Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 7, July 2014, pg.437 – 444

RESEARCH ARTICLE

Result on Enforce Secure and Privacy Preserving Information Brokering in Distributed Information Sharing

Mukesh Kawatghare[1], Pradnya Kamble[2]

[1]Computer Technology & Nagpur University, India

[2]Computer Technology & Nagpur University, India

[1] parthkawatghare@gmail.com; [2] pradnya_kamble@rediffmail.com

Abstract- Today's organizations (e.g. govt. agencies, digital libraries, "Smart" Home) raise increasing needs for information sharing via on-demand information access. An Information Brokering System (IBS) is a peer-to-peer overlay network that comprises diverse data servers and brokering components helping client queries locate the data server(s). Peer-to-peer (P2P) systems are gaining increasing popularity as a scalable means to share data among a large number of autonomous nodes. We study the privacy in Privacy-Preserving Information Brokering in Distributed Information Sharing through an innovative automaton segmentation scheme and query segment encryption and data management issues for processing XM L data in a p2p setting, namely indexing, replication and query routing and processing.

Keywords— automaton segmentation, query segment encryption, privacy, Access control, information sharing

I. INTRODUCTION

Information sharing is becoming increasingly important in recent years, not only among organizations with complementary interests, but also within many fields range from business to other agencies that are becoming ever more globalized and distributed. To provide efficient large-scale information sharing, to accept data heterogeneity and provide interoperability across geographically distributed data sources.

The systems work on two extremes of the spectrum: (1) 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; (2) in the distributed database systems, all the user lost autonomy and are managed by a unified DBMS. However, types of applications often need different forms of information sharing. In particular, while some applications (e.g., stock price updating) would need a publish subscribes framework, the on-demand information access is more suitable for other applications [8].

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. Sequences of DNA and medical histories are stored in databases in autonomous enterprises. [9]

As a data provider, a participant would not assume free or complete sharing with others, its data is legally private or commercially proprietary or both. Instead, it is required to prevent full control over the data and access to the data.

In the sensitive data and autonomous data owners, a more practically and adaptable solution is to construct a data centric overlay [3], [4], including the data sources and a set of brokers helping to locate data sources for queries [6], [7].Mechanisms to route the queries supported their content, that permits users to submit queries while not knowing information or server location. In previous study [7], [8], 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 trust-able. So, system may be abuse by insider or outsider.

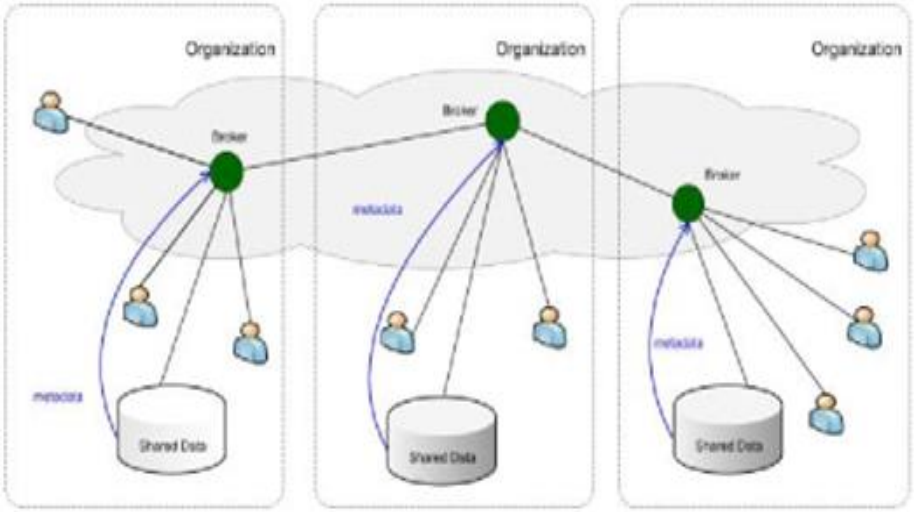

Fig. 1. Overview of the IBS infrastructure

II. PRIVACY-PRESERVING INFORMATION BROKING

Privacy protection is need for the Information Brokering System (novel IBS), named Privacy Preserving Information Brokering (PPIB).PPIB has 2 form of brokering Component: (1) brokers and (2) co-ordinators. The brokering area unit principally to blame for user authentication and question forwarding, the broker performs the role World Health Organization will act between the Co-coordinator and therefore the information Users. The request which is all submitted from the data user will be verified and thus it will be passed to the co-coordinator. 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. Coordinators, each holding a segment of access control automaton and routing guidelines, are mainly responsible for access control and query routing. [8]

PPIB takes an innovator automaton segmentation approach to privacy protection. In particular, two critical forms of privacy, namely data object distribution privacy and query content 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 segment. System can providing full capability to wage in network access control and to path queries to the right data sources, these two schemes ensure that or unservice inquisitional able 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 a data”. PPIB enables wide-ranging security and privacy protection for claimed information brokering, with minor overhead and major scalability.

III.SECURITY AND PRIVACY NEED FOR PPIB

In information brokering scenario, there are three types of entrepreneur data owners, data providers and data requestor. Each entrepreneur has its own privacy: (1) 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 user. (2) Data provider store collected data, and create two types of metadata, access control metadata and routing metadata. (3) 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, curious brokering components and outside attackers. 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 [8], which also strut queries from requestors. Therefore, a curious or corrupted brokering server could: (1) learn query content and query location by impede a local query; (2) learn routing metadata and access control metadata from local data servers and other brokers; (3) learn data location from routing metadata it holds Although attacker may not obtain plaintext data over encrypted data, they can learn query location and data location. The attacks into two major classes: (1) the attribute-correlation attack and (2) inference attack.

A. Attribute-correlation attack

An attacker prevents a query, which typically contains several predicate. Each predicate describes a condition, which is sometimes involves sensitive and private data (e.g. credit card number, etc.).

B. Inference attack

Attacker some techniques and result more than one other type of sensitive information so more sever, further associates to learn implicit and explicit knowledge about entrepreneur

IBS work is designed with user and data privacy. Such type of privacy protection requirements, therefore a novel IBS, named as Privacy Preserving Information Brokering system (PPIB). As shown in Figure, PPIB contains a broker-coordinator 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.

IV. ARCHITECTURE OF PPIB

PPIB has three types of brokering components: (1) Brokers (2) Coordinators and (3) Central authority (CA). The key to defend privacy is to part the work on more than one components in such a way that more than one node can make a meaningful presumption from the information disclosed to it. Figure 2 shows the architecture of PPIB. Through local brokers (green nodes in Fig) Data servers and requestors from different organizations connect to the system

A) *Brokers*
It is intercommunicating through coordinators (white nodes in Fig). A local broker functions as the "entry" to the system. It's responsible for authenticates requestors and hides their. It would also permute query sequence to defend against local traffic analysis.

B) *Coordinators*
It is responsible for content-based query routing and access control actuation. With privacy-preserving idea, coordinator cannot hold any rule in the complete form. Instead, a novel automaton segmentation scheme to divide (i.e. metadata) rules into segments and assign each segment to a coordinator. Coordinators operate collaboratively to enforce secure query routing.

Coordinator prevents from sensitive predicates, a query segment encryption scheme and automaton segmentation scheme, query divide into segment and encrypt it (each segment)

C) *Central Authority (CA)*
It is responsible for key management and metadata maintenance.

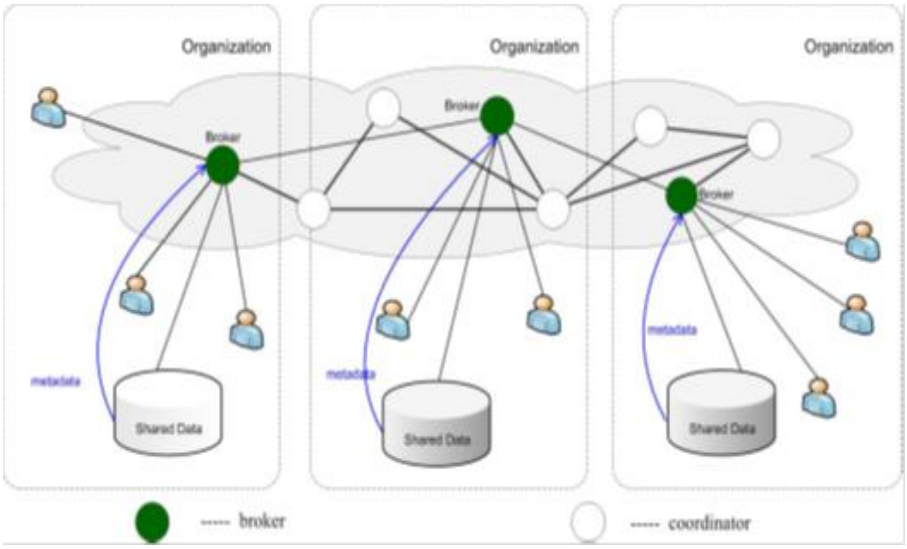

Fig. 2. Architecture of PPIB

The architecture of the privacy preserving information brokering system is shown in Fig. 2, where users and data servers of more than one organizations are communicate via a Broker, coordinator overlay component. User requests for data by sending a XML query to the local broker, which further carry the query to the root of the coordinator tree. The query is processed along a path of the multiple organizations coordinator. The brokering process consists of 4 phases:

1) *Phase 1*: For join the system, a user needs to authenticate to the local broker. And the user submits encrypted segment an XML query by public level keys, and a unique session key Ks, data servers encrypted with the public key, to return data.

2) *Phase 2*: The major task of the broker is metadata preparation: (1) it extracts the role of the user authenticated and attaches it to the encrypted XML query; (2) it make a unique ID for each query, and attaches QID with its own address (as well as < Ks >pkDS) to the query so that the data server can directly return the data.

3) *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.

4) *Phase 4:* In the final phase, 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 Ks, to the broker of the query.

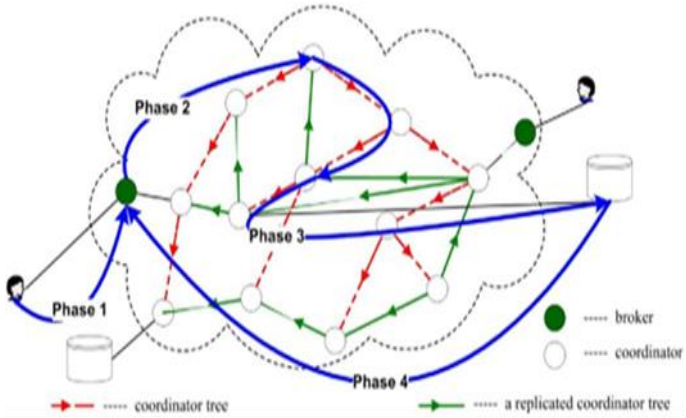

Fig. 3. Query brokering process in 4 phases

V. APPLICATIONS

Information (Data) Brokers collect data and provide data mining services for various organizations, for example in the FBI, Credit Monitoring Services, DoD, etc. The companies are a high value target for social engineers as they contain huge amounts of information that could be used to further elevate. Because of relaxed regulations and federal laws much of our personal information is collected by government agencies and stored or managed by these Information Broker Companies.

Information brokering is suitable for many newly emerged applications, such as information sharing for healthcare or law enforcement, in which organizations share information in ailliberal and controlled manner, not only from business considerations but also due to legal reasons.

A) Healthcare information systems, such as Regional Health Information Organization (RHIO) [1], to facilitate retrieval of clinical data thereon collaborative health providers.

B) Law enforcement, for example young police officers, police academics, researchers agencies use information brokering technologies to share on demand data with other agencies and the public.

VI. EXISTING PROBLEM

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.

VII. EXPERIMENTAL RESULTS

The following modules shows the how PPIB are performed and Secure a data accessing from the different organization. In this LOGIN PAGE four authorized user can access the data or sending the query (i.e. segmentation query). Four users are admin, co-ordinator, broker and user. Each user has a different role. Only one admin control all data, which is create a numbers of organization, registration of user, broker and co-ordinator.

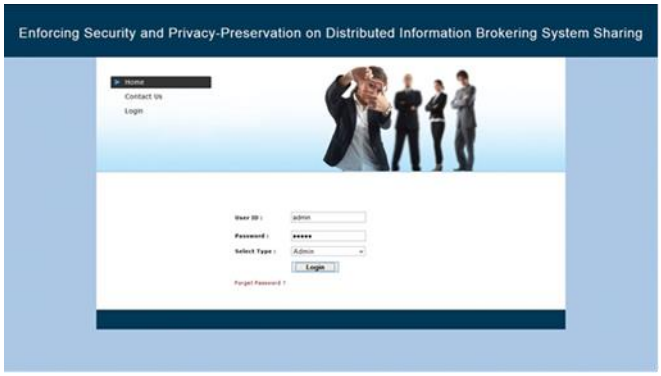

Fig. 3. Admin Page.

A) User Phase: In this phase user send a request to the broker

B) Broker Phase: In broker phase we are send a query (i.e. segmented query) to the different organization for accessing a data.

C) Co-ordinator Phase: In this phase receive a data from the broker (different organization broker) which has a data and send to a user which belongs to same coordinator's organization.

In the figure show the result which is access the data from the different organization database. All data are in encryption form after selected the data is decrypted and user can get a data in plain text form. All the queries are segmented, encrypt and routing until reached the destination organization local database and finally give a data as per user query.

Users stored a data into local database but some time local database is reached at maximum load and other organization database are store minimum load. The PPIB are load balancing concept use for distribute the load to different site.Particular site does not maximum load its distributed the data to other site, PPIB frequently work because of minimum load.

Fig. 3. Coordinator Data

VIII. CONCLUSION

Privacy issues of user and data during the design stage is considered and concluded that existing information brokering systems suffer from a spectrum of vulnerabilities associated with user privacy, data privacy, and metadata privacy. In this paper, PPIB proposed architecture is discussed, a new approach to preserve privacy in XML information brokering. By using automaton segmentation scheme, within network access control and query segment encryption, PPIB put together security enforcement and query forwarding at the same time as providing comprehensive privacy protection. We claim that our analysis is very resistant to privacy attacks. Node-to-node query processing performance and system scalability are also evaluated and the results show that PPIB is efficient and scalable.

REFERENCES

[1] W. Bartschat, J. Burrington-Brown, S. Carey, J. Chen, S. Deming, and S. Durkin, "Surveying the RHIO landscape: A description of current RHIO models, with a focus on patient identification," Journal of AHIMA 77, pp. 64A–D, January 2006.

[2] A. P. Sheth and J. A. Larson, "Federated database systems for managing distributed, heterogeneous, and autonomous databases," ACM Computing Surveys (CSUR), vol. 22, no. 3, pp. 183–236, 1990.

[3] L. M. Haas, E. T. Lin, and M. A. Roth, "Data integration through database federation," IBM Syst. J., vol. 41, no. 4, pp. 578–596, 2002.

[4] 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.

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

[6] N. Koudas, M. Rabinovich, D. Srivastava, and T. Yu, "Routing XML queries," in ICDE '04, p. 844, 2004.

[7] G. Koloniari and E. Pitoura, "Peer-to-peer management of XML data: issues and research challenges," SIGMOD Rec., vol. 34, no. 2, 2005.

[8] 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.

[9] 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.
[10] 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.
[11] D. L. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," Communications of the ACM, vol. 24, no. 2, 1981.
[12] R. Agrawal, A. Evfimivski, and R. Srikant, "Information sharing across private databases," in Proceedings of the 2003 ACM SIGMOD, 2003.
[13] M. Genesereth, A. Keller, and O. Duschka, "Informaster: An information integration system," in SIGMOD, (Tucson), 1997.
[14] I. Manolescu, D. Florescu, and D. Kossmann, "Answering XML queries on heterogeneous data sources," in VLDB, pp. 241–250, 2001.
[15] J. Kang and J. F. Naughton, "On schema matching with opaque column names and data values," in SIGMOD, pp. 205–216, 2003.
[16] 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 Transactions on Networking, vol. 11 of 1, 2003.
[17] R. Huebsch, B. Chun, J. Hellerstein, B. Loo, P. Maniatis, T. Roscoe, S. Shenker, I. Stoica, and A. Yumerefendi, "The architecture of PIER: an Internet-scale query processor," in CIDR, pp. 28–43, 2005.
[18] O. Sahin, A. Gupta, D. Agrawal, and A. E. Abbadi, "A peer-to-peer framework for caching range queries," in ICDE, 2004.
[19] A. Carzaniga, M. J. Rutherford, and A. L. Wolf, "A routing scheme for content-based networking," in Proc. of INFOCOM, 2004.
[20] Y. Diao, S. Rizvi, and M. J. Franklin, "Towards an Internet-scale XML dissemination service," in VLDB Conference, (Toronto), August 2004.
[21] G. Koloniari and E. Pitoura, "Content-based routing of path queries in peer-to-peer systems.," in EDBT, pp. 29–47, 2004.
[22] M. K. Reiter and A. D. Rubin, "Crowds: anonymity for Web transactions," ACM TISS, vol. 1, no. 1, pp. 66–92, 1998.
[23] P. F. Syverson, D. M. Goldschlag, and M. G. Reed, "Anonymous connections and onion routing," in IEEE Symposium on Security and Privacy, (Oakland, California), pp. 44–54, 4–7 1997.
[24] W. Tolone, G.-J. Ahn, T. Pai, and S.-P. Hong, "Access control in collaborative systems," ACM Comput. Surv., vol. 37, no. 1, 2005.
[25] S. Cho, S. Amer-Yahia, L. V. S. Lakshmanan, and D. Srivastava, "Optimizing the secure evaluation of twig queries.," in VLDB, 2002.
[26] M. Murata, A. Tozawa, and M. Kudo, "XML access control using static analysis," in ACM CCS, 2003.
[27] S. Rizvi, A. Mendelzon, S. Sudarshan, and P. Roy, "Extending query rewriting techniques for fine-grained access control," in SIGMOD'04, (Paris, France), pp. 551–562, 2004.
[28] T. Yu, D. Srivastava, L. V. S. Lakshmanan, and H. V. Jagadish, "Compressed accessibility map: Efficient access control for XML," inVLDB, (China), pp. 478–489, 2002.
[29] B. Luo, D. Lee, W. C. Lee, and P. Liu, "Qfilter: Fine-grained runtime XML access control via nfa-based query rewriting enforcement mechanisms," in CIKM, 2004.
[30] A. Berglund, S. Boag, D. Chamberlin, M. F. Fernndez, M. Kay, J. Robie, and J. Simon, "XML path language (XPath) version 2.0." http://www.w3.org/TR/xpath20/, 2003.
[31] E. Damiani, S. Vimercati, S. Paraboschi, and P. Samarati, "A fine-grained access control system for XML documents," ACM TISSEC, vol. 5, no. 2, pp. 169–202, 2002.
[32] E. Damiani, S. di Vimercati, S. Paraboschi, and P. Samarati, "Securing XML documents," EDBT 2000, pp. 121–135, 2000.
[33] H. Zhang, N. Zhang, K. Salem, and D. Zhuo, "Compact access control labeling for efficient secure XML query evaluation," Data & Knowledge Engineering, vol. 60, no. 2, pp. 326–344, 2007.
[34] Y. Xiao, B. Luo, and D. Lee, "Security-conscious XML indexing," Advances in Databases: Concepts, Systems and Applications, 2007.
[35] E. Bertino, S. Castano, and E. Ferrari, "Securing XML Documents with AuthorX," IEEE Internet Computing, vol. 5, no. 3, pp. 21–31, 2001.
[36] E. Damiani, S. Vimercati, S. Paraboschi, and P. Samarati, "Design and implementation of an access control processor for XML documents.," Computer Networks, vol. 33, no. 1-6, pp. 59–75, 2000.
[37] A. Gabillon and E. Bruno, "Regulating access to xml documents," in Proc. DAS, pp. 299–314, 2002.
[38] W. Fan, C.-Y. Chan, and M. Garofalakis, "Secure xml querying with security views," in ACM SIGMOD, pp. 587–598, 2004.
[39] M. Kudo, "Access-condition-table-driven access control for XML databases," ESORICS 2004, pp. 17–32, 2004.
[40] S. Mohan, A. Sengupta, and Y. Wu, "Access control for XML: a dynamic query rewriting approach," in Proc. IKM, pp. 251–252, 2005.

[41] N. Qi and M. Kudo, "XML access control with policy matching tree," ESORICS 2005, pp. 3–23, 2005.
[42] L. Bouganim, F. D. Ngoc, and P. Pucheral, "Client-based access control management for XML documents.," in VLDB, pp. 84–95, 2004.
[43] S. Abiteboul, A. Bonifati, G. Cob´ena, I. Manolescu, and T. Milo, "Dynamic xml documents with distribution and replication," in ACM SIGMOD, pp. 527–538, ACM, 2003.
[44] P. Skyvalidas, E. Pitoura, and V. Dimakopoulos, "Replication routing indexes for xml documents," in DBISP2P Workshop, 2007.
[45] G. Skobeltsyn, Query-driven indexing in large-scale distributed systems. PhD thesis, EPFL, 2009.
[46] P. Rao and B. Moon, "Locating xml documents in a peer-to-peer network using distributed hash tables," TKDE, vol. 21, no. 12, 2009.
[47] C. Clifton, M. Kantarcioglu, J. Vaidya, X. Lin, and M. Zhu, "Tools for privacy preserving distributed data mining," ACM SIGKDD Explorations, vol. 4, no. 2, 2003.
[48] H. Y. S. Lu, "Commutative cipher based en-route filtering in wireless sensor networks," in VTC, vol. 2, pp. 1223– 1227, Sept. 2004.
[49] A. Schmidt, F. Waas, M. Kersten, M. J. Carey, I. Manolescu, and R. Busse, "XMark: a benchmark for XML data management," in VLDB, pp. 974–985, 2002.
[50] 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., vol. 5, no. 1, pp. 154–194, 2005.

