International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

A Secure Approach for Privacy-Preserving in Information Brokering System using Query Forwarding

Madhuri D. Dhayarkar[1], Ram B. Joshi[2]

[1]Post Graduate Student Computer Engineering, MMCOE, Pune, Savitribai Phule Pune University, Maharashtra, India

[2]Asst. Professor Computer Engineering Department, MMCOE, Pune, Savitribai Phule Pune University, Maharashtra, India
PhD scholar JJT University Rajasthan

Abstract: ***There is an increasing need for information sharing via on-demand access in different organizations. Information Brokering Systems (IBSs) have been introduced to connect large-scale lightly-associated data sources. This system includes brokers that are responsible for routing decisions to direct client queries to the requested servers where data is located. Existing IBSs consists of brokers that are trusted and thus only adopt server-side access control for data confidentiality. However the privacy of location of data and information about consumer can still be concluded from metadata (such as query and access control rules) exchanged within the brokering system, thus the protection of the metadata is a major issue in IBS. The proposed scheme presents an overview on information sharing in distributed environment through information brokering system and problems associated with it thus providing scope for healthcare information systems. The objective is to overcome two attacks- attribute correlation attack and inference attack providing security enforcement and to provide two countermeasure schemes namely automaton segmentation and query segment encryption scheme. Thus central authority is necessary for routing decision making responsibility among a selected set brokering servers.***

Keywords: Access Control, Information sharing, Privacy preservation, Automaton segmentation scheme, Query segment encryption scheme

1. Introduction

Today's organizations often operate across organizational boundaries. They raise strong needs for efficient and secure information sharing to facilitate extensive collaborations among organizations. Previous approaches on sharing of information mainly focus on providing transparency and interoperability among heterogeneous system, fall short of satisfying new requirements of these inter-organizational collaborations. The systems work on two extremes of the spectrum: (1) servers are autonomous and system-wide communication is not present while responding to the query; so that user creates one-to-one client-server connections for information sharing; (2) in the distributed systems, all the user lost autonomy and are managed by a unified DBMS. There are different types of applications and they need different forms of information sharing. In particular, while some applications (e.g., stock price updating) use publish subscribe system and the other applications use the system that provides access to the information on-demand.

As an example, medical data is stored in databases in autonomous enterprises. As a data provider, a participant would not assume free or complete sharing of the data with any unauthorized users as this kind of data is legally private or commercially proprietary. It is required that owner should have full control over the data with the help of access control mechanisms. A feasible solution for storing sensitive data is to construct a data centric overlay including the data sources and a set of brokers helping to locate data sources for queries as shown in Figure 1 [1]. Mechanisms are used to route the queries and thus users can submit queries without knowing data or server location. Such a distributed system providing data access through a set of brokers is referred to as Information Brokering System (IBS).

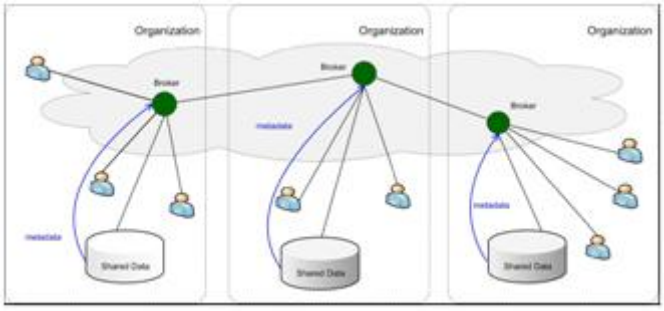

Figure 1: IBS infrastructure

Though the IBS approach provides scalability and server autonomy, there are privacy issues, as brokers are not fully trustable – they may be abused by insiders or compromised by outsiders. In this paper, we present a general solution to the privacy-preserving information sharing problem which preserves the User Privacy, Data Privacy, and Metadata Privacy. First, to address the need for privacy protection, we propose a novel IBS, named Privacy Preserving Information Brokering (PPIB). It is an overlay infrastructure consisting of two types of brokering components: brokers and coordinators. The objective is to overcome two attacks- attribute correlation attack and inference attack providing security enforcement. To prevent curious or corrupted coordinators from inferring private information, we design two novel schemes: (a) Automaton segmentation (b) Query segment encryption scheme.

Licensed Under Creative Commons Attribution CC BY

1.1 Motivation

Nowadays there is an increasing need for inter organizational information sharing to facilitate extensive collaboration. At the same time it is challenging task to handle heterogeneous data and provide interoperability for the same. In many applications we need enforcement while sharing the secret information which can be shared in a conservative and controlled manner due to business considerations or legal reasons among organizations. Here we are providing new emerging technique for healthcare information systems.

In healthcare information systems, confidential reports or documents are made available to the doctors, patient (data owner) as well as in emergency department present in the hospitals in case of accidents. Thus our aim is to facilitate access to and retrieval of clinical data across collaborative healthcare providers that include a number of regional hospitals, outpatient clinics, payers, etc. Also in today's world everything is made online hence it is necessary to preserve privacy of owner's data as it may contain confidential information. But it is very challenging task as privacy should be maintained while transferring the secret documents. Thus two schemes are used namely automaton segmentation and query segment encryption scheme. These schemes help in encrypting the documents so that an unauthorized user or broker or coordinator cannot see the whole content of the document/query.

Because of the growing popularity of XML and XML database systems and the need of the privacy, these databases are used widely as they have the ability to hide data from a group of brokers and coordinators and to make the data available to users in an efficient and friendly manner.

2. Literature Survey

In peer-to-peer (client-server) systems information sharing framework means "sharing everything or nothing". These systems are responsible for sharing files. If the DBMS is centralized then it introduces privacy, and trust issues and is not able to handle heterogeneous data. In IBS, brokers are trusted and thus only adopt server-side access control for data confidentiality.

Peer-to-peer file sharing systems and publish-subscribe systems provide partial solutions to the problem of sharing of data on a large scale. Integrated information provides an integrated view over large numbers of heterogeneous data sources by exploiting the semantic relationship between schemas of different sources. Peer-to-peer systems are designed to share files and data sets (e.g. in collaborative science applications). Further, P2P file-sharing systems may not provide complete set of answers to a request while we need to locate all relevant data.

Distributed hash table technology [12] is adopted to locate replicas based on queries. Though these technologies are recently extended to support range queries, the coarse granularity (e.g. files and documents) still makes them short of our expressiveness needs. Addressing a problem, the XML publish- subscribe systems (e.g. [3]) is probably the closely related technology to the proposed scheme: where we locate relevant data sources for a given query and route the query to these sources of data which are nothing but servers while the publish subscribe systems are responsible for locating consumers for a given document and route the document to these consumers. They have different concerns: they focus on efficiently delivering the same piece of information to consumers located at different sites while we route large volume but small-size queries to fewer sites. In our proposed scheme multicast in pub-sub systems is not applicable. Thus XML overlay architecture is built that supports expressive query processing and security checking. In [8], it uses the pairing-based cryptography mechanism, identity-based encryption in order to support many-to-many interactions between subscribers and publishers.

The specialized data structures are maintained on nodes to route path queries. In [13], a robust mesh has been built to effectively route XML packets by making the use of self-describing XML tags and the overlay networks. Kouds et al. [2] describes a decentralized architecture for ad hoc XPath query routing across a collection of XML databases using the open and agreement cooperation models. In [14], content-based routing of path queries in peer-to-peer systems is studied to serve the purpose as sharing data among a large number of autonomous nodes. In [4], issues for processing XML data in a peer to peer systems, viz indexing of data, replication of XML data and query routing and processing are studied. The main difference between these approaches and our system is that they focus on distributed query routing, while we seamlessly integrate query routing and access control so as to preserve relevant privacy information.

As for security check, research has been proposed on distributed access control. Earlier approaches implement access control mechanisms at the nodes of XML trees and filter out data nodes that users do not have authorizations to access [15]. This processing is handled by XML engines. Creating and maintaining a separate view (e.g. a specific portion of XML documents) for each user is handled by view-based access control mechanisms [16]. However, supporting large number of views causes high maintenance and storage cost. In [7], it uses attribute based encryption (ABE) techniques to achieve fine-grained and scalable data access control for PHRs. PPIB approach adopts a recently proposed NFA-based query re-writing access control scheme [5,17] and extends it to a decentralized manner. Any off the-shelf XML databases can be used due to its query re-writing nature.

3. Comparative Study of Different Methods

Figure 2 shows comparative analysis of different methods used in information brokering systems.

Licensed Under Creative Commons Attribution CC BY

Title	Year	Advantage	Disadvantage
Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption[7]	2013	Focus on the multiple data owner scenario, and divide the users in the PHR system into multiple security domains that greatly reduces the key management complexity for owners and users	Usually assume the use of a single trusted authority (TA) in the system creating a load bottleneck
A Survey on Protecting Information Brokerage System against Intruders[10]	2013	Brokers make use of routing protocols that create hard-to-trace communications by using a chain of proxy servers which are untraceable and mainly responsible for user authentication and query forwarding	The broker functionality may be outsourced to third-party providers and thus vulnerable to be abused by insiders or compromised by outsiders
Design and Implement Privacy Protection for Secure Information Brokering Systems[9]	2014	End-to-end query processing performance and System scalability	Several types of attacks possible
Securing Broker-Less Publish/Subscribe Systems Using Identity-Based Encryption[8]	2014	To ensure that a particular subscriber can decrypt an event only if there is match between the credentials associated with the event and the key	A parent can decrypt every event it forwarded to its children
A Novel Approach to Improve the Privacy of Information Brokering in Semantic Web[11]	2014	Enriches the privacy of data shared within Information Brokering System by using Selective encryption, Vigenere Cipher encryption and Selective Reverse Circle Cipher algorithm	Privacy leakage of data requestor privacy, data privacy and metadata privacy

Figure 2: Comparative study of different methods

4. Privacy Preserving Information Brokering

Privacy Preserving Information Brokering (PPIB) has three type of brokering Component: (1) brokers and (2) co-ordinators (3) central authority(CA)as shown in Figure 3 [1]. The brokers are mainly responsible for user authentication and query forwarding. The broker performs the role who can act between the Co-coordinator and the data Users. The request which is all submitted from the data user will be verified and thus it will be passed to the coordinator. The coordinators are linked in a tree structure enforce access control and query routing based on the embedded nondeterministic finite automata also known as query brokering automata. The coordinators hold a segment of automaton that helps in routing. Central authority is also present for key management and maintaining metadata.

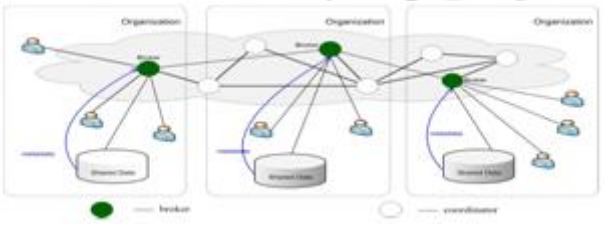

Figure 3: PPIB Architecture

The automaton segmentation and query segment encryption schemes ensure that a curious or corrupted coordinator is not capable to collect enough information to infer privacy, such as “data that is being queried”, “location of data” etc.

5. Two Major Schemes in PPIB

There are two schemes namely (1) automaton segmentation scheme and (2) query segment encryption scheme. An automaton segmentation scheme divides metadata into segments and each segment is then assigned to a coordinator. Coordinators enforce secure query routing thus they operate in collaborative manner. Second scheme is query segment encryption scheme to protect query content and it prevents coordinators from seeing sensitive predicates. In this scheme a query is divided into segments, and each segment is encrypted in such way that no segment apart from the ones needed to enforce secure routing is revealed to the coordinators route.

6. Privacy Vulnerabilities

Information Brokering is dependent on the trust of brokers for query forwarding and leads to harm the privacy of user, data and metadata. The user privacy can be described as identity of user, location of user while sending a query and obtain the purpose of the query. User identity can be assumed by authentication process and information about the user location. Location of data and data object distribution privacy is included in data privacy. It describes which type of data is contained in particular data server. Query indexing and access control rules are two types of metadata. It describes where the data objects are distributed among data server and provides access to authorized users. Data providers push routing and access control metadata to brokers, which also receives queries from users. Thus, corrupted brokering server could: (1) learn query content and query location of a local query; (2) learn routing and access control metadata from local data servers and other brokers; (3) learn location of data from routing metadata. In this type of attack, there is less chances for attacker to obtain plaintext data from the data which is encrypted but they are able to learn location of query and data.

The attacks are classified as (1) attribute correlation attack: when query is routed, compromised broker or external attacker (eavesdropper) may extract the query condition for getting the sensitive information by matching the attributes contained in the query. (2) inference attack: By getting more than one type of sensitive information, the attacker guessing the query location (IP address), query content and identify data owner from the query content. We show that PPIB provides comprehensive privacy protection for on-demand brokering of the information by overcoming these attacks and very good scalability.

7. Conclusion

In this paper, we propose PPIB, which is a new approach to preserve privacy in XML information brokering system. With the help of automaton segmentation and query segment encryption scheme and also access control rules, PPIB integrates security enforcement and query forwarding while providing privacy protection. The analysis shows that PPIB system overcomes the privacy attacks, provides end-to-end query processing performance and scalability. Many directions are ahead for future work. Firstly, site distribution and load balancing in PPIB are conducted in an ad-hoc manner. Several factors can be considered in the scheme such as the workload and trust level at each peer, and privacy vulnerabilities between automaton segments.

Licensed Under Creative Commons Attribution CC BY

References

[1] Fengjun Li, Bo Luo, Peng Liu Dongwon Lee and Chao-Hsien Chu,"Enforcing Secure and Privacy-Preserving Information Brokering in Distributed Information Sharing", IEEE Transactions On Information Forensics And Security Vol:8 No:6 Year 2013.
[2] N. Koudas, M. Rabinovich, D. Srivastava, and T. Yu, "Routing XML queries," in ICDE '04, p. 844, 2004.
[3] A. Carzaniga, M. J. Rutherford, and A. L. Wolf, "A routing scheme for content-based networking," in Proc. of INFOCOM, 2004.
[4] G. Koloniari and E. Pitoura, "Peer-to-peer management of XML data: issues and research challenges," SIGMOD Rec., vol. 34, no. 2, 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] Ming Li, Shucheng Yu, Yao Zheng, Kui Ren, Wenjing Lou, "Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption," IEEE Transactions On Parallel And Distributed Systems, Vol. 24, No. 1, January 2013.
[8] Muhammad Adnan Tariq, Boris Koldehofe, and Kurt Rothermel, "Securing Broker-Less Publish/Subscribe Systems Using Identity-Based Encryption," IEEE Transactions On Parallel And Distributed Systems, VOL. 25, NO. 2,February 2014.
[9] Shaik.Mahaboob Basha, A.Bhaskar, D.V Satish Kaladhar Reddy, "Design and Implement Privacy Protection For Secure Information Brokering Systems," in IJCSMC, Vol. 3, Issue. 8, pg.41–48, August 2014.
[10] Sanchari Saha, Madhusudana H.A, "A Survey on Protecting Information Brokerage System against Intruders," International Journal of Engineering and Innovative Technology (IJEIT) Volume 3, Issue 5, November 2013.
[11] Supriya S. Sankpal, Rupali A. Mahajan, "A Novel Approach to Improve the Privacy of Information Brokering in Semantic Web," International Journal of Science and Research (IJSR),Volume 3, Issue 7, July 2014.
[12] I. Stoica, R. Morris, D. Liben-Nowell, D. Karger, M. Kaashoek, F. Dabek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup protocol for internet applications", in IEEE/ACM Trans. Networking, volume 11 of 1, 2003.
[13] A. C. Snoeren, K. Conley, and D. K. Gifford, "Mesh-based content routing using XML", in Symposium on Operating Systems Principles, pages 160-173, 2001.
[14] G. Koloniari and E. Pitoura. Content-based routing of path queries in peer-to-peer systems. In EDBT, 2004.
[15] M. Murata, A. Tozawa, and M. Kudo. XML access control using static analysis. In ACM CCS, Washington D.C., 2003.
[16] S. Rizvi, A. Mendelzon, S. Sudarshan, and P. Roy. Extending query rewriting techniques for fine-grained access control. In SIGMOD, pages 551{562, Paris, France, 2004.
[17] B. Luo, D. Lee, W.-C. Lee, and P. Liu. QFilter: Fine-grained run-time XML access control via NFA-based query rewriting. In ACM CIKM, Washington D.C., USA, nov 2004.

Licensed Under Creative Commons Attribution CC BY