International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue VI, June 2018- Available at www.ijraset.com

Customized Privacy Preservation Model for Cloud Environment

Jyoti Nagwanshi [1], Mrs. Sarika Jain[2]
[1]*M. Tech, Department of CSE SAIT College, Indore (M.P.)*
[2]*Assistant professor, Department of CSE, SAIT College, Indore (M.P.), India*

Abstract: ***The privacy is primary requirement of growing technology. Maintaining Isolation over sensitive data in public environment is a big challenging task. It becomes more complex when data set becomes very large and number of users reaches to huge figure. Access Control principle help to classify the users according to rights and permission. Integration of Access Control model with knowledge extraction process is proposed to achieve privacy over sensitive data. The project will develop solution to maintain the privacy during the analytics process on smart home data. Subsequently, access control model will help to categorize and priorities the database attributes and data according to access. It will help to maintain data privacy during web mining.***

Keywords: Access Control, Privacy Protection, Sensitive Data, K-mean

I. INTRODUCTION

Cloud computing is a new technology that is a result of wrapping Virtualization, parallel computing and distributed computing into a single unit. The NIST definition of cloud computing "Cloud computing is a model for enabling ubiquitous, convenient, on – demand network access to a shared pool of resources (e.g., networks, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction". This cloud model is composed of five essential characteristics, three service models and four deployment models. The term cloud computing is more than a single product; it is very promising and new technology of our time.

Figure 1: Cloud Computing

Cloud computing is the combination of two fundamental technique utility computing and service oriented architecture [2]. Cloud computing means to deliver everything i.e. software, hardware by internet using any browser. It removes the necessity of setting

higher cost devices for infrastructure for any organization, with the help of cloud computing the organization takes care of its functions work rather than to develop a costly infrastructure.

In cloud environment all the data are outsourced to any external provider and they take care of that data is now a responsibility of the cloud provider and we can access this data on virtual machines or any device. Since the data center of cloud provider is spread to everywhere in the world we can access our data from any corner of the world. Numerous commercial cloud computing systems have been built at different levels, e.g., Amazon's EC2, Amazon's S3, and IBM's Blue Cloud are IaaS systems, while Google App Engine and Yahoo Pig are representative PaaS systems, and Google's App and Salesforce's Customer Relation Management (CRM) System belong to SaaS systems.

II. RELATED WORK

A. Study of Base Paper

Jayashree Agarkhed et al. In[1] proposed an approach using RSA and DES algorithm for security of data in cloud. Protection of data stored in cloud plays an important role which can be achieve through encryption and decryption technique. Which is used by author for the purpose of security.

Figure 2: Types of data security

B. Related Study

Veena1 et al. In [2] proposed that data warehousing represents an ideal vision of maintaining a central repository of all organizational data. Centralization of data is needed to maximize user access and analysis. Dramatic technological advances are making this vision a reality for many companies. And, equally dramatic advances in data analysis software are allowing users to access this data freely. The data analysis software is what supports data mining.

Quang Tran et al. In[3] described about data structure that holds the list of users that the operating system as a whole recognizes, along with a flag which indicates whether each user may read, write, execute, delete, or modify the file (or some combination of these). Whenever a user tries to perform any of these actions on the file, the operating system checks the file's ACL and determines whether the requested action appending data to the file, for example is allowed. If the action is allowed for that user, the data is appended; if not, the append operation fails.

Travis Mayberry et al. In [4] The attributes do not necessarily need to be related to each other, and in fact, the attributes that go into making a decision can come from disparate, unrelated sources. They can be as diverse as the date an employee was hired, to the projects on which the employee works, to the location where the employee is stationed, or some combination of the above. One should also note that an employee's role in the organization can serve as one attribute that can be (and often is) used in making an access control decision.

Tien Tuan et al. In [5] Privacy is a state in which one is not observed or disturbed by other people" Privacy protection policy is an approach to isolate the sensitive information from unauthorized access. The complete work concludes that Map Reduce Framework does not consist proposed security policy and suffering with data leakage problem.

II. PROBLEM STATEMENT

Security threat attack is also possible and malicious framework may give open system access to unauthorized user. Furthermore, Airavat Solution is not efficient solution and does not perform well. The complete phenomena generate a problem to implement security policy with data mining algorithms.

To balance the competing goals of a permissive programming model and the need to prevent information leaks, the untrusted code should be confined. Traditional approaches to data privacy are based on syntactic anonymization, i.e., removal of "personally identifiable information" such as names, addresses, and Social Security numbers. Unfortunately, anonymization does not provide meaningful privacy guarantees. High-visibility privacy fiascoes recently resulted from public releases of anonymized individual data, including AOL search logs and the movie-rating records of Netflix subscribers. [1]

The datasets in question were released to support legitimate data-mining and collaborative-filtering research, but naive anonymization was easy to reverse in many cases. These events motivate a new approach to protecting data privacy. One of the

challenges of bringing security to cloud computing is that users and developers want to spend as little mental effort and system resources on security as possible. Completely novel APIs, even if secure, are unlikely to gain wide acceptance. Therefore, a key research question is how to design a practical system that (1) enables efficient distributed computations, (2) supports a familiar programming model, and (3) provides precise, rigorous privacy and security guarantees to data owners, even when the code performing the computation is untrusted.

III. PROPOSED SOLUTION

Data mining opens new threats to privacy and information security if not done or used properly. The main problem is that to hide sensitive information, including personal information, fact or even patterns which are generated by any algorithm of data mining from the others. In order to focusing on privacy preserving association rule mining, the simplistic solution to address the problem of privacy is presented. To overcome the security problem work concludes a need of role based access control model to provide security in Web Database

Privacy algorithm for frequent item set mining and association rule learning over transactional databases may use to maintain privacy during mining.

The wok concludes that proposed solution will maintain privacy as per access control rules. So wok will be the hybrid solution of data mining with security.

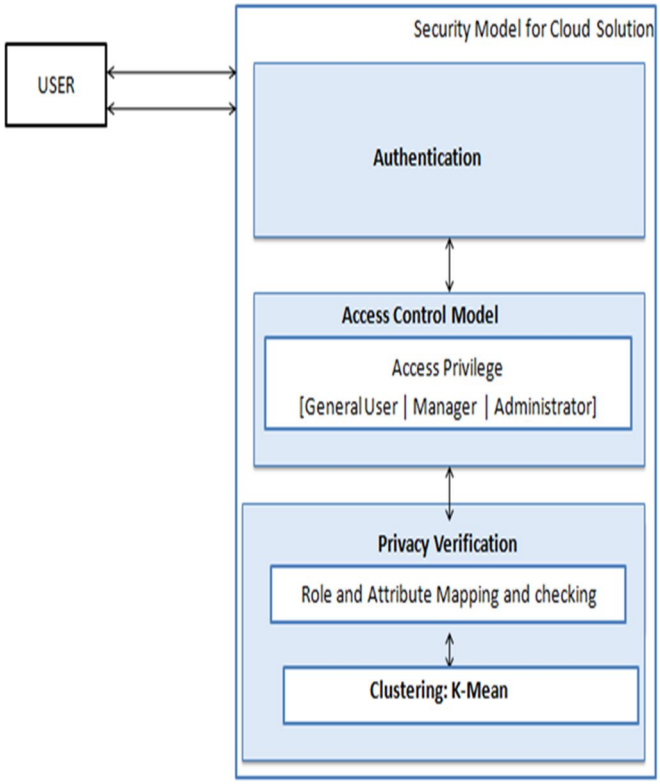

Figure 2: Proposed architecture

Research suggests that tapping the potential of this data can benefit businesses, scientific disciplines and the public sector – contributing to their economic gains as well as development in every sphere. The need is to develop efficient systems that can exploit this potential to the maximum, keeping in mind the current challenges associated with its analysis, structure, scale, timeliness and privacy. There has been a shift in the architecture of data-processing systems today, from the centralized architecture to the distributed architecture.

IV. CONCLUSION

The complete work concludes that proposed work will not only suggest a solution to implement access control mechanism with proposed security model but will help to achieve better performance in large data set. Here a hybrid security model has been

proposed based on ABAC and RBAC to manage table permissions and access. A super market dataset has been considered as the source of information and data schema to implement the proposed solution and evaluate the performance.

REFERENCES

[1] Jayashree Agarkhed, Ashalatha R, "A Privacy Preservation Scheme in Cloud Environment". 3 rd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB17)
[2] Veena1, Devidas, "Data Anonymization Approaches for Data Sets Using Map Reduce on Cloud: A Survey" NMAM Institute of Technology, Nitte, Karnataka, India.
[3] Quang Tran, Hiroyuki Sato, "A Solution For Privacy Protection In MapReduce". Graduate School of Engineering, The University of Tokyo.
[4] Travis Mayberry, Erik-Oliver Blass, Agnes Hui Chan, "PIRMAP: Efficient Private Information Retrieval for MapReduce", Proceedings of Financial Cryptography and Data Security (FC'13), pp. 371—385, Okinawa, Japan,
[5] Tien Tuan Anh Dinh, Prateek Saxena, Ee-Chien Chang, Beng Chin Ooi, Chunwang Zhang, "M 2R: Enabling Stronger Privacy in MapReduce" Computation School of Computing, National University of Singapore.
[6] Indrajit Roy Srinath T.V. Setty Ann Kilzer Vitaly Shmatikov Emmett Witchel, "Airavat: Security and Privacy for MapReduce".The University of Texas at Austin.
[7] https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html#Overview.