Attribute Based Data Distribution In Cloud Using Secure Deduplication

Supriya[1], Leena Giri G[2]
[1]Dept of Computer Science and Engineering
[2]Associate Professor, Dept of Computer Science and Engineering
[1, 2] Dr. AIT, Bengalure

Abstract- *Attribute Based Encryption (ABE) is extensively used in storing secured data with process of deduplication in cloud computing. But ABE does not support secure deduplication, wastes space and decreases the network bandwidth. Cipher text Policy Attribute Based Encryption (CP-ABE) procedure assures an information distribution in the conception of cloud computing. CP-ABE eliminates multiple copies of same data to achieve secure deduplication in cloud. The related data to be shared is under the full authority of the Data owner in accessing policy. Nevertheless, CP-ABE has less ability to security chances which is familiar as key escrow issue, through which the hidden key referring the customers should be delivered through trustworthy keys authorities. Parallel to, many of the current CP-ABE strategy doesn't support the attributes with inconsistent state. In this proposed work, the attribute-based data distributing strategy is considered in finding the solution to key escrow problem and at the same time to also enhance the eloquence or articulation of the attributes. With this, the resulting strategy is found to be most favourable to the applications of cloud computing. An enhanced two-party key providing protocol is proposed to assure that either key authority or cloud service contributor can co-operate with the entire hidden keys of a customer personally. Additionally, in this proposal a new approach of attributes which includes weightage, being presented to improve the interpretation of the attributes that notably expands definitions from doubled to random state, and it reduces complications of accessing policies. Accordingly, together with complexity of encrypted data and storage cost for a cipher text are reassured. The efficiency determination and security confirmation shows that the proposed strategy is capable of performing more efficiently with secured data distribution in cloud computing.*

Keywords- ABE, Storage, Deduplication, CP-ABE

I. INTRODUCTION

With the desire of accessing huge amounts of information which is communicated across the internet world-wide the Cloud repository is emerging as an assuring solution for lending universal, appropriate accessing of information. Nowadays, millions of people share personal data, such as photos and videos, with their relatives or friends with the help of social network websites that is based on cloud repository on everyday basis. Application users are also attracted with cloud repository because of to its various advantages, which includes cheaper cost, greater ability, and improved assets usage.

Research on Cloud computing is newly progressing computation technology or analogy based on advantages and using up of computation resources. Cloud computing includes utilizing class of software application network and remote servers which allows consolidated data repository to provide accessing online the system resources or services. Clouds are classified into private, public and hybrid. Cloud suppliers generally uses "pay as you go" plan.

Organisation of paper: This paper is broadly divided into 6 sub sections: section 2 presents the information collected during literature survey, section3 represents the proposed methodology, section 4 presents the outcome and discussion, section 5 presents the system study, section 6 holds the conclusion and future scope.

II. LITERATURE SURVEY

A literature survey is a process of collecting the information of that which has been already printed on a subject matter or the discussion done by the authorized researchers and scholars. Many of the researchers have studied and evaluated the problems based on cipher text attribute based encryption. In all these studies the information giver defines an accessing equation so a set of customers could perform decryption if the association of the attestations are having successful equation. Encryption is employed by hidden keys splitting after that sharing encryption keys through an allocated public keys.

A. Balu and K. Kuppusamy[2] have proposed Cipher text-Policy Attribute-Based Encryption (CP-ABE) that admits for encrypting the data below an accessing policy, defined as a logical association of the attributes. These cipher texts were decrypted through anybody accompanying group of features which compensate the accessing policies. They have proposed

a Cipher text-Policy Attribute-Based Encryption that includes development upon current hidden distributed technique known Linear Integer Secret Sharing Scheme (LISS). With the proposal, the encryption specifies accessing policies condition to LISS matrix M, across the system features. The proposal was relatively protected below the Decisional Bilinear Diffie–Hellman (DBDH) acceptance [1]. D. Boneh, B. Lynn, and H. Shacham, author introduces a fleeting signature scheme that is established on the Computational Diffie–Hellman acceptance on valid elliptical and hyper elliptical curves. For common privacy parameters, the signature limit was relating to half of that of a DSA signature with identical position of security. These short signature schemes were constructed for systems where the signatures are printed in by a human or were sent across low-bandwidth channels. Chase [3] proposed multi-authority ABE scheme by employing a theory of reliable central authority (CA) and global identifiers (GID). Though, CA development has the authority to perform decryption of single cipher text that appears one way or the other inconsistent to real aim sharing the authority across most probably entrusted power. Additionally, in development, usage of viscosity GID granted power to incorporate their data to construct whole description along every of the customer's features that unwontedly adjust with the confidentiality of the customer. The authors have given a solution that diminishes trustworthy significant authorities, and assures the customers secrecy by restricting the control with combining the data on appropriate users, therefore proving ABE more functional in the proceedings.

J. Bethencourt, A. Sahai, and B. Waters, have presented system for understanding the complex accessing controls on encrypting the data which they have named as Cipher text-Policy Attribute-Based Encryption. By employing method of encrypting information is retained private though the stored server is entrusted; additionally, these techniques were safe opposite to destruction violations. The former Attribute Based Encryption systems employed the features to define the encryption information to develop the policy with the customer keys; even though author prescribed the features of system is employed that define the customer attestations, and group encrypted information defines a policies for one that decrypts the information. Therefore the procedures are hypothetically nearer with the conventional accessing controlling approaches for example Role-Based Accessing Controls (RBAC) [4].M. Chase and S. S. Chow, Attribute based encryption (ABE) [7] defines the decryption agility established on the customers features. In a multiple-authority ABE schemes, the various attribute-authority watches carefully the various group of the attributes and problem equivalent decrypting the keys to the users, and encryption demands the customers, should acquire the keys for suitable features from every control ahead of message decryption.

Liu et al, [6] conferred a condensed accessing authority scenario along features hierarchies that is constructed in different instances, the features were split into various tasks that accomplishes condensed accessing authority through the hierarchical features; instead features expresses isolated binary state. Following, Fan et al, [7] recommended an inconsistent-condition ABE to clarify about problem managing effective solution.

2.1 Existing system:

The Data Owner (DO) is generally wishing to deposit huge volume of information inside cloud for retaining the price in managing narrow information. Beyond any information securing techniques, the Cloud Service Provider (CSP), nevertheless, wholly benefit the accessing every information of customer. Hence, could bring the possible security exposure to the customer, as CSP has to adjust information in consideration of economic gains. Correspondingly, by what method we have to secure the data effectively allot the customer information is the main difficult phase that is faced in clouds computational schemes.

Foremost, every customer hidden key requires has been circulated through wholly trustworthy key authorities (KA). Therefore it leads to a protection danger named as keys escrow issue. When the hidden key of a systems customer is known, KA decrypts every customers cipher text that endures in detail opposite to the privacy of customer. Later, fluency about the features group is a responsibility. Till we get to know, best of current CP-ABE scenario uniquely describes the doubled condition above the characters, considering example, "1 - satisfy" and "0 – dissatisfies", instead of managing along inconsistent-condition attributes.

The proposed work of traditional attributes is splitted to 2 divisions -the attributes or features along with related values. Considering example, the conventional features is defined as {"Professors", "Engineers", "Doctors"}. Features enhanced is defined as: {Careers: "Professors", "Engineers", "Doctors"}, in which "Careers" symbolises an attribute and "Doctors", "Professors" and "Engineers" symbolises the appraisal of an feature "Careers". Relatively, value of computation considering the features are of much higher cost compared to particular earlier scenarios below identical various features.

III. PROPOSED METHODOLOGY

One of the methods that are employed in solving the repetition of data's is data deduplication. The method of Deduplication is mainly employed in cloud computational server for the space reduction of server. To restrict the unauthorized accessing of information and build repeated information upon cloud encryption method for encrypting the information earlier to its storing on cloud server. Cloud repository generally stores functional-exploratory information and methods; therefore excessive protection requirement will be the single answer to maintain powerful and faithful correlation among cloud customers and cloud resource givers.

Let's assume orderly structures of a university, which involves classification of tutors into assistant lecturers, lecturers, associated professors and professors. Here we have distributed the weightage of attributes for every classified kind of lecturers as A, B, C, and D. Hence, the features are designated like "Lecturer: A", "Lecturer: B", "Lecturer: C" and "Lecturer: D", accordingly.

With aforementioned instance, it should indicate along single feature that will have fair various weightages. As specified, the attribute is in inconsistent condition, for example "Lecturer: assistant lecturers, lecturers, associate professors, professors".

3.1 Framework of our Approach

Proposed a protocol for enhancing the key transmission in finding solution for key escrow obstacles of CP-ABE in cloud computation. A feature-related information distribution instance for cloud computation utilization that denotes a cipher text-policies weightage ABE scheme for eliminating escrow problem (CP-WABE-RE). This finds a solution for two kinds of issues: keys escrow and inconsistent-state features interpretation. Therefore improvements in proposed area are, individually have proposed an enhanced key circulating prototype in finding solution for key escrow issue of CP-ABE using cloud computation.

The prototype restricts the KA and CSP across getting to know every one's main hidden keys such that no one can recreate full hidden keys of customer personally. Wholly trustworthy KA will be virtually-trusted. Information security and isolation is presented as features that enhance expressiveness about the features. The weightage of features conveys not only inconsistent-conditioned attributes rather than of the conventional doubled state, it reduces complications with accessing policies. The amount used for storing cipher-text along with manipulating the complexities in encrypting will be minimized. In addition they express huge feature area compared below identical condition.

3.2 System Architecture

Structure and systems prototype of CP-WABE-RE strategy employed using cloud computation is expressed below, that application is composed with 4 variety units such as: Users, DO, KA, and CSP. Here, we also give the complete explanation of CP-WABE-RE strategy. KA is abbreviated as Key Authority. In a cloud system it is a semi-trusted unit. The name itself specifies that KA is very trustworthy-but-surprising, that can truly accomplish the given tasks and obtain the right outcome. Therefore, it collects sensitive contents as many as possible. The entity will be in charge for enrolment of the users in the cloud systems. Meantime, it not only produces many parts of the systems parameters, it also produces many parts of the secrets key for every user.

CSP abbreviated as Cloud Service Provider as supervisor inside cloud servers units that is virtually-reliable that enhances most of the works, for example storage of data computing and transferring the data. To find the solution to problem of key escrow, it produces both divisions of systems parameters and to every user with a secret key. DO abbreviate as Data Owners are data holder of files that is deposited in the cloud system. These are responsible for describing the accessing structures and execution of the data encrypting operations. They will upload the produced cipher text to CSP.

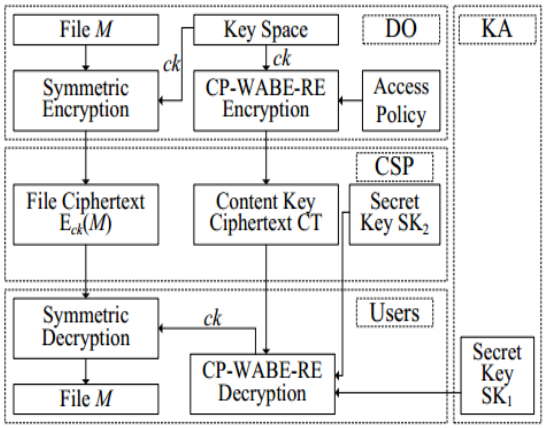

Figure 1. CP-WARE-RE Strategy.

User who requests for accessing the cipher text deposited in the cloud system. These users will have the authority to download cipher text and executes with the equivalent decrypting operations.

Here we discuss the structure of CP-WABE-RE systems, which is inclusion of five process: initializing of the system, creating a new file (encrypting the data), authorization of recent customer (generating customer key), accessing information file (decrypting the file), with deleting of data file.

3.3 Cipher Text Policy Attribute based Encryption

In this proposed work, we foremost specify the designing of a cipher text policy attribute based encryption (CPABE) for tackling the problem; this provides initial building with these instances.

Customer's personal key is correlated having casual numerals features that are conveyed in form of strings. As result, when the user performs an encryption of message in our system, it provides correlated accessed structures over the attributes strings.

The cipher text will be decrypted only by the end user if the user's attributes flows along with the cipher text's accessed structures. At numerical level, accessing the structures in system is defined by a monotone "access trees", in which the nodes of accessing structures are combined of thresholding gates and the leaves describing the features.

A cipher text-policy attribute based encryption scheme incorporates four step algorithms: Setup, Encryption, Key Generation, finally Decryption.

Setup (λ,V): This step yields secured parameters and attributes of universal specification inputs and deliver the public key P and master key M as outputs.

Encryption (P, m, A): This step yields the public key P, a message m, and accessing structure A among the universal features as an inputs. This encrypts m and produces cipher text CT such that customer who owns group of features which gratifies the accessing structures decrypts the message. We have assumed that cipher text absolutely restrains A.

Key Generating (M,F): This step yields master key M and group of features F which defines the key. Private key SK is obtained as output.

Decryption (P,CT,S): The decryption algorithm yields the public parameters PK, a cipher text CT, which contains an access policy A, and a private key S, which is a personal key for group S of features as inputs. If group S of features gratifies the accessing structures A the method performs decryption of cipher text, returns message M.

A protective model for cipher text-policy ABE schemes is proposed. In CP-ABE, cipher texts were defined along accessing structure and the personal keys along features. It is arranged in such a way that security defining the adversary selected to be challenge on encrypting an accessing structure $A*$ and request with any personal key S as it implies that S never assures $A*$.

3.4 Data Deduplication

The most significant method which is used to solve the problem of repetition of uploading of same named files is data deduplication. The technique of deduplication normally uses the cloud server for minimizing the server space. Saving storage space, saving time and increasing the network bandwidth, but several process have been this same concept for deduplication. In this case, if two customers upload the files with same name, the cloud server discriminates identical cipher texts and saves single copy of file. This process indicates that authentication method is available in some issue for security purpose. This process ensures secured deduplication. The owner requires utilizing data from cloud and sharing it with users has to possess some attestations. Attribute Authority provides each customer a decrypting key (for public user) related with users group of features and encryption key (for private user), each time data owner uploads file to cloud for storing purpose. Most of schemes have been proposed to provide data encryption, while still benefiting from a deduplication technique. Every user gets secure key form admin for security purpose. User cannot takes any key and cannot download chipper text file until take the permission from data owner and cloud. After taking the key from data owner as well as cloud they can download only encrypted data. Every detail is managed and maintained by Attribute authority. To keep the representation concise, we employ combination of encrypted and decryption key used to download the corresponding file.

Data deduplication works with the following Figure 2. Firstly separates the input data given into pieces or chunks. Secondly for each and every block of the data file a hash value is calculated. Thirdly employed all these calculated values to identify if any of other file has the same data that has already been stored in data repository. Lastly, if replicated data with the same file name with reference to the file that already exists in the database will be replaced or deleted, and it will store only one copy of the data among all the end users.

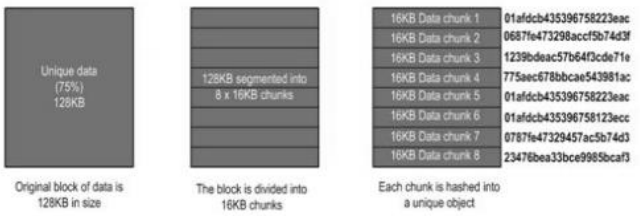

Figure 2. Working of data duplication.

Given data will be chunked, from the outcomes an index of the files will be created, and if any of the duplicates is found the related file is eliminated. Hence a single instance of file will be stored. The data deduplication is process that can be implemented in different ways. Duplicated file is eliminated by comparing two files that is uploaded and decision will be made whether to delete the older one that is already present or no longer needed.

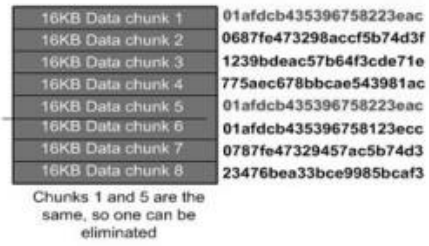

Figure 3. Duplicated data elimination(Deduplication).

Implemented file based comparison system. It is a simple method that reduces the duplicate data at the file level, and makes an comparison with the file system or file based system algorithm that eliminated data duplicates. An example can be comparing of name, size, and type and modified date information of two files with the similar name being saved in the system. If one these parameters match it can be assured that the files are similar copies of data or file can be delete as shown in Figure4.

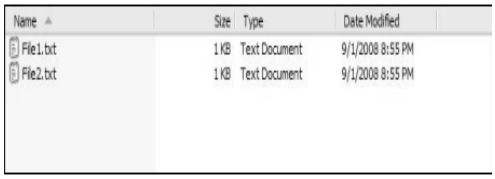

Figure 4. Mostly one the file is duplicated file with the same size and similar time creation.

IV. IMPLEMENTATION

The implementation state includes attentive preparation, investigating the actual application and restricting employment of project, ratifying of procedure attaining change from one system to other and evaluating those changed procedure.

4.1 Different Module

- Data Owners
- Users
- Key Authority
- Cloud Service Provider

4.2 Data Owner

Data Owners are the data holder of different files that is deposited in cloud system. These data owners are responsible of deciding the accessing structures and execution of encrypted data operations. The data owner will upload the cipher text to the CSP.

4.3 User

User desires in accessing the cipher text that is deposited in cloud system. The user downloads the cipher text and the related decrypted operations will be executed. User takes permission from both CSP as well Data Owner to access the file from cloud.

4.4 Unique Key Authority

Key authorization is virtually-trustworthy unit with cloud sever. Key authorization does genuine work-but-surprisingly, executes the specified jobs returning the right results. Nevertheless, it collects most of delicate components as much as possible. The unit selected will be responsible with the users' enrolment, in the cloud system. In meantime, it not only produces some parts of systems parameters, but it also generates almost all parts of unique personal key related to each customer.

4.5 Clouds Service Provider

Clouds service provider is supervisor for the cloud servers which is virtually-trustworthy unit that produces multiple services such as storing of data, computations and transferring of the data. To find the solution of the key escrow issue, it produces both sections of systems parameters and secret key for every user.

V. RESULT & DISCUSSION

The efficiency analysis and protection evidence describes achievement performance with secured data distribution in cloud computing. In encryption the end user and data owner is uploading the file in order to secure the data. In decryption the user wants to access the data from the cloud,

decryption key will be given to the end user. The file transaction includes ID, user name, file name, tasks along with date and time.

The cloud provides two keys one will be the secret key for private user and decryption key for public user. If an unauthorised end user tries in accessing file with in cloud, it shows file id, file name, digital sign generated with date and time will alert the owner of data regarding privacy relating to data.

The application server we have used Tomcat 5 version, and as front end we have employed JAVA and as backend we have used My SQL for the code implementation.

Security Intuition

As comparing to the earlier feature - related encrypting scheme our important goal was concerned to design proposed scheme that restricts opposition intrusion from colluding users. Our scheme provides randomizing of customers' personal key so these can never be united; nevertheless, regarding result the personal sharing is embedded into cipher text instead to the personal keys. In order to decrypt an attacker clearly must recover P $(A,A)^{as}$. For completing this, attacker must pair D from the cipher text with the C component from some user's private key. This results in the suitable value P $(A,A)^{qs}$, but blinded by some value $P(A,A)^{rs}$. This value is blinded if and only if enough the user has the right key components that satisfies the secret sharing scheme embedded in to cipher text.

Efficiency

The efficiency of generating key with encryption methods is simple. The encrypting method requires two exponents for every leaf in cipher text's accessing tree. The cipher text size will include two group elements for each tree leaf.

The key generation method requires two exponents for every feature provided to customer, the private key consists of two group elements for every feature. In simplest form, the decryption method requires two pairings for every leaf of the accessing tree being matched by a private key feature and (at most2) one exponentiation for each node along a path from such a leaf to the root. As we have used the cloud computing for uploading file and automatically the key will be generated as per our proposed scheme provides more security and efficiency. Some of the results snapshots are given below.

Shows the deduplication of the encrypted data in cloud.

Figure 5.1: Represents the deduplication of the encrypted cloud.

Represents the user requesting an permission to grant an secret key.

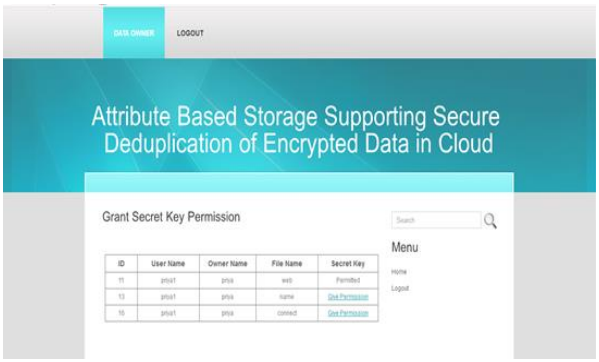

Fig 5.2: Represents the granting secret key permission.

Represents the set of files requesting an secret key for decryption.

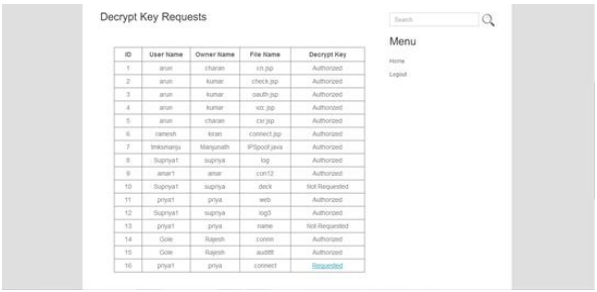

Figure 5.3: Represents the set of decrypt key requests.

Representing to dowload the decrypted file requested by the end user.

Figure 5.4: Represents that the user can download the decrypted file.

VI. CONCLUSION & FUTURE SCOPE

User's encrypted is deployed to cloud service provider, shared with customers possessing precise attribute (or credentials). A ciphertext policy which is attribute based and with secure deduplication in hybrid cloud setting is explained. Anew methodology to modify a ciphertext over one access policy into other access policies without disclosing the underlying plain text is proposed.

It strengthens the maintenance of confidential data and secretes in cloud system in opposition to the administrator of KA and CSP in addition to malevolent system strangers, in which CSP and KS were semi trusted.

Along with this approach, the weighted characteristic attributes were presented to enhance the attribute expression, that illustrates not only the inconsistent categorical features, but it reduces complexities in accessing policies, so expenses in storing the ciphered text and cost of time in performing encryption is saved.

Ultimately, here we present the efficiency and reliability examination for this proposed methodology. The result demonstrates high performance and reliability of our proposed scheme. Many recent challenges have appeared with the fast growth of adaptable cloud services. The most important problem is how we will firmly eliminate the repetition data obtained from outsider or outside supplier that is stored in the cloud severs.

REFERENCES

[1] A. Balu and K. Kuppusamy, "An Expressive and Provably Secure Ciphertext-Policy Attribute-Based Encryption," Information Sciences, 276(4):354–362, 2014.

[2] D. Boneh, B. Lynn, and H. Shacham, "Short Signatures from the Weil Pairing," JournalofCryptology, 17(4):297–319,2001.

[3] M. Chase, "Multi-authority attribute based encryption," Proceedings of the 4th Conference on Theory of Cryptography, pages 515–534, 2007.

[4] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-Policy Attribute Based Encryption," IEEE Symposium on Security and Privacy, pages 321–334, 2007.

[5] S. S. Chow," Removing Escrow from Identity-Based Encryption," Proceedings of the 12th International Conference on Practice and Theory in Public Key Cryptography, pages256–276,2009.

[6] M. Chase and S. S. Chow, "Improving privacy and security in multi authority attribute-based encryption," Proceedings of the 16th ACM Conference on Computer and Communications Security, pages 121–130, 2009.

[7] S. Lai, J. K. Liu, K.-K.R. Choo, and K. Liang, "Secret Picture: An Efficient Tool for Mitigating Deletion Delay on OSN" Information and Communications Security, pages467–477,2015.

[8] C. Fan, S. Huang, and H. Rung, "Arbitrary-State Attribute-Based Encryption with Dynamic Membership," IEEE Transactions on Computers, 63(8):1951–1961, 2014.

[9] J. Baek, Q. H. Vu, J. K. Liu, X. Huang, and Y. Xiang, "A Secure Cloud Computing Based Framework for Big Data Information Management of Smart Grid," IEEE Transactions on Cloud Computing, 3(2):233–244, 2015.

[10] M. Belenkiy, J. Camenisch, M. Chase, M. Kohlweiss, A. Lysyanskaya, and H. Shacham, "Randomizable Proofs and Delegatable Anonymous Credentials," Proceedings of the 29th Annual International Cryptology Conference, pages 108-125,2009.

[11] L. Cheung and C. Newport, "Provably Secure Cipher text Policy ABE," Proceedings of the 14th ACM conference on Computer and communications security, pages 456–465,2007.

[12] C. K. Chu, W. T. Zhu, J. Han, J. K. Liu, J. Xu, and J. Zhou, "Security Concerns in Popular Cloud Storage Services," IEEE Pervasive Computing, 12(4):50–57,2013.

[13] A. De Caro and V. Iovino, "JPBC: Java Pairing Based Cryptography," IEEE Symposium on Computers and Communications, 22(3):850–855, 2011.

[14] H. Deng, Q. Wu, B. Qin, J. Domingo-Ferrer, L. Zhang, J. Liu, and W. Shi, "Ciphertext-Policy Hierarchical Attribute-Based Encryption with short Ciphertexts," InformationSciences,275(11):370–384,2014.

[15] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-Based Encryption for Fine-Grained Access Control of Encrypted Data," Proceedings of the 13th ACM conference on Computer and communications security, pages 89–98, 2006.

[16] J. Hur, "Improving Security and Efficiency in Attribute-Based Data Sharing," IEEE Transactions on Knowledge and Data Engineering, 25(10):2271– 2282, 2013.

[17] L. Ibraimi, M. Petkovic, S. Nikova, P. Hartel, and W. Jonker, "Mediated Ciphertext-Policy Attribute-Based Encryption and its Application," Proceedings of the 10th International Workshop on Information Security Applications, pages 309–323, 2009.

[18] T. Jiang, X. Chen, J. Li, D. S. Wong, J. Ma, and J. K. Liu, "Towards Secure and Reliable Cloud Storage Against Data Re-outsourcing," Future Generation Computer Systems,52:86–94,2015.

[19] K. Liang, M. H. Au, J. K. Liu, W. Susilo, D. S. Wong, G. Yang, T. V. X. Phuong, and Q. Xie, "A DFA-Based Functional Proxy Re-Encryption Scheme for Secure Public Cloud Data Sharing," IEEE Transactions on Information Forensics and Security, 9(10):1667–1680, 2014.

[20] K. Liang, M. H. Au, J. K. Liu, W. Susilo, D. S. Wong, G. Yang, Y. Yu, and A. Yang, "A Secure and Expressive Ciphertext-Policy Attribute-Based Proxy Re-Encryption for Cloud Data Sharing," Future Generation Computer Systems, 52(C):95–108, 2015.