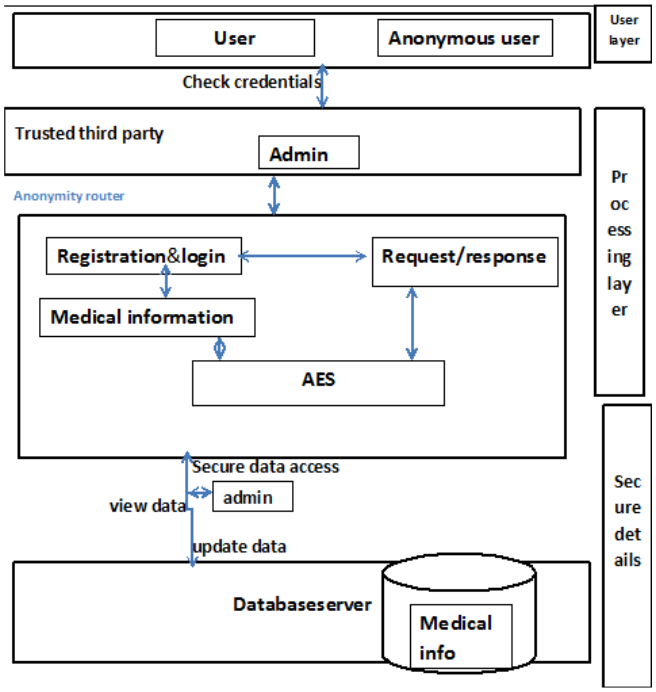


Fig. 1 System Architecture

3.2 PROPOSED WORK

The Newton identities greatly decrease communication overhead. This will enable the use of a larger number of slots with a consequent reduction in the number of rounds required. Private communication channels, our algorithms are secure in an information theoretic sense. Apparently, this property is very fragile. The very similar problem of mental poker had shown to have no such solution with two players and three cards. To bounds on completion time developed in the previous work, our formulae give the expected completion time exactly.

We conjecture the asymptotic formula of Corollary, based on computational experience, to be a true upper bound. All of the non cryptographic algorithms have been extensively simulated, and we can say that the present work does offer a basis upon which implementations can beconstructed. The communications requirement of thealgorithms depends heavily on the underlyingimplementation of chosen secure sum algorithm.

3.3.Advantages of Proposed System

Increasing parameters in the algorithm will reduce the number of expected rounds. However, our central algorithm requires solving the polynomial with coefficient taken from finite field of integers modulo a prime. That task restricts the level to which can be practically raised. We show in detail how to obtain the average number of required rounds and in the Appendix

details a method for solving the polynomial, which can bedistributed to all the participants

3.4 Algorithm:

Algorithm 1(Secure Sum):- Given nodes n_1 ...n_N each holding an data item d_i from a finitely representable abelian group , share the value T = $\sum d_i$ among the nodes without revealing the value d_i .

1) Each node$r_{i,}$ i=1 . . . ,N chooses random values $r_{i,1}$ . . . . . $r_{i,N}$ such that
$$r_{i,1} + \ldots . + r_{i,N} = d_i$$
2) Each "random" value $r_{i,j}$is transmitted from node n_i to node n_j .The sum of all these random numbers $r_{i,j}$ is of course, the desired total T.
3) Each node $n_j$totals all the random values received as:
$$s_j = r_{1,j} + \ldots + r_{N,j}$$
4) Now each node n_i simply broadcasts s_i to all other nodes so that each node can compute:
$$T = s_1 + \ldots . . + s_N$$

Algorithm 2(Anonymous Data Sharing With Power Sums):- Given node n_1 , ... , n_N each holding a data item d_i from a finitely re-presentable field F , make their data items public to all nodes without revealing their sources.

1) Each node n_i computes d_i^m over the field F for n = 1,2,. . . . , N. The nodes then use secure sum to share knowledge of the power sums:

$P_1 = \sum_{i=1}^{N} d_i^1$	$P_2 = \sum_{i=1}^{N} d_i^2$	……	$P_N = \sum_{i=1}^{N} d_i^N$

2) The power of sums P_1 , ... , P_N are used to generate a polynomial which has d_1,...,d_N as its roots using Newton's Identities as developed in . Representing the Newton polynomial as
$$p(x) = c_N x^N + \ldots + c_1 x + c_0$$
The values c_0, .. …. , c_N are obtained from the equations:
$$c_N = -1$$
$$c_{N-1} = -\frac{1}{1}(c_N P_1)$$
$$c_{N-2} = -\frac{1}{2}(c_{N-1} P_1 + c_N P_2)$$
$$c_{N-3} = -\frac{1}{3}(c_{N-2} P_1 + c_{N-1} P_2 + c_N P_3)$$
$$c_{N-4} = -\frac{1}{4}(c_{N-3} P_1 + c_{N-2} P_2 + c_{N-1} P_3 + c_N P_4) \ldots$$
$$c_{N-m} = -\frac{1}{m}\sum_{k=1}^{m} c_{N-m+k}\, P_k$$

3) The polynomial p(x) is solved by each node , or by a computation distributed among the nodes, to determine the roots d_1, . . . , d_N.

Algorithm3 (find AIDA):- Given node n_1 , . . . , n_N use distributed computation (without central authority) to find an anonymous indexing permutation s: {1 . . . N} →{1 . . . N}.

1) Set the number of assigned nodes A = 0.
2) Each unassigned node n_i chooses a random number $r_i$in the range 1 to S. A node assigned in a previous round chooses r_i = 0.
3) The random numbers are shared anonymously. One method for doing this was given in Section III. Denote the shared values by q_1, ... , q_N.
4) Let q_1 , ... $q_k$denote a revised list of shared values with duplicated and zero values entirely removed where k is the number of unique random values. The nodes n_i which drew unique random numbers then determine their index $s_i$from the position of their random number in the revised list as it would appear after being sorted:
$$s_i = A + Card\{q_j : q_j \le r_i\}$$
5) Update the number of nodes assigned: A = A + k.

If A < N then return to step (2).

3.5Mathematical Model:

Input = {users id, user information}
Output= {Secure data sharing}
Process:

1. Secure Sum
2. Anonymous data sharing with power sum

Power sum calculation:
$$P_N = \sum_{i=1}^{N} d_i^N$$
Polynomial generation:
P(x) = $C_{Nx}{}^N$ +...+c1x+c0
The values of c0 to c_N are obtained from the equations:
$$C_{N-m} = -\frac{1}{m}\sum_{k=1}^{m} C_{N-m} + \quad {}_k P_k$$

IV. Practical Results and Environment:

In this section we are presenting practical environment, dataset used, and metrics computed.

4.1. Input Dataset:

In implementation user id and their personal information who register to web server are used as input set.

4.2. Hardware and Software Configuration used is Pentium –IV processor, 256 MB RAM and 20 GB Hard Disk, Operating system used for this project is Windows XP/7/8, Programming Language used is Java withNet beans Tool

4.3 Results of Practical Work:

Following Fig.2 shows User vs Exposer Performance Comparison Graph

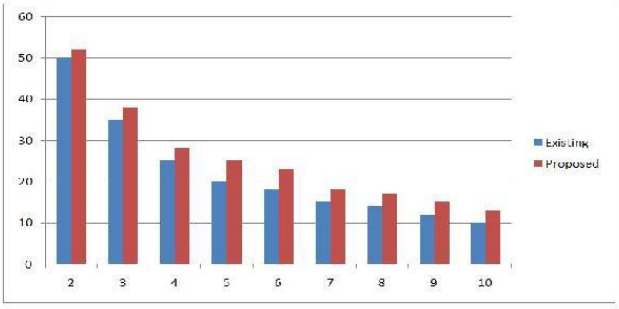
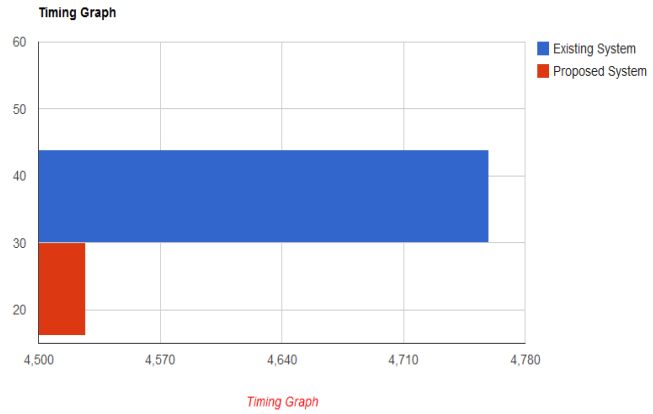


Fig.2: User vs Exposer Performance Comparison Graph



V.CONCLUSION

Implementation of Newton identities greatly decreases communication overhead. This can enable the use of a larger number of "slots" with a Consequent reduction in the number of rounds required. All of the cryptographic algorithms have been extensively simulated, and we can say that the present work does offer a basis upon which implementations can be constructed. The communications requirements of the algorithm are depend on the underlying implementations of the chosen secure sum algorithm. In some cases, merging the two layers could result in reduced overhead.

VI. REFERENCES

1) Sarbanes–Oxley Act of 2002, Title 29, Code of Federal Regulations, Part 1980, 2003.
2) White Paper—The Essential Guide to Web Analytics Vendor Selection, IBM[Online].Available: http://measure.coremetrics.com/corem/getform/reg/wp-evaluation-guide
3) A. Shamir, "How to share a secret," Commun . A C M, vol. 22, no. 11, pp. 612–613, 1979.
4) A. Friedman, R. Wolff, and A. Schuster, "Providing k-anonymity in data mining," VLDB Journal, vol. 17, no. 4, pp. 789–804, Jul. 2008.
5) F. Baiardi, A. Falleni, R. Granchi, F. Martinelli, M. Petrocchi, and A. Vaccarelli, "Seas, a secure e-voting protocol: Design and implementation," Comput. Security, vol. 24, no. 8, pp. 642–652, Nov. 2005.
6) D. Chaum, "Untraceable electronic mail, return address and digital pseudonyms," Commun. A C M, vol. 24, no. 2, pp. 84–88, Feb. 1981.
7) Q. Xie and U. Hengartner, "Privacy-preserving matchmaking for mobile social networking secure against malicious users," in Proc. 9th Ann. IEEE Conf. Privacy, Security and Trust, Jul. 2011, pp. 252–259.
8) O. Goldreich, S. Micali, and A. Wigderson, "How to play any mental game," in Proc. 19th Ann. ACM Conf. Theory of Computing, Jan. 1987, pp. 218–229, ACM Press.
9) A . Yao, "Protocols for secure computations, " in Proc. 23rd Ann. IEEE Symp. Foundations of Computer Science, 1982, pp. 160–164, IEEE Computer Society.
10) C. Clifton, M. Kantarcioglu, J. Vaidya, X. Lin, and M. Y. Zhu, "Tools for privacy preserving distributed data mining," ACM SIGKDD Explorations Newsletter, vol. 4, no. 2, pp. 28–34, Dec. 2002.
11) 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.
12) J. Smith, "Distributing identity [symmetry breaking distributed access protocols]," IEEE Robot. Autom. Mag. , vol. 6, no. 1, pp. 49–56, Mar. 1999.
13) 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.
14) D. M. Goldschlag, M. G. Reed, and P. F. Syverson, "Hiding routing information," in Proc. Information Hiding, 1996, pp. 137–150, SpringerVerlag
15) L. Willenborg and T. Waal, Elements of Statistical Disclosure Control, ser. Lecture Notes in Statistics. New York: Springer, 2001, vol. 155.
16) S. S. Shepard, R. Dong, R. Kresman, and L. unning, "Anonymous id assignment and opt-out," in Lecture Notes in Electrical Engineering, S. Ao and L. Gleman, Eds. New York: Springer, 2010, pp. 420–431.
17) A. Karr, "Secure statistical analysis of distributed databases, emphasizing what we don't know," J. Privacy Con fidentiality, vol. 1, no. 2, pp. 197–211, 2009.
18) D. Angluin, "Local and global properties in networks of processors (extended abstract)," in Proc. 12th Ann. ACM Symp. Theory of Computing (STOC '80), New York, 1980, pp. 82–93.
19) W. Fokkink and J. Pang, "Variations on itai-rodeh leader election for anonymous rings and their analysis in prism," J. Universal Comput. Sci. , vol. 12, no. 8, pp. 981–1006, Aug. 2006.
20) J. W. Yoon and H. Kim, "A new collision-free pseudonym scheme in mobile ad hoc networks," in Proc. 7th Int. Conf. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT'09), Piscataway, NJ, 2009, pp. 376–380, IEEE Press.
21) J. W. Yoon and H. Kim, "A perfect collision-free pseudonym system," IEEE Commun. Lett. , vol. 15, no. 6, pp. 686–688, Jun. 2011.
22) A. Shamir, R. L. Rivest, and L. M. Adleman, Mental Poker Massachusetts Institute of Technology, Tech. Rep. MIT-LCS-TM-125, 1979.
23) J. Castellà-Roca, V. Daza, J. Domingo-Ferrer, and F. Sebé, "Privacy homomorphisms for e-gambling and mental poker," in Proc. IEEE Int. Conf. Granular Computing, 2006, pp. 788–791.
24) R. Canetti, "Security and composition of multi-party cryptographic protocols," J. Cryptol. , vol. 13, no. 1, pp. 143–202, 2000.
25) U. Maurer, "Secure multi-party computation made simple," in Proc. 3rd Int. Conf. Security in Communication Networks (SCN'02), Berlin, Heidelberg, 2003, pp. 14–28, Springer-Verlag.