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

A Framework to maintain Access Control Mechanism for Relation Database

Pranali A. Khatode[1], Jyoti N. Nandhimath[2]

[1]ME student, Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Pune, India

[2]Assistant Professor, Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Pune, India

Abstract: ***Now a days due to the prominent progress of internet and increase of web applications, number of transactions are happening. This in turn it affects the database too .So when the applications are designed to share and update the database of the relational database then accuracy becomes a huge matter of concern. As number of requesters including the hackers can keep asking for the data then it can become security threat for the data. So implementing a proper access control mechanism becomes the great need of the present need over the rational databases.***

Keywords: PIR, Access Control Mechanism, k-anonymity, privacy preserving

1. Introduction

In the recent years, organization facilitates the storage, access and sharing of data which is analyzed to improve the service. To ensure the privacy control over the information available to the users, it is compulsory to apply the Access Control Mechanism. One more important aspect is maintain the privacy of the data, which discloses the identity and satisfy requirement of privacy. This paper presents Framework which deals with security by providing dynamic access control and privacy preservation in controlled manner for relational database using ontology and fuzzy classification well weaved key management system.

2. Literature Survey

PIR has to date been the primary approach to the problem of preserving access privacy for Internet users. For an n-bit database X that is organized into r b-bit blocks, Beimel et al. [1,2] shows that standard PIR schemes cannot avoid a computation cost that is linear in the database size because each query for block Xi must necessarily process all database blocks Xj , j [r]. They proposed a preprocessing model of PIR that computes and stores some extra bits of information, which is polynomial in the number of bits n of the database.

Several hardware-assisted PIR schemes [3,15,16] rely on the preprocessing model. With the exception of all secure coprocessor-based PIR schemes require periodic database reshuffles (i.e., repeats of the preprocessing stage). The reshuffling cost of [4], for example, is O(log4(n)), but when amortized, it is O(log3(n)) per query. Nonetheless, the paper [4].shows how to achieve improvements in the communication and computational complexity bounds of hardware-assisted PIR to O(log2 n) per query, provided that a small amount of temporary storage, on the order of O(p n), is available on the secure coprocessor. An initial suggestion to base PIR on a subset of a database as a means of reducing the high computational overhead in some application areas was made by Chor et al[5].A similar suggestion to improve the performance of PIR-based techniques for location-based services by basing PIR on a restricted subset of the data space was left as an open problem by Ghinita [6]. Olumofin et al. [17] addressed the open problem identified by Ghinita in the specific context of location-based services. In contrast to these prior works, this paper uniquely addresses the problem of preserving access privacy over a large database in a generic way and provides a concrete system for querying a large database.

Wang et al.[18] Proposed a bounding-box PIR (bbPIR) which combines the concept of k-anonymity with the single-server computational PIR scheme by Kushilevitz and Ostrovsky allow users to define a “bounding box” portion of the database matrix and basing PIR on that smaller portion. Their extension also allows the user to specify both the privacy and the service charge budget for PIR queries. The bbPIR work overlaps our work in some areas, but there are several differences. First, bbPIR defines rectangular bounding boxes within a PIR matrix at runtime, whereas our work considers both runtime and offline approaches to defining database portions. The way we define database portions at runtime also differs from that of bbPIR; we consider the sensitive constants in the input query, statistical information on the distribution of the data, and past query disclosures, which allow for logical or non-contiguous database portions. This is unlike bbPIR, which is agnostic to logical data contents. Second, the bbPIR charge-budget model is based on the number of blocks retrieved (typically the square root of the bounding box area). We model the user’s budget in terms of her delay tolerances, which has more generic interpretations (e.g., response time, number of blocks, computations). Third, bbPIR is restricted specifically to one particular PIR scheme, whereas this approach is generic and can use any underlying PIR scheme. Fourth, bbPIR is limited to the retrieval of numeric data by address or by key using a histogram, whereas we support retrieval using any of three data access models— by index, keyword, or SQL. Our approach also involves an explicit intermediate stage for transforming input query q to an equivalent privacy-preserving query Q and requires minimal user intervention.

Howe and Nissenbaum [8] developed a browser extension known as TrackMeNot which tries to solve the problem of preserving access privacy during web searches. TrackMeNot tries to hide a user's request to a search engine in a cloud of dummy queries that are made at specified time intervals. The privacy guarantee is not as strong as our technique that is based on PIR because the server is still able to observe the content of every query made. Trackmen utilizes a significant amount of constant bandwidth for generating decoy queries, which can potentially slow down the network and the search engine In addition, the adversary might be able to distinguish actual queries from dummy queries by considering their submission timings or other meta information.

Domingo-Ferrer et al. [9] considered a scenario where the holders of a database (e.g., a search engine corpus) are uncooperative in allowing the user to obtain access privacy. In other words, the holders are unwilling to support any PIR protocol, and yet the user desires reasonable access privacy over the large data set. They proposed h(k)-PIR which embellishes the user's query keywords with some other k bogus keywords. After the server returns a response, the client filters the response to remove items related to the bogus keywords, and finally displays the result to the user. They defined an access privacy scheme as satisfying h(k)-PIR if the adversary can only view the user's query as a random variable Q0 satisfying H(Q0) * h(k), where h(k) is a function, k is a non-negative integer, and H(Q0) is the Shannon entropy of Q0. The security of the scheme relies on using a large set of k bogus keywords with identical relative frequencies as the query keywords. However, the accuracy of the query result degenerates with higher values of k, which is their point of tradeoff, unlike our approach where the tradeoff is between privacy and computational efficiency. In addition, their approach relies on the availability of a public thesaurus of keywords and their relative frequencies. It is somewhat misleading for the label of PIR to be used for this approach as its privacy guarantee is not as strong as standard PIR; the adversary can still observe the content of every query made by users.

Verykios et al. [19] describe privacy preserving data mining methods while they took consideration of five dimensions as data distribution, data modification, data mining algorithm, rule hiding, preserving privacy. There are some another issues which are present in PPDM approaches. For example, Oliveira and Zaïane [20] says violation type that can be neglected can be define early before applying PPDM algorithm. Most of the PPDM methods focus more on the attacks that can be done on the data mining or extracted results.Consider example where algorithm gives the anonymous dataset K., Friedman et al. [21] discuss the possibility of building data mining models based on this K dataset. Such types of model preserve the individual privacy for usage of such models for instance classification tree. Another most important issue is to define the privacy of the algorithm. For instance, for reconstructing-based techniques can be used to add noise, thus privacy can be defined by using value range that contains the original value.

The widely accepted PPDM techniques use k-anonymity [22] as a base for achieving privacy in data mining. K-anonymity method is proposed by the Sweeny [23] which says that there is no individual is linked with rows having count less than k. This protocol is set to ensure that minimum k rows can be used for violation of individual privacy. This k rows have the same combination of values in the attribute. This care guarantees that the chances of individual identification which are based on the released data should not more than 1/k. Such model is useful for getting accurate privacy definition that basically considers the linking attack. k-anonymity model his more simple and easy to understand ,in spite of having this advantages it rely on two assumptions i.e. the owner of data should know the attributes in advance that are useful for user identity and the no of levels (K) are sufficient enough to preserve the privacy of the data. The most useful method to follow with k-anonymity is to find certain values that are not more specific but having generalized semantic and then replace such values. But there may be possibility that there is the possibility of not releasing some of the values at all.

It has been proved that the method that finds the minimum j-anonymous dataset using suppression is NP-hard. Thus to implement this, there is diverse need of heuristic algorithms. One group of such approach is explained by various heuristic measures that minimizes the data loss. In such cases the quality of data is totally depends on the difference of attribute values from the original value after applying the said process. Li present a generalized study that gives lots of guidance for the study of generalization schemes. In [11,12,13], the authors says the MASK methodology to preserve privacy for mining of most frequent item set and described the issue of efficiency to calculate the estimated nearby values. The results conclude that the high degree of privacy for both users and mining system can be achieved simultaneously. an analytical formula has been proposed by them to state the privacy metrics and to evaluate the obtained privacy.[14] moves one step advance to describe the issue of giving accuracy in privacy preserving mining. they addressed the issue of how the accuracy of mined randomized data is affected for each association rule.

3. Conclusion

Proposed System Successfully extracts the features of the user requests to identify its score and type of the access parameter. Then by using of ontology in our system it identifies the access protocol policy which is assigned while sharing the data from the data owner. Our system is enhanced with fuzzy based classification protocol to maintain the accurate access control mechanism.

References

[1] A. Beimel, Y. Ishai, and T. Malkin. Reducing the servers' computation in private information retrieval: PIR with preprocessing. In Advances in Cryptology - CRYPTO 2000, 20th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2000, Proceedings, Lecture Notes in Computer Science, pages 55–73. Springer, 2000.
[2] A. Beimel, Y. Ishai, and T. Malkin. Reducing the servers' computation in private information retrieval:

Pir with preprocessing. J. Cryptol., 17(2):125–151, 2004.
[3] D. Asonov. Querying Databases Privately: A New Approach To Private Information Retrieval. pringerVerlag, 2004.
[4] P. Williams and R. Sion. Usable PIR. In NDSS'08: Proceedings of the 16th Annual Network and Distributed System Security Symposium, 2008.
[5] B. Chor and N. Gilboa. Computationally private information retrieval (extended abstract). In STOC'97: Proceedings of the 29th Annual ACM Symposium on Theory of Computing, pages 304–313, New York, NY, USA,1997.
[6] G. Ghinita. Understanding the privacy-efficiency trade-off in location based queries. In SPRINGL '08: Proceedings of the SIGSPATIAL ACM GIS 2008 International Workshop on Security and Privacy in GIS and LBS, pages 1–5, New York, NY, USA, 2008.
[7] S. Wang, D. Agrawal, and A. El Abbadi. Generalizing PIR for practical private retrieval of public data. In DBSec'10, pages 1–16, Rome,Italy,2010.
[8] D. C. Howe and H. Nissenbaum. TrackMeNot: Resisting surveillance in web search. In I. Kerr, V. Steeves, and C. Lucock, editors, Lessons from the Identity Trail: Anonymity, Privacy, and Identity in a Network.
[9] J. Domingo-Ferrer, M. Bras-Amorós, Q. Wu, and J. Manjón. User-private information retrieval based on a peer-to-peer community. Data Knowl. Eng., 68:1237–1252, November 2009.
[10] D. C. Howe and H. Nissenbaum. TrackMeNot: Resisting surveillance in web search. In I. Kerr, V. Steeves, and C. Lucock, editors, Lessons from the Identity Trail: Anonymity, Privacy, and Identity in a Network.
[11] S. Agrawal and J. Haritsa. A framework for high-accuracy privacy-preserving mining. In Proceedings of the 21st IEEE International Conference on Data Engineering, pages 193–204, 2005.
[12] S. Agrawal, V. Krishnan, and J. Haritsa. On addressing efficiency concerns in privacypreserving mining. Proc. of 9th Intl. Conf. on Database Systems for Advanced Applications (DASFAA), pages 113–124, 2004.
[13] S. Rizvi and J. Haritsa. Maintaining data privacy in association rule mining. In Proceedings of the 28th international Conference on Very Large Data Bases, 2002.
[14] Ling Guo, Songtao Guo, and Xintao Wu," On Addressing Accuracy Concerns in Privacy Preserving Association Rule Mining".
[15] S. W. Smith and D. Safford. Practical server privacy with secure coprocessors. IBM Syst. J., 40(3):683–695, 21 2001.
[16] P. Williams and R. Sion. Usable PIR. In NDSS'08: Proceedings of the 16th Annual Network and Distributed System Security Symposium, 2008.
[17] F. Olumofin, P. K. Tysowski, I. Goldberg, and U. Hengartner. Achieving efficient query privacy for location based services. In PETS'10: Proceedings of the 10th Privacy Enhancing Technologies Symposium, Berlin, 2010.
[18] S. Wang, D. Agrawal, and A. El Abbadi. Generalizing PIR for practical private retrieval of public data. In DBSec'10, pages 1–16, Rome, Italy, 2010.
[19] V.S. Verykios, E. Bertino, I.N. Fovino, L.P. Provenza, Y. Saygin, Y. Theodoridis, State-of-the-art in privacy preserving data mining, ACM SIGMOD Record 3 (1) (2004) 50–57.
[20] E. Kushilevitz and R.Ostrovsky. Replication is not needed: single database, computationally-private information retrieval. In FOCS'97: Proceedings of the 38th Annual Symposium on Foundations of Computer Science, page 364, Washington, DC, USA, 1997.
[21] A. Friedman, A. Schuster, R. Wolff, Providing k-anonymity in data mining, VLDB 17 (4) (2008) 789–804.
[22] Slava Kisilevich, Lior Rokach, Yuval Elovici, Bracha Shapira, "Efficient Multi-Dimensional Suppression for K-Anonymity", in proceedings of IEEE Transactions on Knowledge and Data Engineering, Vol. 22, No. 3. (March 2010), pp. 334-347, IEEE 2010.
[23] L. Sweeney, "k-Anonymity: A Model for Protecting Privacy", in proceedings of Int'l Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 2002.
[24] Chang Choi, Junho Choi, Byeongkyu Ko, Kunseok Oh and Pankoo kim, "A Design of Onto-ACM(Ontology based Access Control Model) in Cloud Computing Environments", in proceedings of Journal of Internet Services and Information Security,Vol.2No.2.pp54-64,2012.

Author Profile

Pranali A.Khatode, ME student, Department of Computer Engineering, Smt.Kashibai Navale College of Engineering, Pune, India

Jyoti N. Nandhimath, Assistant Professor, Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Pune, India