Access Control Mechanisms for Outsourced Data in Cloud

Purushothama B R
Computer Science and Engineering
National Institute of Technology
Warangal, INDIA
Email: puru@nitw.ac.in

B B Amberker
Computer Science and Engineering
National Institute of Technology
Warangal, INDIA
Email: bba@nitw.ac.in

***Abstract*—Cloud Computing poses new security and access control challenges as the users outsource their sensitive data onto cloud storage. The outsourced data should be protected from unauthorized users access including the honest-but-curious cloud servers those hosts the data. In this paper, we propose two access control mechanisms based on (1) Polynomial interpolation technique and (2) *Multilinear map*. In these schemes, the authorized user need to store only a single key material irrespective of number of data items to which he has authorized access.**

I. Introduction

Cloud computing offers Infrastructure (storage, computing, networking) [1], Platform (development tools, runtime environment) [2] and Software (application software like google docs) as services. This prompts the organizations to outsource their storage and computing needs. Private cloud and Public cloud are the two categories of cloud infrastructure. The infrastructure in public cloud is owned and managed by the organization and is located on-premise, whereas in public cloud the service provider owns and manages the infrastructure and the data is located off-premise.

The customer (organization) can avoid the cost of building and maintaining private storage infrastructure by outsourcing the data to the cloud infrastructure which provides availability, reliability at relatively low cost. As evident, the public cloud poses significant security (confidentiality and integrity) risks to the customer's sensitive data.

The data owners often need to share their data selectively among the authorized users. Employing the data owner to enforce the access control hampers the system performance as the owner is required to involve in the query and the response phase. The owner should filter the response to avoid unauthorized users from accessing data. Also, enabling the server to enforce the authorization is not possible as the access control policy may be sensitive and cannot be disclosed or may depend on the content of the data and the servers must be trusted.

In several schemes, all the data items are encrypted with the single key and in some, each data items are encrypted with different keys. These methods have inherent demerits in the selective data access scenario. The scheme in [4] combines the multi-key techniques with the key derivation method.

Our proposed schemes are based on the intuitive idea that the key used to encrypt the data is regulated by the authorizations holding the data. There is no involvement of data owner and the unauthorized users (including the servers) cannot access the data if they do not have access permission. Each user stores only single key material irrespective of the number of data items to which he has authorized access.

978-1-4673-0298-2/12/$31.00 © 2012 IEEE

II. Notations

- Let $U = \{u_1, \ldots, u_n\}$ be the set of users and $D = \{d_1, \ldots, d_m\}$ be the set of data objects.
- A be the access control matrix of order is $n \times m$ whose entries are defined as below.

$$A[u_i, d_j] = \begin{cases} 0 & \text{if } u_i \text{ is not authorized to access } d_j; \\ 1 & \text{if } u_i \text{ is authorized to access } d_j. \end{cases}$$

- Let pac_{u_i} be the set of data objects to which u_i is authorized to access and acl_d be the set of users who are authorized to access the data object d.

III. Access control mechanism based on polynomial interpolation

Suppose $pac_{u_i} = \{d_{i_1}, \ldots, d_{i_k}\}$ be the set of data objects to which u_i is authorized to access. In other words, these are the data objects whose entries are 1's in the row corresponding to u_i in A.

For each user $u_i \in U$ perform the following.

- Construct a polynomial of degree $k-1$ as below.

$$P_{u_i}(x) = d_{i_k} + d_{i_1}(x) + d_{i_2}x^2 + \ldots + d_{i_{k-1}}x^{k-1}$$

- Evaluate $P_{u_i}(x)$ at $j = 1, \ldots, k$. Let $s_{u_i,d_j} = P_{u_i}(j)$ for $j = 1, \ldots, k$.
- Give the data share s_{u_i,d_k} to the user u_i. This share is the secret information which is *key* to the user u_i to access the data objects in pac_{u_i}.

- Let $0 < t < k-1$. Outsource t shares to the server. The parameter t is introduces to balance the number of shares to be kept on cloud server and *public catalog*.
- Store rest of the $k-t-1$ shares in *public catalog*.
- Suppose the rows corresponding to the users u_i and u_j have identical entries in A. Let l be the number of data objects to which these users have authorized access. Only one polynomial of degree $l-1$ is constructed on behalf of these two users and shares are distributed as defined above. By doing this, the number of shares produced will be l instead of $2l$. This is extended to any number of users with identical entries in A.

The user u_i accesses the data objects in pac_{u_i} as follows.

- Retrieves the shares s_{u_i,d_j} for $j = 1, \ldots, k-1$ from the server and the *public catalog*.
- The user constructs the polynomial $P_{u_i}(x)$ of degree $k-1$ using *Lagranges's polynomial interpolation method* with $k-1$ shares retrieved from the server and *public catalog* and one secret share that is with the user.
- The coefficients of the polynomial are the data objects.

IV. Multilinear map based Access control

The definitions and notations are referred from [3].

- **Setup(t,A):** Let $D = \{d_1, \ldots, d_m\}$ such that $\forall i, d_i \in G_2$ be the data set to be outsourced. Set

$$k = max_i\{|acl_{d_i}|\}$$

Run the randomized *multilinear map generator algorithm* $G = G(t,k)$ to obtain (Γ, g, q). Let $e : G_1^k \rightarrow G_2$ be the *k-multilinear map* defined by Γ. G_1 and G_2 are multiplicative groups. The generator of G_1 is g and q is the order of G_1 and G_2. We assume *discrete log problem* in G_1 is intractable. Set

$$l = min_i\{|acl_{d_i}|\}$$

Choose $k - l + 1$ numbers randomly from $\mathbb{Z}_q$. Let these numbers be $\{s_1, \ldots, s_{k-l+1}\}$. Output the public parameters $P_{pub} = (\Gamma, g, q, g^{s_1}, \ldots, g^{s_{k-l+1}})$.

- **User Key Generation,** $KGU(g,q)$: A user $u's$ public and private key pair is of the form $pk_u = g^\alpha$ and $sk_u = \alpha$, where $\alpha \leftarrow \mathbb{Z}_q \backslash \{0\}$. Let Alice be the owner of the data to be outsourced. The key pairs corresponding to the Alice (data owner) be $(sk_O, pk_O) = (a, g^a)$.
- **Outsourced Data Encryption Key Generation,** $EKG(A, k, l, d_i, sk_O, P_{pub})$: Suppose $|acl_{d_i}| = k$. W.l.o.g let $acl_{d_i} = \{u_{i_1}, \ldots, u_{i_k}\}$. Let $\{g^{\alpha_{i_1}}, \ldots, g^{\alpha_{i_k}}\}$ be corresponding public keys of the users in acl_{d_i}. Compute key K_i,

$$\begin{aligned} K_i &= (e(g^{\alpha_{i_1}}, \ldots, g^{\alpha_{i_k}}))^a \\ &= e(g, \ldots, g)^{a\alpha_{i_1} \ldots \alpha_{i_k}} \end{aligned}$$

Suppose $|acl_{d_i}| < k$ and $acl_{d_i} = \{u_{i_1}, \ldots, u_{i_j}\}, j < k$. Compute K_i as,

$$K_i = e(g^{\alpha_{i_1}}, \ldots, g^{\alpha_{i_j}}, g^{s_1}, \ldots, g^{s_{k-j}})^a$$

Output K_i.

- **Data Encryption,** $E(d_i)$**:** the data owner runs $EKG(A, k, l, d_i, sk_O, P_{pub})$ to get K_i. Encrypt d_i using K_i as,

$$C_i = K_i d_i.$$

- **Outsourced Data Decryption Key Generation,** $DKG(A, k, l, C_i, sk_{u_j}, P_{pub})$: W.l.o.g let $acl_{d_i} = \{u_{i_1}, \ldots, u_{i_{k-1}}, u_j\}$. Let $\{g^{\alpha_{i_1}}, \ldots, g^{\alpha_{i_{k-1}}}\}$ be the corresponding public keys of the users in $\{u_{i_1}, \ldots, u_{i_{k-1}}\}$. Note that C_i is the encrypted object of the data d_i. Suppose $|acl_{d_i}| = k$ Compute the key K_i as,

$$K_i = e(g^a, g^{\alpha_{i_1}}, \ldots, g^{\alpha_{i_{k-1}}})^{\alpha_j}$$

Suppose $|acl_{d_i}| < k$ and $acl_{d_i} = \{u_j, u_{i_1}, \ldots, u_{i_s}\}, s + 1 < k$. Compute K_i as,

$$\begin{aligned} K_i &= (e(g^a, g^{\alpha_{i_1}}, \ldots, g^{\alpha_{i_s}}, g^{s_1}, \ldots, g^{s_{k-s-1}}))^{\alpha_j} \\ &= e(g, \ldots, g)^{a\alpha_j\alpha_{i_1} \ldots \alpha_{i_j} s_1 \ldots s_{k-s-1}} \end{aligned}$$

Output K_i.

- **Data Decryption,** $D(C_i)$: Run Outsourced Data Decryption Key Generation, $DKG(A, k, l, C_i, sk_{u_j}, P_{pub})$ to get the key K_i used to decrypt C_i. Compute as below to get the d_i.

$$d_i = C_i / K_i$$

A. Security Analysis of the scheme

Security of the scheme using *multilinear map* is based on *Multilinear Diffie-Hellman Assumption* and intractability of solving *discrete log problem* in the groups considered [3]. We hope to build the scheme with grant & revoke of permission.

V. Conclusion

The mechanisms to enforce the access restrictions are proposed. The scheme based on the polynomial interpolation can be used in distributed cloud infrastructure. Also, an elegant scheme based on *multilinear map* is proposed to provide access control whose security depends on the *multilinear Diffie-Hellman assumption*. In both the schemes user has to store only single key material irrespective of the number of the data items to which he has access privilege.

Acknowledgment

This work was supported by *Ministry of Human Resource Development,* Government of India.

References

[1] "Amazon elastic compute cloud (amazon ec2)." [Online]. Available: http://aws.amazon.com/ec2/.

[2] "Microsoft windows azure platform." [Online]. Available: http://www.microsoft.com/windowsazure.

[3] D. Boneh and A. Silverberg, "Applications of Multilinear Forms to Cryptography," *Contemporary Mathematics, American Mathematical Society*, Vol. 324, pp. 71-90, 2003.

[4] E. Damiani and S. De Capitani di Vimercati and S. Foresti and S. Jajodia and S. Paraboschi and P. Samarati, "An experimental evaluation of multi-key strategies for data outsourcing," *in Proceedings of the 22nd IFIP TC-11 International Information Security Conference*, 2007.