Securing Privacy of Data in E-Marketing Against Malicious User

S. R. Jadhao, Sampada Bhonde

Abstract: ***Internet is at its best for personal as well as professional use as long as it is involved in anonymous communication. There are many technologies which are evolving and growing consistently in the field of computer one of them is cloud computing. But, the security issues and threats associated with it still serve as hindrances. The focal point of this paper is privacy preserving of data in cloud. There are different approaches for preserving privacy of data. Our main concentration would be securing privacy of data in cloud by assigning ID's (further referred as token) which are unique. The goal of unique ID's is to eliminate the privacy risk by modifying the dataset in such a way that only owner can access the original data. Preferably, any authority, server or an adversary alone should not know any client's personal information.***

This paper analyses and discusses various approaches for securing data like adopting cryptographic methods, writing access rights and policies, anonymising data, assigning unique ID's or token .Finally, the approach is made as why anonymity technique is used. Algorithms are discussed for anonymous sharing of private data among N parties. A technique is used so that ID numbers are used ranging from 1 to N.This assignment is anonymous such that when the identities are received at other end these are unknown to the other members of the group.

Keywords: Anonymization and de-anonymization, cloud and distributed computing systems, multiparty computation, privacy preserving data mining, privacy protection.

I. INTRODUCTION

With the advance in information technology (IT) concerns have been raised about the risks to data associated with weak IT security, including vulnerability to viruses, malware and attacks. Inadequate IT security may result in compromised confidentiality, integrity, and availability of the data due to unauthorized access. This data may present on web, on cloud or server.

The world is becoming more interconnected with the adve nt of the Internet and new networking technology. There is a large amount of personal, commercial, military,government informationon networking infrastructures worldwide. Netw ork security is becoming of great importance because of intellectual property that can be easily acquired through the internet. So it has become very important to preserve the privacy of the data while accessing it on the internet as well as preserving integrity while sharing the data to other nodes of the network. Some efficient ways should be found out so that the data becomes secure while accessing and sharing. In the field of technology cloud computing is evolving and growing consistently over the world. In the field of technology cloud computing is evolving and growing consistently over the world.

Revised Version Manuscript Received on July 04, 2015.

Prof. S. R Jadhao, Computer Science and engineering, BabasahebNaik College of Engineering, Pusad, MH, India.

Sampada Bhonde, M.E. Student, Computer Science and engineering, Babasaheb Naik College of Engineering, Pusad, MH, India.

Cloud computing is a model which gives Feasible and On-demand network access to shared resources where millions of users share an infrastructure. Cloud computing is a clean effective solution with following advantages [5]:

- Allows organizations to migrate their data to a cloud.
- Promises high speed access.
- 99.99% availability.

Issue related with cloud computing is Privacy. The protection of information from unauthorized disclosure is a long-standing concern of computer system design.

Our work deals with efficient algorithms as well as techniques for accessing data and assigning unique identifiers (IDs) or secure key in such a way that the IDs are anonymous using a distributed computation with no central authority. Given N nodes, this assignment is essentially a permutation of the integers {1....N} with each secure key being known only to the node to which it is assigned. Our main algorithm is based on a method for anonymously sharing simple data and results in methods for efficient sharing of complex data. There are many applications that require dynamic unique IDs for network nodes [3]. Such IDs can be used as part of schemes for sharing/dividing communications bandwidth, data storage, and other resources anonymously and without conflict.

Iterative technique is used so that ID numbers are used ranging from 1 to N.This assignment is anonymous in that the identities received are unknown to the other members of the group.

II. EXISTING TECHNIQUE

The existing system consists of the protected service which helps in securing privacy of the data. But these systems are not reliable. There are security breaches which harms the integrity of the data.

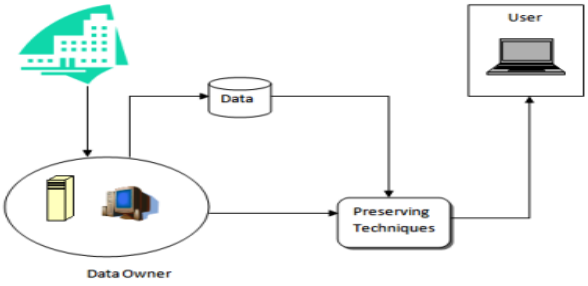

Figure 1: Process for preserving privacy of Data

A. Issues Related with cloud computing

A data breach is an incident that involves the unauthorized or illegal viewing, access or retrieval of data by an individual,

application or service. It is a type of security breach specifically designed to steal and/or publish data to an unsecured or illegal location. A data breach is also known as a data spill or data leak[8].

The below mentioned are some of the causes of the data breach:

1. Theft or loss of computer drive
2. Insider theft
3. Fraud
4. Unknown
5. Hackers
6. Data made accidently public

A data breach may result in data loss, including financial, personal and health information. A hacker also may use stolen data to impersonate himself to gain access to a more secure location. For example, a hacker's data breach of a network administrator's login credentials can result in access of an entire network [8].

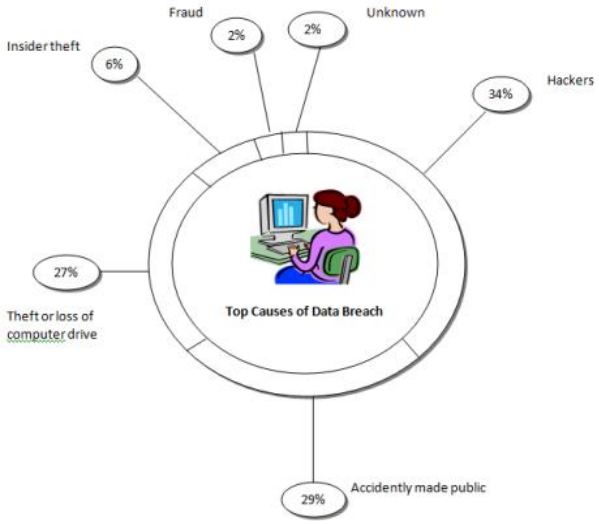

Figure 2: Causes of data Breach

The data is often not protected when used. Regardless of where the data is stored or transferred, Data security is important. Whenever an individual, a business, a government agency or other entity shares information, privacy or confidentiality question may arise. In such cases risks are produced such as someone's personal data is disclosed in public domain[8].

According to Data Protection Act,1998 (DPA) 'Personal Data' can be defined as data related to living individual who can be identified from that data or from that data and other information which includes expression of opinion about the individual[6].

There are different methods in preserving privacy of data in cloud [1].

1. Anonymity-based Method
2. A privacy-preserving Architecture
3. Privacy-Preserved Access Control
4. A Privacy Preserving Authorisation System
5. A Privacy Preserving Data Outsourcing.
6. PccPModel for Cloud
7. Dynamic Metadata Reconstruction

If we consider the module for cloud computing there are some issues related to privacy of data in cloud[1]:

Table 1: Issues in attaining privacy in cloud computing

Issues	Description
Insufficient user control	Owner of the data lacks control over their data in the cloud, especially when their data are accessed or processed in the cloud[1].
Information disclosure	Disclosure of sensitive data while data moves across the cloud. Sensitive information may be user's identity, usage data etc.[1]
Unauthorized secondary storage	Possibility of accessing and retrieving the sensitive information and backing up[1]
Uncontrolled data proliferation	Data flow in the cloud is unpredictable and uncontrollable.[1]
Dynamic Provision	Legally responsible entity in the cloud to assure privacy remains unclear, due to the dynamic nature of the cloud.[1]

III. IMPLEMENTATION

A. Analysis

Our system is consisting of E-marketing module where on site there would different purchase items. Customers can buy goods. There are different departments like Accounts which manages account section of market, HR responsible for human resource, Sales people etc. Whereas our main idea would be providing secure access rights to the users belonging to different departments so that each of them has different access rights and no one could view each other's data.

The proposed system would be consisting of following modules:

- *Web Application for Demonstration*: This is a demonstrative module for web application.
- *Data sharing with Client*: In this module, the data is shared with the client with anonymous ID's.
- *Data Security using Anonymizaton and Deanonymizaton*: In this module, security is provided to the data by applying Anonymization techniques like algorithms.

B. Objective

- Sharing privately held data so that the individuals who are the subjects of the data cannot be identified.
- Deal with efficient techniques for assigning identifiers (IDs) to the nodes of a network in such a way that the IDs are anonymous using a distributed computation with no central authority.
- It is necessary that every user must be guaranteed that his data is stored, processed, accessed and audited in a secured manner at any time.

- Dealing with algorithms which ensure privacy with anonymous data sharing.
- Using techniques which help in hiding and using cryptographic methods.
- To investigate some or all of the issues in privacy preserving data sharing.
- To analyze the existing literature related to privacy preserving data sharing.
- To study the connection between sharing secrets in an anonymous manner, distributed secure multiparty computation and anonymous ID assignment.

C. System Architecture and System Flow

The general architecture of the system is as shown below:

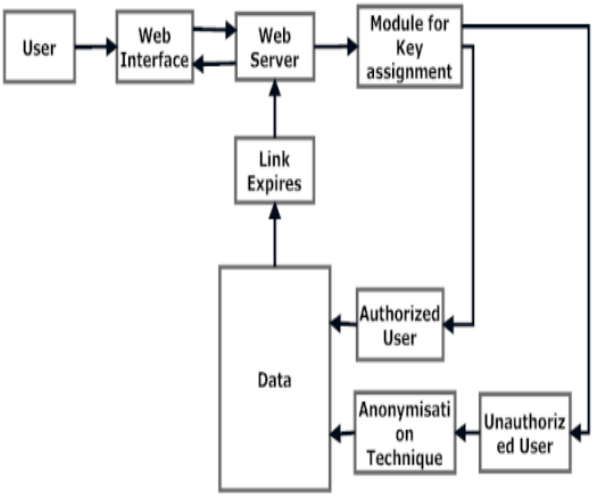

Figure 3: **Architecture of the system.**

As shown in the figure once the user has logged in he has to go through the key assignment module where a unique random key is generated and sent to authorize user. If the key matches' user gets access to the dataset else gets anonymized view of dataset. This all happens in the confined time period, when session expires user loses the access.

IV. SYSTEM DESIGN

This model is the perfect example of security of data in cloud.

Web application for demonstration:

We have two users:

1. Admin: Responsible for adding and deleting customers, users. Providing access rights to users
2. User/Client: Has access to the data as per rights given by Admin.

While executing, admin provides unique key to the user via email or SMS. This key is unique and is needed when authenticated user who has ID and password gets entry and after entering this key user gets access to the required data.

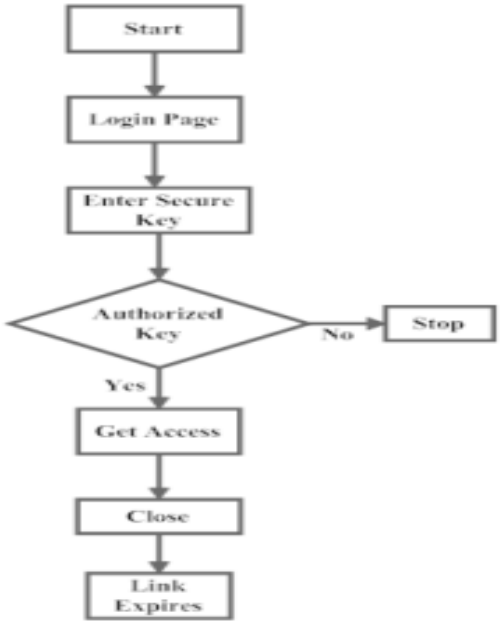

Figure 4: System flow

V. ALGORITHM

A. Algorithm for creating random number:

Given nodes$n_1, \ldots \ldots n_N$, use distributed computation (without central authority) to find an anonymous indexing permutation $s: \{1, \ldots, N\} \rightarrow \{1, \ldots, N\}$[2].

1. Set the number of assigned nodes$A = 0$
2. Each unassigned node $n_i$chooses a random number $r_i$in the range 1 to s. A node assigned in a previous round chooses$R_i = 1$.
3. The nodes n_i drew unique random numbers where random no is allotted at each round. Update the number of nodes assigned:$A = Ak$. (where "k" is random no allotted at each round and A is complete no. allotted in the previous rounds and at each round new value of "k" joins the values in previous round.
4. If $A < 9$ then return to step (2).

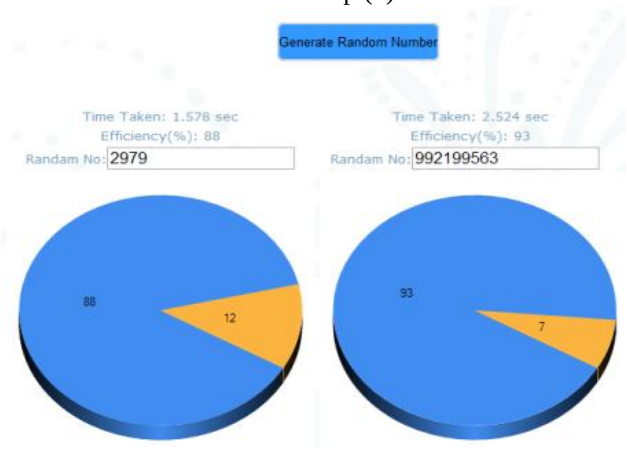

Figure 5: Comparison of two algorithms

So here implementing any random algorithm would give efficiency less than our algorithm. This algorithm takes care of the security breaches.

We took help of the following algorithms to achieve anonymity.

The problem of sharing privately held data so that the individuals who are the subjects of the data cannot be identified has been researched extensively [2]. There are some algorithms proposed like:-

B. Secure sum

Suppose the users of departments with individual databases wish to compute and share only the total of the data item such as total purchase value without revealing the value of the data item for any member of the group.

Thus N nodes n1,n2….,nN has data items d1,d2,…dN and wish to compute and share only the total value: T = d1+d2+…+dN, A secure algorithm allows the sum total T to be collected with some guarantee of anonymity[2].

1) Each node n_i, where $i = 1, \ldots . N$ chooses random values $r_{i,1}, \ldots r_{i,N}$ such that

$$r_{i,1}, \ldots r_{i,N} = d_i$$

2) Each "random" value $r_{i,j}$ is transmitted from node n_i to node n_j. The sum of all these random numbers $r_{i,j}$ is, of course, the desired total T.

3) Each node n_j totals all the random values received $s_j = r_{1,j} +, \ldots + r_{N,j}$

4) Now each node n_i simply broadcasts s_i to all other nodes so that each node can compute:

$$T = s_1, \ldots . s_n$$

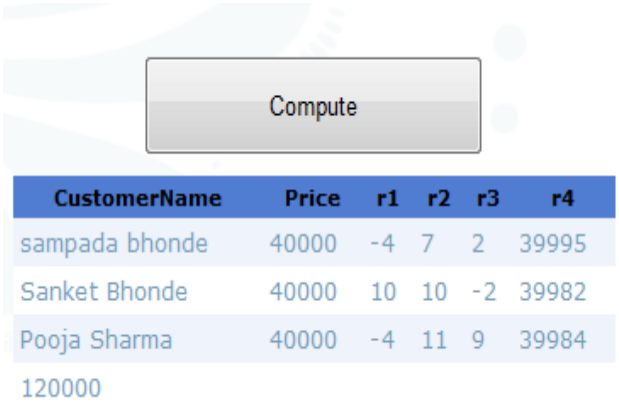

Figure 6: Implementation of Secure Sum

In this T=120000. The total value of price is calculated by implementing secure sum

C. Anonymous Data Sharing With Power Sums

Suppose that the group of nodes wishes to share actual data values from their databases. That is, each member ni of the group of N nodes has the data item di which is to be communicated to all other members of the group. However the data is to remain anonymous.

Given nodes $n_1, \ldots \ldots n_N$ each holding a data item d_i.from a finitely representable field F, make their data items public to all nodes without revealing their sources[2].

1) Each node n_i computes d_i^n over the field F for $n = 1, \ldots . N$. The nodes then use secure sum to share knowledge of the power sums:

$P_1 = \sum_{i=1}^{N} d_i^1$	$P_2 = \sum_{i=1}^{N} d_i^2$	$\cdots$	$P_N = \sum_{i=1}^{N} d_i^N$

2) The power sums $P_1, \ldots . P_N$ are used transfer at each node $n_1, \ldots \ldots n_N$ which has $D_1, \ldots . D_N$ as its output.

Thus the data items are used shared anonymously.

Figure 7: Implementation of Anonymous Data Sharing With Power Sums

Here data is shared but the sources are not revealed. Data is shared by Id's.

VI. HACKER VIEW

Till now we have observed what would happen if user is authenticated but what if the user ID and password is hacked. But attempt to access data would be forbidden when the unauthorized user wants have access over the data.

In this we have used two methods [4]:

1. Hiding

   Hiding is a very simple method where the value of the data item is made "0" and thus the data is hidden from its original format.

   It can be viewed in following format:

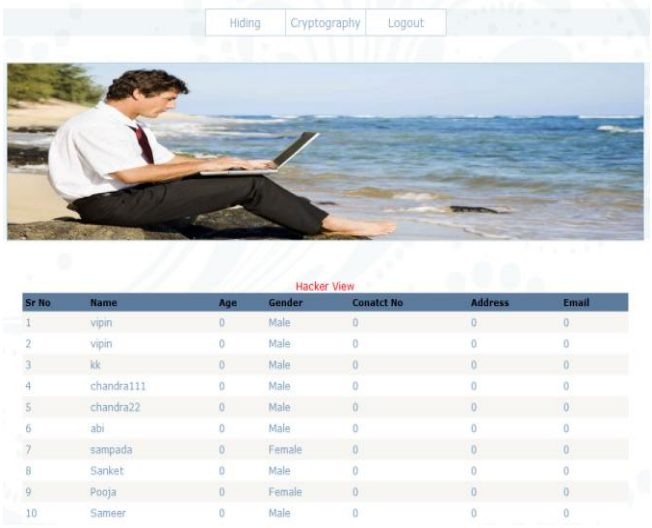

Figure 8: Hiding View

2. Cryptography:

Here we have used Cryptographic hash function. A Cryptographic hash function is a hash function which is considered practically impossible to invert, that is, to recreate the input data from its hash value alone. These one-way hash functions have been called "the workhorses of modern cryptography"[7]. The input data is often called the *message*, and the hash value is often called the *message digest* or simply the *digest*.

i. Encrypt a string using dual encryption method. Return a encrypted cipher Text
ii. Get a string to be encrypted.
iii. Use Hash? Send to for extra security.
iv. DeCrypt a string using dual encryption method. Return a DeCrypted clear string.
v. Get the encrypted string.
vi. Did you use hashing to encrypt this data? Pass true is yes.

And we get the following output:

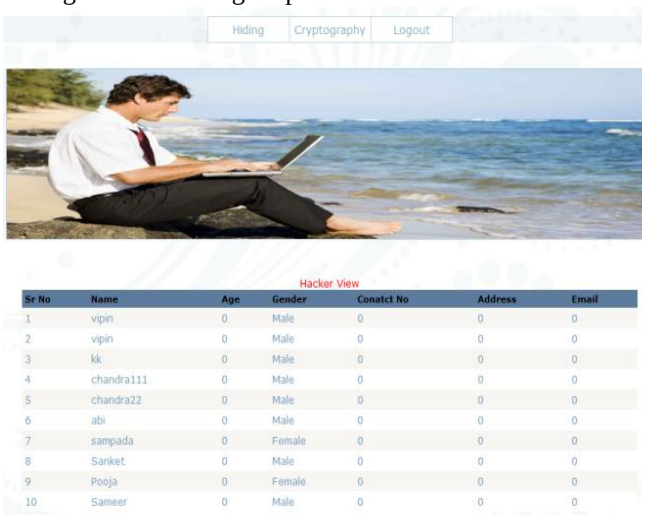

Figure 9: Cryptographic View

VII. CONCLUSION

The proposed system would be to secure privacy of shared data by Anonymous ID Assignment, by implementing discussed algorithms. This technique effectively preserves both information utility and individual's privacy. Privacy preserving is growing field of research. It is clear that there are many privacy preserving techniques available but still they have shortcomings. Anonymity technique gives privacy protection and usability of data. This technique will secure anonymous sharing of private data by anonymous ID assignment.

ACKNOWLEDGMENT

First and Foremost, I must offer my profoundest gratitude to my guide Prof. S. R. Jadhao for his useful comments, remarks and engagement through the learning process of this master thesis. Last but not the least I would like to thank my parents, HOD Dr. S.Y. Amdani, senior teaching faculty and all my friends who helped me directly or indirectly in my endeavor and infused their help for the success of this project.

REFERENCES

1. Wang J, Zhao Y et al. (2009). Providing Privacy Preserving in cloud computing, International Conference on Test and Measurement, vol 2, 213–216.
2. Larry A. Dunning, *Member, IEEE,* and Ray Kresman *"Privacy Preserving Data Sharing With Anonymous ID Assignment",* IEEE Transactions on information forensic and security, vol. 8, no. 2, February 2013
3. Jana, A. Chaudhuri, and B. B. Bhaumik, "Privacy and anonymity protection in computational grid services," *Int. J. Comput. Sci. Applicat.*, vol. 6, no. 1, pp. 98–107, Jan. 2009.
4. Friedman, R. Wolff, and A. Schuster, "Providing k-anonymity in data mining," *VLDB Journal*, vol. 17, no. 4, pp. 789–804, Jul. 2008.[5]Cloud
5. computing:http://www.south.cattelecom.com/Technologies/CloudComputing/0071626948_chap01.pdf
6. Information Commissioner's office, "Anonymization: managing data protection risk, code of practice", 2012
7. Cryptographic hash function: https://en.wikipedia.org/wiki/Cryptographic_hash_function
8. Data breach, causes of data breach: http://www.techopedia.com/definition/13601/data-breach