ISSN (Print) : 2320 – 9798
ISSN (Online) : 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 1, March 2013

PRIVACY CONSERVING APPROACH to ANONYMOUS DATABASE

Jaimin Marfatia[1], Nainish Modi[2], Niraj Lad[3], Vaishali Patel[4], Jignasa Patel[5]

Student, Dept. of Information Technology, Shri S'ad Vidhya Mandal Institute of Technology, Bharuch, India[1,2,3]

Professor, Dept. of Information Technology, Shri S'ad Vidhya Mandal Institute of Technology, Bharuch, India[4,5]

Abstract: **Suppose bank has some private data which has important details of every employee of the bank. Now bank wants to send these details to its head office for some specific purpose with a primary concern that privacy of each employee should not violate by disclosing his data to other people. If we permit each employee to add data directly into the database then database confidentiality will be broken and if we permit database owner to read every stuff of employee then privacy of employee will break. So to preserve privacy and confidentiality we have proposed approach named as suppression based method so as to maintain the privacy of the employee. The meaning of anonymity is to remove identifying entity from the database.**

Keywords: Anonymous, Confidentiality, Privacy, Suppression

I. INTRODUCTION

The amount of sensitive information about citizens accumulated in the databases of government agencies and private organizations, such as Social Security Administration, banks, and health care providers, has been increasing steadily in the past decades. While it has long been realized that there is a need to protect the information both in storage and transition, it has recently become apparent that the information needs to be properly guarded from unauthorized disclosure during the Process of testing newly developed applications that employ the databases.Now-a-days there is a great need of privacy of the users in the society. As we know that the use of computers is increasing in great amount, the need of privacy of each user and the confidentiality of the database are of the prime importance to the respective organization. There are large numbers of databases stored in the system and therefore by correlating these databases we can get private information of any specified user. Therefore, the database confidentiality is a big issue and here we have proposed a method by which we will maintain the privacy of each and every individual and simultaneously will devise a method to maintain the confidentiality [1].

Now let us stress up on *privacy* and *confidentiality.*Privacy is the data that can be safely shown to the valid owner without leaking the sensitive information from the database. Data confidentiality is the difficulty experienced by the third party to know any sensitive information stored in the database [1]. Let us get into the discussion of *anonymity.* Anonymization is the technique to preserve identifying or the private information of the user. There are many techniques to achieve anonymization but we will move to the technique named as *k-anonymization* technique. Data anonymization enables transferring information between two organizations, by converting text data into non-human readable form using encryption method. Also the non-anonymized version of the information which is stored at the sender side will be deleted after it is being sent to the receiver side. This is one of the most important concept of this technique [2]. This technique protects privacy of original data by modification. So problem arises at this point where database needs to be updated. So when tuple is to be inserted in the database problem occurs relating to privacy and confidentiality that is database owner decide that whether database preserve privacy without knowing what new tuple to be inserted. In this paper, we propose a protocol called Suppression based approach to solve the above problem.

It is also important to note that the work of maintaining the privacy of each user should be carried out by the organization that owns the database. It is therefore the thing of concern not only for the user to maintain his/her privacy but also for the organization who owns the database. According to the current ongoing rules and regulations, organizations collecting data about individuals are under the obligation of assuring individual privacy. It is thus in their interest to check that data that are entered in their databases do not violate privacy, and to perform such a verification without seeing any sensitive data of an individual [2].

II. RELATED WORK

There are various techniques like data perturbation, data reduction, query processing and SMC (Secure Multi-party Communication) which came into existence some time before and were told that these techniques provide confidentiality and privacy to anonymous database.The first technique deals with designing certain algorithms for

ISSN (Print) : 2320 – 9798
ISSN (Online) : 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 1, March 2013

achieving data anonymization. Researchers tried to find a way through data suppression and data perturbation for anonymity and they have also come out with some complex results with respect to k-anonymity. But the main problem is that these results do not deal with the updates that were made in the database which then results in the privacy break. Therefore none of this work resulted in achieving privacy.The second technique is related to Secure Multi-party Communication (SMC). It is widely known technique which is carried out and is very much investigated by the researchers in the field of cryptography.There are certain techniques which are available for performing computations securely. But these techniques are not that much efficient. Therefore it is needed to do more research in devising an efficient protocol so that the computations can be performed securely. Therefore we have proposed a protocol in this paper to solve the above mentioned problem and to gain some good amount of efficiency [3], [4].

The third technique is related with the fetching of private information which is nearer to the application of the SMC in the field of data management. Here the ideas of imposing expressive queries on the database without letting the database know the actual queries. Again we experience here the same problem of updating the database at the time each tuple or entry is inserted into the database.At last, the fourth technique is related to query processing for the encrypted data. There is no connection of this approach with the concept of k-anonymity since the objective of this technique is to encrypt the data. So here while encrypting the data, the original data is revealed in front of the database owner which is not at all accepted in our case as our main aim is to protect the privacy of the user [5].

III. PROPOSED TECHNOLOGY

In this paper we have proposed a protocol named *suppression protocol* that allows the owner of anonymous database to anonymise the tuple t, without disclosing the content of the tuple to the person who is managing the database. To achieve this goal two party exchange message by encrypting the data. Here the connection between the two parties will be anonymous so as to secure the connection. It is needed because someone can easily know the IP address and can get the sensitive information. Therefore it is mandatory to authenticate the user so that only the authorized user can have access to the data [6], [7].

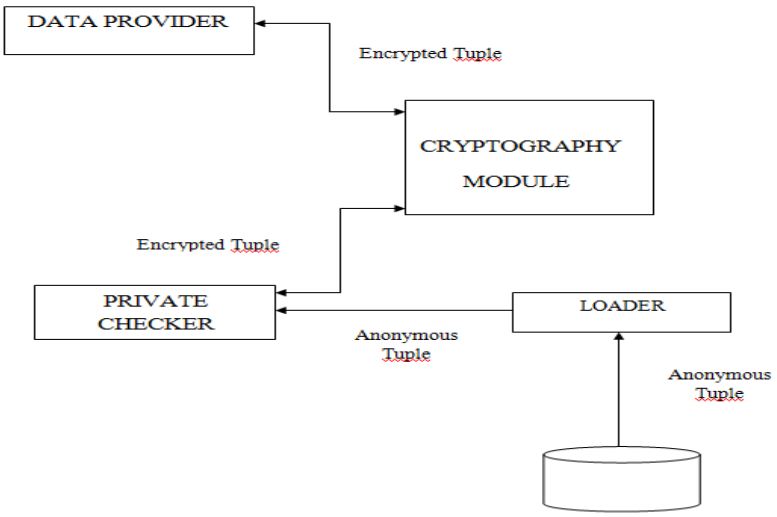

Fig.1 System Prototype

The information flow of fig. 1 across the above mentioned modules is as follows: after an initial setup phase in which the user and the Private Checker prototype exchange public values for correctly performing the subsequent cryptographic operations, the user sends the encryption of her/his tuple to the Private Checker; the loader module reads from the k-anonymous DB the first chunk of tuples to be checked with Encrypted tuple by the user. Such tuples are then encrypted by the crypto module. The checker module performs the check one tuple at time in collaboration with the user, according Protocol (in the case of suppression based anonymization). If none of the tuples in the chunk matches the User tuple, then the loader reads another chunk of tuples from the k-anonymous DB.

ISSN (Print) : 2320 – 9798
ISSN (Online) : 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 1, March 2013

Table 3.1 Original Dataset

AREA	DESIGNATION	SALARY
Account	Cashier	20000
Loan dept	Head	25000
Loan dept	Assistant	12000

Table 3.2 Suppressed Data with k =1

AREA	DESIGNATION	SALARY
Account	Cashier	*
Loan dept	Head	*
Loan dept	Assistant	*

Fig.2 Example

In suppression based method, every attribute is suppressed by *.So third party cannot differentiate between any tuples. Figure 2.1 shows Suppressed attributes or data with k= 1.As shown in the table above, the column named as "salary" is converted into "*" format so the party at the other end will not be able to see that column. This is the basic example of suppression protocol.

IV. IMPLEMENTATION

In this part, the snapshots of the whole implemented system are given step by step. Let us now look at each one of them.

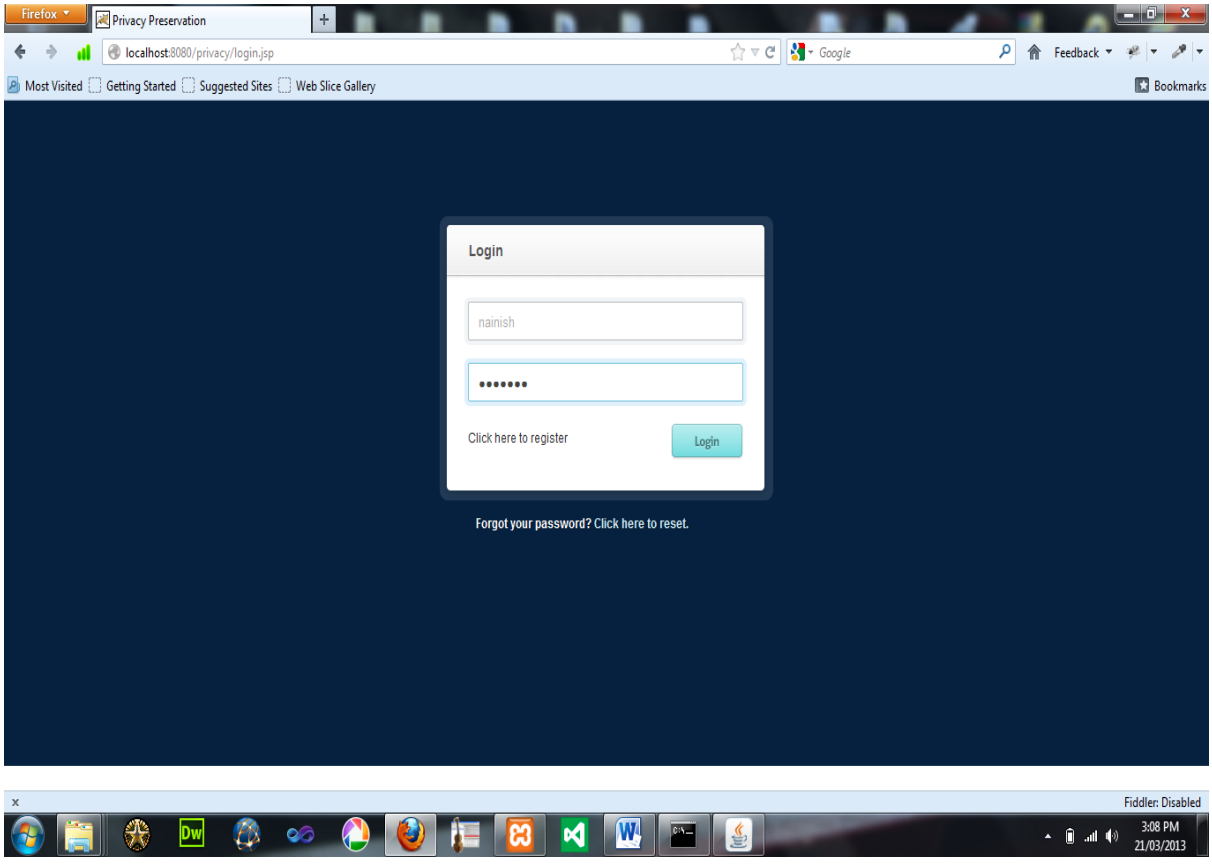

Fig.3 Login Page

ISSN (Print) : 2320 – 9798
ISSN (Online) : 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 1, March 2013

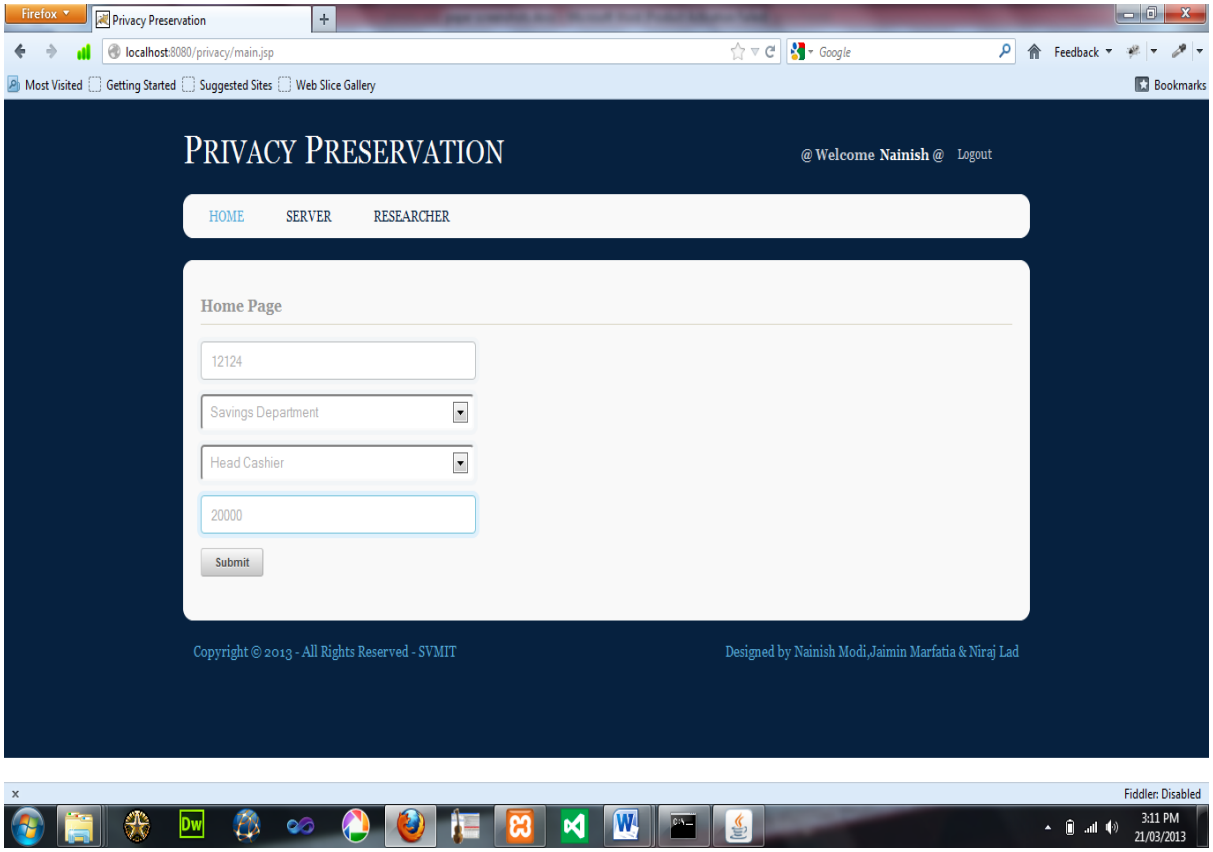

Fig.4 Home Page

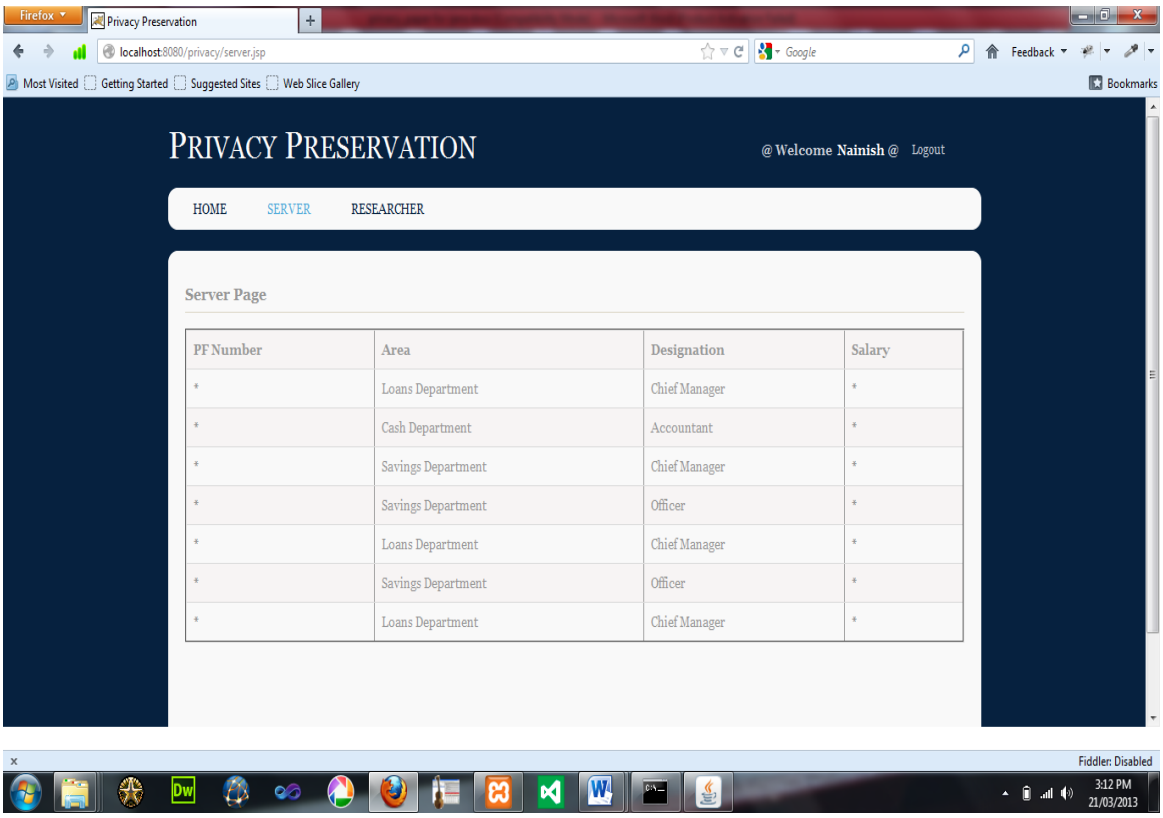

Fig.5 Server Page (k=2)

ISSN (Print) : 2320 – 9798
ISSN (Online) : 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 1, March 2013

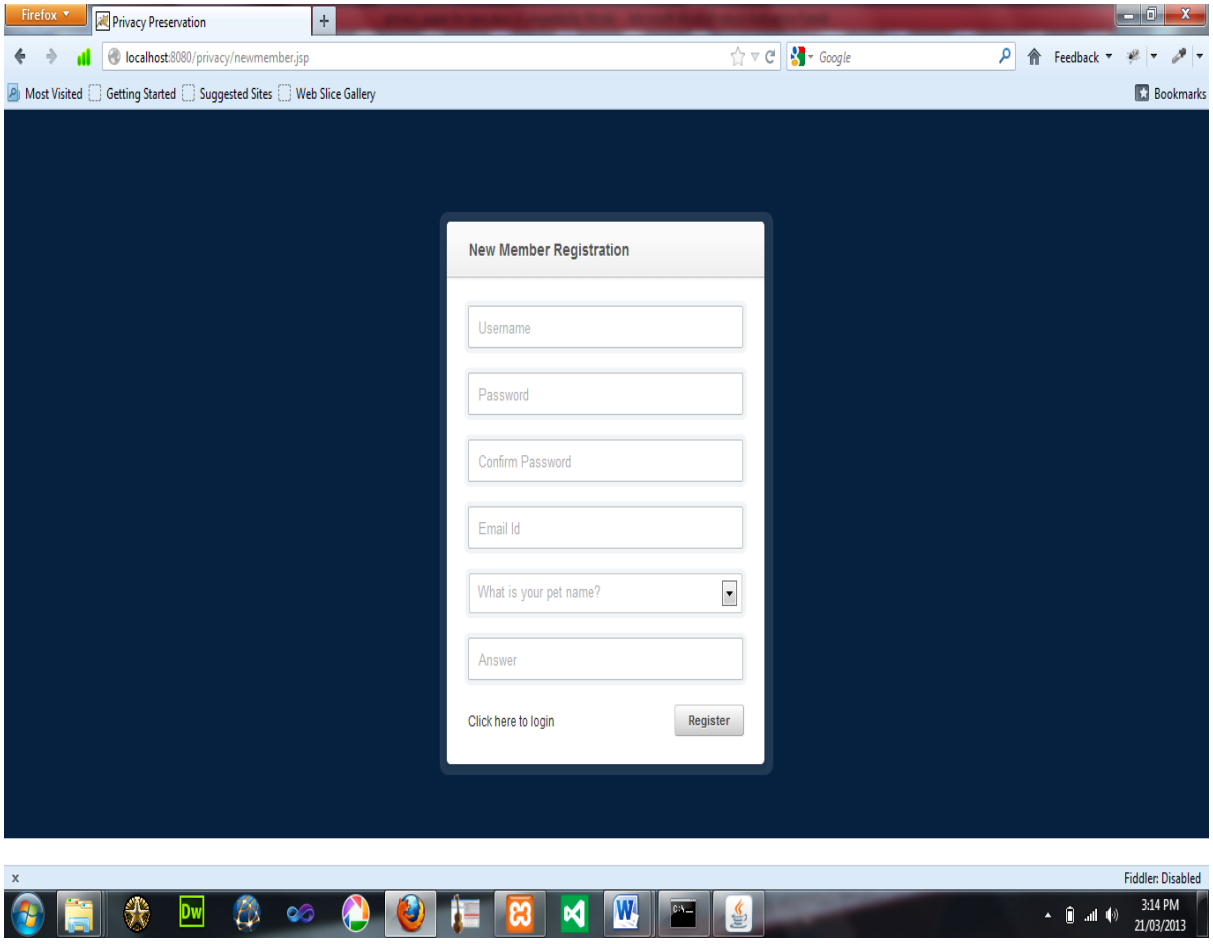

Fig.6 Registration Page

V. CONCLUSION

In this paper, we have proposed a secure protocol which will check the anonymity of the database without disclosing the entered tuple to the database owner. This indicates that when we enter the new tuple into the database, it will maintain its anonymity and will also insert the tuple. Here the user only has to send the non-suppressed attributes to the database which is k-anonymous. So here our main aim that is to maintain privacy and confidentiality will be achieved.

VI. ACKNOWLEDGEMENT

Before penning a single word for the Paper, we take this opportunity to thankMrs. Vaishali Patel and Mrs. Jignasa Patel from bottom of our heart who guided us as much as possible and for giving us valuable information regarding the paper. This was our first professional step toward the high careers in IT field. It was a great experience of exposing as well as learning lot of new things in Information Technology. We are indebted to all those who provided reviews our tasks and we apologize to anyone if we may have failed to mention.

REFERENCES

[1]. N.R. Adam and J.C. Worthmann, "Security-Control Methods for Statistical Databases: A Comparative Study," ACM Computing Surveys (CSUR), vol. 21, no. 4, pp. 515- 556, 1989.
[2]. Privacy-Preserving Updates to Anonymous and Confidential Databases ,Alberto Trombetta, Wei Jiang, Elisa Bertino and Lorenzo Bossi, Department of Computer Science and Communication, University of Insubria, Italy.
[3]. Generalization Based Approach to Confidential Database Updates , Neha Gosai, S H Patil, Department of ComputerScience,pune,Maharashtra,2012
[4]. www.wikipedia.com/wikifiles/.
[5]. R. Agrawal and R. Srikant, "Privacy preserving data mining," in Proceedings of the ACM SIGMOD Conference on Management of Data, Dallas, TX, May 2000, pp. 439–450.
[6]. Andrew C. Yao, Protocols for secure computations, University of California Berkeley, California 94720, 1982
[7]. Privacy preserving database application testing, Xintao Wu, CS Department, University ofNorth Carolina at Charlottexwu@uncc.edu.