Volume 5, Issue 2, February 2015 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Study of Privacy Preserving Data Sharing with Anonymous ID Assignment using AIDA Algorithm

[1]Lekshmi Priya S*, [2]Fitha Naseem Salim, [3]Sushitha Susan Joseph

[1], [2]B.Tech Student, Department of CSE, Mar Baselios Christian College of Engineering and Technology, Kerala, India

[3]Assistant Professor, Department of CSE, Mar Baselios Christian College of Engineering and Technology, Kerala, India

Abstract—The popularity communication medium whether for personal or business requires anonymous communication in various ways of internet as a. Businesses also have legitimate reasons to make communication anonymous and avoid the consequences of identity revelation. The problem of sharing privately held data so that the individuals who are the subjects of the data cannot be identified has been researched extensively. An algorithm for anonymous sharing of private data among parties is developed. This technique is used iteratively to assign these nodes ID numbers ranging from 1 to N. This assignment is anonymous in that the identities received are unknown to the other members of the group. The assignment of serial numbers allows more complex data to be shared and has applications to other problems in privacy preserving data mining, collision avoidance in communications and distributed database access. The required computations are distributed without using a trusted central authority

Keywords— Anonymity,Data mining,Collision avoidance.

I. INTRODUCTION

The Multiparty data sharing deals with how we can secure multiparty data sharing. There are efficient algorithms for assigning identifiers (IDs) to the nodes of a network such that the IDs are anonymous by using a distributed computation with no central authority. In order to have complex secure data sharing AIDA can be used so that the computation will be easier than the existed one. The main algorithm is based on a technique for anonymously sharing simple data and results in methods for efficient sharing of complex data.

There are many applications that require dynamically generated unique IDs for network nodes. Such IDs can be used as part of schemes for sharing/dividing communications bandwidth, data storage, and many other resources without conflict. An application where IDs need to be anonymous is mainly grid computing where one may seek services without exposing the identity of the service requestor.

This paper presents a study on various algorithms based on the data sharing with anonymous IDs t. Also, this paper provides a marginal overview for future research and improvements.

II. STUDY OF ALGORITHMS

A. SECURE SUM

Should all pairs of nodes have a secure communication channel available, a simple, but resource intensive, secure sum algorithm can be constructed. In the following algorithm, it is useful to interpret the values as being integer on first reading.

Table 1 Random numbers transmitted by a secure sum execution

Nodes		$r^{\wedge}_{i,1}$	$r_{i,1}$	$r_{i,2}$	$r_{i,3}$	$r_{i,4}$	d_i	$d^{\wedge}_i$
n_i=1	13-6+8=	15	13	-10	6	-3	6	8
n_i=2	7-10+9=	6	7	3	-5	5	10	9
n_i=3	-8-6+5=	-9	-8	11	12	-9	6	5
n_i=4		6	6	-8	-5	9	2	2
s_i		18	18	-4	8	2	T=24	24

Algorithm 1 (Secure Sum): Given nodes $n_1 \ldots\ldots n_N$ 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 values d_i

1.Each node n_i , i=1…N chooses random values $r_{i,1} \ldots r_{i,N}$ 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: $S_j = r_{i,j}+\ldots .+ r_{N,j}$

4. Now each node n_i simply broadcasts S_i to all other nodes so that each node can compute:

$T = S_1+\ldots+S_N$

Example1 (A Secure Sum Computation): In Table 1 two examples are shown, one for later use. The reader can ignore the columns labelled and need not attribute any significance to the boldface type. In the example, the initial data items held by nodes n_1,n_2,n_3 and n_4 are d_1=6,d_2=10,d_3 =10 and d_4=2 respectively. For example,node n_2 would transmit 7,3,-5 ,and 5 to nodes n_1 ,n_2 (itself),n_3 and n_4 respectively. Node would receive -10, 3, 11, and -8 from nodes n_1 ,n_2 (itself),n_3 and n_4 respectively. Then node n_2 would compute and transmit the total s_2 =-4 of the values received to all nodes.Finally, n_2 would compute the total of all the second round transmissions received, 24=18+-4+8+2.

(Secure Sum Hides Permutations): The secure sum method of Algorithm 1 is input permutation resistant to the collusion of any subset of the participating nodes. Other secure sum algorithms certainly can be used with physically or cryptographically secured communications channels. For example, it is easy to see that secure sum using a single Hamiltonian cycle is input permutation collusion resistant provided that the coalition is trapped in a connected region of the cycle. Such results canal so be extended to provide privacy guarantees for the algorithms in subsequent sections should they utilize, e.g., a Hamiltonian cycle based secure sum. The secure sum technique can be employed with finite abelian groups

B.TRANSMITTING SIMPLE DATA WITH POWER SUMS

Suppose that a group of nodes wishes to share actual data values from their databases rather than relying on only statistical information. That is, each member of the group of nodes has a data item which is to be communicated to all the other members of the group. However, the data is to remain anonymous. A collusion resistant method is developed for this task using secure sum as the underlying communication mechanism. The data items d_i are taken from a,typically finite, field F.In the usual case, each d_i will be an integer value and will be the field GF(P) where P is a prime number satisfying $P>d_i$ for all i . Thus, arithmetic will typically be performed using modulus P, but other fields will also be used.

Given nodes n_i ...n_N each holding a data item d_i from a finitely representable field , make their data items public to all nodes without revealing their sources.

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

$P_1 = \sum_{i=1}^{N} di^1$	$P_2 = \sum_{i=1}^{N} di^2$	....	$P_N = \sum_{i=1}^{N} di^N$

2) The power sums are P_1...P_N used to generate a polynomial which has d_1...d_N as its roots using Newton's Identities as developed. Representing the Newton polynomial as

$$P(x)=c_N x^N+\ldots+c_1 x+c_0 \qquad (1)$$

The values $c_{0\ldots}c_N$ are obtained from the equations:

c_N=-1

$c_{N-1}=-\frac{1}{1}(c_N p_1)$

$c_{N-2}=-\frac{1}{2}(c_{N-1} p_1 + c_N p_2)$

$c_{N-3}=-\frac{1}{3}(c_{N-2} p_1 + c_{N-1} p_2 + c_N p_3)$

$c_{N-4}=-\frac{1}{4}(c_{N-3} p_1 + c_{N-2} p_2 + c_{N-1} p_3 + c_N p_4)\ldots$

$$c_{N-m} == -\frac{1}{m}\sum_{k=1}^{m} c(n-m+k)\ p^k \qquad (2)$$

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

C. SHARING COMPLEX DATA WITH AIDA

Now consider the possibility that more complex data is to be shared amongst the participating nodes. Each node has a data item of length N-bits which it wishes to make public anonymously to the other participants. As the number of bits per data item and the number of nodes becomes larger, the method of the previous section becomes infeasible. Instead, to accomplish this sharing, we will utilize an indexing of the nodes. Methods for finding such an indexing are developed in subsequent sections.

Assume that each node has a unique identification (ID) or serial

HOW TO FIND AN AIDA

We present a simple algorithm for finding an AIDA which has several variants depending on the choice of the data sharing method at step (3) below. At one step, random integers or slots between 1 and are chosen by each node. A node's position will be determined by its position among the chosen slots, but provisions must be made for collisions. The parameter should be chosen so that .S≥N.

Given nodes $n_1...n_n$, 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.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 the irrandom number in the revised list as it would appear after being sorted:
$$S_i=A=card\{q_j : q_j<=r_i\}$$
5) Update the number of nodes assigned: A=A=k.
6) If then return to step (2).

COMPARISON OF AIDA VARIANTS

In the previous section the algorithm to find an AIDA required that the random numbers be shared anonymously at step (3). We now look at three methods which are variants of that procedure. The parameter must be chosen in each case. The expected number of rounds depends only on the selection of and not on the variant chosen.

III. PROPOSED SYSTEM

An algorithm for anonymous sharing of private data among parties is developed. This technique is used iteratively to assign these nodes ID numbers ranging from 1 to N. This assignment is anonymous in that the identities received are unknown to the other members of the group. Resistance to collusion among other members is verified in an information theoretic sense when private communication channels are used. This assignment of serial numbers allows more complex data to be shared and has applications to other problems in privacy preserving data mining, collision avoidance in communications and distributed database access. The required computations are distributed without using a trusted central authority.

Advantage:

1. The anonymity of DB is not affected by inserting the records.
2. We provide security proofs and experimental results for both protocols

IV. CONCLUSIONS

This paper presents a survey on various techniques and algorithms that was proposed earlier by researchers for the better privacy-preserving data access. The proposed AIDA algorithm is foolproof in allocating ID to users and the anonymous identity is maintained, thus providing ample proof for the sets of users in multiparty environments. Even under difficult situations the communications and bandwidth is not affected in any manner. So unlike cryptographic measures and traditional systems AIDA proves to be secure for distributed architecture keeping the user safe from attacks in different segments. In future the scheme may be extended as a web service so that any interconnected user of the network can utilize it to the maximum without the need to implement the code. Also mobile web services are an area of interest for future extensions to AIDA.

ACKNOWLEDGMENT

We would like to thanks all the reference authors for the completion of this paper.

REFERENCES

[1] Andreas Jakoby and Maciej Li´skiewicz (2005), "Revealing Additional Information in Two-Party Computations " , Advances in Cryptology - ASIACRYPT 2005 Lecture Notes in Computer Science Volume 3788, 121-135.

[2] Dr. Durgesh Kumar, Neha Koria, Nikhil Kapoor, Ravish Bahety (2009), "A Secure Multi-Party Computation Protocol for Malicious Computation Prevention for Preserving Privacy During Data Mining", International Journal of Computer Science and Information Security,Vol. 3.

[3] Akheel Mohammed, Sajjad Ahmed Md , Ayesha (2013), "Confidentiality And Anonymity Strengthening in Computational Services", IJRRECS,Volume-1,Issue-6,1006-1011.

[4] Swathi, P.Jyothi, and Anil Kumar(2014), "Assigning Privacy Ids For Each Data That Have Been Sharing In Wireless Networks", International Journal of Communication Network and Security (IJCNS) ISSN: Volume-2, Issue-3.

[5] Ms. R. Kalaivani, Ms. R. Kiruthika (2014), "Automated Anonymous Id Assignment For Maintaining Data Privacy", International Conference on Science,Engineering and Management,Srinivasan Engineering College, India.

[6] Ayswarya R Kurup, Simi Lukose (2014), "Security Enhanced Privacy Preserving Data sharing With Random ID Generation",IJSRE Volume 2 Issue 8.

[7] Larry A. Dunning, And Ray Kresman (2013), "Privacy Preserving Data Sharing With Anonymous ID Assignment ",IEEE Transactions on Information Forensics and Security, Vol. 8, NO. 2.

BIOGRAPHY

Lekshmi Priya S is a B-Tech student in the department of of Computer Science and Engineering, MBCCET, Peermade,Kerala. Her main research interests focus on data base management system, ,object oriented programming,OS.

Fitha Naseem Salim is a B-Tech student in the department of of Computer Science and Engineering, MBCCET, Peermade,Kerala. Her main research interests focus on database management system,object oriented programming.

Sushitha Susan Joseph (M.E.,2012, Anna University,Chennai; B.Tech.,2009, Mahatma Gandhi University,Kottayam) is an assistant professor in the Department of Computer Science and Engineering, MBCCET, Peermade,Kerala. Her main research interests focus on artificial intelligence, data mining,wireless sensor networks and data privacy.

