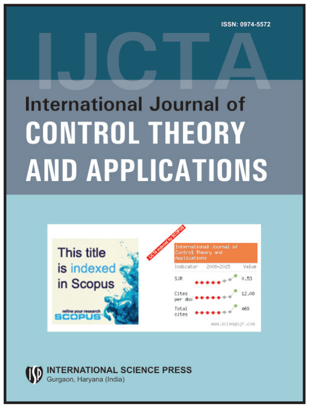

ISSN: 0974-5572

IJCTA

International Journal of

CONTROL THEORY

AND APPLICATIONS

This title is indexed in Scopus

INTERNATIONAL SCIENCE PRESS

Gurgaon, Haryana (India)

International Journal of Control Theory and Applications

ISSN : 0974-5572

© International Science Press

Volume 10 • Number 25 • 2017

Make: Manifold Array Keyword Encryption Over Cloud Data Search

S. Sujitha[a] and G.V. Sriramakrishnan[b]

[a]*M.Phil Research Scholar, Department of IT, Vel's University, Chennai. Email: suji_mca11@yahoo.com*
[b]*Asst. Prof, Department of IT, Vel's University, Chennai. Email: greatsri8@gmail.com*

Abstract: Executing query over multi encrypted data is a critically important technique in cloud computing, encryption before outsource is a basic solution to protecting user data privacy in the unasserted cloud server environment. Many security search method were proposed on the single contributor scenario, where the outsourced dataset of the dataset are encrypted and managed by a single owner, typically based on multiple cryptography techniques. In this article, we focus on a different yet more challenging scenario where the outsourced dataset can be contributed from multiple owners and are searchable by multiple user case. Inspired by attribute-based encryption (ABE), we present the new method of manifold array keyword encryption search (MAKE).

Manifold array keyword encryption search (MAKE) which enables scalable fine-grained search authorization. Our scheme allows multiple owners to encrypt and outsource data to the cloud server independently. Users can generate their own search capabilities without relying on an always online trusted authority. Fine-grained search authorization is also implemented by the owner-enforced access policy on the index of each file. In future, by incorporating re-encryption schemes techniques, we are able to delegate heavy system update workload during user revocation to the resourceful semi-trusted cloud server. We formalize the security definition and prove the proposed MAKE scheme selectively secure against chosen-keyword attack. To build confidence of data user in the proposed secure search system, we also design a search result authentication system. Finally, performance evaluation shows the efficiency of our scheme.

Keywords: Manifold array keyword encryption search, Multi Encryption scheme, Unsecured Cloud data, Cloud Computing.

1. INTRODUCTION

Nowadays, as an emerging and efficient computing model, cloud computing has attracted widespread attention and support in many fields. In the cloud computing environment, many services such as resource renting, application hosting, and service outsourcing show the core concept of an on-demand service in the IT field. In recent years, many IT tycoons are developing their business cloud computing system, e.g. Amazon's EC2, S3, Google AppS and Microsoft Azure etc. Cloud computing can provide flexible computing capabilities, reduce costs and capital expenditures and charge according to usage.

Although the cloud computing standards brings many benefits, there are many unavoidable security problems caused by its inbuild characteristics such as the dynamic complexity of the cloud computing environment, the openness of the cloud platform and the high concentration of resources. One of the important problems is how to ensure the security of user data. Security problems, such as data security and privacy protection in cloud computing, have become serious obstacles which, if not appropriately addressed, will prevent the development and wide application of cloud computing in the future. In 2016, a few serious security incidents with cloud service occurred at many IT companies, including Google, Microsoft, and Amazon. These incidents affected the information services to millions of consumers. Therefore, it is important that security problems in cloud computing receives significant attention.

In cloud computing, users store their data files in cloud servers. Thus, it is crucial to prevent unauthorized access to these resources and realize secure resource sharing. In traditional access control methods, we generally assume data owners and the storage server are in the same secure domain and the server is fully trusted. However, in the cloud computing environment, cloud service providers may be attacked by malicious attackers. These attacks may leak the private information of users for commercial interests as the data owners commonly store decrypted data in cloud servers. How to realize access control to the encrypted data and ensure the confidentiality of data files of users in an untrusted environment are problems that must be solved by cloud computing technologies and applications. Moreover, since the number of users is large in a cloud computing environment, how to realize scalable, flexible and fine-grained access control is strongly desired in the service-oriented cloud computing model.

This article featured on the problem of search over encrypted data, which is an important enabling procedure for the encryption before outsourcing privacy protection among cloud computing, or in general in any networked information system where servers are not fully trusted. Much work has been done, with majority focusing on the single contributor scenario, i.e., the dataset to be searched is encrypted and managed by a single entity, which we call owner or contributor in this paper. Under this setting, to enable search over encrypted data, the owner has to either share the secret key with authorized users or stay online to generate the search trapdoors, i.e., the "encrypted" form of keywords to be searched, for the users upon request. The same symmetric key will be used to encrypt the dataset (or the searchable index of the dataset) and to generate the trapdoors. These schemes seriously limit the users' search flexibility.

Consider a file sharing system that hosts a large number of files, contributed from multiple owners and to be shared among multiple users (e.g., 4shared.com, mymedwall.com). This is a more challenging multi-owner multi-user scenario. How to enable multiple owners to encrypt and add their data to the system and make it searchable by other users? Moreover, data owners may desire fine-grained search authorization that only allows their authorized users to search their contributed data. By fine-grained, we mean the search authorization is controlled at the granularity of per file level. Symmetric cryptography based schemes are clearly not suitable for this setting due to the high complexity of secret key management. Although authorized keyword search can be realized in single-owner setting by explicitly defining a server-enforced user list that takes the responsibility to control legitimate users search capabilities search can only be executed by the server with the assistance of genuine users with duplicate keys on the user list, these method did not realize secured search authorization and thus are unable to provide differentiated access privileges for different users within a dataset. Asymmetric cryptography is better suited to this dynamic setting by encrypting individual contribution with different public keys. However, extending such user list approach to the multi-owner setting and on a per file basis is not trivial as it would impose significant scalability issue considering a potential large number of users and files supported by the system. Additional challenges include how to handle the updates of the user lists in the case of user enrollment, revocation, etc., under the dynamic cloud environment.

In this paper, we address these open issues and present a Manifold array keyword encryption search (MAKE) over encrypted cloud data with efficient user revocation in the multi-user multi-data-contributor scenario. Specifically, the data owner encrypts the index of each file with an access policy created by him, which defines what type of users can search this index. The data user generates the trapdoor independently without relying on an always online trusted authority (TA). The cloud server can search over the encrypted indexes with the trapdoor on a user's behalf, and then returns matching result if and only if the user's attributes associated with the trapdoor satisfy the access policies embedded in the encrypted indexes. We differentiate attributes and keywords in our design. Keywords are actual content of the files while attributes refer to the properties of users. The system only maintains a limited number of attributes for search authorization purpose. Data owners create the index consisting of all keywords in the file but encrypt the index with an access structure only based on the attributes of authorized users, which makes the proposed scheme more scalable and suitable for the large scale file sharing system.

2. RESEARCH METHODOLOGY

2.1. Searching Models

A. **Plaintext Searching Mode:** In practice, to realize effective data retrieval on large amount of documents, it is necessary to perform relevance ranking on the results. Ranked search can also significantly reduce network traffic by sending back only the most relevant data. In ranked search, the ranking function plays an important role in calculating the relevance between files and the given searching query.

The most popular relevance score is defined based on the model of TF IDF, where term frequency (TF) is the number of times a term (keyword) appears in a file and inverse document frequency (IDF) is the ratio of the total number of files to the number of files containing the term. There are many variations of TF IDF-based ranking functions.

B. **Ciphertext Searching Model:** Due to the special background of cloud computing, unlike traditional plaintext information retrieval, there are usually three entities in cloud data retrieval as shown in Figure 1.

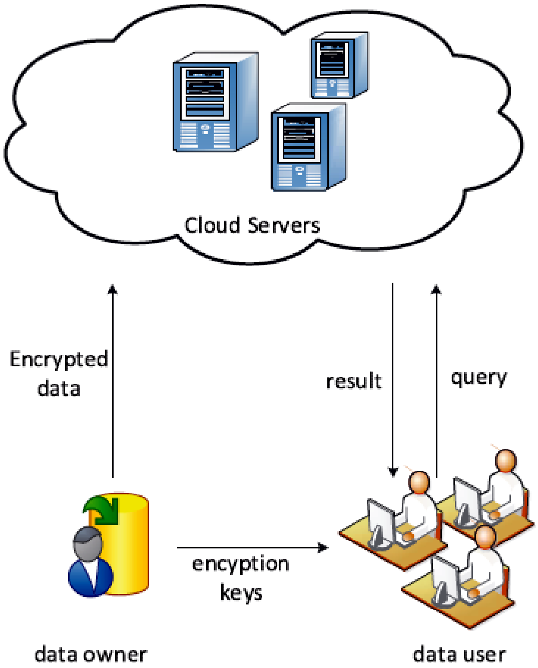

Figure 1: Cipertext Searching Model

Data owner, remote cloud server and users. A data owner can be an individual or a corporation, i.e., it is the entity that owns a collection of documents $D_c = \{D_1, D_2, ..., D_{Nd}\}$ that it wants to share with trusted users. The keyword set is marked as $W = \{w_1, w_2, \ldots, w_{Nw}\}$.

2.2. Keyword Search over Encrypted Data

A. **Secret Key Versus Public Key:** In Existing design the first searchable encryption scheme to enable a full text search over encrypted files. Since this seminal work, many secure search schemes have been proposed to boost the efficiency and enrich the search functionalities based on either secret key cryptography (SKC) or public-key cryptography (PKC). Also they presented an efficient single keyword encrypted data search scheme by adopting inverted index structure. To enrich search functionalities, they proposed the first privacy-preserving multi-keyword ranked search scheme over encrypted cloud data using "coordinate matching" similarity measure. Later on, a secure multi-keyword text search scheme in the cloud enjoying more accurate search result by "cosine similarity measure" in the vector space model and practically efficient search process using a tree based secure index structure. Compared with symmetric search techniques, PKC-based search schemes are able to generate more flexible and more expressive search queries. The first PKC-based encrypted data search scheme supporting single keyword query. The scheme supports search queries with conjunctive keywords by explicitly indicating the number of encrypted keywords in an index.

B. **Authorized Keyword Search:** From the Study on references, multiple users authorization over search capabilities, user authorization should be enforced. But these SKC-based schemes only allow one data contributor in the system. Where public-key setting presented a conjunctive keyword search scheme in multi-user multi-owner scenario. But this scheme is not scalable under the dynamic cloud environment because the size of the encrypted index and the search complexity is proportional to the number of the authorized users, and to add a new user, the data owner has to rewrite all the corresponding indexes. By exploiting hierarchical predicate encryption a file-level authorized private keyword search (APKS) scheme over encrypted cloud data proposed. However, it makes additional communication cost, since whenever users want to search, they have to resort to the attribute authority to acquire the search capabilities. Moreover, this scheme is more suitable for the structured database that contains only limited number of keywords. The search time there is proportional to the total number of keywords in the system, which would be inefficient for arbitrarily-structured data search.

2.3. Attribute-Based Encryption

There has been a great interest in developing attribute based encryption due to its fine-grained access control property. First key policy attribute-based encryption scheme, where ciphertext can be decrypted only if the attributes that are used for encryption satisfy the access structure on the user private key. Under the reverse situation, CP-ABE allows user private key to be associated with a set of attributes and ciphertext associated with an access structure. CP-ABE is a preferred choice when designing an access control mechanism in a broadcast environment. Since the first construction of CP-ABE, many works have been proposed for more expressive, flexible and practical versions of this technique.

3. DESIGN OF MANIFOLD ARRAY KEYWORD ENCRYPTION

A. **System Design:** The cloud service provider manages a cloud to provide data storage service. Data owners encrypt their data files and store them in the cloud for sharing with data consumers. To access the shared data files, data consumers download encrypted data files of their interest from the cloud and then decrypt them. Each data owner/consumer is administrated by a domain authority. A domain authority is managed by its parent domain authority or the trusted authority. Data owners, data consumers, domain authorities, and the trusted authority are organized in a hierarchical manner as shown below:

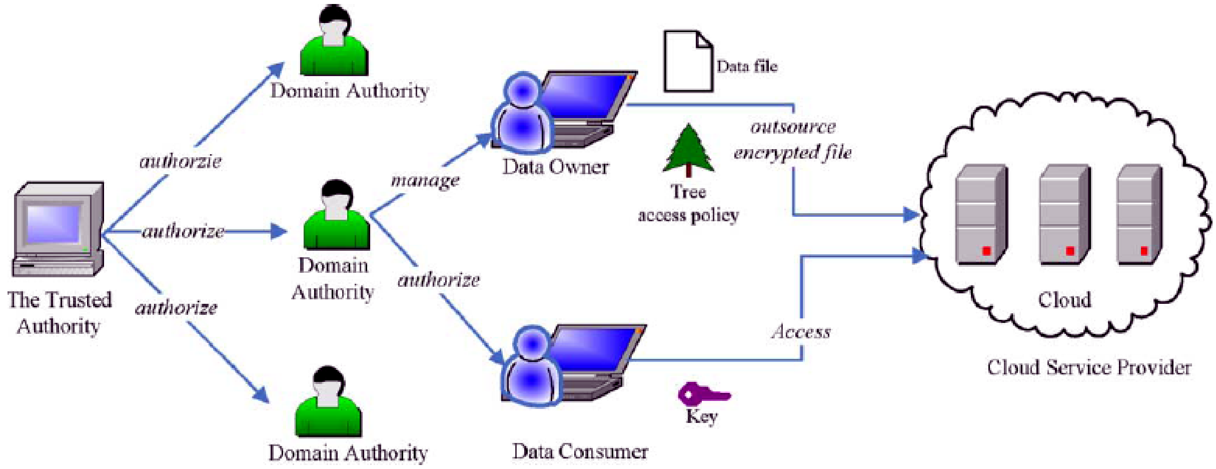

Figure 2: Design of MAKE scheme

The trusted authority is the root authority and responsible for managing top-level domain authorities. Each top-level domain authority corresponds to a top-level organization, such as a federated enterprise, while each lower-level domain authority corresponds to a lower-level organization, such as an affiliated company in a federated enterprise. Data owners/consumers may correspond to employees in an organization. Each domain authority is responsible for managing the domain authorities at the next level or the data owners/consumers in its domain.

In our system, neither data owners nor data consumers will be always online. They come online only when necessary, while the cloud service provider, the trusted authority, and domain authorities are always online. The cloud is assumed to have abundant storage capacity and computation power. In addition, we assume that data consumers can access data files for reading only.

B. **Security Model:** We assume that the cloud server provider is untrusted in the sense that it may collude with malicious users (short for data owners/data consumers) to harvest file contents stored in the cloud for its own benefit. Addition, we assume that communication channels between all parties are secured using standard security protocols, such as SSL.

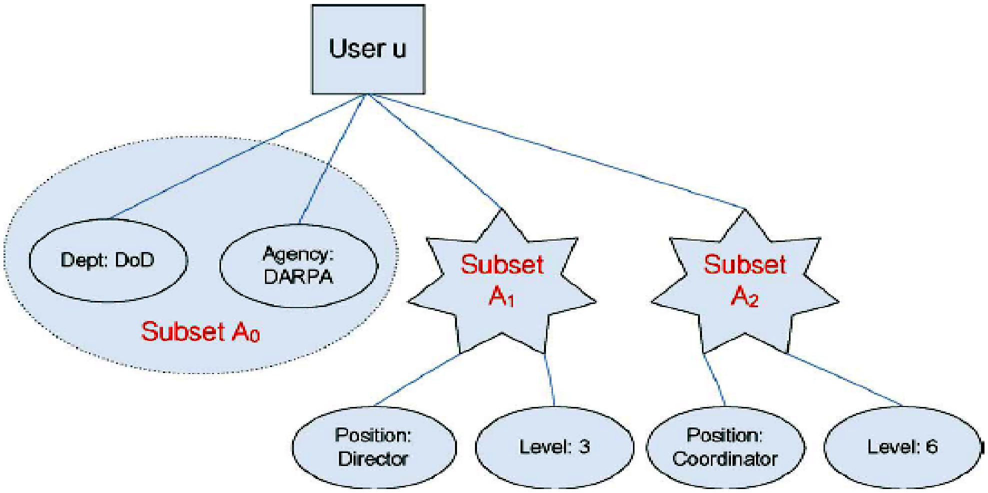

Figure 2: Design of Security Model

C. **Threat:** We consider the cloud server was secured but its overwriteable, which is also employed by related works on secure search over encrypted data. We assume that the cloud server honestly follows the designated protocol, but curiously infers additional privacy information based on the data available to him. Furthermore, malicious data users may collude to access files beyond their access privileges by using their secret keys. As we delegate most of the system update workload to the cloud server, we assume that the cloud server will not merge with the revoked malicious users to help them gain unauthorized access privileges.

The entire design was focused to achieve the following results and security levels,

1. Keyword search based on secured authorization
2. Effective revocation schemes over unanimous user
3. Adding feature to collaborate more data contributors with much data users
4. Validating search results
5. Enhanced security goals which discussed on literature survey on reference study.

4. EXPERIMENTAL ANALYSYS AND RESULTS

We proposed the Manifold array keyword encryption search (MAKE) technique to achieve scalable fine-grained authorized keyword search over encrypted cloud data supporting multiple data owners and data users. Specifically, for each file, the data owner generates an access-policy-protected secure index, where the access structure is expressed as a series of AND gates. Only authorized users with attributes satisfying the access policies can obtain matching result. Moreover, we should consider user membership management carefully in the multiuser setting. A native solution is to impose the burden on each data owner. As a result, data owner is required to be always online to promptly respond the membership update request, which is impractical and inefficient. By using proxy re-encryption, the data owner can delegate most of the workload to the cloud without infringing search privacy.

The system level operations include following functions:

1. System Arrangement
2. User Registration
3. Secure Index Generation
4. Wormhole Creation
5. Keyword Search
6. User Revocation.

Here *i* listed some important functions of our system which contains many sub functions capable of invoke users.

System Arrangement: Here we generate the random key using some existing algorithm and share the key securely between owners and users.

User Registration: When receiving a registration request from a new unknown user trusted authority first selects a random key as a new component. Then, the trusted authority generates a new public keyword component and publishes it with its signature. After that, the KeyGen algorithm is called to create secret key for this user. For every secret key trusted authority select randomly.

Secure index generation: Before outsourcing a file to the cloud server, the data owner calls Index algorithm to generate a secure index for this file. In particular, set of value to be given an access for each secret key and public key. For each secret key some attribute and a keyword, the data owner sets keywords where without loss of generality, attribute is assumed to be positive encrypted index.

Wormhole Creation: Every genuine user in the system is able to generate a wormhole for any keyword of interest. Specifically, user selects random secret key for the same keyword which in secure index generation phase.

Keyword Search: Upon receipt of a wormhole and the user identity the client server which finds out if Keyword exists on the user list of the target dataset. If not, the user is not allowed to search over the dataset. Otherwise, the client server continues the Search algorithm with the input of wormhole encrypted index from the user list. We call this process dataset search authorization. We can achieve Manifold array keyword encryption search (MAKE) by data-owner-enforced manifold array based access structure on the index of each file. The search complexity is linear to the number of attributes in the system rather than the number of authorized users. Hence, this one-to-many authorization mechanism is more suitable for a large scale system, such as cloud. Moreover, the dataset search authorization by using a per-dataset user list may accelerate the search process, since the cloud server can decide whether it should go into a particular dataset or not. Otherwise, the cloud server has to search every file at rest.

User Revocation: To revoke a user from current system, we re-encrypt the secure indexes stored on the server and update the remaining genuine users secret keys. Note that these tasks can be delegated to the client server using proxy re-encryption technique so that user revocation is very efficient. In particular, the trusted authority adopts the reversible key generation algorithm to generate the re-encryption key set.

After receiving secret key from the trusted authority, the server checks whether the version number is equal to current version of the system. If not, it discards this re-encryption key set. Then, the server calls the Re Encryption Index algorithm to re-encrypt the secure indexes in its storage with valid secret key. Furthermore, the server is able to update the remaining genuine users secret keys by the Re-generative Key algorithm. Suppose that secret key is a list stored on the cloud server containing all the partial secret keys of all the genuine users in the system. Partial Secret key is defined as invalid. Note that the cloud server cannot generate a valid warmhole with partial secret key. Finally, the server may eliminate ID information of the revoked user from all the corresponding user lists.

To handle file index update efficiently, we could adopt the lazy re-encryption technique which mentioned in the reference paper. The cloud server stores the re-encryption key sets and will not re-encrypt indexes until they are being accessed. Specifically, the cloud server could “aggregate”

5. RESULTS

Data users may desire the authenticated search result to boost their confidence in the entire MAKE search process, especially when the result contains errors that may come from the possible storage corruption, software malfunction, and intention to save computational resources by the server, etc. We are able to assure data user of the authenticity of the returned search result by checking its correctness, completeness and freshness. The main idea of the verification scheme is to allow the cloud server to return the auxiliary information containing the authenticated data structure other than the final search result, upon which the data user is capable of doing result authenticity check. In what follows, we elaborate on the concrete scheme.

In order to check if he is a genuine user for a particular dataset, the data owner can simply sign the corresponding user list or to avoid disclosing other user’s membership information, the trusted authority may generate the keywords secret key value for each authorized user. The data owner can insert the key values into a filter based on these users membership, and then signs it to filter. Next, the data owner prepares another filter for the keywords appearing in the dataset to enable the data user quickly find out the existence of the intended keyword. Specifically, the trusted authority generates a secret key and gives it to the data owner. He then encrypts it with manifold array encryption for each genuine user.

Coming to search part the cloud server returns the search result along with the auxiliary information for result authenticity check later by the data user. The auxiliary information includes all the user list filters of the datasets stored on the server, the keyword filters of the datasets that the user is authorized to access, the file list for the intended keyword if the search result contains files from this dataset, if the search result does not contain files from this dataset, it is not necessary to return the corresponding file list.

We create an authenticated sharing structure using encryption technique to organize the outsourced data in the server. The data user can search over this structure to verify the returned search result, since all the signatures can only be generated by data contributors. By checking verified filters, the user is assured of the existence and integrity of all the returned files, and search result does not exclude any qualified matching files. Hence, we can achieve the verification design goals. Freshness can be simply realized by adding time stamp into the corresponding signatures. Thus, we make the MAKE scheme verifiable and the authenticity of the returned search result is guaranteed.

6. CONCLUSION

In this article, we design the Manifold array keyword encryption search over the cloud environment, which enables standard and secured encrypted data search supporting multiple data owners and data users. Compared with existing authorized keyword search scheme, our scheme could achieve system standard and security at the same time. Different from search scheme with predicate encryption, our scheme enables a flexible authorized keyword search over un structured data. This proposed scheme is better suited to the cloud outsourcing model and enjoys efficient user revocation. On the other hand, we make the whole search process verifiable and data user can be assured of the authenticity of the returned search result. We also formally prove the proposed scheme semantically secure in the selective model.

Acknowledgment

This work was done on fulfillment of my final year project work on guidance of staff members of Department of Computer Science and Engineering, Vel’s University, Chennai. Thanks to Mr. XXXX For valued guidance to complete this Journal.

REFERENCES

[1] P. Mell and T. Grance, “The NIST definition of cloud computing,” NIST special publication, 800(145): 7, 2011.

[2] S. Subashini and V. Kavitha, “A survey on security issues in service delivery models of cloud computing,” Journal of Network and Computer Applications, 34(1): 1-11, 2011.

[3] D. Song, D. Wagner and A. Perrig, “Practical techniques for searches on encrypted data,” Security and Privacy, 2000. Proceedings. 2000 IEEE Symposium on. IEEE, pp. 44-55, 2000.

[4] D. Boneh, G. Di Crescenzo and R. Ostrovsky, “Public key encryption with keyword search,” Advances in Cryptology-Eurocrypt, 2004. Springer Berlin Heidelberg, pp. 506-522, 2004.

[5] Y.-C. Chang and M. Mitzenmacher, “Privacy preserving keyword searches on remote encrypted data,” Applied Cryptography and Network Security. Springer Berlin Heidelberg, pp. 442-455, 2005.

[6] W. Sun, S. Yu, W. Lou, Y.T. Hou, and H. Li, “Protecting your right: Attribute-based keyword search with fine-grained owner enforced search authorization in the cloud,” in Proc. IEEE Conf. Comput. Commun., 2014, pp. 226–234.

[7] S. Yu, C. Wang, K. Ren, and W. Lou, “Achieving secure, scalable, and fine-grained data access control in cloud computing,” in Proc. IEEE Conf. Comput. Commun., 2010, pp. 1–9.

[8] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, “Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption,” IEEE Trans. Parallel Distrib. Syst., Vol. 24, No. 1, pp. 131–143, Jan. 2013.

[9] S. Kamara and K. Lauter, “Cryptographic cloud storage,” in Proc. 14th Int. Conf. Financial Cryptography Data Security, 2010, pp. 136–149.

[10] R. Buyya, C. Shin Yeo, J. Broberg, and I. Brandic, “Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility,” Future Generation Comput. Syst., Vol. 25, pp. 599–616, 2009.

[11] E. Angel Anna Prathiba and B. Saravanan “HASBE for Access Control by Separate Encryption/Decryption in Cloud Computing” International Journal of Emerging Trends in Electrical and Electronics (IJETEE) Vol. 2, Issue. 2, April-2013.

[12] Sultan Ullah, Zheng Xuefeng and Zhou Feng “TCLOUD: A Multi – Factor Access Control Framework for Cloud Computing” International Journal of Security and Its Applications Vol. 7, No. 2, March, 2013.

[13] Rajanikanth aluvalu, lakshmi Muddana, “A Survey on Access Control Models in Cloud Computing”-in Springer International Publishing, Advances in Intelligent Systems and Computing 337, DOI: 10.1007/978-3-319-13728-5_7.

[14] Sonam Chugh, Sateesh Kumar Peddoju “Access Control Based Data Security in Cloud Computing” International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 2, Issue 3, May-Jun 2012.

[15] Jawahar Thakur, Nagesh Kumar “DES, AES and Blowfish: Symmetric Key Cryptography Algorithms Simulation Based Performance Analysis” International Journal of Emerging Technology and Advanced Engineering (IJEATE)- (ISSN 2250-2459, Volume 1, Issue 2, December 2011.