



EFFICIENT RETRIEVAL OF DATA FROM CLOUD USING DATA PARTITIONING METHOD FOR BANKING APPLICATIONS [RBAC]

Rajalakshmi V., Jothi Nisha V. and Dhanalakshmi S.
Faculty of Computing, Sathyabama University, Chennai, India
E-Mail: rajalakshmi.it@sathyabamauniversity.ac.in

ABSTRACT

Cloud computing is a pioneering area, where abundant services are provided for storing and handling data. Though the cloud usage reduces the cost of its customers, there is a severe drawback towards its privacy. As access to the data is provided open, there is a tendency for the lack of privacy of data. The existing methods use different encryption algorithms to handle this issue. Encryption algorithms being costlier and depends on the chosen key value the method becomes less secured. Hence, in this paper a simple yet secured perturbation procedure using integer partitioning which uses the keyword itself for generating the key value. The procedure is explained with the algorithm and the results are compared with the existing method.

Keywords: data partitioning, privacy preservation, data storage in cloud, encryption.

1. INTRODUCTION

Cloud computing is a technique in which large groups of remote servers are connected to allow centralized data storage and online access to computer services or resources. The present availability of high-capacity networks, low-cost computers and storage devices as well as the efficient adoption of hardware virtualization, service-oriented architecture, and autonomic and utility computing have led to the development in cloud computing. The goal of cloud computing is to allow users to utilize all of these technologies, without the need for deep knowledge about or expertise with each one of them. The cloud aims to reduce costs, and help the users focus on their core business instead of being impeded by IT obstacles.

With cloud computing, multiple users can access a single server to retrieve and update their data without purchasing licenses for different applications. Cloud resources are usually not only shared by multiple users but are also dynamically reallocated on demand [4].

Cloud computing exhibits Application programming interface (API), Cost reductions, Device and location independence, Multi tenancy etc., usually in a cloud, loosely coupled architectures are constructed using web services as the system interface.

With these advantages, Security over data becomes a huge concern. The complexity of security is greatly increased when data is distributed over a wider area or over a greater number of devices, as well as in multi-tenant systems shared by unrelated users [12]. Cloud computing providers offer their services according to several fundamental models [9]:

1. Infrastructure as a service (IaaS)
2. Platform as a service (PaaS)
3. Software as a service (SaaS)

Across all forms of deployment, architecture and

service models the basic concept of cloud computing remains to be the abstraction of computation over the used hardware/resources [6]. If considering various groups of stakeholders, a three-tier setup can be considered: i) users of virtual services ii) tenants who provide services iii) providers who provide the infrastructure. The key to gaining trust in cloud computing is assuring the user or tenant of security being applied and maintained in all components contributing to his virtual service.

Several deterrents to the widespread adoption of cloud computing remain. Among them are reliability, availability of services and data, security, complexity, costs, regulations and legal issues, performance, migration, reversion, the lack of standards, limited customization and issues of privacy[8][10]. The provider of such services lies in a position such that with the greater use of cloud computing services has given access to a plethora of data. This access has the immense risk of data being disclosed either accidentally or deliberately.

In a multi user shared cloud environment users are only physically isolated, but their data are stored in a same physical equipment [5]. For example, the details of bank transactions are stored in a cloud. The request of every customer to their corresponding account should be provided in a less time and none should be given others data. For every data stored in the cloud, security is a definite concern required for it. It is usually acquired by various conventional encryption algorithms. These algorithms rely on a key value - private or public. If the key is revealed or identified by a brute force method the entire system collapses [1].

2. RELATED WORK

In [2], multi keyword based data retrieval is implemented to rank the received records. Ranking the records based on keywords is done for the encrypted records after decrypting them. As multi keywords are not possible for all forms of access like bank account number,



this method is disadvantageous. In [3], multiple layers of security is built and it increases the cost of the system and also increases the time taken for storing and retrieving the data. In [13], a certificate less aggregate signature scheme is proposed, which provides different signatures on different messages being compressed into one, however, those corresponding messages cannot be compressed. Compression leads to loss of information and provides inefficiency during data retrieval.

[11] Uses a separate layer of encryption to increase the security of data. In addition, they have used the basic public/private key encryption systems. This increases the overall processing time for storing and retrieving data. In [7], a multi authority hierarchical attribute based encryption is proposed and it is compared with key policy and cipher text policy attribute based encryption techniques. The disadvantages of attribute based encryption are the specific identification of attributes are required, which further hinders the security.

Hence all these existing procedures have a drawback to implement an efficient secure method for a cloud computing environment. If a third party who monitors the storage and maintenance of the cloud is not trustworthy, the person should not be given the right to handle the data.

In this paper, we have proposed RBAC procedure which first stores the data in cloud using integer partitioning based encryption. The keyword is also encrypted and only the encrypted data are compared. The third party administrator also gets the encrypted data.

3. PROBLEM DEFINITION

An efficient encryption algorithm The proposed algorithm handles the problem using an efficient encryption using Integer partitioning method which calculates a Unique Key for every Keyword. Hence there is no chance for a secured record given to a vulnerable user. The following issues in the field of data security in a

cloud are handled.

1. The third party who verifies and maintains the cloud should be trustworthy.
2. The execution time for encrypting huge amount of data should be reduced.
3. There should be an efficient encryption algorithm which ensures security.
4. The encryption algorithm should not lead to symmetric reconstruction problem.
5. The method should allow authenticated data acquirement in an easier and efficient manner.

4. RBAC PROCEDURE

Security of a data that is present in a common location can be handled by using an encryption method which does not have a single key that encrypts all the data in a similar way. Such procedures are vulnerable to symmetric disclosure of key value. If a key generated for every possible keyword, the security of the data is highly increased. Hence procedure of Integer Partitioning is used that generates a key for every keyword provided. The Figure-1, explains the architecture of the system.

The Figure-1 shows the system model of RBAC procedure. All the data stored in the cloud are encrypted through RBAC procedure, i.e., for every numeric attribute like customers account number integer partitioning is applied and then stored in cloud. The actual parameters of partitioning are known only to the RBAC procedure, which can be altered only by a responsible person. When a consumer requires his account's details by providing the account number, first the validity of the customer is verified by an admin, along with sending the encrypted password to RBAC procedure. The RBAC procedure compares every record with the encrypted KEYWORD and returns only the valid, authorized record belonging to the corresponding customer.

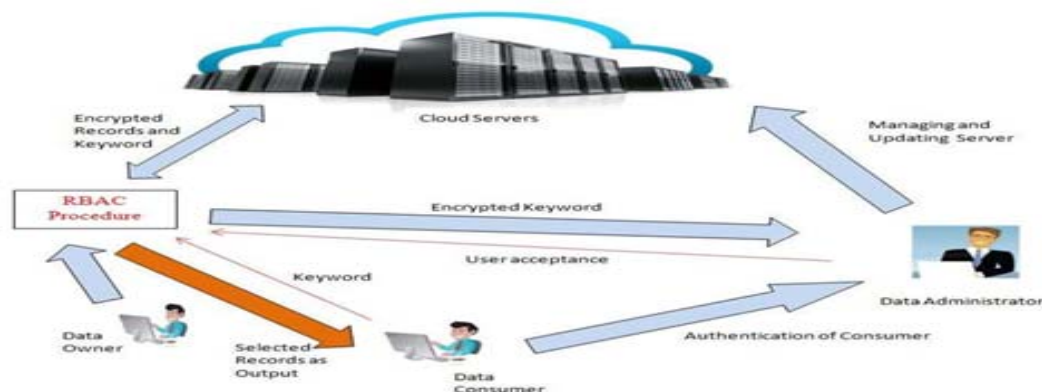


Figure-1. System Model of RBAC.

For implementing this procedure a banking database with the attributes specified in Table-1 is chosen. In the



given database “Account number” is the primary key which is unique for every individual customer. The remaining attributes can remain unaltered as they cannot be used to identify a particular record.

Table-1. Details of database used.

S. No.	Attribute name	Attribute Type
1	Account number	Numeric, Continuous
2	Name	String, Continuous
3	Transaction Date	Date
4	Transaction amount	Numeric, Continuous
5	Credit/Debit	Boolean
6	Balance	Numeric, Continuous
7	Mode of transaction	Categorical

A. Integer Partitioning

A partition of a positive integer n is defined to be a sequence of positive integers whose sum is n. the order of the partitions are usually written in a non- increasing order. The partitions can be any positive integer, but in our work we have restricted to prime numbers.

The procedure is reversible and hence can decrypt the data when required. The method also do not relate by any means with other records like clustering the records. The algorithm for the procedure is given in Figure-2.

```

Input: Keyword, K from the user.
Output: Target result from cloud
Method for Data storage:
1. Key is generated for the given keyword using Integer Partitioning method
   Key=IP(accountnumber)
2. For every numeric attribute i,
   If gender = 'M'
     Ai = Ai + Key
   Else
     Ai = Ai - Key
Method for Data Retrieval:
1. Using the account number the keyword is generated.
2. Using the keyword, all the related records in the cloud are compared.
3. The record which matches the keyword is retrieved.
4. The retrieved record is decrypted using the same generated keyword and provided to the user.
    
```

Figure-2. Algorithm of RBAC Procedure.

The algorithm has two sections, one for data storage and the other is for data retrieval. Only if the data is stored in the specified method, it can be retrieved in the same method. For the account number, its integer partitioning value is calculated and it remains a unique value as we have chosen them to be minimum prime numbers.

Another level of variation is given with respect to the value of the attribute gender. The value is either added or subtracted from the calculated key value. Since the key value is generated by the input keyword, it need not be

Example:

The partitions of 5 are

- 5+0
- 4+1
- 3+2
- 3+1+1
- 2+2+1
- 2+1+1+1
- 1+1+1+1+1

Thus $p_5=7$.

A partition is uniquely described by the number of 1s, number of 2s, and so on, that is, by the repetition numbers of the multi set. We devote one factor to each integer:

$(1+x+x^2+x^3+\dots)(1+x^2+x^4+x^6+\dots)(1+x^k+x^{2k}+x^{3k}+\dots)$ given by,

$$P_n = \prod_{i=1}^n \sum_{j=0}^{\infty} x^{ij} \tag{1}$$

For every integer, it has been proved that there is a unique partition with prime numbers. Hence this method can be adopted for anonymizing a data.

stored anywhere and also it is not susceptible to any kind of insecurity.

Example:

- 100 = 97 + 2 + 1
- 101 = 97 + 3 + 1
- 102 = 97 + 3 + 2
- 103 = 97 + 5 + 1
- 104 = 97 + 5 + 2

There are many combinations for the same number. But if a single constraint like maximum prime number added with lesser prime numbers is specified, it produces a unique result every time. Such encrypted data present in cloud can be made available to public users and they can receive their required details by providing their unique details. Figure-3 specifies the number of possible combinations for number “40”.

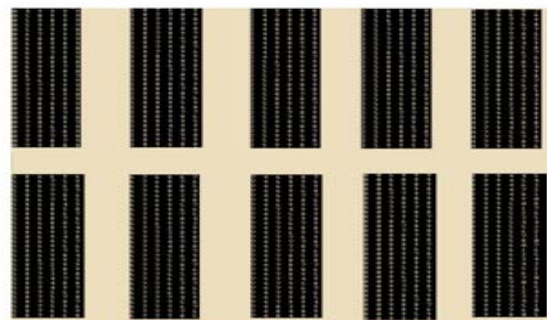


Figure-3. Possible combinations for number “40”.

The RBAC algorithm can be tuned by controlling the



variable m , which specifies the highest prime number used for partitioning. As m is increased the amount of privacy is increased, but the computation time slightly increases. The following graph shows the performance of RBAC over a standard encryption algorithm.

The graph in Figure-4 shows that the execution time of RBAC is less than that of a conventional encryption based algorithm. As the number of shifts and time consuming rotations present in an encryption algorithm is not there, the execution time reduces. As cloud is used mainly for storing huge amount of data, the performance over execution time is highly important. Hence RBAC is considered to be better than any conventional encryption algorithm.

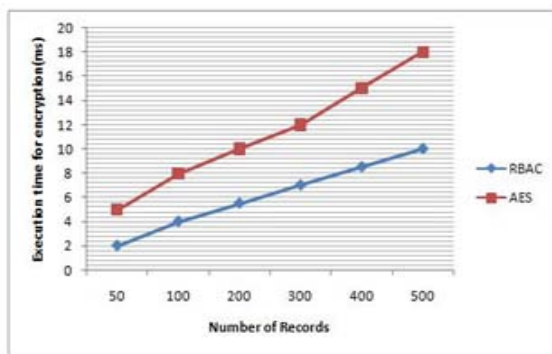


Figure-4. Performance of RBAC over AES encryption.

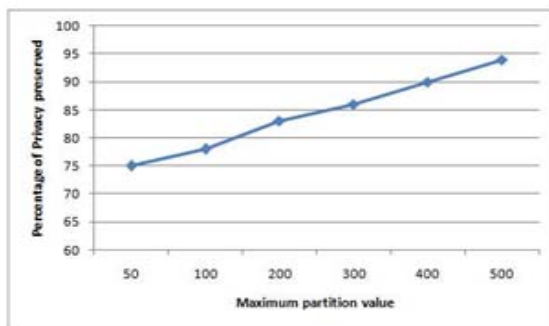


Figure-5. Tuning of RBAC using key size.

Figure-5 shows the tuning of system by selecting a maximum prime number for partitioning. By choosing the maximum number, there is only one unique partitioning for a numeric value and also the amount of privacy required can be fixed. Though the requirement is always the maximum privacy, increase in the maximum partition value increases the execution time. Hence an optimal value is chosen for the maximum partition value.

5. CONCLUSIONS AND FUTURE WORK

Thus the procedure for RBAC is explained with the steps. The performance of the procedure is also compared with a conventional encryption algorithm. A tuning parameter is also set for the tuning between privacy and execution time. The method computes the encrypted

value in a lesser time and also useful even when the administrator is a third party. The main drawback of this method is it cannot be used to encrypt non-numerical data. In future, this method can be used to encrypt non-numerical attributes.

REFERENCES

- [1] Bhanumathi S. and S. Sakthivel. A new model for privacy preserving multiparty collaborative data mining. Circuits, Power and Computing Technologies (ICCPCT), 2013 International Conference on. IEEE, 2013.
- [2] Cao Nin. *et al.* Privacy-preserving multi-keyword ranked search over encrypted cloud data. Parallel and Distributed Systems. IEEE Transactions on 25.1 (2014): 222-233.
- [3] Chandramohan D. *et al.* A novel framework to prevent privacy breach in cloud data storage area service. Green High Performance Computing (ICGHPC), 2013 IEEE International Conference on. IEEE, 2013.
- [4] Chen Deyan. and Hong Zhao. Data security and privacy protection issues in cloud computing. Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference on. Vol. 1. IEEE, 2012.
- [5] Juels Ari. and Alina Oprea. New approaches to security and availability for cloud data. Communications of the ACM 56.2 (2013): 64-73.
- [6] Li Ming *et al.* Authorized private keyword search over encrypted data in cloud computing. Distributed Computing Systems (ICDCS), 2011 31st International Conference on. IEEE, 2011.
- [7] Manjusha R. and R. Ramachandran. Comparative study of attribute based encryption techniques in cloud computing. Embedded Systems (ICES). 2014 International Conference on. IEEE, 2014.
- [8] Ren Kui., Cong Wang. and Qian Wang. Security challenges for the public cloud. IEEE Internet Computing. 16.1 (2012): 69-73.
- [9] Sri Venkatesh S. Bui: Browser User Interface And Extensions With Cloud Storage. Information Sciences & Computing. Vol. 7, Issue 1, pp. 47-52, 2013.
- [10] Takabi Hassan. James BD Joshi. and Gail-Joon Ahn. Security and Privacy Challenges in Cloud Computing Environments. IEEE Security & Privacy 8.6 (2010): 24-31.



www.arpnjournals.com

- [11] Vanitha Muthusamy, Kavitha C. Secured Data Deletion In Cloud Based Multi-Tenant Database Architecture. *Information Sciences & Computing*, Vol. 6, Issue 2, pp. 73-76, 2012.
- [12] Xiao Zhifeng. and Yang Xiao. Security and privacy in cloud computing. *Communications Surveys & Tutorials*. IEEE 15.2 (2013): 843-859.
- [13] Zhang L. and Zhang F. A New Certificateless Aggregate Signature Scheme. *Computer Communications*, vol. 32, no. 6, pp. 1079–1085, 2009.