International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

Access control using k-anonymity on Sensitive Information by query evaluation unauthorized users outsourced database (ACM) Role-predicated access control (PPM)

K.V.L Prasanthi[1]& Ms.Veena Trivedi[2]

[1]M Tech (S.E), Dept of IT, Gokaraju Rangaraju Institute of Engineering and Technology, Mail Id: - prasanthi.kvl@gmail.com

[2]Associate Professor, Dept of IT, Gokaraju Rangaraju Institute of Engineering and Technology, Mail Id: - veenatrivedi@hotmail.com

Abstract

Micro data refers to series of records, each record with information on an individual unit like a patient or an organization. Access Control Mechanisms (ACM) bulwarks the sensitive information from unauthorized users. Even sanctioned users may misuse the data to reveal the privacy of individuals to whom the data refers to. Privacy Auspice Mechanism (PPM) anonymize the relational data to avert identity and attribute disclosure. It is achieved by generalization or suppression. Role-predicated access control gives users the sanctions to access the data predicated on their roles. The access control policies define cull predicates available to roles while the privacy requisite is to satiate the k-anonymity or l-diversity. Imprecision bound constraint is assigned for each cull predicate. Top Down Cull Mondrian (TDSM) algorithm is utilized for query workload-predicated anonymization algorithm is constructed utilizing acquisitive heuristics and kd-tree model. Query cuts are culled with minimum bounds in Top-Down Heuristic1 algorithm (TDH1). The query bounds are updated as the partitions are integrated to the output in Top-Down Heuristic 2 algorithm (TDH2). The cost of reduced precision in the query results is utilized in Top-Down Heuristic 3 algorithm (TDH3). Repartitioning algorithm is utilized to reduce the total imprecision for the queries. The privacy preserved access control framework is enhanced to provide incremental mining features utilizing R+-tree. Data insert, expunge and update operations are associated with the partition management mechanism.

Keywords: Access control; k-anonymity; Sensitive Information; query evaluation; unauthorized users; outsourced database; Access Control Mechanisms (ACM); Role-predicated access control**;** Privacy Auspice Mechanism; (PPM); Anonymization

1. Introduction

Current ecumenically networked society greatly demands the sharing and propagation of information. While information relinquished in the past was in tabular and precompiled statistical form (macro data), there is a desideratum for the relinquishment for categorical data to perform statistical analysis on them (micro data). Micro data refers to series of records, each with information on an individual unit like a person or an industrial unit. It sanctions the recipient to perform a whole incipient analysis on them as needed. In order to bulwark the identity of individuals to whom the data refers to, when relinquishing micro data, data holders often abstract or encrypt explicit identifiers, such as names and gregarious security numbers.

Organizations accumulate and analyze consumer data to amend their accommodations. Access Control Mechanisms (ACM) is acclimated to ascertain that only sanctioned

information is available to users. However, sensitive information can still be misused by sanctioned users to compromise the privacy of consumers. The concept of privacy-preservation for sensitive data can require the enforcement of privacy policies or the bulwark against identity disclosure by gratifying some privacy requisites [1]. In this paper, we investigate privacy-preservation from the anonymity aspect. The sensitive information, even after the abstraction of identifying attributes, is still susceptible to linking attacks by the sanctioned users [2]. This quandary has been studied extensively in the area of micro data publishing [3] and privacy definitions, e.g., k-anonymity [2].

The other quasi-identifiers reveal the privacy. The "linking attack" [4] should be managed to secure privacy of individuals. These linking attacks can be managed by anonymizing data in tables. Anonymization is the process of abstracting the identity particulars by obnubilating or transmuting the data. An in nominate table is the one which is composed after transmuting the data that does not distinguish the individual characteristics. There are sundry anonymization methods that avail in holding privacy.

2. Related Work

Existing System:

ORGANIZATIONS amass and analyze consumer data to ameliorate their accommodations. Access Control Mechanisms (ACM) are habituated to ascertain that only sanctioned information is available to users. However, sensitive information can still be misused by sanctioned users to compromise the privacy of consumers. The concept of privacy-preservation for sensitive data can require the enforcement of privacy policies or the auspice against identity disclosure by gratifying some privacy requisites. Subsisting workload cognizant anonymization techniques minimize the imprecision aggregate for all queries and the imprecision integrated to each sanction/query in the anonymized micro data is not kenned. Making the privacy requisite more stringent (e.g., incrementing the value of k or l) results in adscititious imprecision for queries.

Disadvantages:

- There is no privacy for users.
- The sensitive information, even after the removal of identifying attributes, is still susceptible to linking attacks by the authorized users.

Proposed System:

The heuristics proposed in this paper for precision-constrained privacy-preserving access control are additionally pertinent in the context of workload-cognizant anonymization. The anonymization for perpetual data publishing has been studied in literature. In this paper the focus is on a static relational table that is anonymized only once. To exemplify our approach, role-predicated access control is surmised. However, the concept of precision constraints for sanctions can be applied to any privacy-preserving security policy, e.g., discretionary access control.

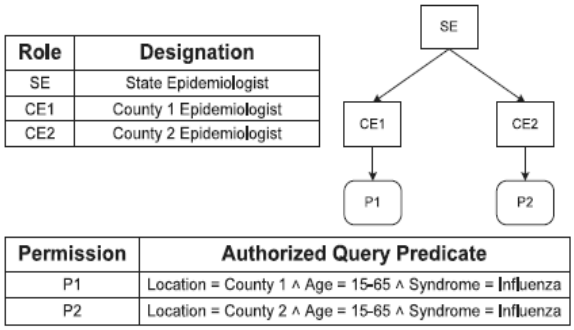
Advantages:

- Accuracy constrained privacy preserving access.
- It's maintained data's in secure manner.

3. Implementation

Access control policy:

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org



Role	Designation
SE	State Epidemiologist
CE1	County 1 Epidemiologist
CE2	County 2 Epidemiologist

Permission	Authorized Query Predicate
P1	Location = County 1 ∧ Age = 15-65 ∧ Syndrome = Influenza
P2	Location = County 2 ∧ Age = 15-65 ∧ Syndrome = Influenza

Fig 1: Access control policy.

Syndromic surveillance systems are utilized at the state and federal levels to detect and monitor threats to public health. The department of health in a state accumulates the emergency department data (age, gender, location, time of advent, symptoms, etc.) from county hospitals circadianly. Generally, each daily update consists of a static instance that is relegated into syndrome categories by the department of health. Then, the surveillance data is anonymized and shared with departments of health at each county. An access control policy is given in Fig. 1 that sanctions the roles to access the tuples under the sanctioned predicate, e.g., Role CE1 can access tuples under PermissionP1. The epidemiologists at the state and county level suggest community containment measures, e.g., isolation or quarantine according to the number of persons infected in case of a flu outbreak. According to the population density in a county, an epidemiologist can advise isolation if the number of persons reported with influenza are more preponderant than 1,000 and quarantine if that number is more preponderant than 3,000 in a single day. The anonymization integrates imprecision to the query results and the imprecision bound for each query ascertains that the results are within the tolerance required. If the imprecision bounds are not slaked then nonessential mendacious alarms are engendered due to the high rate of erroneous positives.

Anonymity:

Age	Gender	Zip	Disease
22	Male	500016	Fever
33	Female	500038	Cold
44	Male	500016	Fever
55	Female	500038	Cancer
10	Male	500012	Flu

Age	Gender	Zip	Disease
1-30	Male-2	500016-1, 500012-1	Fever
31-60	Male-1 Female-2	500038-2, 500016-1	Cold
31-60	Male-1 Female-2	500038-2, 500016-1	Fever
31-60	Male-1 Female-2	500038-2, 500016-1	Cancer
1-30	Male-2	500016-1, 500012-1	Flu

Fig 2: Sensitive & Anonymity tables.

Anonymity is prone to homogeneity attacks when the sensitive value for all the tuples in an equipollence class is equipollent. To contravene this shortcoming, l-diversity has been proposed and requires that each equipollence Fig. 1. Access control policy. Class of T contains at least l distinct values of the sensitive attribute. For sensitive numeric attributes, an l-diverse parity class can still leak information if the numeric values are proximate to each other. For such cases, variance diversity has been proposed that requires the variance of each parity class to be more preponderant than a given variance diversity parameter. The table in Fig. 2a does not slake k-anonymity because kenning the age and zip code of a person sanctions associating a disease to that person. The table in Fig. 2b is a 2-innominate and 2-diverse version of table in Fig. 2a. The ID attribute is abstracted in the anonymized table and is shown only for identification of tulles. Here, for any accumulation of cull predicates on the zip code and age attributes, there are at least two tuples in each parity class.

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

Accuracy-Constrained Privacy-Preserving Access Control:

A precision-constrained privacy-preserving access control mechanism. (Arrows represent the direction of information flow), is proposed. The privacy bulwark mechanism ascertains that the privacy and precision goals are met afore the sensitive data is available to the access control mechanism. The sanctions in the access control policy are predicated on cull predicates on the QI attributes. The policy administrator defines the sanctions along with the imprecision bound for each sanction/query, utilizer-to-role assignments, and role-to sanction assignments. The designation of the imprecision bound ascertains that the sanctioned data has the desired level of precision. The imprecision bound information is not shared with the users because kenning the imprecision bound can result in infringing the Privacy requisite. The privacy aegis mechanism is required to meet the privacy requisite along with the imprecision bound for each sanction.

Top-Down Heuristic:

In TDSM, the partitions are split along the median. Consider a partition that overlaps a query. If the median withal falls inside the query then even after splitting the partition, the imprecision for that query will not transmute as both the incipient partitions still overlap the query as illustrated. In this heuristic, we propose to split the partition along the query cut and then optate the dimension along which the imprecision is minimum for all queries. If multiple queries overlap a partition, then the query to be utilized for the cut needs to be culled. The queries having imprecision more preponderant than zero for the partition are sorted predicated on the imprecision bound and the query with minimum imprecision bound is culled. The intuition abaft this decision is that the queries with more diminutive bounds have lower tolerance for error and such a partition split ascertains the defragmentation in imprecision for the query with the most diminutive imprecision bound. If no feasible cut slaking the privacy requisite is found, then the next query in the sorted list is utilized to check for partition split. If none of the queries sanction partition split, then that partition is split along the median and the resulting partitions are integrated to the output after compaction.

4. Experimental results

Fig 3: Administrator home Page.

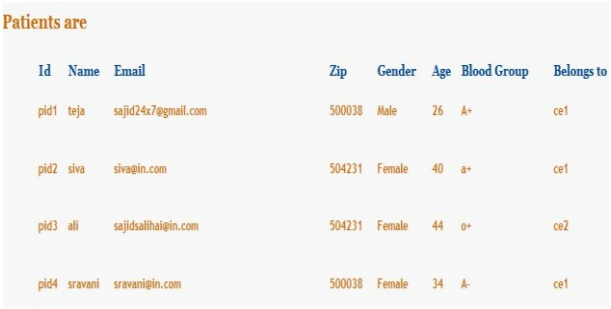


Fig 4: Patients sensitive data.

Fig 5: Sensitive Data.

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

P.Id	Name	Email	Zip	Gender	Age	Disease	RBAC
pid1	teja	sajid24x7@gmail.com	500038-1; 504231-1;	Female-1; Male-1;	1-30	cold	ce1
pid2	siva	siva@in.com	500038-1; 504231-1;	Female-1; Male-1;	1-30	cold	ce1
pid3	ali	sajidsalihal@in.com	500038-1; 504231-1;	Female-1; Male-1;	1-30	fever	ce2
pid4	sravani	sravani@in.com	500038-1; 504231-1;	Female-1; Male-1;	31-60	fever	ce1
pid5	sajid	cloudtechnologiesprojects@gmail.com	500038-1; 504231-1;	Female-1; Male-1;	1-30	Headach	ce2
pid6	swamy	swamy123@in.com	500038-1; 504231-1;	Female-1; Male-1;	31-60	Cancer	ce1

Fig 6: Data Anonymization Result Page.

5. Conclusion

Access control mechanism for relational data is constructed with the privacy preservation predicated model. Role Predicated Access Control (RBAC) scheme provides security to the data by sanctioning access predicated on sanctions. K-Anonymity model is integrated with minimum imprecision predicated data access control mechanism. Partitioning utilizing R+-trees results in less number of overlapping partitions. Hence precision is ameliorated and time involution is reduced in the system. Privacy preserved data access control mechanism is ameliorated with incremental mining model. The system reduces the imprecision rate in query processing. Access control mechanism is acclimated for incremental mining model.

6. References

[1] S. Chaudhuri and Sudarshan, "Fine Grained Authorization through Predicated Grants," Proc. IEEE 23rd Int'l Conf. Data Eng.,, 2007.

[2]R.Agrawal,P.Bird,T.Grandison,J.Kiernan,S.Logan and W.Rjaibi,"Extending Relational Database Systems to automatically Enforce Privacy Policies,"*Proc.21st Int'l Conf. Data Eng.,*pp.1013-1022,2005.

[3] S. Chaudhuri, Kaushik and R. Ramamurthy, "Database Access Control & Privacy: Is There a Common Ground?" Proc. Fifth Biennial Conf. Innovative Data Systems Research, 2011.

[4] G. Ghinita, P. Karras, P. Kalnis and N. Mamoulis, "Fast Data Anonymization with Low Information Loss," Proc. 33rd Int'l Conf. Very Large Data Bases, pp. 758-769, 2007.

[5] N.Li,W.Qardaji, and D.Su,"Provably Private Data Anonymization:Or,k-Anonymity Meets Differential Privacy," *Arxio preprintarXiv:1101.2604*,2011.

[6] X. Xiao, G. Bender, M. Hay and J. Gehrke, "Ireduct: Differential Privacy with Reduced Relative Errors," Proc. ACM SIGMOD Int'l Conf. Management of Data, 2011.

[7] Zahid Pervaiz,Walid G.Aref,Arif Ghafoor,Nagabhushana Prabhu,"Accuracy-constrained Privacy Preserving Access Control Mechanism for Relational Data,"IEEE Trans.Knowledge and Data Engineering,vol.26,no.4,pp. 795-807,2014.

[8] K. LeFevre, D. DeWitt and R. Ramakrishnan, "Workload-Aware Anonymization Techniques for Large-Scale Datasets," ACM Trans. Database Systems, vol. 33, no. 3, pp. 1-47, 2008.