Survey On A New Approach To Preserve Privacy In Information Brokering

Abstract

Now a day’s sharing information across different organizations often operate across organizational boundaries. So needs for effective and secure information sharing to facilitate extensive collaborations organizations. Data sharing is provided in distributed file sharing environment. Information Brokering System (IBS) is a peer-to-peer overlay network to share information in loosely federated data sources. We present the problem occurs in privacy protection in information brokering process and introduce privacy preserving information sharing via on-demand information access. However increasing on protecting the private and sensitive data, the organizations sharing data in a more secure and privacy-preserving manner. A formal presentation of the threat models with a focus on mainly two attacks: attribute-correlation attacks and inference attacks. Query and access control rules are maintained under metadata with shared data details. Automata segmentation and query segment encryption schemes are used in a new approach Privacy Preserving Information Brokering (PPIB) system for security enforcement and privacy protection.

1. Introduction

In distributed file system sharing and access data local file as well as remote file. In traditional distributed file system use centralized data and common admin manager. But large distributed system which involves many organization fail to this old system. Now a day’s internet introduce worldwide sharing data and collaboration multiple organization. Distributed or Networked file system to access
and share data using file sharing mechanism. The internet and various communication networks provide data communication channel, so main goal is achieving secure, protected, privacy also handling the heterogeneous data through different organization management systems. However, different types of applications often need different forms of information sharing. Government organization such as the military are in an analogous situation to share sensitive information about potential military targets, suspicious activities, difficult technical problems, or vulnerabilities with partners at differing levels of trust. Misuse of this information may result in harm. But harm is also possible if information is not shared, as the information could be necessary to prevent loss of life, assets, or advantage.

Ex[1] To understand such requirements, let us consider Regional Health Information Organization (RHIO) model, are being developed to share medical information (e.g., patient records). Since the records are highly sensitive and private, intensive privacy and security enforcement is desired and collected by collaborative health providers (e.g., hospitals). Each health provider as data source is authorized by its patients to collect medical information. Since the data is protected the health providers are responsible for not leaking patient records to irrelevant parties. Diverse users (e.g., doctors, assistants, pharmacists, and administrators) are to access local or remote patient data according to certain access control policies. Finally, the RHIO should be able to maintain a large number of data. In general, such interorganization collaboration application requires an information sharing system that offers full autonomy to underlying databases preserves data security and privacy comprehensively, and provides good scalability [5].

In existing information broking systems IBS [2] work is designed with user and data privacy. To satisfy such privacy protection requirements, therefore, a novel IBS, named as Privacy Preserving Information Brokering system (PPIB) is proposed. As shown in fig 1, PPIB contains a broker-coordinator overlay network where the brokers are responsible for forwarding user queries to coordinators 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. PPIB takes an innovative 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. This scheme preserves privacy without sacrificing functionality. While providing full capability to do in-network access control and to route queries to the right data sources, this scheme ensures the information that a (curious, corrupted or broken) coordinator can gather is far from being enough to infer either “which data is being queried” or “where the data is located”. Second, the automaton segmentation scheme can also provide high-quality privacy protection to metadata (e.g., access control policy). Third, user location privacy is protected by multilateral security, a design principle of PPIB.

To the best of this work, (1) PPIB is the first system that uses automaton segmentation to do privacy-preserving in-network access control. (2) PPIB is the first system that integrates automaton segmentation, in-broker access control, and query routing. (3) PPIB provides the most comprehensive privacy protection for information brokering systems, and its performance degradation is insignificant compared with traditional IBS systems. (4) The evaluation results show that PPIB is a scalable privacy solution. Brokers and Coordinator are linked in a peer-to-peer fashion that makes PPIB a scalable system [1].
2. Privacy Preserving Information Brokering System

As shown fig 2, it is an overlay architecture consisting of two types of brokering components, brokers and coordinators. The brokers, acting as mix anonymizer, are mainly responsible for user authentication and query forwarding. The coordinators, concatenated in a tree structure, enforce access control and query routing based on the embedded nondeterministic finite automata the query brokering automata. To prevent curious or corrupted coordinators from inferring private information, we design two novel schemes to segment the query brokering automata and encrypt corresponding query segments so that routing decision making is decoupled into multiple correlated tasks for a set of collaborative coordinators, while providing integrated in-network access control and content-based query routing. Our automaton segmentation scheme first divides the global access control automaton into several segments. Each accept state of the global automaton is specially partitioned as a separate segment. Then we assign each segment to one independent site [7].

![Figure 1.1: Overview of the PPIB Infrastructure](image1)

![Figure 2.1: System architecture of a privacy preserving information brokering system.](image2)

I. Segmentation: The atomic unit in the segmentation is an NFA state of the original automaton. Each segment is allowed to hold one or several NFA states.

II. Deployment: We employ physical brokering servers, called coordinators, to store the logical segments. To reduce the number of needed coordinators, several segments can be deployed on the same coordinator using different port numbers. Therefore, the tuple uniquely identifies a segment.

III. Replication: Since all the queries are supposed to be processed first by the root coordinator, it becomes a single point of failure and a performance bottleneck. For robustness, we need to replicate the root coordinator as well as the coordinators at higher levels of the coordinator tree. Replication has been extensively studied in distributed systems.
3. Literature Survey
In [2] F. Li, B. Luo, P. Liu, D. Lee, P. Mitra, 2006, W. Lee, and C. Chu authors define in Broker Access Control for Information Brokerage Systems, 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. 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. 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 this paper, challenge taken-for-granted access control deployment methodology, and show that query brokering and access control are not two orthogonal issues because access control deployment strategies can have significant impact on the “whole” system’s end-to-end performance. 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.

In [4] S. Rizvi, A. Mendelzon, S. Sudarshan, and P. Roy define model Query Rewriting Techniques for Fine Grained Access Control. In this model, users are members of appropriate roles; and an access control policy consists of a set of role based 5-tuple access control rules (ACR): R = {subject, object, action, sign, type}, where (1) subject is a role to whom an authorization is granted; (2) object is a set of XML nodes specified by XPath; (3) action is one of “read,” “write,” and “update”; (4) sign 2 {+, −} refers to access “granted” or “denied,” respectively; and (5) type 2 {LC,RC} refers to either “Local Check” (i.e., authorization is only applied to attributes or textual data of context nodes—“self::text()” or “self::attribute()”), or “Recursive Check” (i.e., authorization is applied to context nodes and propagated to all descendants—“descendant-or-self::node()”). When an XML node does not have either explicit (via LC rules) or implicit (via RC rules) authorization, it is considered to be “access denied.” It is possible for an XML node to have more than one relevant access control rule. If conflict occurs between “+” and “−” rules, “−” rules take precedence.

In [6] G. Koloniari and E. Pitoura present Content Based Routing (CBR) of Path Queries in Peer-to-Peer Systems. Where peer-to-peer (P2P) systems are gaining increasing popularity as a scalable means to share data among a large number of autonomous nodes. The nodes in a P2P system store XML documents. A fully decentralized approach to the problem of routing path queries among the nodes of a P2P system based on maintaining specialized data structures, called filters this will efficiently summarize the content. Building a hierarchical organization of nodes by clustering together nodes with similar content. Similarity between nodes is related to the similarity between the corresponding filters. The existing CBR System follows hierarchical organization, which is time-consuming job to process the data between the agent and the remote user.

4. Existing System
Today’s organizations raise associate increasing want for info sharing via on-demand access. Info brokering systems (IBSs) are planned to attach large-scale loosely federated knowledge sources via a brokering overlay, within which the brokers make routing selections to direct shopper queries to the requested data servers. Several existing IBSs assume that broker’s area unit sure and so solely adopt server-side access management for knowledge confidentiality. However, privacy of information,
location and data client will still be inferred from data (such as question and access management rules) changed inside the IBS, however very little attention has been place on its protection.

Consider an information brokerage system where sensitive information is shared among geographically distributed participants (e.g., users and data sources). In general information brokering process, XML queries created by a user are forwarded to data sources by intermediate brokers. Since multiple data sources may be relevant to one XML query, replies from all relevant data sources will be merged to provide an aggregate view to the user. In this process, brokers perform as a bridge connecting users and data sources, so they are necessary to know who holds the required data and where they are located. To make the exposition simple, we assume that each broker has a full knowledge of whereabouts of stored data. Therefore each broker may direct an inquiry to relevant data sources without consulting others (i.e., single-hop brokering).

Conventional information sharing system has trust assumption on the use of servers. This trust or semi-trust may not hold for brokers. So, system may be abused by insider or outsider. Information Brokering depends on the trust of brokers for query forwarding 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 user location information. Data privacy contains data location and data object distribution 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 provide access to authorized users. Although attacks may not obtain the plaintext data over encrypted data, they can still learn the query location and data location from eavesdrop. The attacks are classified as (1) attribute correlation attack: when query is routed, compromised broker or external attacker (eavesdropper) may be extract the query condition for getting the sensitive information by matching the attributes contained in the query. (2) Inference

Figure 4.1: Information Brokering System.
attack: By getting more than one type of sensitive information, the attacker guessing the query location (IP address) and query content and identify data owner from the query content.

5. Conclusion

We have prepared a survey paper on existing information brokering system suffers from vulnerabilities associated with the user privacy, data privacy and metadata privacy. These are addressed by the privacy preserving information brokering system. By using query brokering automata, PPIB achieves privacy conservation and security enforcement while information is exchanged in distributed system. It provides the security against eavesdropper, malicious broker, and malicious coordinators.

References