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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

Security of Big Data in Cloud Environment Using Dynamic Policy Updating System

R.S.Manimegalai, Dr.K.Thamodaran,

MPhil Research Scholar, Dept. of Computer Science, Marudupandiyar College, Thanjavur, Tamilnadu, India

Professor, Dept. of Computer Science, Marudupandiyar College, Thanjavur, Tamilnadu, India

ABSTRACT: Big Data are playing vital role in everyday activity to manage the usage of high volume and velocity of data. Cloud computing is an efficient choice to store Big Data and the cloud computing has potential of storing Big Data and processing high volume of user access requests. Cloud computing has more scalability, reliability, and quality of services. To offer the security and confidentiality for Big Data, the Attribute-Based Encryption (ABE) is used in the cloud environment. On the other hand, the policy updating has always been a challenging matter when ABE is used to construct access control schemes. An insignificant accomplishment is to permit data owners to retrieve the data and decrypt it through the new access policy, and then send it back to the cloud. In this system, a high communication overhead and heavy computation workload are created for data owners. In this paper, a new security system is devised that enabling competent access control with dynamic policy updating for big data in the cloud. The new ABE based outsourced policy updating system is implemented to avoid the transmission of encrypted data and minimize the computation work of data owners. The data owners are verifying that whether the cloud server has updated the ciphertexts correctly or not with help of the proposed security system.

KEYWORDS: ABE, Access Control, Big Data, Cloud Computing, Data Mining, Policy Updating.

I. INTRODUCTION

The huge volume of data outsourced every day by individuals or each enterprise for various purposes. Practically it is very difficult to manage or to store this complex data at individual level, as the chances of crash the system is more, and the system becomes the single point of failure. The necessity of storing and managing the data in such a way that the cloud computing is the best choice to the store the data with better flexibility and cost savings. The encryption systems playing significant role to maintain the privacy of data, as the data might be confidential or sensitive. Data mining has attracted a great deal of attention in the information industry and in society as a whole in recent years, due to the wide availability of huge amounts of data and the imminent need for turning such data into useful information and knowledge. The information and knowledge gained can be used for applications ranging from market analysis, fraud detection, and customer retention, to production control and science exploration. Data warehouses have been defined in many ways. Loosely speaking, a data warehouse refers to a database that is maintained separately from an organization's operational databases. Data warehouse systems allow for the integration of a variety of application systems. They support information processing by providing a solid platform of consolidated historical data for analysis.

Cloud Computing is a technical and social reality and, at the same time, it is an emerging technology. The main attraction of cloud computing is the ability to use as many servers as necessary to optimally respond to the cost and the timing constraints of an application; this is only possible if the workload can be partitioned in segments of arbitrary size and can be processed in parallel by the servers available in the cloud. At this time one can only speculate how the infrastructure for this new paradigm will evolve and what applications will migrate to it. The economical, social, ethical, and legal implications of this shift in technology, when the users rely on services provided by large data centers and store private data and software on systems they do not control, are likely to be significant. The main attraction of cloud computing is the ability to use as many servers as necessary to optimally respond to the cost and the timing constraints of an application; this is only possible if the workload can be partitioned in segments of arbitrary size and can be processed in parallel by the servers available in the cloud.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

The rest of this paper is organized as follows. The Section II contains the details about related work. In Section III, Data Mining and Security of Big Data concepts are given. The Section IV describes the Significant of Cloud Computing and features. The Section V presents the proposed Security of Big Data in Cloud Environment Using Dynamic Policy Updating System. The results and discussion are offered in Section VI and Section VII concludes this paper.

II. RELATED WORK

Cloud services are having different benefits parallels some security issues are available. The process of outsourcing sensitive information (such as e-mails, personal health records, company finance data, government documents, etc.) to remote servers gets privacy issues. The cloud service providers (CSPs) that keep the data for users may access users' sensitive information without authorization. A general approach to protect the data confidentiality is to encrypt the data before outsourcing. Hadoop, like many open source technologies was created with security uses the MapReduce facility and a distributed file system with no built-in security. The Hadoop community realized that more robust security controls were needed, and decided to focus on security by applying technologies including Kerberos, firewalls, and basic HDFS permissions [19], [21]. Some components of the Hadoop ecosystem have applied their own security as a layer over Hadoop; for example, Apache Accumulo [24] provides cell-level authorization, and HBase [24] provides AC at the column and family level [22]. Some of them configured Hadoop to perform AC based on user and group permissions by ACLs, but this may not be enough for every organization, because many organizations use flexible and dynamic AC policies based on security attributes of users and resources and business processes, so the ACL approach is certainly limited [22].

Attribute-Based Encryption(ABE) [2],[4],[5],[6],[7],[8] has emerged as a promising technique to ensure the end-to-end data security in cloud storage system. It allows data owners to define access policies and encrypt the data under the policies, such that only users whose attributes satisfying these access policies can decrypt the data. In [8], Sahai and Waters (SW) propose ABE as follows: Given a secret key on a set of attributes , one candecrypt a ciphertext encrypted with a public key based on a set of attributes ω', only if the sets ω and ω' overlap sufficiently as determined by a threshold value t. The SW scheme also proposes the use of an access tree-based policy to decide on the attributes required to decrypt a message. An example for access tree could be: Class2005Λ (MyCollege v MyTeacher) implying whichever user who graduated in the class of 2005 either under MyTeacher or from MyCollege satisfies the policy.As an extension of the ABE scheme, two variants are proposed in the literature: the Key-Policy based ABE (KP-ABE) scheme and the Ciphertext-Policy based ABE (CP-ABE) scheme. In KPABE [9], the ciphertext is associated with a set of attributes and the secret key is associated with the access tree. The encrypting party has no control over who has access to the data and can only define the set of descriptive attributes necessary to decrypt the ciphertext. There is a trusted authority that generates the secret key, provided the user submits the appropriate values for the attributes that constitute the access tree. In CP-ABE [1], the ciphertext is associated with the access tree and the encrypting party determines the policy under which the data can be decrypted, while the secret key is associated with a set of attributes. In [2], the CP-ABE scheme has been leveraged towards an efficient implementation of the Permission as a Service model to provide users (content owners) with a single point of access control to set permissions on data belonging to multiple services.

III. DATA MINING AND BIG DATA SECURITY

A. *DATA MINING :*

The process of extracting or mining knowledge from bulky quantity of data is known as Data Mining. For example, the mining of gold from rocks or sand is referred to as gold mining rather than rock or sand mining. Therefore, data mining should have been more appropriately named "knowledge mining from data," or "Knowledge mining". Data mining is an interdisciplinary field, the confluence of a set of disciplines, including database systems, statistics, machine learning, visualization, and information science. Moreover, depending on the data mining approach used, techniques from other disciplines may be applied, such as neural networks, fuzzy and/or rough set theory, knowledge representation, inductive logic programming, or high-performance computing. Depending on the kinds of data to be mined or on the given data mining application, the data mining system may also integrate techniques from spatial data analysis, information retrieval, pattern recognition, image analysis, signal processing, computer graphics, Web technology, economics, business, bioinformatics, or psychology. Because of the diversity of disciplines contributing to

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

data mining, data mining research is expected to generate a large variety of data mining systems. Data mining systems are classified based on various criteria such as i.)Kinds of Databases, ii.) Kinds of Knowledge, iii.)Kinds of Techniques Utilized, iv.)Applications Adapted [3], [15]. The figure 1. Shows the Architecture of Data Mining.

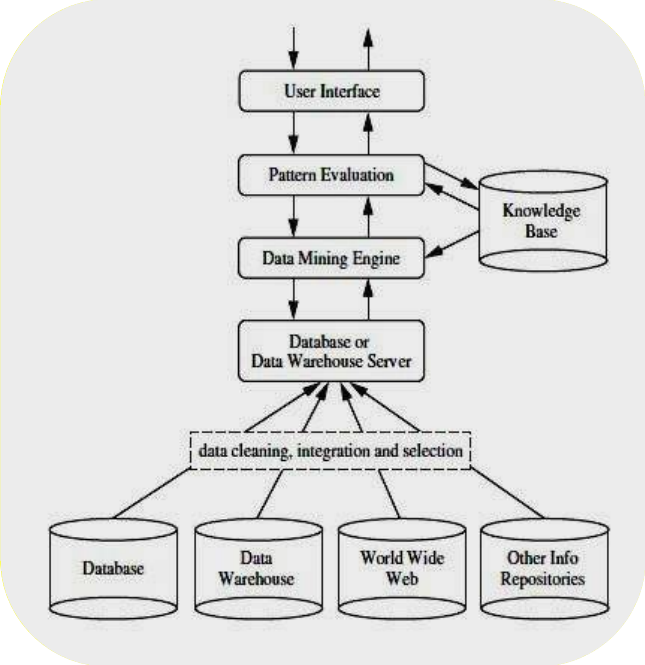

Figure 1. : Architecture of Data Mining

B. ***BIG DATA AND ITS SECURITY:***

In 1998, very first time Mr.John has illuminated the term “Big Data” in a Silicon Graphics Mashey with the title of "Big Data” and the Next Wave of Infra Stress. Every day the huge amount of data is derived. This data is known as Big Data [17]. Big Data is a cross-disciplinary concept which contains more data than making sense out of. Big Data is a call for us computer scientists to once again provide even better methods to crunch even more diverse, even more complex, even more dynamic, even more fine-grained, even larger data. Big Data brings new opportunities for institutions of higher education, as institutions continue to face unprecedented challenges in their environment. Big Data has denser and higher resolutions such as media, photos, and videos from sources such as social media, mobile applications, public records, and databases; the data is either in static batches or dynamically generated by machine and users by the advanced capacities of hardware, software, and network technologies. Examples include data from sensor networks or tracking user behavior. Rapidly increasing volumes of data and data objects add enormous pressure on existing IT infrastructures with scaling difficulties such as capabilities for data storage, advance analysis, and security[11], [13]. Big Data includes five 5 major characteristics. They are Volume, Variety, Velocity, Veracity and Value.

The value of big data to an organization falls into two categories: analytical use, and enabling new products. Big data analytics can reveal insights hidden previously by data too costly to process, such as peer influence among customers, revealed by analysing shoppers transactions, social and geographical data. Big data is increasingly becoming a factor in production, market competitiveness and, therefore, growth. Cutting-edge analysis technologies are making inroads into all areas of life and changing our day-to-day existence. Sensor technology, biometric identification and the general trend towards a convergence of information and communication technologies are driving the big data movement [13], [14], [16], [17],[18], [20], [23]. The figure 2. shows the Big Data and OT, IT Universal Risk Management

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

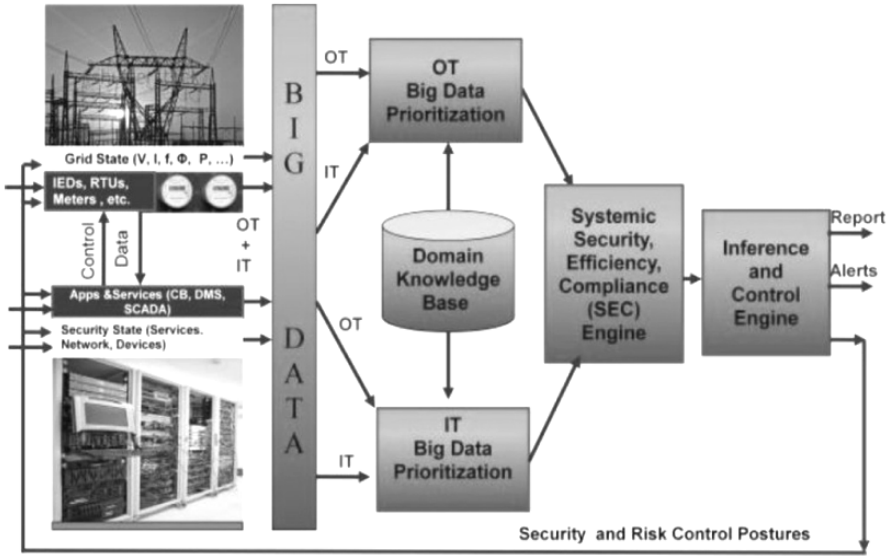

Figure 2. : The Big Data and OT, IT Universal Risk Management

In present day environment the security of Big Data is incompetent for exposing answers to many challenging problems. The revolutionary alterations in the types of security attacks can now be developed using big data and big data analytics. Big data analytics can be used to study transactions, log files as well as network traffic to recognize anomalies and suspicious activities as well as present a one dimensional view of the combined data set [25] ,[26]. Richard Zuech et al. offered the review process on the scope of works considering the problem of heterogeneous data and in particular Big Heterogeneous Data. Explained about the specific issues of Data Fusion, Heterogeneous Intrusion Detection Architectures, and Security Information and Event Management (SIEM) systems, as well as presenting areas where more research opportunities exist. Overall, both cyber threat analysis and cyber intelligence could be enhanced by correlating security events across many diverse heterogeneous sources. A significant aspect for intrusion detection is long-term storage of certain security data. Fundamentally, there are two main objectives for the archival of security data. The first objective is to improve intrusion detection capabilities even in real-time with offline data mining operations and Security Analytics. This offline data mining operation on security data can further try to identify previously unknown cyber threats, and then update the real-time detection capabilities with additional new signatures or behavior traits. The second objective is to provide forensic capabilities with this data so that in the event of a security breach, forensic evidence is available to assist the investigation [23].

IV. SIGNIFICANT OF CLOUD COMPUTING

Cloud Computing is a technical and social reality and, at the same time, it is an emerging technology. Owing to the growing popularity of cloud computing, more and more data owners are motivated to outsource their data to cloud servers for great convenience and reduced cost in data management. Cloud computing is revolutionizing many of our ecosystems, including healthcare. Compared with earlier methods of processing data, cloud computing environments provide significant benefits, such as the availability of automated tools to assemble, connect, configure and reconfigure virtualized resources on demand. These make it much easier to meet organizational goals as organizations can easily deploy cloud services[10],[21]. The figure 3 shows the Pyramid Model of Cloud Computing Paradigms.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

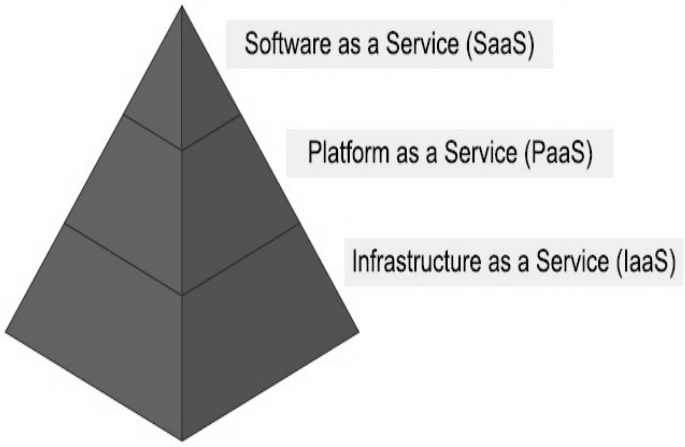

Figure 3. : The Pyramid Model of Cloud Computing Paradigms.

The Pyramid model of cloud computing paradigms are having the following features.

- The infrastructure provides the basic resources,
- The platform adds an environment to facilitate the use of these resources through software,
- The software allows direct access to services [21].

The following entities are involved in cloud computing:

- *Service Consumer* - entity that maintains a business relationship with, and uses service from, service providers;
- *Service Provider* - entity responsible for making a service available to service consumers; *Carrier* - the intermediary that provides connectivity and transport of cloud services between providers and consumers;
- *Broker* - an entity that manages the use, performance and delivery of cloud services, and negotiates relationships between providers and consumers;
- *Auditor* - a party that can conduct independent assessment of cloud services, information system operations, performance and security of the cloud implementation.
- *Audit* is a systematic evaluation of a cloud system by measuring how well it conforms to a set of established criteria. For example, a security audit evaluates cloud security, a privacy-impact audit evaluates the privacy-impact, while a performance audit evaluates the cloud performance.

The Cloud Computing system having the following features:

- Cloud computing uses Internet technologies to over scalable and elastic services; the term \elastic computing refers to the ability to dynamically acquire computing resources and to support a variable workload.
- The resources used for these services can be metered and the users can be charged only for the resources they used.
- The maintenance and security are ensured by service providers.
- The service providers can operate more efficiency due to specialization and centralization.
- Cloud computing is cost-effective because of the multiplexing of resources; lower costs for the service provider are past to the cloud users.
- The application data is stored closer to the site where it is used in a device and location independent manner; potentially, this data storage strategy increases reliability, as well as security and lowers communication costs[26].

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

V. PROPOSED BIG DATA SECURITY SYSTEM

A. *DYNAMIC POLICY UPDATING :*

In order to update the access policy of the encrypted data in the cloud, we delegate the ciphertext update from the data owner to the cloud server, such that the heavy communication overhead of the data retrieval can be eliminated and the computation cost on data owners can also be reduced. When the data owner wants to update the ciphertext from the previous access policy A to the new access policy A0, it first generates an update key UKm by running the update key generation algorithm UKGen, and then sends the update key UKm to the cloud server. Upon receiving the update key from the data owner, the cloud server will run the ciphertext updating algorithm CTUpdate to update the ciphertext from the previous access policy A to the new one A0. However, the update key generation algorithm UKGen and the ciphertext updating algorithm CTUpdate are related to the structure relationship between the previous access policy A and the new access policy A0. For different types of updating operation, we have different design of UKGen and CTUpdate, which will be described in detail in the next section.

Any access policy can be expressed by either LSSS structure or Access Tree Structure, which are defined in the Supplemental File. In this section, we only consider monotonic structures, and non-monotonic structures can be similarly achieved by taking NOT operation as another attribute. Specifically, we first design the policy updating algorithms for monotonic boolean formulae. Then, we present the algorithms to update LSSS structures. Finally, we consider general threshold access tree structures by designing algorithms of updating a threshold gate.

The grand challenge of outsourcing policy updating to the cloud is to guarantee the following requirements:

- Correctness: Users who possess sufficient attributes should still be able to decrypt the data encrypted under new access policy by running the original decryption algorithm.
- Completeness: The policy updating method should be able to update any type of access policy.
- Security: The policy updating should not break the security of the access control system or introduce any new security problems.

B. *ATTRIBUTE-BASED ACCESS CONTROL:*

Attribute-based encryption (ABE) is more suitable (compared to the traditional public-key infrastructure based or identity-based encryption) to protect the privacy and secrecy of data in a cloud computing environment. ABE is useful when the source of the data knows neither the identity of the recipient nor their public key; but only knows certain attributes of the recipient. For example, imagine user Alice wishing to communicate with her former classmates, but she does not know their email addresses. ABE identifies a user with a set of attributes [12],[15]. The volume of big data is high and its velocity is increasing in high level in the age of big data. The proposed attribute-based access control (ABAC) method is fairly appropriate for controlling big data than traditional access control methods due to the following features:

1) Policy Checking Entity Free: In ABAC, access policies are defined by data owners but do not require any entity (e.g., the server) to check these policies. Instead, access policies in ABAC are enforced implicitly by the cryptography. Due to this key feature, ABAC is widely applied to control big data in cloud environments, where cloud servers are not trusted to enforce access policies.

2) Storage Efficiency: In traditional Public Key Cryptography, for each data, multiple copies of ciphertexts are produced whose number is proportional to the number of users. Considering the high volume of big data, it incurs a huge storage overhead even when only doubling the volume of big data. Fortunately, in ABAC, only one copy of ciphretext is generated for each data, which can reduce the storage overhead significantly.

3) Dynamic Policies but Same Keys: Data owners can use the same public key to encrypt data under different access policies, and users do not need to change their secret keys either. What's more, data owners can change access policies of existing ciphertexts by simply sending a request to the cloud server, and let the server do the policy change without leaking out any sensitive information of the data as well as the keys.

C. *PROPOSED SECURITY MODEL:*

The proposed system focuses on solving the policy updating problem in ABE systems, and proposes a secure and verifiable policy updating outsourcing method. Instead of retrieving and re-encrypting the data, data owners only send policy updating queries to cloud server, and let cloud server update the policies of encrypted data directly, which means that cloud server does not need to decrypt the data before/during the policy updating. This scheme can not only satisfy

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

all the above requirements, but also avoid the transfer of encrypted data back and forth and minimize the computation work of data owners by making full use of the previously encrypted data under old access policies in the cloud. The figure 4 shows the architecture of Security System for Big Data Access Control in Cloud.

- Formulate the policy updating problem in ABE systems and develop a new method to outsource the policy updating to the server.
- Propose an expressive and efficient data access control scheme for big data, which enables efficient dynamic policy updating ,
- Design policy updating algorithms for different types of access policies.

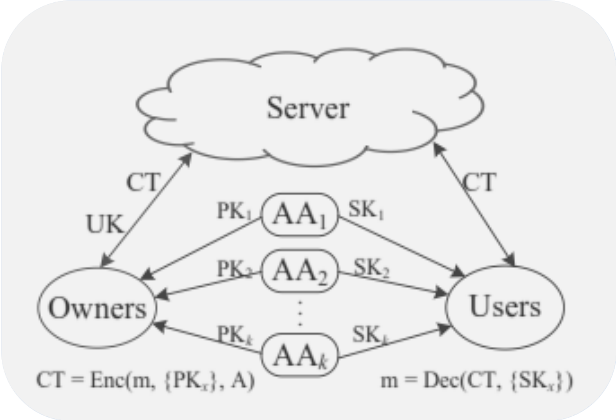

Figure 4. : Security System for Big Data Access Control in Cloud

The figure 4 shows the architecture of Security System for Big Data Access Control in Cloud computing. The figure 5 contains procedure for Comparison of Policy, DNF2LSSS, DNF2SSS, SSS2MSP.

- **Procedure for Key Generation**

In order to generate the key, on behalf of each one user GID, each authority AID will first assign a set of attributes $S_{GID,\,AID}$ to this user. It then runs the secret key generation algorithm SKGen to generate a set of secret keys as

$SK_{GID;AID} = \{K_{x,GID} = g^{\alpha x} H(GID)^{\beta x}\}\ \forall x \in S_{GID;AID}$

- **Procedure for Data Encryption**

In order to encrypt the plain text the owner of data encrypts the data m with help of the encryption algorithm. The algorithm gets as inputs a set of public keys (PK) for relevant authorities, the global parameters, the data m and (n x l) access matrix M with ρ mapping its rows to attributes. It chooses a random encryption exponent. s ϵ Zp and a random vector $\vec{v} = (s', y_{2,\ldots,}y_l) \in z_p^l$, where $y_{2,\ldots,}y_l$ are used to share the encryption exponent s. For i = 1 to n, it computes $\lambda_i = M_i . \vec{v}$, where M_i is the vector corresponding to the i-th row of M. It also chooses a random vector $\vec{w} \in z_p^l$ p with 0 as its first entry and computes $w_i = M_i . \vec{w}$. For each row i of M, it chooses a random $r_i \in$ Zp and computes the ciphertext as CT = (C = m . $e(g;g)^s$; □i = 1 to n : $C_{1;i} = e(g;g)^{\lambda i} e(g;g)^{\alpha\rho(i) ri}$; $C_{2,i} = g^{ri}$; $C_{3,i} = g^{\alpha\rho(i) ri} g^{wi}$):
The encryption information EnInfo(m) of the data m contains all the random numbers r_i, i.e., EnInfo(m) = {r1,…, r_n}.

- **Procedure for Data Decryption**

In order to decrypt a cipher text, the user gets the H(GID) from the random prediction. If the user has secret keys $\{K_{\rho(i),GID}\}$ for a subset of rows i of M such that (1,0, … ,0) is in the span of these rows, then the user proceeds as follows.

For each such i, the user computes $[(C_{1,i} . e(H(GID),C_{3,i})) / (e(K_{\rho(i),GID}, C_{2,i}))] = e(g;g)^{\lambda ie}(H(GID), g)^{wi}$:
The user then chooses constants ci 2 Zp such that åi ciMi = (1,0,…,0) and computes

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

$\Pi(e(g;g)^{\lambda i}\ e(H(GID), g)^{wi}\ )^{ei} = e(g,g)^{s}$, consider $\lambda_i = M_i \cdot \vec{v}$ and $M_i \cdot \vec{w}$, where $\vec{v} \cdot (1,0, \ldots ,0) = 0$ and $\vec{w} \cdot (1,0, \ldots ,0) = 0$. The data can then be decrypted as m =$(C/e(g, g)^{s})$.

Procedure1 : Comparison of Policy
Input: new policy (M', p') with $l' x k'$ matrix
Input: previous policy (M,p) with l x k matrix
Output: I_1, M' ; I_2, M' ; $I_3; M' \rightarrow$ three subsets of row indexes in M'
$I_M \leftarrow$ index set of rows in M
for j = 1 to l' do
if p' (j) in M then
if I_M! = Ø & ∃i ϵρ I_M s.t. p' (i) == p' (j) then
add (j, i) into I_1,M0
delete i from I_M
else
find any i 2 [1, l] s.t. p'(i) == p'(j)
add (j, i) into I_2, M'
end if
else
add (j,0) into I_3, M',
end if
end for

Procedure2 : DNF2LSSS
Input: A . an access structure A
Input: s . the secret to be shared
Output: M . Monotone Span Program
let sss be DNF2SSS(s;A);
let M be SSS2MSP(s; sss);
return s, M;

Procedure3: DNF2SSS
Input: A_n . a node n from an access structure A
Input: s . the secret to be shared
Output: $\{s_1,\ldots, s_l\}$. a set of shares
I_M index set of rows in M;
if Node Type = OR then
for each child c of A_n do
DNF2SSS(s, c); pass the s to child nodes
end for
end if
if Node Type = AND then
let gate have t inputs;
select $r_{1,\ldots,}r_{t-1}$ ϵ_R Zp;
for all r_i do
DNF2SSS(r_i, i^{th} child of A_n);
end for
let r_t = s-r_1-…r_{t-1} (mod p);
DNF2SSS(r_i, t^{th} child of A_n);
end if

Procedure4 : SSS2MSP
Input:$\{\vec{v}\}$. Set of piece vectors for each attribute $Attr_i$
Input: s . the secret to be shared
Output: M . Monotone Span Program
1: let $\vec{z}$ be a vector and set $\vec{z}$ (0) = s;
2: let M be a matrix;
3: let p be a labelling function;
4: for all $Attr_i$ do
5: for each piece vector $\vec{v}$ for $Attr_i$; do
6: apppend each random value in $\vec{v}$ to $\vec{z}$;
7: construct the position vector $\vec{v}$ for $Attr_i$;
8: append $\vec{v}$ to M;
9: let r(M_{vi}) to $Attr_i$;
10: end for
11: end for
12: pad M with the same row size;
13: return ($\vec{z}$,M,p);

Figure 5. : Procedure for Comparison of Policy, DNF2LSSS, DNF2SSS, SSS2MSP

VI. RESULTS AND DISCUSSION

In the proposed security system, the data owner only needs to send the update keys to the cloud server, instead of the whole encrypted big data. Therefore, our method can significantly reduce the communication cost during the policy updating. Suppose jpj is the element size in the G;GT ;Zp. Table 1 shows the size of update keys in our scheme. We can see that Type1 operation incurs the smallest size of update keys. When updating an access policy to a new one, the most common operation is Type1 operation, such that our scheme incurs a small communication cost. The proposed scheme makes full use of the previous ciphertexts encrypted under the old access structure. That is if an attribute in the new access policy has ever appeared in the previous access policy, the new ciphertext component of this attribute can be derived from the previous ciphertext component with the update key. The data owner only needs to compute ciphertext components for new attributes. Moreover, in our scheme, we also delegate all the pairing operations to the server, such that the workload of the data owner can be further reduced.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

Table 1. size of update key

Operation	Attr2OR	Att2AND	Type1	Type2	Type3
Size(UK)	4| p |	5 | p |	2 | p |	3 | p |	3 | p |

When compares the computation time between generating an update key (e.g., Type1, Type2 and Type3 in our scheme) versus generating a ciphertext component (if the owner choose to re-encrypt the ciphertext using a new secret) corresponding to an attribute. It is more efficient for data owners to only generate an update key than generate a ciphertext component for each attribute.

VII. CONCLUSION

The proposed Big Data Security System is used to perform the investigation to solve the policy updating problem in big data access control systems and formulated some challenging requirements of this problem. The efficient method has been developed to outsource the policy updating for the cloud server, which can satisfy all the requirements. In addition to that the proposed attribute-based access control scheme is important for security of big data in the cloud. The policy updating algorithms are designed for different types of access policies. Furthermore, the proposed method enables the data owners to check the correctness of the ciphertext updating. Then the schemes are analyzed in terms of correctness, completeness, security and performance. Although the policy updating algorithms were designed based on Lewko and Waters' scheme, our ideas and methods of outsourced policy updating can also be applied to other ABE systems. The proposed scheme guarantees that the actual data owner could pass the cloud server's authentication and legally update the cipher text corresponding to the owner's data. Also designed policy updating algorithms with authentication for access policy expressed and also given the analysis of the scheme on the security, authentication and performance. Since the cloud will learn nothing of the data owner except that the owner could open the commitment, the scheme supports anonymous authentication. The access control scheme is constructed on prime order groups, because the group operations on prime order groups are much faster than the ones on Composite order groups. A dynamic policy access control scheme is secure in the generic bilinear group model. Public key encryption also called as asymmetric encryption involves a pair of keys, public key and private key associates with an entity. Ensure the data confidentiality in the cloud.

REFERENCES

1. E. Bertino, P. A. Bonatti and E. Ferrari, "TRBAC: A Temporal Role-based Access Control Model," ACM Transactions on Information and System Security, vol. 4, no. 3, pp. 191-233, August 2001.
2. R. Bhatti, J. B. D. Joshi, E. Bertino, A. Ghafoor, "Access Control in Dynamic XML-based Web-Services with XRBAC," Proceedings of the 1st International Conference on Web Services, LasVegas, June 23-26, 2003.
3. Jiawei Han, Micheline Kamber, "Data Mining: Concepts and Techniques", Second Edition, Morgan Kaufmann publications, 2006 .
4. V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in CCS'06. ACM, pp 89–98, 2006.
5. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute based encryption," in S&P'07. IEEE, pp. 321–334,2007.
6. A.B.Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters, "Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption," in EUROCRYPT'10. Springer, pp. 62–91,2010.
7. A.B.Lewko and B. Waters, "Decentralizing attribute-based encryption," in EUROCRYPT'11. Springer, 2011, pp. 568–588.
8. B. Waters, "Ciphertext-policy attribute-based encryption: An expressive, efficient, and provably secure realization," in PKC'11. Springer, pp. 53–70,2011.
9. K-Y. Chen, C-Y. Lin and T-W. Hou, "The Low-Cost Secure Sessions of Access Control Model for Distributed Applications by Public Personal Smart Cards," roceedings of the 17th IEEE International Conference on Parallel and Distributed Systems, pp. 894-899, December 2011.
10. Dan C. Marinescu, "Cloud Computing and Computer Clouds", 2012.
11. IBM," Extending Security Intelligence with Big Data solutions" Jan 2013.
12. V. Hu, D. Ferraiolo, R. Kuhn, A. Schnitzer, K. Sandlin, R. Miller, and K. Scarfone, "Attribute Based Access Control Definition and Consideration," NIST Special Publication 800-162, Gaithersburg, MD, USA, 2013.
13. J Oltsik. Defining big data security analytics. Network world, 1 April 2013.
14. J. Hurwitz, "Big Data for Dummies," Wiley, 2013.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

15. K. Yang, X. Jia, and K. Ren, "Attribute-based fine-grained access control with efficient revocation in cloud storage systems," in AsiaCCS'13. ACM, pp. 523–528,2013.
16. Alvero A. Cardenas, Pratyusa K. Manadhata, Sreeranga P.Rajan, "Big Data Analytics for Security" IEEE Security & Privacy, vol.11 no.6, Nov.-Dec., pp. 74-76, 2013.
17. D. Che, M. Safran, and Z. Peng, "From Big Data to Big Data Mining: challenges, issues, and opportunities," in Database Systems for Advanced Applications, Springer, Berlin, Germany, pp. 1–15, 2013.
18. Michael Minelli, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses,"Wiley, 2013
19. "The Big Data Security Gap: Protecting the Hadoop Cluster," White Paper, Zittaset,http://www.zettaset.com/wp-content/uploads/2014/04/zettaset_ wp_ security_0413 .pdf, 2014.
20. Thomas F. Dapp , "Big data : The untamed force" , May 5, 2014 .
21. K. Yang and X. Jia, "Expressive, efficient, and revocable data access control for multi-authority cloud storage," IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 7, pp. 1735–1744, July 2014.
22. K..T.Smith,"Big Data Security: The Evolution of Hadoop's Security Model," InfoQ, http://www.infoq.com/articles/HadoopSecurityModel, Aug. 2014.
23. Richard Zuech, Taghi M Khoshgoftaar and Randall Wald , Intrusion detection and Big Heterogeneous Data: a Survey Springer Journal of Big Data, pp 1-41, 5015.
24. "Apache Accumulo," https://accumulo.apache.org
25. Cloud Security Alliance Big data Analytics for Security Intelligence. http://cloudsecurityalliance.org/research/bigdata
26. J.Varia.\Cloudarchitectures.",http://jineshvaria.s3.amazonaws.com/public/ cloudarch itectures-varia.pdf.

BIOGRAPHY

R.S. Manimegalai is a M.Phil research scholar in the Department of Computer Science Marudupandiyar College, Thanjavur, Tamilnadu, India. She has received her B.Sc Computer Science degree and MSc Computer Science degree from Bharathidasan University, Trichy, TamilNadu, India. She is having more than 3 years of experience in office and hospital management system. Her areas of research interest are Data Mining and Cloud Computing.