International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249 – 8958, Volume-3, Issue-2, December 2013

Secure Anonymization for Privacy Measure

Ishwarya M.V, K.Ramesh Kumar

AbstractIn this paper we are going to discuss about the privacy preservation of Patients details in a medical centre . The medical centre may have various login for various people like Administrator,Doctor,Analyst and Receptionist .We design a model such that the patients entire details are not known to everyone who logins with their id.It is available in a suppressed form to each and everyone who logs in .All the patients data are being split using Slicing algorithm and shuffled and stored in different databases in encryption side.The data are realigned and deshuffled and the original data are retrieved in the decryption side.

Keywords: Privacy preservation, Authentication, Security, Slicing, Shuffling

I. INTRODUCTION

Authentication is a process of determining whether a particular individual or a device should be allowed to access a system or an application or merely an object running in a device. This is an important process which assures the basic security goals, viz. confidentiality and integrity. Also, adequate authentication is the first line of defense for protecting any resource. It is important that the same authentication technique may not be used in every scenario. For example, a less sophisticated approach may be used or accessing a "chat server" compared to accessing a corporate database. Most of the existing authentication schemes require processing both at the client and the server end. Thus, the acceptability of any authentication scheme greatly depends on its robustness against attacks as well as its resource requirement both at the client and at the server end.The resource requirement has become a major factor due to the proliferation of mobile and hand-held devices. Nowadays with the use of mobile phones, users can access any information including banking and corporate database. In this paper, we specifically target the mobile banking domain and propose a new and intelligent authentication scheme. However, our proposal can also be used in other domains where confidentiality and integrity are the major securityrequirement

II. PROPOSED SYSTEM

In this paper we are going to discuss about the privacy preservation of Patients in a medical centre . The medical centre may have various login for various people like Administrator,Doctor, Analyst and Receptionist .We are going to design a model such that the patients entire details are not known to everyone who logins with their id.It is available in a suppressed form to each and everyone who logs in .

The Receptionist can login and can register the new Patients details .If any need he can make an enquiry about the patient,but cannot view his disease history .

If a Doctor logins the home page, he can view the Patients details .He can enter the patients details and renew the changes if the patients disease has any changes.

If an administrator logins the home page,he can View the medical centre's Employee details.He can register new Employee details and can view the patient details and search any patient for details.

If any Analyst logins the home page,he can view the entire report of the patient and the data and available in suppressed form.

All the patients data are being split using Slicing algorithm and shuffled and stored in different databases in encryption side.The data are realigned and deshuffled and the original data are retrieved in the decryption side.

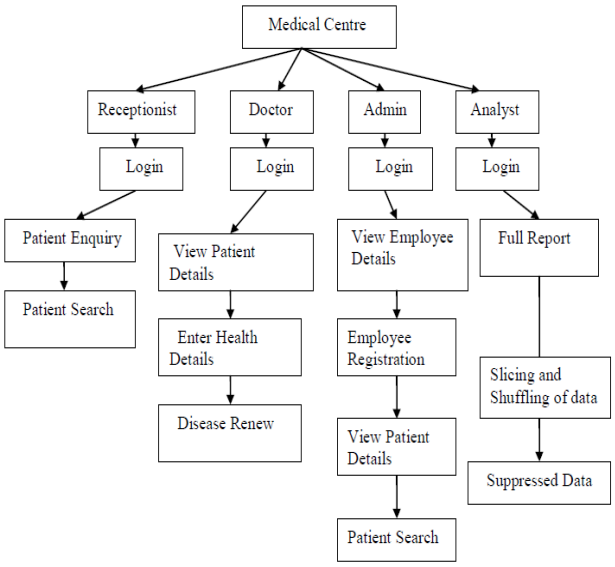

FIG 1 .Architechture Diagram of Proposed System

III. SLICING ALGORITHM:

The algorithm consists of three components:

1) Choosing a dimension on which to partition;
   Find Number of rows in patient_enq
2) Choosing a value to split; and
   //start Suppression

Here we suppress using zipcode. This zipcode is having 5 digits like 47983.

The variable inc is the value to split

If we set inc=4;

The zipcode is displayed as first 4 digit numbers like 4798**.

And we set threshold value t=0.5F;

And n is the second highest value of table age_count according to patients age in that table.

For example our patient table containing this data,

age	Count
2*	6
3*	4
4*	10

Manuscript received December, 2013

Ishwarya M.V, Research Scholar,HITS,Padur,Assistant Professor,Sri Sairam Engg.College,

Dr K.Ramesh Kumar , Associate Professor,HITS,Padur

We simply get the second highest value n=6;

3) Checking if the partitioning violates the privacy requirement.

After that we check this following calculation.

i.e.

$$\frac{\text{(row count)}}{n} <= t$$

If Each row of our table satisfied this condition, our privacy requirement is satisfied

Else

We decrement our inc value and again we check this condition satisfied by each row or not till condition is satisfied.

IV. SLICING ALGORITHM

input: P is partitioned into r partitions $\{P_1, P_2, ..., P_r\}$
output: true if (n,t)-closeness is satisfied, false otherwise

1. **for** every P_i
2. **if** P_i contains less than n records
3. find=false
4. **for** every $Q \in Parent(P)$ and $|Q| \geq n$
5. **if** $D[P_i, Q] \leq t$, find=true
6. **if** find==false, **return** false
7. **return** *true*

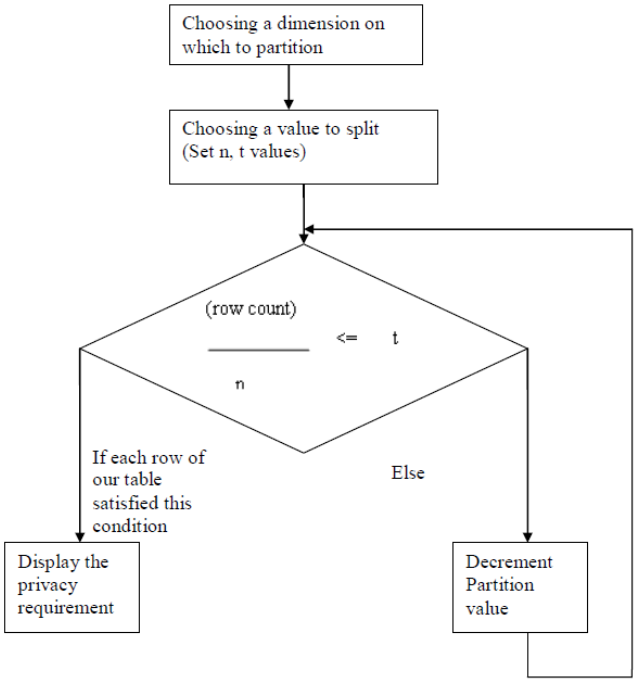

Fig .2 Flowchart For Slicing The Data

V. CONCLUSION

Thus the data is secure and the data is available in an Unknown format to the unauthorized users.Thus we can improve the data security and this concept can be used in any monitoring system in any model for privacy preservation of data.

REFERENCES

[1] C. Aggarwal, "On k-Anonymity and the Curse of Dimensionality," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 901-909, 2005.

[2] G. Aggarwal, T. Feder, K. Kenthapadi, S. Khuller, R. Panigrahy, D.Thomas, and A. Zhu, "Achieving Anonymity via Clustering," Proc. ACM Symp. Principles of Database Systems (PODS), pp. 153-162, 2006.

[3] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin, Network Flows: Theory, Algorithms, and Applications. Prentice-Hall, Inc., 1993.

[4] R.J. Bayardo and R. Agrawal, "Data Privacy through Optimal k-Anonymization," Proc. Int'l Conf. Data Eng. (ICDE), pp. 217-228, 2005.

[5] F. Bacchus, A. Grove, J.Y. Halpern, and D. Koller, "From Statistics to Beliefs," Proc. Nat'l Conf. Artificial Intelligence (AAAI), pp. 602- 608, 1992.

[6] J.-W. Byun, Y. Sohn, E. Bertino, and N. Li, "Secure Anonymization for Incremental Datasets," Proc. VLDB Workshop Secure Data Management (SDM), pp. 48-63, 2006.

[7] B.-C. Chen, K. LeFevre, and R. Ramakrishnan, "Privacy Skyline: Privacy with Multidimensional Adversarial Knowledge," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 770- 781, 2007.

[8] G.T. Duncan and D. Lambert, "Disclosure-Limited Data Dissemination," J. Am. Statistical Assoc., vol. 81, pp. 10-28, 1986. LI ET AL.: CLOSENESS: A NEW PRIVACY MEASURE FOR DATA PUBLISHING 955

[9] B.C.M. Fung, K. Wang, and P.S. Yu, "Top-Down Specialization for Information and Privacy Preservation," Proc. Int'l Conf. Data Eng. (ICDE), pp. 205-216, 2005.

[10] C.R. Givens and R.M. Shortt, "A Class of Wasserstein Metrics for Probability Distributions," Michigan Math J., vol. 31, pp. 231-240, 1984.

[11] V.S. Iyengar, "Transforming Data to Satisfy Privacy Constraints," Proc. ACM SIGKDD, pp. 279-288, 2002.

[12] D. Kifer and J. Gehrke, "Injecting Utility into Anonymized Datasets," Proc. ACM SIGMOD, pp. 217-228, 2006.

[13] N. Koudas, D. Srivastava, T. Yu, and Q. Zhang, "Aggregate Query Answering on Anonymized Tables," Proc. Int'l Conf. Data Eng. (ICDE), pp. 116-125, 2007.

[14] S.L. Kullback and R.A. Leibler, "On Information and Sufficiency," Annals of Math. Statistics, vol. 22, pp. 79-86, 1951.

[15] D. Lambert, "Measures of Disclosure Risk and Harm," J. Official Statistics, vol. 9, pp. 313-331, 1993.