ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Privacy Protection for Information Brokering System

Akba Zoheer Mohammed
Master of Science (Information System),
Nizam College (Autonomous),O.U,
Basheer Bagh, Hyderabad.

T. Ramdas Naik
Assistant Professor, Dept of Computer Science (PG),
Nizam College (Autonomous),O.U,
Basheer Bagh, Hyderabad.

ABSTRACT:

Information Brokering Systems are attracting and increasing attention as an efficient means of sharing data among large, diverse and dynamic sets of user. The peer from logical over lay network by establishing links to some other peers they know are discover. A user in a peer-to-peer system in issues quires the describe data of interest the quires are propagated throw the overlay network to locate peer that provide data relevant to the query and only matching results are returned to the user. Information Brokering System (IBS) atop a peer-to-peer overlay has been proposed to support information sharing among loosely federated data sources. In existing IBSs adopt server side access control deployment and honest assumptions on brokers, and shack little observation on privacy of data and metadata stored and exchanged within the IBS.

This paper studies the problem of privacy protection in information brokering process (PPIB). Then, this paper propose a broker-coordinator face, as well as two schemes, automaton segmentation scheme and query segment encryption scheme, to share the secure query routing use among a set of brokering servers. With comprehensive survey on privacy, end-to-end performance, and scalability, we show that the proposed system can combine security enforcement and query routing while preserving system-wide privacy with reasonable overhead. Further enhanced by Information Brokering System using Data Encryption Standard (DES), Digital Signature and XOR swap algorithm.

Keywords:

Privacy, XML, Access Control, load balancing, information sharing, peer to peer, PPIB.

INTRODUCTION:

Information sharing is becoming increasingly important in recent years, not only among organizations with common or complementary interests, but also within many field ranging from business to other agencies that are becoming ever more globalized and distributed. To provide efficient large-scale information sharing, to reconcile data heterogeneity and provide interoperability across geographically distributed data sources. The systems work on two extremes of the spectrum:

in the query-answering model, peers are fully autonomous but there is no system-wide communication; so that user creates one-to-one client-server connections for information sharing; in the distributed database systems, all the user lost autonomy and are managed by a unified DBMS. However, different types of applications often need different forms of information sharing. In particular, while some applications (e.g., stock price updating) would need a publish subscribe framework, the on-demand information access is more suitable for other applications . As an example, imagine a future where many people have their DNA sequenced.

A medical researcher wants to validate a hypothesis connecting a DNA sequence D with a reaction to drug G. People who have taken the drug are partitioned into four groups, based on whether or not they had an adverse reaction and whether or not their DNA contained the specific sequence; the researcher needs the number of people in each group. DNA sequences and medical histories are stored in databases in autonomous enterprises. As a data provider, a participant would not assume free or complete sharing with others, since its data is legally private or commercially proprietary, or both. Instead, it is required to retain full control over the data and access to the data.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

In the sensitive data and autonomous data owners, a more practical and adaptable solution is to construct a data centric overlay including the data sources and a set of brokers helping to locate data sources for queries . Mechanisms to route the queries based on their content, which allows users to submit queries without knowing data or server location. In previous study such a distributed system providing data access through a set of brokers is referred to as Information Brokering System (IBS).This system provide scalability and server autonomy. In IBS infrastructure given broker and coordinator, broker are no longer fully trustable. So, system may be abuse by inside or outside.

PRIVACY- PRESERVING INFORMATION BROKERING:

Privacy protection is need for the Information Brokering System (novel IBS), named Privacy Preserving Information Brokering (PPIB).PPIB has two type of brokering Component: brokers and co-ordinators. The brokering 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 co-coordinator. The coordinators which 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, each holding a segment of access control automaton and routing guidelines, are mainly responsible for access control and query routing. PPIB takes an innovator automaton segmentation approach to privacy protection. In particular, two critical forms of privacy, namely query content privacy and data object distribution privacy (or data location privacy), are enabled by a novel automaton Segmentation scheme, with a "little" help from an assisting query segment encryption scheme. To prevent inquisitive or unserviceable coordinators from inferring private information, we design two novel schemes: (a) to segment the query brokering automata, and (b) to encrypt corresponding query segments. System will providing full capability to wage in network access control and to path queries to the right data sources, these two schemes ensure that inquisitive or unserviceable coordinator is not capable to collect sufficient information to guess privacy, like "which data need to be queried, where located and what are the policies to access data".

Privacy Preserving Information Brokering (PPIB) enables wide-ranging security and privacy protection for claimed information brokering, with minor overhead and major scalability.

SECURITY AND PRIVACY NEED FOR PPIB

In information brokering scenario, there are three types of entrepreneur, namely data owners, data providers, and data requestors. Each entrepreneur has its own privacy the privacy of a data owner (e.g. a patient) is identifiable data and the information keep together by this data (e.g. medical records). Data owners usually sign stiff privacy agreements with data providers to protect their privacy from unauthorized disclosure/user. Data providers store collected data, and create two types of metadata, namely routing metadata and access control metadata. Data requestors divulge identifiable and private information in the querying process. For example, a query process about AIDS or DNA treatment reveals the (possible) disease of the requestor. Assume that for the brokers, two types of enemy, outside attackers and curious or corrupted brokering components.

Outside attackers passively eavesdrop communication channels. Curious or corrupted brokering components follow the protocols be seemingly to accomplish their functions, others' private information from the information disclosed in the querying process Data providers push routing and access control metadata to brokers , which also strut queries from requestors. Therefore, a curious or corrupted brokering server could: learn query content and query location by impede a local query; learn routing metadata and access control metadata from local data servers and other brokers; learn data location from routing metadata it holds Although attacker may not obtain plaintext data over encrypted data, they can still learn query location and data location from eavesdrop. The attacks into two major classes: the attribute-correlation attack and inference attack. Attribute-correlation attack: An attacker prevents a query, which typically contains several predicates. Each predicate describes a condition, which sometimes involves sensitive and private data (e.g. name, credit card number, etc.). Inference attack: Attacker some techniques and result more than one other type of sensitive information so more sever, and further associates to learn explicit and implicit knowledge about entrepreneur IBS work is designed with user and data privacy.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Such privacy protection requirements, therefore a novel IBS, named as Privacy Preserving Information Brokering system (PPIB). As shown in Figure, PPIB contains a broker-coordinator overlay network, in which the brokers are amenable for onus transmission user queries to coordinators concatenated in tree structure while preserving privacy. The coordinators, each holding a segment of access control automaton and routing guidelines, are mainly responsible for access control and query routing.

Project Definition with Example:

As an example, healthcare information systems, such as Regional Health Information Organization (RHIO), aim to facilitate access to and retrieval of clinical data across collaborative health providers.

An RHIO is formed with multiple stakeholders, including hospitals, outpatient clinics, payers, etc. As a data provider, a participant would not assume free or complete sharing with others, since its data is legally private or commercially proprietary, or both. Instead, it is required to retain full control over the data and access to the data.

Meanwhile, as a consumer, a health provider requesting data From other providers expects to protect private information (e.g. requestor's identity, interests) in the querying process.An overview of the IBS infrastructure.

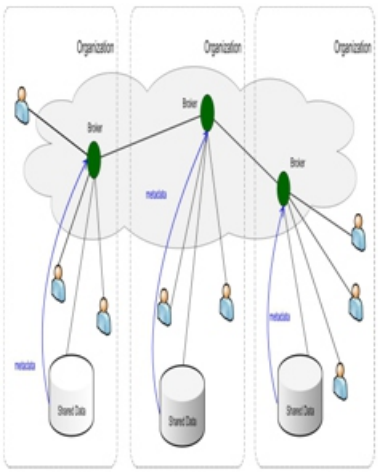

Figure 1.1 - Shared Data

In such scenarios, sharing a complete copy of the data with others or "pouring" data into a centralized repository becomes impractical. To address the need for autonomy, federated database technology has been proposed to manage locally stored data with a federated DBMS and provide unified data access. However, the centralized DBMS still introduces data heterogeneity, privacy, and trust issues. Meanwhile, the peer-to-peer information sharing framework is often considered a solution between "sharing nothing" and "sharing everything". In its basic form, every pair of peers establishes two symmetric client-server relationships, and requestorsSend queries to multiple databases.This approach assumes 2n relationships for n peers, and is not scalable.In the context of sensitive data and autonomous data owners, a more practical and adaptable solution is to construct a data centric overlay including the data sources and a set of brokers helping to locate data sources for queries.Such infrastructure builds up semantic-aware index mechanisms to route the queries based on their content, which allows users to submit queries without knowing data or server location.

In our previous study [9], [10], such a distributed system providing data access through a set of brokers is referred to as Information Brokering System (IBS). As shown in Figure 1, applications atop IBS always involve some sort of consortium (e.g. RHIO) among a set of organizations.Databases of different organizations are connected through a set of brokers, and metadata (e.g. data summary ,server locations) are "pushed" to the local brokers, which further advertise" (some of) the metadata to other brokers .Each query is sent to the local broker, and routed according to the metadata until reaching the right database(s). In this way, alarge number of information sources in different organizations are loosely federated to provide an unified

LITERATURE SURVEY:

An XML brokerage system is a distributed XML database system that comprises data sources and brokers, which, respectively, hold XML documents and document distribution information [8]. However, all existing information brokerage systems view or handle query brokering and access control as two orthogonal issues: query brokering is a system issue that concerns costs and performance, while access control is a security issue that concerns information confidentiality.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

As a result, access control deployment strategies (in terms of where and when to do access control) and the impact of such strategies on end-to-end system performance are neglected by existing information brokerage systems. In addition, data source side access control deployment is taken for granted as the —right thing to do We challenge this traditional, taken-for-granted access control deployment methodology, and argue that query brokering and access control are not two orthogonal issues because access control deployment strategies can have a significant impact on the —whole system's end-to-end performance. We propose the first in-broker access control deployment strategy where access control is —pushed from the boundary into the —heart of the information brokerage system.

The PPIB study assumes that a global schema exists within the consortium, therefore, information integration is out of our scope. Peerto-peer systems are designed to share files and data sets(e.g., in collaborative science applications).

Distributed hash table technology is adopted to locate replicas based on keyword queries. However, although such technology has recently been extended to support range queries , the coarse granularity (e.g., files and documents) cannot meet the expressiveness needs of applications focused in this work. Furthermore,P2P systems often returns an incomplete set of answers while we need to locate all relevant data in the IBS.

Addressing a conceptually dual problem, XML publish-subscribe systems (e.g., [19], [20]) are probably the closely relate technology to the proposed research problem: while PPIB aims to locate relevant data sources for a given query and route the query to these data sources, the pub/sub systems locate relevant consumers of a given document and route the document to these consumers. However, due to this duality, we have different concerns.

PROBLEM STATEMENT:

In this system has some existing problem as like site distribution and load balancing. In PPIB, site distribution and load balancing are conducted in an ad-hoc manner.

PPIB can suffer from certain load imbalances due to data storing and query routing, load imbalance caused by these factors can be efficiently tackled without substantial performance degradation. However, no load balancing is considered and no explicit results showing query processing costs are reported. [11]. Load balancing of the load caused by resolving queries from caches is more crucial due to the high traffic it creates to supply query results compared to the metadata-index lookup. Another problem is drawing an automatic scheme which performs dynamic site distribution.

There is a need o consider several other factors such as the workload and trust level of each peer, and privacy disagreement between automaton segments. A scheme that can strike a balance among these factors is a point of consideration. Second, we would like to quantify the level of privacy protection achieved by PPIB. A plan to minimize or eliminate the participation of the administrator, whose role is decide some issues such as automaton segmentation granularity will also worked out. A primary intention is to build PPIB self-reconfigurable.

Disadvantage:

1.The database with the tuple data does not be maintained confidentially.
2.The existing systems another person to easily access database.

PROBLEM DEFINITION:

In the current paper, we present two efficient protocols, one of which also supports the private update of a generalization-based anonymous database. We also provide security proofs and experimental results for both protocols. So far no experimental results had been reported concerning such type of protocols; our results show that both protocols perform very efficiently.

Advantage:

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

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

IMPLEMENTATION:

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

Interaction Model:

1.Client-driven interventions:

Client-driven interventions are the means to protect customers from unreliable services. For example, services that miss deadlines or do not respond at all for a longer time are replaced by other more reliable services in future discovery operations.

2.Provider-driven interventions:

Provider-driven interventions are desired and initiated by the service owners to shield themselves from malicious clients. For instance, requests of clients performing a denial of service attack by sending multiple requests in relatively short intervals are blocked (instead of processed) by the service.Co-Ordinator Module:-In this module, the co-coordinator performs the global service between the two end users. Initially the Data Owner needs to submit the details of the patient in the server.Data Users needs to search the data which is stored in the servers and they give request for the data and the co-Ordinator sends the key to the Data users and the Data will be passed by the broker Way.

Broker Module:-In this module, the broker performs the role who can act between the Co-coordinator and the data Users. The request which are all submitted from the data user will be verified and thus it will be passed to the co-coordinator.The data will be passed from the co-coordinator and thus it will be submitted to the End Users(Data Users).

User Module:- In this module, the Users are classified into two types they are, Data Users and Data Owner Depends on the restriction the data will be passed to the Co-coordinator.The co-coordinator pass the details via broker and the data will be checked with the secret key and thus it will display for the users.

Admin Module:

In this module, to arrange the database based on the patient and doctor details and records. The admin needs to register and register the Organization and Users Forms.

CONCLUSION:

In this paper, we have presented With little attention drawn on privacy of user, data, and metadata during the design stage, existing information brokering systems suffer from a spectrum of vulnerabilities associate with user privacy, data privacy, and metadata privacy. In this paper, we propose PPIB, a new approach to preserve privacy in XML information brokering. Through an innovative automaton segmentation scheme, in-network access control, and query segment encryption, PPIB integrates security enforcement and query forwarding while providing comprehensive privacy protection. Our analysis shows that it I very resistant to privacy attacks.

End-to-end query processing performance and system scalability are also evaluated and the results show that PPIB is efficient and scalable. Many directions are ahead for future research. First, at present, site distribution and load balancing in PPIB are conducted in an ad-hoc manner. Our next step of research is to design an automatic scheme that does dynamic site distribution.

Several factors can be considered in the scheme such as the workload at each peer, trust level of each peer, and privacy conflicts between automaton segments. Designing a scheme that can strike a balance among these factors is a challenge. Second, we would like to quantify the level of privacy protection achieved by PPIB. Finally, we plan to minimize (or even eliminate) the participation of the administrator node, who decides such issues as automaton segmentation granularity. A main goal is to make PPIB self-reconfigurable.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

REFERENCES:

[1] Fengjun Li, Bo Luo, Peng Liu, Anna Squicciarini, Dongwon Lee, and Chao-Hsien Chu. "Defending against Attribute-Correlation Attacks in Privacy-Aware Information Brokering", Proceedings of the 4th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), November, 2008. Orlando, FL.

[2] Noe Elisa , K. Suresh Babu " Survey on Protecting privacy and security in xml information brokering", IJC SMC, Vol.3 Issue.4, April- 2014,

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

[4] Xiong, L., Liu, L., A Reputation-Based Trust Model for Peer-toPeer eCommerce Communities, IEEE International Conference on E-Commerce (CEC), 2003

[5] Yu, B., Singh, M. P., A social mechanism of reputation management in electronic communities, 4th Intl Workshop on Cooperative Information Agents, 2000

[6] Frank Dabek, Frans Kaashoek, David Karger, Robert Morris, and Ion Stoica, "Wide-area Cooperative Storage with CFS," in Proc. ACM SOSP, Banff, Canada, 2001.

[7] P. Resnick, R. Zeckhauser, E. Friedman, and K. Kuwabara. Reputation systems. Communications of the ACM, 43(12):45-48, December 2000.

[8] I. Stoica, R. Morris, D. Karger, F. Kaashoek, and H. Balakrishnan. Chord: A scalable Peer-To-Peer lookup service for internet applications. Proceedings of the 2001 ACM SIGCOMM Conference, 2001.

[9] J. Kang and J. F. Naughton. On schema matching with opaque column names and data values. In SIGMOD, pages 205–216, 2003.

[10] G. Koloniari and E. Pitoura. Content-based routing of path queries in peer-to-peer systems. In EDBT, 2004.

[11] G. Koloniari and E. Pitoura. Peer-to-peer management of xml data: issues and research challenges. SIGMOD Rec., 34(2):6–17, 2005.

[12] N. Koudas, M. Rabinovich, D. Srivastava, and T. Yu. Routing xml queries. In IEEE ICDE, page 844, 2004.

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

[14] H. Lu, J. X. Yu, G. Wang, S. Zheng, H. Jiang, G. Yu, and A. Zhou. What makes the differences: benchmarking xml database implementations. ACM Trans. Inter. Tech., 5(1):154–194, 2005.

[15] B. Luo, D. Lee, W.-C. Lee, and P. Liu. QFilter: Fine-grained runtime XML access control via NFA-based query rewriting. In ACM CIKM, Washington D.C., USA, nov 2004.

[16] I. Manolescu, D. Florescu, and D. Kossmann. Answering xml queries on heterogeneous data sources. In VLDB, pages 241–250, 2001.

[17] M. Murata, A. Tozawa, and M. Kudo. XML access control using static analysis. In ACM CCS, Washington D.C., 2003.

[18] S. Park, A. Khrabrov, D. M. Pennock, S. Lawrence, C. L. Giles, and L. H. Ungar. Static and dynamic analysis of the internet's susceptibility to faults and attacks. In IEEE Infocom, 2003.

[19] M. K. Reiter and A. D. Rubin. Crowds: anonymity for Web transactions. ACM Transactions on Information and System Security, 1(1):66–92, 1998.

[20] 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.

AUTHORS BIOGRAPHY:

Akba Zoheer Mohammed, pursuing his Master of Science in Information System, from Nizam College (Autonomous),O.U, Basheer Bagh, Hyderabad, India.

T. Ramdas Naik,Assistant Professor, Dept of Computer Science (PG),Nizam College (Autonomous),O.U, Basheer Bagh, Hyderabad.