ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

An Effective privacy preserving and Accuracy Constrained Method for Relational Data

Dr.M.Sreedevi[1], P.Chytanya lakshmi[2]

Associate Professor, Department of computer science, MITS

Department of computers, MITS

Abstract: Now-a-days, the usage of www has increased and huge amount of data is getting generated and managed in every second, 20-30% of total volume of data contains sensitive information which needs to be managed with data security mechanism. The data should be exposed through proper access control mechanism. To manage sensitive data more over many Privacy Protection mechanisms are implemented which might not claim complete durability every time. In this situation, to pass proper authentication user needs to share different identity disclosure parameters which also are maintained as generalized data. Privacy protection method use suppression and generalization of relational data to anonymized against attribute disclosure. Normally below processes are used to manage these requirements k-anonymity and l-diversity against identity and attribute disclosure. Anyway the data security and privacy is managed based on only the authorized information. The access control paradigm follows k-anonymity and l-diversity satisfying the data privacy. Imprecision bound validation by PPM should pass to manage privacy control. Technique of workload-aware anonymization is one concern for this assignment, but as multiple role based privacy protection does not bring any significance with this approach.

Keywords: Access control, privacy, k-anonymity, privacy preservation.

I. INTRODUCTION

Organizations collect and analyze consumer data to improve their services. Access Control Mechanisms 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 anonymity techniques can be used with an access control mechanism to ensure both security and privacy of the sensitive information. The anonymization for continuous data publishing has been studied in literature.

Anonymization algorithms use suppression and generalization of records to satisfy privacy requirements with minimal distortion of micro data. The anonymity techniques can be used with an access control mechanism to ensure both security and privacy of the sensitive information. In this system the focus is on a static relational table that is anonymized only once. However, the concept of accuracy constraints for permissions can be applied to any privacy-preserving security policy, e.g., discretionary access control.

II.EXISTING SYSTEM

Anonymity method preserves privacy of individuals against identity disclosure attack alone. But Attribute disclosure attack makes compromise this method. Limitation of k-anonymity is fulfilled through l-diversity method.

DISADVANTAGES OF EXISTING SYSTEM:

To ensure that privacy of no individual is being put in danger due to the released data by protecting released information against inference and linking attacks.

III.PROPOSED SYSTEM

The efficiency of t-closeness method is better than l-diversity and K-anonymity. But the complexity of Computation is more than other proposed methods. A new anonymization based method is proposed for preserving the security of sensitive attribute values against identity and attribute disclosure attacks.

ADVANTAGES OF PROPOSED SYSTEM:

- In the proposed method providing the privacy of individuals sensitive information from attribute attack and identity disclosure attacks.
- In this paper propose one reduce the information loss occur by using anonymity techniques and providing privacy in new mode.

WRONG APPROACH:

Removing the unique identifiers such as Name, Employee Id from a table cannot guarantee privacy. Other attributes like Date of Birth, Sex, PIN Code when combined together can also reveal the identity of an individual. Re-identification is possible by using a set of attributes and another database containing the same set of attributes. Sometimes this approach can also leak sensitive information about an individual

To overcome this we propose a privacy based anonymization algorithm:

In existing anonymity techniques have drawbacks like attribute disclosure and identity disclosure and also no privacy for individual attributes in a Database. The proposed method overcome the problems and provide privacy for individual and reduce the limitations of anonymity techniques with privacy in new mode.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

Proposed method performs generalization operation with suppression technique and generate a anonymity table T. The suppression technique is applied over selected Quasi identifiers QI. After this technique the records in anonymity table T are placed and sort in groups G1,G2,G3,G4…..Gn. In each group is ordered by suppressed value of QI. Among Quasi identifiers QI, one with more different is selected and find the nearby numeral value Li and the next maximum Numeral value Mi in group.

Step 1: Place the records in a group G1,G2,G3,G4…….Gn in the table T by value si of Quasi Identifier Bi. where i= 1, 2, … k.
Step 2: Repeat steps 3 to 5 variable j from 1 to k
Step 3: Let Lj = qj.
Catch the nearby numeral value Lj less than qj in group Gi where i=1,2,3,…n and if found, Let Lj = qj
Step 4: Let Mj = qj.
Catch the nearby numeral Mj greater than qi in group Gi where i=1,2,3,…n and if found, Let Lj = qj
Step 5: If Lj and Mj are found in the same group Gm, the generalization condition is set as set qj = Lj<=Mj
Step 6: If Lj and Mj are not found in the same group Gm, the generalization condition is set asset qj= <=Qj
Output: Anonymized table T*

The performance of proposed privacy algorithm calculated in terms of two data metrics namely information loss and privacy gain. performance analysis measure the information loss.By using this below formula.

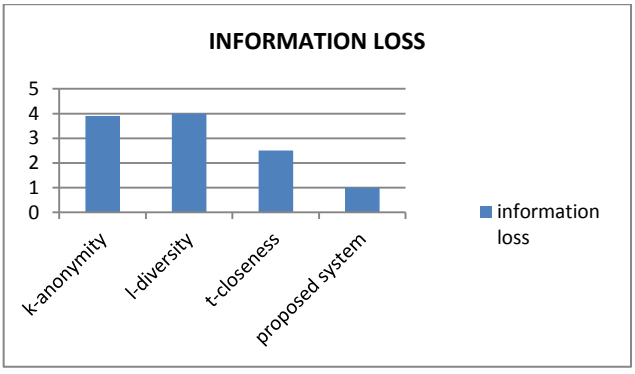
$$ILOSS(v_g) = \frac{|v_g| - 1}{|D_A|}$$

Where
|Vg| is the number of domian values of Vg.
|Da| is the number of domian values in the attribute A of Vg.
ILOSS(Vg)=0 if Vg is an original data value ina table.

The Principle of information/privacy trade-off can also be used to select a generalization g, in the which case it will minimize.

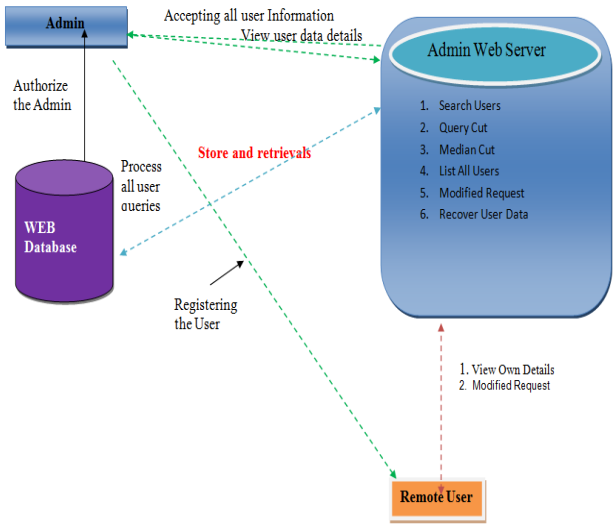
$$ILPG = \frac{IL(g)}{PG(g)}$$

Where *IL*(g) denotes the information loss and *PG*(g) denotes the privacy gain by performing.



IV.IMPLEMENTATION

ARCHITECTURE DIAGRAM



MODULES:

- **Admin**
  In this module, the Admin has to login by using valid user name and password. After login successful he can do some operations such as search users, query cut, median cut, list users, view attackers, data recovery and logout.

Search Users
In this module, the admin can view the two types of data first one is sensitive data and second one is anonymous data. The sensitive data means we can view the particular disease, pin code, age and Id. The anonymous data means we can view the diseases between ages (eg: 0-10) and pin codes (eg: 40-60). In this system we are hiding the information about patient details and showing the anonymous records about patient.

Query Cut
In this module, the admin can search the diseases details based on the key words such as enter age and enter disease name, then server will search the details related to key words, then response will send to particular user.

Median cut
In this module, the admin can search the diseases based on the age and blood group, then server will mine the all data and send the related data to particular user.

List of users
In this module, the Admin can view list of all users. If the admin clicks on users button, then it will show all registered users with their tags such as user ID, user name, blood group, diseases, E mail ID, mobile no, Location, date of birth, address and pin code

Modify Request
In this module, Admin can check the updated details of each single user and modify the details if it is unauthorized.

Approve Nominee
In this module, the register nominee acceptance should be approval by admin and user acceptance. admin generate a unique password for each nominee.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

Data Recovery
In this module, the admin will recover the modified data. After attacking a data the admin will recover the attacked data and again upload to the database.

- **User**
  In this module, User should register before doing some operations. After registration successful he has to login by using authorized user name and password. user will do some operations like attack user details, view my details and logout. If user clicks on my details button, then the server will give response to the user with their tags such as user ID, name, mobile no, address, pin code and email ID. The user want attack the particular user information, then click on attack user details button, then enter user name to attack and submit. The server will display the user details, and then you can edit the user information, submit and server will give response to user. After modifying a data, the user will be considered as an attacker. The attacker details will be stored in an attacker module.

- **Query & Median cut age limit Result**
  In this module, we can view the Query and Median cut results for different age user. This graph will increase based on the age limit and diseases. 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.

V.CONCLUSION AND FUTURE WORK

Data sharing and publishing are increasing in every day. The usage of www has increased and huge amount of data is getting generated and managed in every second, 20-30% of total volume of data contains sensitive information which needs to be managed with data security mechanism. In this situation, to pass proper authentication user needs to share different identity disclosure parameters which also are maintained as generalized data.

In this paper a proposed method which provides the privacy and reduce the information loss. The proposed method is achieved the preservation of individuals of sensitive information in anonymized Database. future work is to find the more anonymity techniques to reduce the loss of information and solution for preserving security of non-numerical Quasi identifers.

REFERENCES

[1] L. Sweeney, "k-Anonymity: A Model for Protecting Privacy," Int'l J. Uncertainty, Fuzziness and Knowledge Based Systems, vol. 10, no. 5, pp. 557-570, 2002.
[2] A.Machanavajjhala, J.Gehrke, et al., *ℓ-diversity: Privacy beyond k-anonymity*, In Proc. of ICDE, Apr.2006.B
[3] N. Li, T. Li, and S. Venkatasubramanian, *t-Closeness: Privacy Beyond k-anonymity and l-Diversity*, In Proc. Of ICDE, 2007, pp. 106-115.
[4] E. Bertino and R. Sandhu, "Database Security-Concepts, Approaches, and Challenges," IEEE Trans. Dependable and Secure Computing, vol. 2, no. 1, pp. 2-19, Jan.-Mar. 2005.
[5] P. Samarati, "Protecting Respondents' Identities in Microdata Release," IEEE Trans. Knowledge and Data Eng., vol. 13, no. 6, pp. 1010-1027, Nov. 2001.
[6] B. Fung, K. Wang, R. Chen, and P. Yu, "Privacy-Preserving Data Publishing: A Survey of Recent Developments," ACM Computing Surveys, vol. 42, no. 4, article 14, 2010.
[7] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "L-Diversity: Privacy Beyond k-anonymity," ACM Trans. Knowledge Discovery from Data, vol. 1, no. 1, article 3, 2007.
[8] Qiang Wang,Zhiwei Xu and Shengzhi Qu,(2011) "An Enhanced K-Anonymity Model against Homogeneity Attack", Journal of software, Vol. 6, No.10, October 2011;1945-1952