SECURE DE-DUPLICATION WITH ATTRIBUTE-BASED ENCRYPTION IN CLOUD

Dhananjay Kandekar, Aviraj Ghavate, Akshay Ekshinge, Rohan Kukade , Prof. B.B.Gite
Computer Engineering Department, Sinhgad Academy of Engineering, Pune, India

Abstract: Cloud is an important source of data storage that can be maintained, managed and backed up via remotely anytime. These cloud services not only provide easy access but also protects user's data to a greater extent from the third party. Due to rise of 'Big Data', Cloud computing has become one of the most significant and essential rising application platforms to solve the expanding of data exchange. Cloud networks are not so secure, to protect data from exposure; users need to encrypt their data before sharing it to other users. So, to achieve this few terms Attribute-based encryption (ABE) has been widely used in cloud computing where data providers outsource data in encrypted format to cloud and can share the encrypted data with different users possessing specific attributes which will indeed get decrypted using security measures. It allows a user with limited computational resources to outsource their large data processing workloads to the cloud, and economically enjoy the computational bandwidth, better power, storage, and even appropriate software that can be shared in a pay|per|use manner. To the other side, the outsourced computation workloads often contain highly-confidential and sensitive data, such as the proprietary research data, business statistics, or personally identifiable information, university student data records, stock market data etc. To combat against unauthorized data leakage and exposure, this confidential and sensitive data have to be encrypted before outsourcing so as to provide end-to-end data confidentiality assurance in the cloud and beyond that .Still, ordinary data encryption methods, in essence, to prevent the cloud from performing any important operation of the underlying cipher-text policy, making the computation over encrypted data a very hard problem for the system. The proposed system achieves scalability due to its hierarchical structure as well as efficiency and easiness of data flow in cloud computing.

I. INTRODUCTION

Data Security is a primary concern for every communication platform. The relentless growth of the Internet and communication technologies has made the extensive use of images unavoidable. Our present ensures the avoid deduplication storage in the cloud and an authenticated person only allowed to view or download the file.

The standard attribute-based encryption system doesn't support secure Deduplication, which is crucial for eliminating duplicate copies of identical data in order to save network bandwidth and storage space. In this paper, we present an attribute-based encryption and storage system with secure deduplication in a hybrid cloud setting, where a private cloud is responsible for duplicate detection and a public cloud manages the access rights of data files in the cloud storage.

Compared with the prior ABE and data deduplication systems, our system has two advantages. Firstly, it can be used to confidentially exchange data with users by specifying access policies rather than sharing decryption keys. Secondly, it achieves the standard notion of semantic security for data confidentiality while existing systems only achieve it by defining a weaker security notion. In addition, we put forth a methodology to modify a cipher-text over one access policy into cipher-text of the same plaintext but under other access policies without revealing the underlying plaintext.

Cloud computing greatly facilitates data providers who want to outsource their data to the cloud without disclosing their private data to other parties and would like users with certain protocols and policies to be able to access the data. This requires data to be stored in encrypted forms with access control policies such that no one except users with attributes of specific forms can decrypt the encrypted data outsourced by the data provider. An encryption technique that we follow is called attribute-based encryption (ABE), where a user's secure key is associated with an attribute set, a message is encrypted under an access policy (or access policy structure) over a set of attributes, and a user can decrypt a cipher-text with his/her private key if his/her set of attributes satisfies the access policy associated with this cipher-text. On the other hand, to the best of our knowledge, existing constructions for secure deduplication are not built on attribute-based encryption. The default ABE system fails to follow secure de-duplication which is a method to save network and space bandwidth by eliminating unnecessary double copies of the encrypted data stored in the cloud. We consider the following scenario in the design of an attribute-based storage system supporting secure deduplication of encrypted data in the cloud, in which the cloud will not store a file more than once even though it may receive multiple copies of the same file encrypted under different access policies.

A data provider, Ali, intends to upload a file D to the cloud, and share D with users having certain credentials. In order to do so, Ali encrypts D under an access policy P over a set of attributes, and uploads the corresponding cipher-text to the cloud, such that only users whose sets of attributes satisfying the access policy can decrypt the cipher-text. Later, another data provider, Atik, uploads a cipher-text for the same underlying file D but ascribed to a different access policy P0. Since the file is uploaded in an encrypted form, the cloud is not able to discern that the plaintext corresponding to Atik's cipher-text is the same as that corresponding to Ali's, and will store D twice. Obviously, such duplicated storage wastes storage space and communication bandwidth.

II.EXISTING SYSTEM

- In storage system with secure deduplication, to store a file in the cloud, a data provider generates a tag and a ciphertext. A data provider uploads the tag and the ciphertext to the cloud. Receiving request from a data provider for uploading a ciphertext and an associated tag, the cloud runs a so-called equality checking algorithm, which checks if the tag in the incoming request is identical to any tag in the storage system. If there is a match, then the underlying plaintext of this incoming ciphertext has already been stored and the new ciphertext is discarded**.** It is apparent that such a system with a tag appended to the ciphertext does not provide the standard notion of semantic security for data confidentiality, because if the plaintexts can be predicated from their tags, an adversary can always make a correct guess by computing the tag of a plaintext and then testing it against the tag in the challenge phase in the semantic security game.

    - To store the data in encrypted form.
    - To save the storage space and network bandwidth by eliminating redundant copies of the encrypted data stored in the cloud.
    - To upload a file P to the cloud and share P with user having certain credentials.
    - To modify a cipher text over one access policy into cipher texts of the same plaintext but under any other access policies without revealing the underlying plaintext.
    - To achieve data consistency in the system.

DISADVANTAGES OF EXISTING SYSTEM:

- The CP-ABE construction, but it is secure under the generic group model.
- Cheung and Newport presented a CP-ABE is proved to be secure under the standard model, but it only supports the AND access structures.
- A CP-ABE system under more advanced access structures is based on the number theoretic assumption. Overcome the limitation that the size of the attribute space is polynomially bounded in the security parameter and the attributes are fixed ahead, Waters built a large universe CP-ABE system under the prime-order group

III.PROPOSED SYSTEM:

- In this paper, we present an attribute-based storage system which employs ciphertext-policy attribute-based encryption (CP-ABE) and supports secure deduplication. Our main contributions can be summarized as follows.
- Achieves the standard notion of semantic security for data confidentiality in attribute-based deduplication systems by resorting to the hybrid cloud architecture.
- We put forth a methodology to modify a ciphertext over one access policy into ciphertexts of the same plaintext but under any other access policies without revealing the underlying plaintext..
- Thirdly, we propose an approach based on two cryptographic primitives, including a zero-knowledge proof of knowledge and a commitment scheme, to achieve data consistency in the system.

ADVANTAGES OF PROPOSED SYSTEM:

- We bring in our system a hybrid cloud architecture, which consists of a private cloud responsible for tag checking and ciphertext regeneration and a public cloud storing the ciphertexts.
- Approaches producing such a proof makes use of the randomness reuse technique in the generation of the tag and the ciphertext with additional zero-knowledge proof of knowledge on the shared random coin in the ciphertext. So, it is impossible for an adversary to perform duplicate faking attacks unless the adversary casually obtains the content of the plaintext hidden in the ciphertext.

IV.DESIGN OF SYSTEM ARCHITECTURE

Modules:

1. Data Provider Module.
2. Attribute Authority (AA) Module.
3. Cloud Module.
4. Users Module.

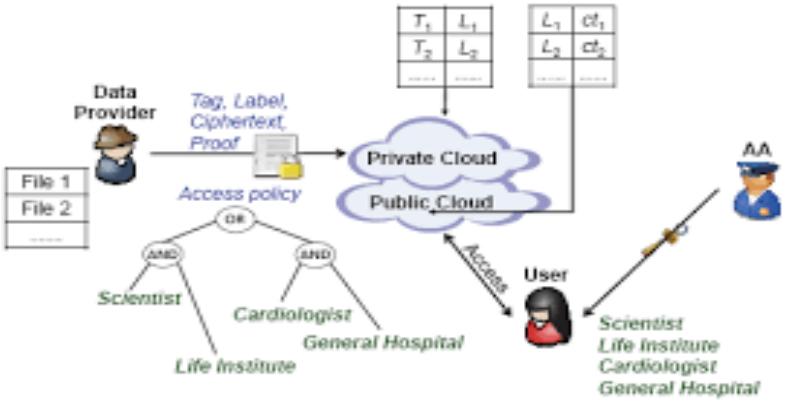


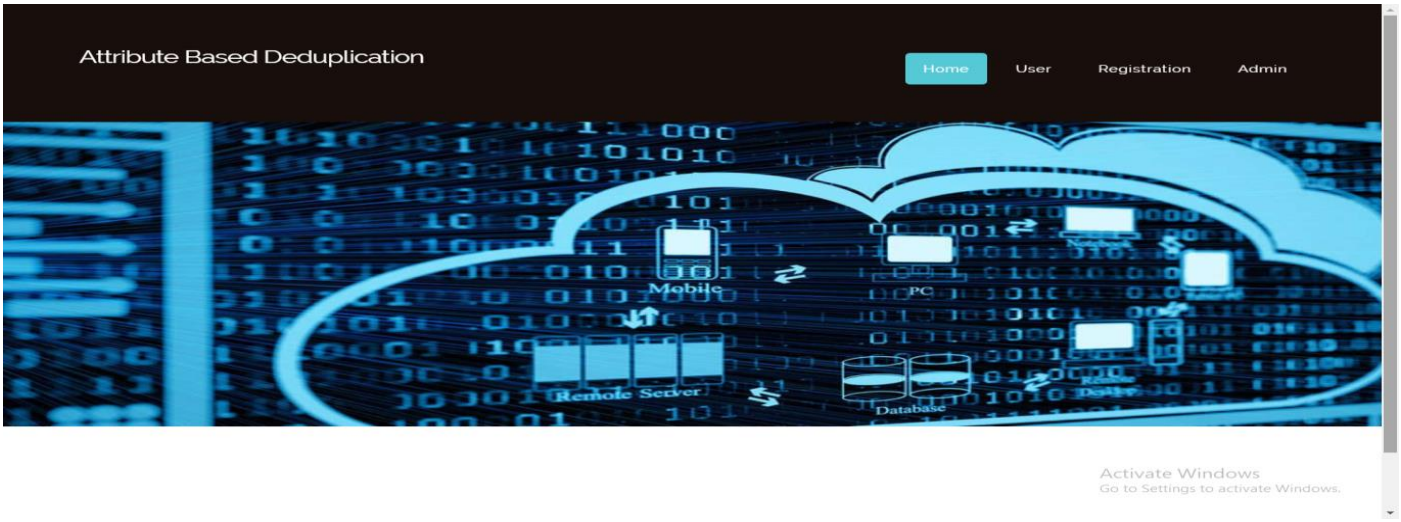
Figure: The architecture of attribute-based storage with secure de-duplication.

The architecture of our attribute-based storage system with secure deduplication is shown in Fig. 1 in which four entities are involved: data providers attribute authority (AA), cloud and users. A data provider wants to outsource his/her data to the cloud and share it with users possessing certain credentials. The AA issues every user a decryption key associated with his/her set of attributes. The cloud consists of a public cloud which is in charge of data storage and a private cloud which performs certain computation such as tag checking. When sending a file storage request, each data provider firstly creates a tag T and a label L associated with the data and then encrypt the data under an access structure over a set of attributes. Also, each data provider generates a proof pf on the relationship of the tag T, the label L, and the encrypted message ct, but this proof will not be stored anywhere in the cloud and is only used during the checking phase for any newly generated storage request. After receiving a storage request, the private cloud first checks the validity of the proof pf and then tests the equality of the new tag T with existing tags in the system. If there is no match for this new tag T, the private cloud adds the tag T and the label L to a tag-label list and forwards the label and the encrypted data, (L, ct) to the public cloud for storage. Otherwise, let ct' be the ciphertext whose tag matches the new tag and L' be the label associated with ct', and then the private cloud executes as follows. • If the access policy in ct is a subset of that in ct0, the private cloud simply discards the new storage request; else, if the access policy in ct0 is a subset of that in ct, the private cloud asks the public cloud to replace the stored pair (L0, ct0) with the new pair (L, ct) where L = L0. • If the access policies in ct and ct0 are not mutually contained, the private cloud runs the ciphertext regeneration algorithm to yield a new ciphertext for the same underlying plaintext file and associated with an access structure which is the union of the two access structures, and forwards the original label and the resulting ciphertext to the public cloud.
At the user side, each user can download an item, and decrypt the ciphertext with the attribute based private key generated by the AA if this user's attribute set satisfies the access structure. Each user checks the correctness of the decrypted message using the label and accepts the message if it is consistent with the label. Concerning the adversarial model of our storage system, we assume that the private cloud is "curious-but-honest" such that it will attempt to obtain the encrypted messages but it will honestly follow the protocols, whereas the public cloud is distrusted such that it might tamper with the label and ciphertext pairs outsourced from the private cloud (note that such a misbehavior will be detected by either the private cloud or the user via the accompanying label.

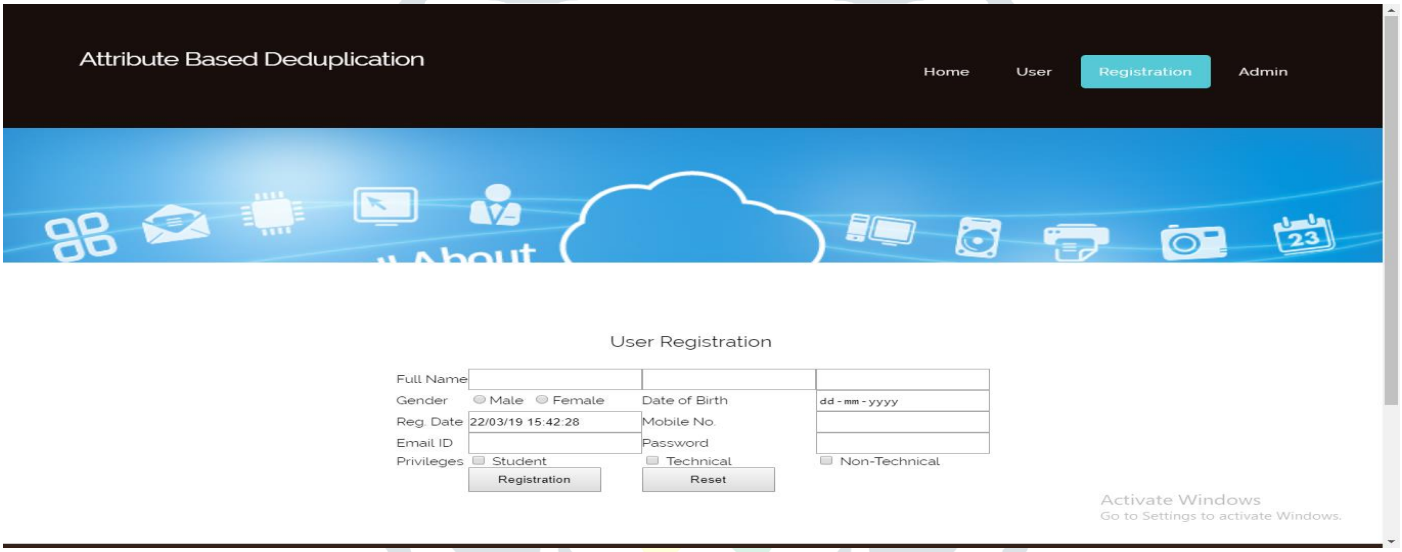
V.RELATED WORK

Distributed computing is a progressive registering worldview which empowers adaptable, on-request and minimal effort utilization of figuring assets. Those points of interest, unexpectedly, are the reasons for security and protection issues, which rise in light of the fact that the information claimed by various clients are put away in some cloud servers rather than under their own control. The security issue of distributed computing is yet to be settled. To manage security issues, different plans in light of the Attribute-Based Encryption have been utilized. From one perspective, the outsourced figuring workloads often contain sensitive information, for instance, the business money related records, prohibitive research data, or eventually identifiable prosperity information et cetera. To fight against unapproved information spillage, sensitive data must be mixed before outsourcing so as to offer end-to-end data protection affirmation in the cloud and past. Regardless, normal data encryption procedures by and large shield cloud from playing out any critical operation of the essential figure content game plan, making the count over encoded data a troublesome issue. The proposed plot does not simply achieve flexibility due to its dynamic structure. We give the protection secure out in the open social distributed computing. In our venture, we actualize progressive property base security the pecking orders are Cloud specialist, Domain expert, and clients. Cloud expert can just have benefit to make or expel the domain (private cloud specialist) in the cloud and they can keep up every one of the points of interest in general cloud Domain expert can make or evacuate the clients inside the area this client is called private clients. Clients are two sorts of private cloud client and open cloud client's Private cloud clients are depending on the space Public clients under cloud specialist. Clients can transfer the documents in two ways: Public and Private. On the off chance that the private client transfers general society document, the record permeability and availability is just inside area itself and same space clients can get to that document with no security validation If the general population client transfer people in general document, the record permeability, and openness is constantly open any cloud client can get to that document. For Private transfer If private client transfers the private document implies that record permeability is just inside space yet document openness is who have the emit key (OTP) implies who have benefit to get to the record If general society client transfers the private document implies that document permeability is open anybody can obvious the document yet who have a benefit (OTP) to get to they just can get to the document.

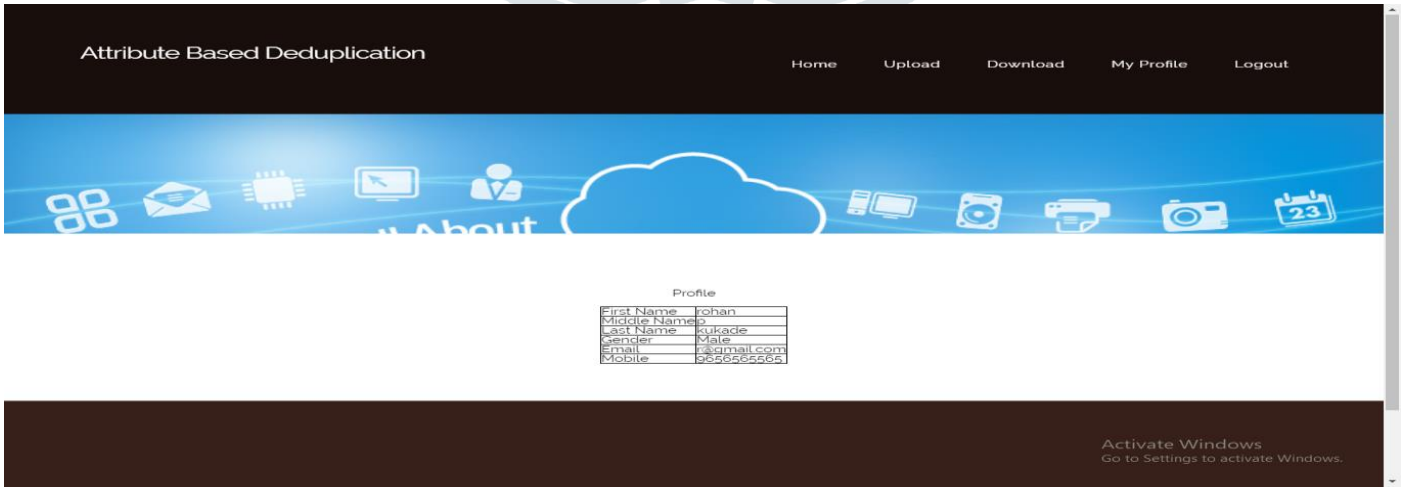
VI.REPREENTATION OF RESULT



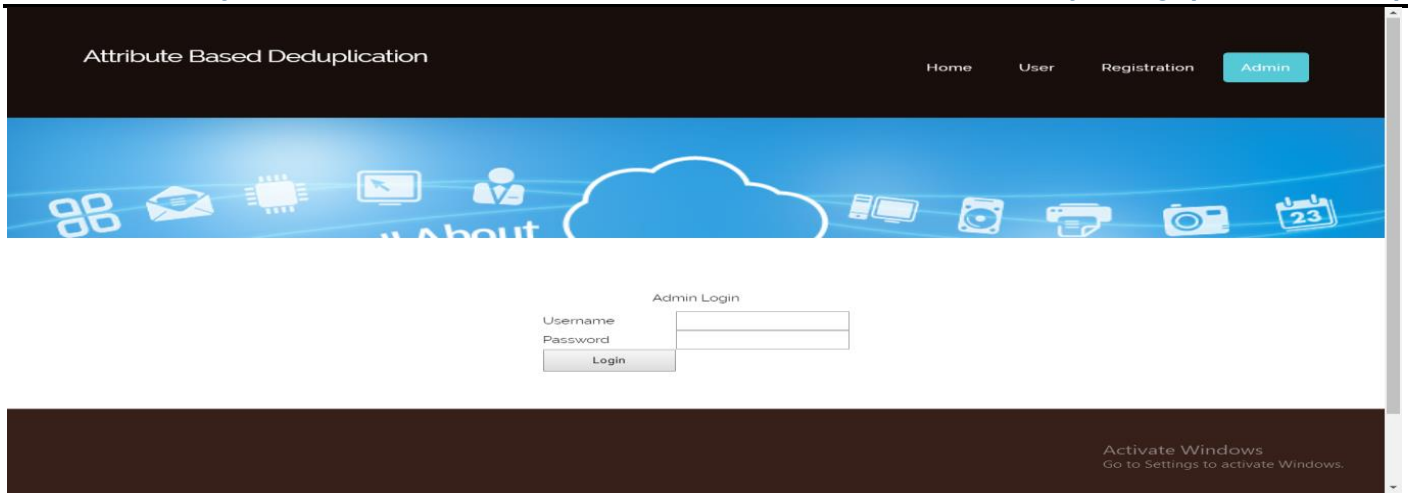
Home page



Registration Page

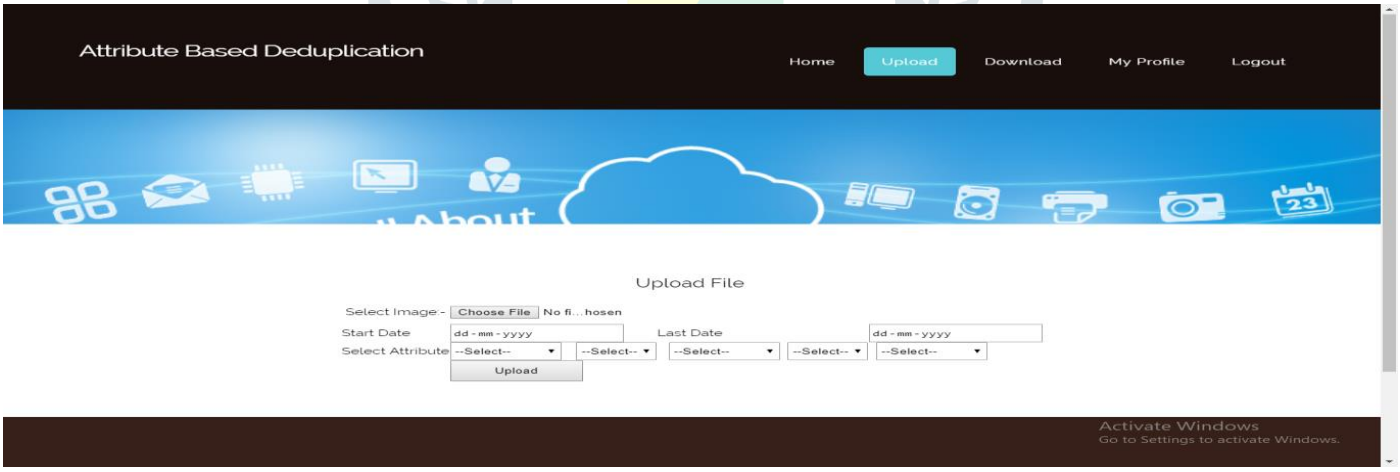


User profile page

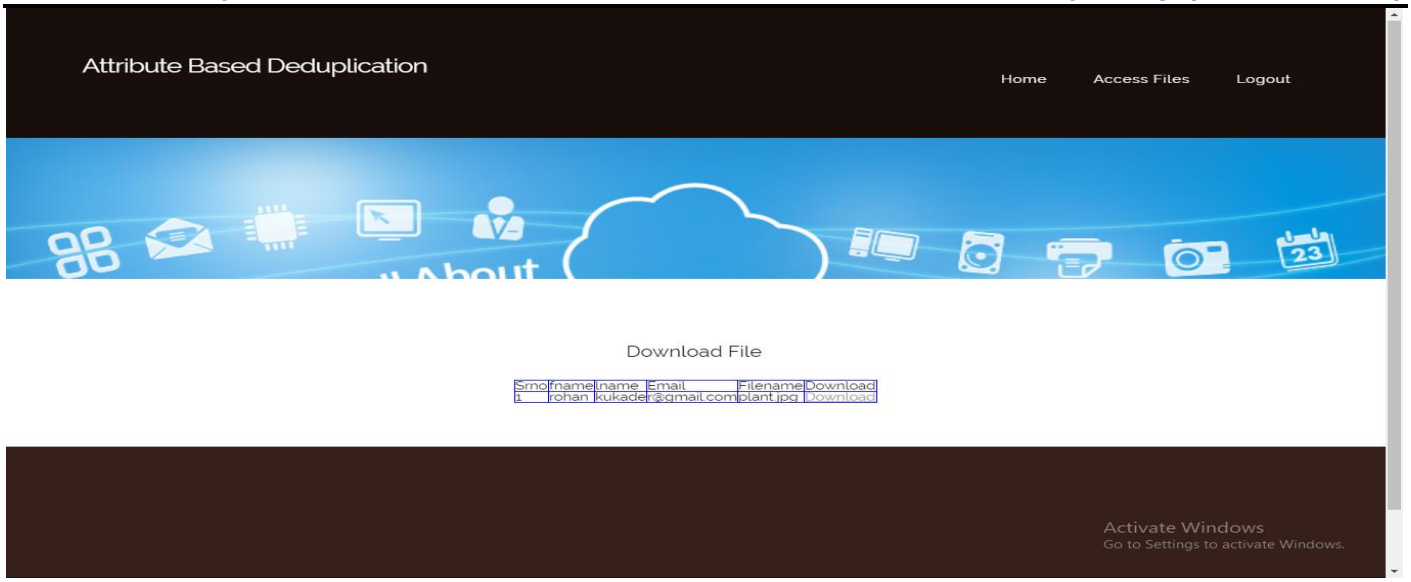


Admin login page

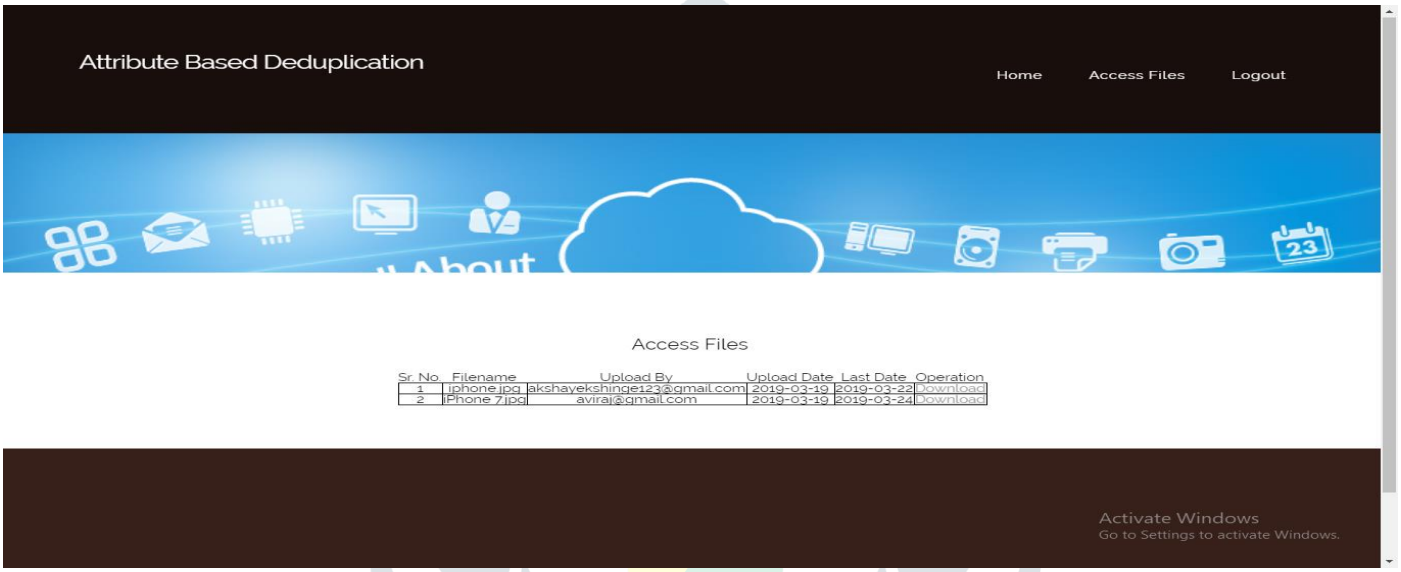
User Activation page



Upload file page



Download file page



Access file page

VII.CONCLUSION

Attribute-based encryption (ABE) has been widely used in cloud computing where data providers outsource their encrypted data to the cloud and can share the data with users possessing specified credentials. On the other hand, deduplication is an important technique to save the storage space and network bandwidth, which eliminates duplicate copies of identical data. Here we provided the reason that our proposal framework information deduplication of given data is done approves way and securely. The user (End customer) needs to present the benefit alongside the unique key as a proof of possession.

However, the standard ABE systems do not support secure deduplication, which makes them costly to be applied in some commercial storage services. We presented a novel approach to realize an attribute-based storage system supporting secure deduplication. Our storage system is built under a hybrid cloud architecture, where a private cloud manipulates the computation and a public cloud manages the storage. The private cloud is provided with a trapdoor key associated with the corresponding ciphertext, with which it can transfer the ciphertext over one access policy into ciphertext of the same plaintext under any other access policies without being aware of the underlying plaintext. After receiving a storage request, the private cloud first checks the validity of the uploaded item through the attached proof. If the proof is valid, the private cloud runs a tag matching algorithm to see whether the same data underlying the ciphertext has been stored. If so, whenever it is necessary, it regenerates the ciphertext into a ciphertext of the same plaintext over an access policy which is the union set of both access policies. The proposed storage system enjoys two major advantages. Firstly, it can be used to confidentially share data with other users by specifying an access policy rather than sharing the decryption key. Secondly, it achieves the standard notion of semantic security while existing deduplication schemes only achieve it under a weaker security notion. A proposed routine guarantees the data duplication and encryption with all the security measures.

REFERENCES

[1]. El Maraghy M, Hesham S and Abd El Ghany M.A, “Real-time Efficient FPGA Implementation of AES Algorithm”, IEEE International SOC Conference (SOCC), page 203-208, Sept 2013.
[2]. Kamali S.H, Shakerian R, Hedayati M, and Rahmani M, "A new modified version of Advanced Encryption Standard based algorithm for image encryption", (ICEIE) International Conference on Electronics and Information Engineering, volume 1, page 1250-1255, Aug 2010.
[3]. M. Bellare, S. Keelveedhi, and T. Ristenpart, “Message-locked encryption and secured duplication,” in Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings, ser. Lecture Notes in Computer Science, vol. 7881. Springer, 2013, pp. 296–312.
[4]. K. R. Choo, J. Domingo-Ferrer, and L. Zhang, "Cloud cryptography: Theory, practice, and future research directions," Future Generation Comp. Syst., vol. 62, pp. 51–53, 2016.
[5]. K. R. Choo, M. Herman, M. Iorga, and B. Martini, “Cloud forensics: State-of-the-art and future directions,” Digital Investigation, vol. 18, pp. 77–78, 2016.
[6]. Y. Yang, H. Zhu, H. Lu, J.Weng, Y. Zhang, and K. R. Choo, "Cloud-based data sharing with fine-grained proxy re-encryption," Pervasive and Mobile Computing, vol. 28, pp. 122–134, 2016.