International Journal of Emerging Trend in Engineering and Basic Sciences (IJEEBS)
ISSN (Online) 2349-6967 *Volume 2 , Issue 1(Jan-Feb 2015), PP1-5*

Securing And Maintaining PrivacyInformation Brokering In Distributed DataSharing

SanketDivate[1], PratapGhayal[2], SaurabhMamidwar[3], MayurNarawade[4],Meghna Lokhande [5]
[1 2 3 4 5]*(Department of Computer, JSPM's RajarshiShahu College Of Engineering ,Savitribai Phule Pune University,india)*

Abstract :- Some organization work in collaboration with other organizations, sometimes this organizations need to share the information using on demand access from the require database. For this purpose Information brokering system (IBS) which has peer to peer overlay structure is used for sharing the information among all the distributed data sources. This IBS contains some components like distributed data servers and brokering components to locate the database servers for requested queries. The existing IBSs use server side access control and make assumption that brokers are honest with giving little bit attention to privacy of data along with metadata which is stored in and shared from distributed database servers using IBS. We consider the problem of privacy of information in Information brokering process. Here we focus on the two types of attacks which are often happen with distributed information sharing; they are "Inference Attack" and "Attribute-Correlation Attack". Then we provide broker-coordinator working structure with two proposed schemes query segment encryption scheme and automaton segmentation scheme for sharing the secure query routing function between the set of brokering servers. On analysis of performance, analysis on privacy and scalability we show that with the reasonable overhead the privacy is preserved along with enforcing the security in distributed information sharing using information brokering

Keywords:- Access control, IBS, automaton segmentation, broker, Attribute-Correlation Attack, Inference Attack.

I. INTRODUCTION

Along with explosion on information collected by organization in many fields such as medical, Business to government agencies, there is an need for communication between various organization. It is challenging to reconcile data heterogeneity and provide interoperability. Most of the existing systems work on two spectrums, either adopting either query answering model or the distributed database model, where peers managed by unified DBMS. There is no model suitable for newly emerged.

Regional Health Information Organization access and retrieves the clinical data across collaborative healthcare providers that includes hospitals from particular regions, outpatients clinics etc. The data is private so, participating organization would not complete sharing or free data sharing. It is necessary to retain full control over data and access particular data. It is expected from providers that consumer must posses privacy. In querying process, sharing of complete copy of data with others in centralized system becomes impractical. Consider a solution from"sharing everything", "sharing nothing " the peer to peer framework needs to establish pairwise client-server relationship between each pair of peers. For sensitive data and autonomous data providers, it is necessary to construct a data-centric overlay ([2],[3]). There are brokers that make routine decisions. That depends on contents of queries [4],[5]. This infrastructure builds on semantic-aware index mechanism that helps to route the queries base on contents which it possess. In our previous study [5],[6], such as distributed system provides data access through a set of brokers. Database is connected to asset of brokers. Local brokers contain the metadata, which further"advertise".the metadata to other brokers. Queries are sent to the brokers according to the metadata to reach the data server. While the IBS approach provides scalability and reliability against privacy. The broker maybe outsourced to the third-party providers. In this article, we are going to implement a solution to the privacy preserving information brokering. It contains two types of brokers, 1. Brokers and 2.Coordinators. The broker is responsible for user authentication and query forwarding. The coordinators follow atree like structure, and perform transferring of segments to next. Inference access control and query routing based on embedded non-deterministic finite automata. Here we design two schemes to segment the query broker

brokering automata and encrypt corresponding query segments so that it will help in routing decisions making is decoupled into multiple correlated tasks. For a set of collaborative coordinators PPIB provides comprehensive brokering with good scalability.

II . RELATED WORK

In distributed database system there are collection of various different types of databases. This type of system provides interfaces for the users/client or requesters for the storing their data and share the data from distributed databases as per his/her interest.Firstly peer –to peer framework is used for data sharing purpose. Peer-to-peer is decentralized approach. To establish this kind of framework point to point connection is created in between each peer of the system. For sharing the information. This type of framework is not applicable for large scale collaborative sharing. This type of decentralized approach is used for routing path queries among all the peers. Here, XML Overlay structure is used which is supportive for processing the query and checking security. Here some specialized data structures are also used on nodes for routing purpose of XML queries [3]. After that there is centralized DBMS used for distributed data sharing but this system has privacy related issues as well as some trust issues also there. To avoid above problem in international journal of intelligent control and systems[5]. The XML brokerage system is used.

It is distributed database system which has the components:1] Databases/Database servers- This database servers contains XML documents. 2] Brokers- This broker contains the document distribution information of XML documents which are stored in database servers. In this system data is requested by user by sending queries to database through brokers. There are two issues with this approach related with performance of the system and query brokering is costly. Access control is another issue this system. The access control issue is fixed in this brokerage system by using in-broker access scheme[5]. To avoid problems which are occurred in information Brokerage system the new “automaton segmentation scheme” introduced in 2007 by Fengjun Li, Bo Lou. With this automaton segmentation scheme distributed access control enforcement is there.

Query segmentation scheme also introduced which is required for preserving the privacy of the query at the time of forwarding the query. There is also automaton segmentation algorithm is expressed. There are vulnerabilities related with user, data, and metadata. Main advantage of given scheme is, the segments are created by dividing access control information (metadata). XML schema is used for access control rules. The single automation is used for expressing same XML schema and different XML schema represented using independent automatons and they are merged by combining route coordinators. Some PPIB maintains is required. Xpath expression used for indexing rule. Automata based access control is multilateral security concept [10]. Now, in this paper we are introducing novel approach Privacy Preserving Information Brokering System (PPIB) to preserve privacy of query as well as data. PPIB is broker- coordinator overlay structure in which multiple users and different organizations are connected. Automaton segmentation, query segment encryption is used for preserving. All the previous approaches build the system by making the assumption that the broker is honest. But in PPIB no any type of assumption on brokers are made. In PPIB Central Authority (CA) is used for maintenance of all the components of PPIB. It adds only trusted brokers to the system by proper authentication. Xpath query expression is used for expressing index rules in PPIB. Access control rules are used for managing access to the databases and represented by XML structure. The automaton segmentation algorithm is implemented here which is given in previous papers. For content based query routing

“Automaton Segmentation Scheme”[10] used here. For preserving privacy of segments we are going to use “query segment encryption scheme”[6].

III. PROPOSED METHODOLOGY

3.1 Architecture of PPIB:

For addressing all the privacy related problems in present information brokering system we propose new system which is called as Privacy Preserving and Information Brokering (PPIB) System. The PPIB has a overlay architecture which normally consist of 3 different types of brokering components. 1] Central Authority, 2] Brokers and 3] Coordinators. These all the components are organized in overlay structure as shown in figure. The brokers are act as mix anonymizers [7]. The reason behind why PPIB is divided into these components is no any single component is capable for disclosing the information which is holding by itself. Because in PPIB every component did not have any type of info in complete form.In this PPIB architecture the data servers and all the data requestors/users are connected to the system through the Brokers (the nodes shown in figure). All the Brokers are connected with each other using the Coordinator's tree like structure. The Brokers are responsible brokering component for sequencing and forwarding the users requested query to next brokering component with considering local traffic analysis. Brokers are nothing but "Entry Point" for the data requestors in the system. Brokers are also providing authentication to the users and hide their identity from the others. Coordinators are responsible for access control enforcement and also for content based query routing. We can't let the single coordinator to hold a complete rule for maintaining privacy. Instead of this we are presenting here a novel approach "automaton segmentation scheme" to divides rules (metadata) into several parts and these parts are known as "segments". Then each segment is assigned to a single Coordinator. For the secure query routing Coordinators must operate collaboratively. We are introducing another scheme for preventing Coordinators from seeing sensitive information and this scheme is called as "query segment encryption scheme" which is used to encrypt the created segments for maintaining privacy. PPIB uses another Brokering server which is known as "Central Authority". The key management as well as metadata maintenance done by Central Authority.

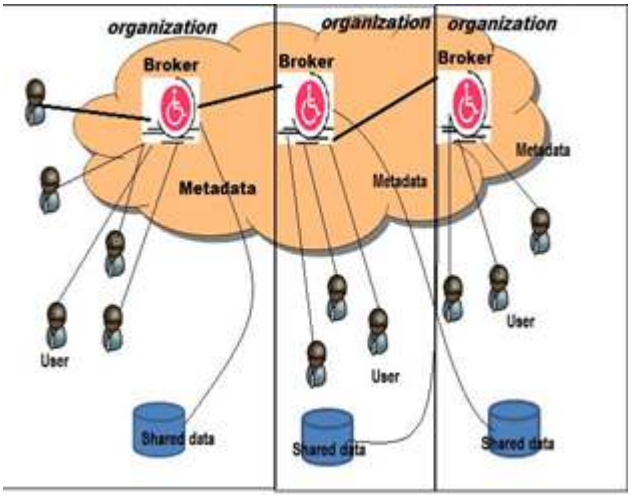

Fig 1 . Brokering Structure Architecture

3.2 Working Of PPIB:

PHASE 1: When user joins the system, it isrequire that the user must authenticate himself to the local broker. The user submits the query which is in an XML, with each segment encrypted by corresponding public level key and session key. Data server encrypted with public key, to return data. The input will be as the file details and user details. And output will be broker authentication and session key generation.

PHASE2**:** In this phase the major task ofbroker is preparation of metadata and also done with the authentication. It extracts the role of the user authenticated and attaches it to the encrypted XML query. It makes a unique ID

for each query and attaches QID with its own address to the query so that the data server can directly return the data. The input will be as the XML file details and file details. And output will be communication of data.

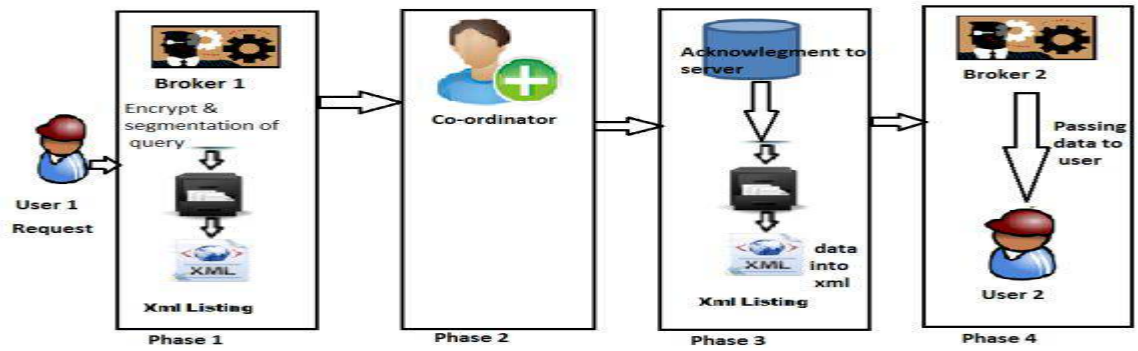

Fig 2:overall working of PPIB

PHASE 3: When the root of the coordinator tree receives the query and its metadata from a local broker, it follows schemes i.e. the automata segmentation scheme for segment the XML query and the query segment encryption scheme to perform access control and to route the query within the coordinator tree, until it reaches a leaf coordinator, which forwards the query to the related data servers. If query is not reach at leaf coordinator or it will denied access then the failure message with QID given to broker. The input will be as the XML file details and file details. And output will be communication of data.

PHASE 4: It is the final phase in that; the data server gets a safe query in an encrypted form. The data server evaluates the query and returns the data after decryption, encrypted by session key, to the broker of the query. The input will be as the XML file details and file details. And output will be as retrieve data.

3.3 Segmentation**:-**

Segmentation is used for making segments of metadata and each segment is assigned to the coordinator for preserving privacy of data. How segmentation works is given in figure. Consider the example shownin figure to understand how access control is enforced by the decentralized automaton.

Segmentation Algorithm:

Input: Automaton State S Output: Site Address: addr
for each symbol m in S:StateTransTable
do
addr=Depoly(S:StateTransTable(m):NextState)
DS=CreateDummyAcceptState()
DS:NextState addr ←
S:StateTransTable(m).NextState DS ←
Site = CreateSite()
Site:addSite(S)
Coordinator=GetCoordinator()
Coordinator:AssignSite(Site)
Return Coordinator:address

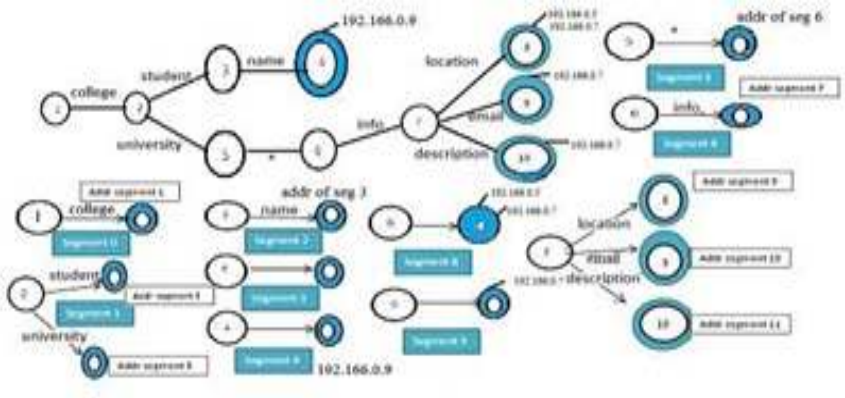

Fig 3: Detailed segmentation

Let user requested query is: "college/university/pune/info[clg_name='RSCOE']/location" When query arrives at segment 0,te first Xpath phase "/segment" is accepted. As a dummy accept state of segment 0 points to segment 1. Query is forwarded to segment 1 then the second Xpath step "/university" is accepted and corresponding dummy accept state directs to the remaining part of the query to the segment 5 There segment 5 accepts "/pune" (*" matches the any input token) and forwards query to segment 6. At segment 6 element name \info" is first accepted since the automaton segment does not carry any predicate state. The predicate from query kept as it is. Finally name is accepted at segment 7 and segment 10 forwards the query to the server at 192.166.0.7.In this way the privacy of the query along with the privacy of related data and metadata is preserved by dividing the metadata into different segments using above illustrated "Automaton segmentation Scheme.

IV . CONCLUSION

Existing brokering system associated with user, metadata and data related privacy issues. We have introduced a new approach to preserve privacy with XML information brokering. It includes Automaton segmentation scheme, network access control, query segmentation and encryption, security enforcement and query forwarding. We concluded that it is very resistant to privacy attack. Result shows that PPIB is efficient and scalable and provide end to end processing.

Recently for future research many directions are ahead. Now in recent, PPIB is conducted in an ad-hoc manner, provides site distribution and load balancing. Now our next aim is to implement scheme that does dynamic site distribution. Some factors we are introducing are work load at each peer, privacy conflicts between automaton segments and trust level at each peer. Administrator performs the job of automaton segmentation and we plan to reduce that participation. To make PPIB self-reconfigurable.

REFERENCES

[1] Fengjun Li, Bo Luo, peng Liu DongwonLee and Chano-Hsien Chu," EnforcingSecure and Privacy Preserving DistributedInformation Sharing".

[2] X. Zhang, J. Liu, B. Li, and T.-S.P.Yum, "CoolStreaming/DONet: A data-driven overlay network for efficient live media streaming," in Proceedings of IEEE INFOCOM, 2005.

[3]A. C. Snoeren, K. Conley, and D. K. Gifford, "Mesh-based content routing using XML," in SOSP, pp. 160–173, 2001.

[4]M. Franklin, A. Halevy, and D. Maier, "From databases to dataspaces: a new abstraction for information management," SIGMOD Rec., vol. 34, no. 4, pp. 27–33, 2005.

[5]F. Li, B. Luo, P. Liu, D. Lee, P. Mitra, W. Lee, and C. Chu, "In-broker access control: Towards efficient end-to-end performance of information brokerage systems," in Proc. IEEE SUTC, 2006.

[6] F. Li, B.Luo, P. Liu, D. Lee, and C.-H.Chu, "Automaton segmentation: A new approach to preserve privacy in XML information brokering," in ACM CCS '07, pp. 508–518, 2007.

[7] D. L. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," Communications of the ACM, vol. 24, no. 2, 1981.