IJCSEC- International Journal of Computer Science and Engineering Communications.
Vol.3, Issue 2, 2015, Page.899-902, ISSN: 2347–8586

Application specific Anonymization and Privacy – Preserving Access Control Mechanism for Relational data

T.Monika[1] S.JayaPrakash[2]

[1]PG Student [2]Asst. Prof Dept. of CSE,
[1&2]Dept. of Computer Science and Engineering
Idhaya Engineering College for Women
madhumoni.lena@gmail.com

Abstract: Access Control Mechanisms (ACM) are used to ensure that only authorized information is available to users. However, sensitive information can still be misused by authorized users to compromise the privacy of consumers. The concept of privacy-preservation for sensitive data can require the enforcement of privacy policies or the protection against identity disclosure by satisfying some privacy requirements. Anonymity techniques can be used with an access control mechanism to ensure both security and privacy of the sensitive information. The privacy is achieved at the cost of accuracy and imprecision is introduced in the authorized information under an access control policy. The proposed system is an integrated framework of achieving both privacy and security is proposed though the integration of Access Control Mechanism with Privacy Preservation Technique to prevent the authorized user from misusing the sensitive information with some accuracy loss of data due to the application of privacy preserving technique. The access control policies define selection predicates available to roles while the privacy requirement is to satisfy the k-anonymity or l-diversity. An additional constraint that needs to be satisfied by the PPM is the imprecision bound for each selection predicate. The proposed system applies the application specific anonymization.

Index Terms: Access control, privacy, k-anonymity, query evaluation, application specific anonymization .

Introduction

AS organizations increase their adoption of database systems as the key data management technology for day-to-day operations and decision making, the security of data managed by these systems becomes crucial. Damage and misuse of data affect not only a single user or application, but may have disastrous consequences on the entire organization. The recent rapid proliferation of Web based applications and information systems have further increased the risk exposure of databases and, thus, data protection is today more crucial than ever. It is also important to appreciate that data needs to be protected not only from external threats, but also from insider threats, the proposed system uses the concept of imprecision bound for each permission to define a threshold on the amount of imprecision that can be tolerated. Existing workload aware Anonymization techniques. In this proposed system the focus is on a static relational table that is anonymized only once. To exemplify the proposed approach, role-based access control is assumed. However, the concept of accuracy constraints for permissions can be applied to any privacy-preserving security policy, e.g., discretionary access control.

Organizations collect and analyze consumer data to improve their services. Access Control Mechanisms (ACM) are used to ensure that only authorized information is available to users. However, sensitive information can still be misused by authorized users to compromise the privacy of consumers. The concept of privacy-preservation for sensitive data can require the enforcement of privacy poli- cies or the protection against identity disclosure by satisfy- ing some privacy requirements [1]. In this paper, we investigate privacy-preservation from the anonymity aspect. The sensitive information, even after the removal of identifying attributes, is still susceptible to linking attacks by the authorized users [2]. This problem has been studied extensively in the area of micro data publishing [3] and pri- vacy definitions, e.g., k-anonymity [2],

l-diversity [4], and variance diversity [5]. Anonymization algorithms use sup- pression and generalization of records to satisfy privacy requirements with minimal distortion of micro data.

Proposed System

The existing methods focus on a universal approach that exerts the same amount of preservation for all persons, without catering for their concrete needs. The consequence is that the system may be offering insufficient protection to a subset of people, while applying excessive privacy control to another subset. Contributed technique performs the minimum generalization for satisfying everybody's requirements, and thus, retains the largest amount of information from the microdata. The core of the solution is the concept of ***application specific anonymity,*** i.e., a Administrator can specify the degree of privacy protection for her/his sensitive values.

Predicate Evaluation and Imprecision

In this module, the query predicate evaluation semantics have been discussed. For query predicate evaluation over a table, say T, a tuple is included in the result if all the attribute values satisfy the query predicate. Here, proposed system only considers conjunctive queries (The disjunctive queries can be expressed as a union of conjunctive queries), where each query can be expressed as a d-dimensional hyper-rectangle. The semantics for query evaluation on an anonymized table T_ needs to be defined. When the equivalence class partition (Each equivalence class can be represented as a d-dimensional hyper-rectangle) is fully enclosed inside the query region, all tuples in the equivalence class are part of the query result. Uncertainty in query evaluation arises when a partition overlaps the query region but is not fully enclosed.

Anonymization with imprecision bounds:

For anonymization the proposed system selects the, quasi identifier Attributes, e.g., gender, zip code, birth date, that can potentially identify an individual based on other information available to an adversary. QI attributes are generalized to satisfy the anonymity requirements. And selects the corresponding selected attribute vector values, and set the upper bounds and the lower bound for the corresponding vector values for the selected attribute.

Accuracy-Constrained Privacy-Preserving Access Control

The exact tuple values in a relation are replaced by the generalized values after the anonymization. In this case, access control enforcement over the generalized data needs to be defined.

1. Relaxed. Use overlap semantics to allow access to all partitions that are overlapping the permission.
2. Strict. Use enclosed semantics to allow access to only those partitions that are fully enclosed by the permission.

Both schemes have their own pros and cons. Relaxed enforcement violates the authorization predicate by giving access to extra tuples but is beneficial for applications where low cost of a false alarm is tolerable as compared to the risk associated with a missed event. Examples include epidemic surveillance and airport security. On the other hand, strict enforcement is suitable for applications where a high risk is associated with a false alarm as compared to the cost of a missed event. An example is a false arrest in case of shoplifting. In this system, the focus is on relaxed enforcement. However the proposed methods for anonymization are also valid for strict enforcement because the proposed heuristics reduce the overlap between partitions and queries. It can further assume that under relaxed enforcement if the imprecision bound is violated for a permission then that permission is not assigned to any role.

Heuristics for partitioning

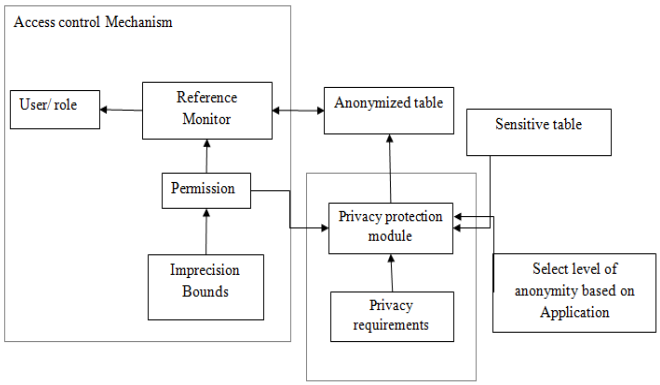
Starting with the whole tuple space the nodes in the kd-tree are recursively divided till the partition size is between k and 2k. The leaf nodes of the kd-tree are the output partitions that are mapped to equivalence classes in the given table**.** In the partitions are split along the median. Consider a partition that overlaps a query. If the median also falls inside the query then even after splitting the partition, the imprecision for that query will not change as both the new partitions still overlap the query.

In this heuristic, the proposed system proposes to split the partition along the query cut and then choose the dimension along which the Imprecision is minimum for all queries. If multiple queries overlap a partition, then the query to be used for the cut needs to be selected. The queries having imprecision greater than zero for the partition are sorted based on the imprecision bound and the query with minimum imprecision bound is selected. The intuition behind this decision is that the queries with smaller bounds have lower tolerance for error and such a partition split ensures the decrease in imprecision for the query with the smallest imprecision bound. If no feasible cut satisfying the privacy requirement is found, then the next query in the sorted list is used to check for partition split. If none of the queries allow partition split, then that partition is split along the median and the resulting partitions are added to the output after compaction.

Improving the number of Queries satisfying the imprecision bound

In module, the query imprecision slack is defined as the difference between the query bound and query imprecision. This query imprecision slack can help satisfy queries that violate the bounds by only a small margin by increasing the imprecision of the queries having more slack. The margin by which queries violate the bounds .In this repartitioning step, It considers only the first two groups of queries that fall within 10 percent and 10-25 percent of the bound only and these queries are added to the Candidate Query set (CQ), while all queries satisfying the bounds are added to the query set SQ. The output partitions are all the leaf nodes in the kd-tree. For repartitioning, it only considers those pairs of partitions from the output that are siblings in the kd-tree and have imprecision greater than zero for the queries in the candidate query set.

Architecture diagram of proposed system



Related work

Gabriel Ghinita [4] Data anonymization does not constrain the privacy of information **. Surajit Chaudhuri [2]** Data privacy is inadequate for supporting data. **Ninghui Li [5]** Privacy-preserving microdata does not provide access control for different roles. **Elisa Bertino [1]** Role-based access control does not tells about scalability. **Shariq Rizvi [7]** Fine Grained Access Control does not support privacy related database.

Conclusion

The proposed system proposes an accuracy-constrained privacy-preserving access control framework for relational data has been proposed. The framework is a combination of access control and privacy protection mechanisms. The access control mechanism allows only authorized query predicates on sensitive data. The privacy preserving module anonymizes the data to meet privacy requirements and imprecision constraints on predicates set by the access control mechanism. The proposed system proposes the application specific anonymization.

Future Work

The proposed system plan to extend the proposed privacy-preserving access control to incremental data and cell level access control**.**

REFERENCES

[1] Bertino E. and Sandhu .(2005),"Database Security-ConceptsApproaches, and allenges,"IEEE Trans.Dependable and Secure Computing, vol. 2, no. 1, pp. 2-19.

[2] Chaudhuri S. et al (2011), "Database Access Control & Privacy: Is There a Common Ground?" Proc. Fifth Bien- nial Conf. Innovative Data Systems Research (CIDR), pp. 96-103.

[3] Fung B. et al (2010), "Privacy-Preserving Data Publishing: A Survey of Recent evelopments," ACM Computing Surveys, vol. 42, no. 4, article 14, 2010.

[4] Ghinita G. et al (2009),"A Framework for Efficient Data Anonymization Under Privacy and Accuracy Constraints,"ACM Trans. Database Systems, vol. 34, no. 2, article 9.

[5] Li N. et al (2011), "Provably Private Data Anonymiza- tion: Or, k-Anonymity Meets Differential Privacy," Arxiv preprint arXiv:1101.2604.

[6] LeFevre K. et al (2008), "Workload-Aware Anonymization Techniques for Large-Scale Datasets," ACM Trans. Database Systems, vol. 33, no. 3, pp. 1-47.

[7] Rizvi S. et al (2004), "Extending Query Rewriting Techniques for Fine-Grained Access Control," Proc. ACM SIGMOD Int'l Conf. Management of Data, pp. 551-562.

[8] Zahid Pervaiz and Walid G. Aref (2014), "Accuracy - Constrained Privacy-Preserving Access Control Mechanism for Relational Data" IEEE Transactions On Knowledge And Data Engineering, Vol. 26, No. 4.