International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST

ABE and Bloom Filter Based Searchable Attribute-Based Encryption Scheme over Cloud Data

Vishal Jalindar Gondil[1], Prof. H. A. Hingoliwala[2]
M. Tech Scholar, Department of Computer Engineering, JSCOE, Pune, Maharashtra, India[1]
Professor, Department of Computer Engineering, JSCOE, Pune, Maharashtra, India[2]

ABSTRACT

Article Info
Volume 8, Issue 3
Page Number : 208-215

Publication Issue
May-June-2021

Article History
Accepted : 07 May 2021
Published : 14 May 2021

Searchable Attribute based encryption is a promising technique that achieves flexible and fine-grained data access control over encrypted data, which is very suitable for a secure data sharing environment such as the currently popular cloud computing. However, Information Security has remained a primary concern and today when most of the sensitive data is stored on cloud with client organization having lesser control over the stored data, the fundamental way to fix this issue is to encrypt such data. So, a secure user imposed data access control system must be given, before the users outsource any data to the cloud for storage. Attribute Based Encryption (ABE) system is one such asymmetric key based cryptosystem that has received much attention that provides fine-grained access control to data stored on the cloud. In this paper, we propose a more proficient and richer type of Attribute Based Encryption technique that considers the Outsourced ABE construction and removes the burden of data owner to be always online. In multi keyword search; data owners and users can generate the keywords index and search trapdoor, respectively, without relying on always online trusted authority, also we are using bloom filter for de-duplication checking of file is performed before uploading any encrypted file at cloud server which reduces memory utilization at cloud server. Experimental results show that the performance of the proposed system is better than existing system in terms of security, data availability, time consumption and memory utilization.

Keywords : Attribute-Based Encryption, Bloom Filter, Fine-Grained Access Control, Multi Keywords Search, Cloud Computing.

I. INTRODUCTION

Cloud Computing is received as another option to conventional data innovation due to its intrinsic resource-sharing and low-maintenance attributes. In cloud registering, the cloud specialist co-ops (CSPs, for instance, Amazon, can send diverse administrations to cloud customers with the help of extreme datacentres. By combining the local data management frameworks into cloud servers, clients can appreciate top notch services and recovery huge speculations on their nearby infrastructures. Information storage is a fundamental administration given by cloud framework. By utilizing the cloud, the clients can be totally discharged from the troublesome nearby information storage and support. Additionally, it likewise has a huge hazard to the secrecy of those put away documents. In particular, the cloud servers overseen by cloud suppliers are not trusted absolutely by clients while the information records put away in the cloud might be touchy and secret, for example, marketable strategies. To give information security, as a fundamental arrangement is to scramble information records, and after that transfer the encoded information into the cloud. Sadly, structuring proficient and secure information sharing a philosophy for gatherings in the cloud isn't a simple assignment because of the accompanying testing issues.

To begin with, personality security is a significant defeat for the improvement of cloud processing. With no security of personality protection, clients might be reluctant to participate in cloud registering frameworks on the grounds that their genuine characters could be effectively revealed to cloud suppliers and aggressors. Second, it is exceedingly suggested that any part in a gathering can almost certainly utilize the information putting away and sharing administrations given by the cloud, which is defined as the various proprietor way. Contrasted and the single-proprietor way, in which just the gathering director can store and adjust information in the cloud, the numerous proprietor way is increasingly adaptable progressively applications. To wrap things up, bunches are dynamic practically speaking. The alterations of participation make secure information sharing extremely troublesome. Toward one side, the mysterious framework provokes new conceded clients to get familiar with the substance of information records put away before their cooperation, because of its unrealistic for new allowed clients to contact with obscure information proprietors and get the relating unscrambling keys. At the opposite end, an effective enrollment revocation system without refreshing the mystery keys of different clients are additionally wanted to limit the unpredictability of key administration.

To settle this issue, data which is to be put away is encoded in mixed structure. Anyway, such encoded information must be pleasing to the sharing and access control. Different private and open key cryptographic procedures are not receptive to versatile access control. So as to settle this issue Searchable Attribute-Based Encryption method was proposed. Attribute-Based Encryption (ABE) has increased much consideration in the research network. Attribute-Based Encryption is a lopsided key based cryptographic method which improves the skill-fullness of access control systems.

In a Searchable ABE system, a client's keys just as ciphertext are marked with sets of engaging attributes and a specific key can decode a specific ciphertext just if there is a match in the attributes of the ciphertext and the client's critical. Be that as it may, a defect in the standard ABE framework is the immense size of the ciphertext and the computational complexities in the decoding stage are very saddling. In this way, there is a need to upgrade the capability of ABE. To unravel this issue, an improved bloom filter based searchable ABE for the mobile cloud storage is proposed. Watchword search is additionally upheld, in which information proprietors and clients can produce the keywords list and search trapdoor, separately, without depending on constantly online confided in power.

Literature review is described in the section II. Section III presents the proposed system implementation details which includes searchable encryption, bloom filter algorithm. Section IV presents experimental analysis, results and discussion of proposed system. Section V concludes our proposed system. While at the end list of references paper are presented.

II. LITERATURE REVIEW

The neutral-network based NLP processing method - Doc2Vec model which uses word's and paragraph's context information to extract documents' features is proposed by authors at el. Dai, Xuelong & Dai, Hua & Yang, Geng & Yi, Xun & Huang, Haiping in paper [1]. The features contain latent semantics information and can measure the similarity between documents. They also adopt the Doc2Vec model to achieve a semantic-aware multi-keyword ranked search scheme. Their proposed Doc2Vec model uses the distributed representation of words and documents with a modest dimensionality of vectors while trained on a dataset with a few hundreds of millions of words. Documents' distributed representations are extracted as documents feature vector by Doc2Vec model and utilized as the search index. The features of the queried keywords are also extracted as the query feature vector, and the secure inner product operation is adopted to achieve privacy-preserving semantic search with the query feature vector and index. Their work scheme supports dynamic update on the document set with Doc2Vec model.

Authors Patil Rashmi, Gandhi Yatin, Sarmalkar Vinaya, Pund Prajakta & Khetani Vinit at el. In [2] solves various security issues that are concerned with cloud storage. As Cloud service providers or storage servers are not completely trustworthy, to solve this issue their work utilises homomorphic hash algorithm. Further it supports dynamic operations such as insert, update, delete and modify at block level, for data dynamics Merkle Hash Tree is used which helps to find the location of each dynamic operation. Third party auditor checks the user's data for correctness and gives the accuracy of the data that is stored in cloud server. Also deduplication technique is used to check whether the file that user need to store in cloud storage is already exist at cloud server or not.

In paper [3] authors Shangping Wang, Duo Zhang, Yaling Zhang, And Lihua Liu at el heighlight the capacity of attribute revocation is productively accomplished by assigning the update of mystery key and ciphertext to the ground-breaking cloud server. Catchphrase search is additionally bolstered, in which information proprietors and clients can create the keywords list and search trapdoor, individually, without depending on constantly online confided in power. Moreover, a redistributed decoding innovation is utilized to diminish the computational heap of unscrambling on the client side.

In [4] a Secure Encryption is such a cryptographic crude, that empowers clients to search keywords over the scrambled information without spilling keywords data. In this paper, the catchphrase search is upheld and afterward the access structure is in part covered up to ensure protection data in ciphertexts is proposed.

In this paper [5], the creator proposed a dynamic searchable encryption plot. In their development, recently included tuples are put away in another database in the cloud, and erased tuples are recorded in a revocation list. The last search result is accomplished through barring tuples in the revocation list from the ones recovered from unique and recently included tuples. However, Cash et al. dynamic search plot do not understand the multi-watchword positioned search usefulness.

In this paper [6] the creators considered another need of ABE with redistributed unscrambling that is the certainty of changes. Casually, it ensures that a client can effectively check if the change is done precisely or not. Their framework exhibit that the new plan is both secure and certain, without relying upon arbitrary forecasts. In their work, they propose an alternate view for ABE that, everything considered, clears out the overhead for customers. Anyway, their development does not consider overhead calculation at the attribute specialist associated with the key-issuing process.

Here in [7], Green et al. proposed an ABE framework with re-appropriated unscrambling that, all things considered, take out the decoding overhead for customers. In such a framework, a client gives an untrusted server, state a cloud specialist organization, with a change key that allows the cloud to interpret any ABE ciphertext satisfied by that client's attributes or access arrangement into a basic ciphertext, and it just achieves somewhat computational overhead for the client to recuperate the plaintext from the changed ciphertext. Security of an ABE framework with redistributed unscrambling guarantees that a foe (Including a vindictive cloud) won't have the ability to get the hang of anything about the encoded message; regardless, it doesn't guarantee the rightness of the change performed by the cloud.

In this paper [8], Yu et al. consider the issue of client revocation which includes re-scrambling the information that is accessible to the client leaving the framework and refreshing the private keys of clients staying in the framework. They have proposed a plan that empowers the proprietor of the information to re-appropriate the errand of re-encryption and private key updates to an outsider without uncovering the substance and the client data. They have great accomplished the finely grained and versatile access in cloud processing. Anyway, the unpredictability in client revocation increments with the expansion in the number of clients which makes the framework complex. What's more, their plan does not bolster client responsibility.

Cheung et al. in [9] have proposed yet each other kind of Attribute-Based Encryption plot known as ciphertext approach attribute-based encryption (CP-ABE) where each mystery key is named with attributes, and each ciphertext is set with an access strategy. Unscrambling is done if and just if the customers characteristic set fulfills the ciphertext access structure. This gives fine-grained access control on shared information in different settings, including secure databases and secure multicast. In this paper, they consider CP-ABE designs in which access structures are AND doors on positive and negative attributes. Their essential arrangement has been ended up being picked plaintext assault (CPA) secure under the decisional bilinear Diffie-Hellman presumption however the utilization of autonomous occurrences of CP-ABE encryption, and furthermore, the security of this proposition stays as an open issue.

In this paper [10], the creators proposed a cryptosystem that gives fine-grained access control to scrambled data that they called Key-Policy Attribute-Based Encryption (KP-ABE). In their cryptosystem, ciphertext is marked with sets of attributes and private keys are set with access structures that control which ciphertext a client can translate. They have connected their development in the measurable examination and communicate encryption. Anyway, their frameworks neglect to conceal the attributes that do the encryption. Thus the issue of attribute stowing away is left open.

Here Curtmola et al. [11] proposed two plans (SSE-1 and SSE-2) which accomplish the ideal search time. Their SSE-1 plot is secure against picked catchphrase assaults (CKA1) and SSE-2 is secure against versatile picked watchword assaults (CKA2). These early works are single watchword boolean search plans, which are exceptionally straightforward as far as usefulness. A short time later, plenteous works have been proposed under various risk models to accomplish different search usefulness, for example, single watchword search, similitude search, multi-catchphrase Boolean search, positioned search, and multi-watchword positioned search, and so forth.

The thought of ABE was proposed in this paper [12] as a fluffy form of Identity-Based Encryption (IBE). In Fuzzy IBE, Sahai et al. see the way of life as a lot of sensible characteristics. A Fuzzy IBE course of action thinks about a private key for a personality, to interpret a ciphertext blended with a character w, if and just if the personalities w and w' are near one another made a decision by some measurement. A Fuzzy IBE course of action can be united with secure encryption utilizing biometric contributions as characters; the break opposition property of a Fuzzy IBE plan is exactly what considers the utilization of biometric personalities, which ordinarily will have some confusion each time they are explored. Furthermore, they demonstrate that Fuzzy-IBE can be utilized for a kind of use that the term" attribute-based encryption". In this paper, they show two progressions of Fuzzy IBE orchestrates. Their progressions can be viewed as an Identity-Based Encryption of a message under two or three attributes that make a (delicate) character. Their IBE plans are both oversights tolerant and secure against plot assaults. Plus, the key progression does not utilize self-assertive prophets. Maker displays the security of their game plans under the Selective-ID security demonstrate.

Searchable encryption plans empower the customers to store the encoded information to the cloud and execute a catchphrase search over the ciphertext area. Because of various cryptography natives, searchable encryption plans can be developed utilizing open key based cryptography or symmetric key based cryptography [13].

III. PROPOSED WORK

A. System Architecture

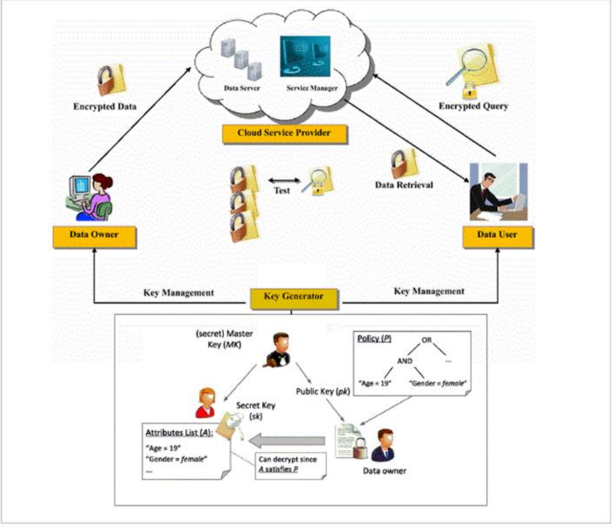

Figure 1. System Architecture

1. Attribute Based Key Generation

Initially user needs a key for file encryption or decryption before storing or downloading from cloud server. To minimize the key management overhead a separate outsourcing Key Generation Service Provider (KGSP) is used. And for data security purpose, Attribute Authority (AA) is introduced in the system. In this, upon receiving the key request from user, AA verifies the attributes of users, if they are valid then and only then KGSP generate and distribute the key for that user.

2. File Encryption and upload

For data confidentiality purpose, each and every data files are stored on server in encrypted format. After verification of attributes, user get key. Upon receiving this key, user encrypt their files and upload on cloud server.

3. Deduplication Checking

Bloom Filter which holds a fixed number of elements can represent the set with a large number of elements. To check if an element is already present in the Bloom Filter, we must again hash the search query and check if the bits are present or not. If bits are presents, then we can say its duplicate file. This process is performed at the time of uploading files at cloud server.

4. Searchable Encryption

Searchable encryption (SE) allows a server to perform search over encrypted data according to a search token submitted by a data user.

5. File Download and Decryption

After applying query search user gets list of file, from that any file can be downloaded. After downloading the file which is in encrypted format user need to perform file decryption operation to view original content of file. User decrypt the file using key obtained from KGSP.

B. Algorithms

1. Build Index Tree

Input: the document collection F = {f_1, f_2... f_n} with the identifiers FID = {FID—FID = 1, 2... n}.
Output: the index tree T

```
for each document FID in F do
    Construct a leaf node u for FID,
    Insert u to CurrentNodeSet;
end for
while the number of nodes in CurrentNodeSet is larger than 1 do
    if the number of nodes in CurrentNodeSet is even, i.e. 2h then
        for each pair of nodes $u_0$ and $u_{00}$ in CurrentNodeSet do
            Generate a parent node u for $u_0$ and $u_{00}$ ,
            Insert u to TempNodeSet;
        end for
    else
        for each pair of nodes u0 and u00 of the former (2h - 2) nodes in CurrentNodeSet do
            Generate a parent node u for u0 and u00 ;
            Insert u to TempNodeSet;
        end for
        Create a parent node u1 for the (2h - 1)-th and 2h-th node, and then create a parent node u for u1 and the (2h + 1)-th node;
        Insert u to TempNodeSet;
    end if
    Replace CurrentNodeSet with TempNodeSet and then clear TempNodeSet;
end while
return the only node left in CurrentNodeSet, namely, the root of index tree T;
```

2. ABE Algorithm

ABE Setup: The setup algorithm takes as input a security parameter I. It outputs a public key PK and a master key MK.

KeyGen : For each users private key request, the initialization algorithm for delegated key generation takes as input an access policy (or attribute set) and the master key MK. It outputs the key partial transformation key.To achieve the same results with less time.

Encrypt: The encryption algorithm takes as input a message M and an attribute.

Decrypt: the decryption algorithm takes as input the ciphertext (ct) and the private key sk. It outputs the original message M.

IV. RESULTS AND DISCUSSION

Experimental Setup

For implementation the system required the software as: JDK with version 1.8, window platform is used. Netbeans 8.0.2 IDE tool is used, for the development. The experiment system is implement in Java language. Here experimentation is performed on enron email data set.

C. Results

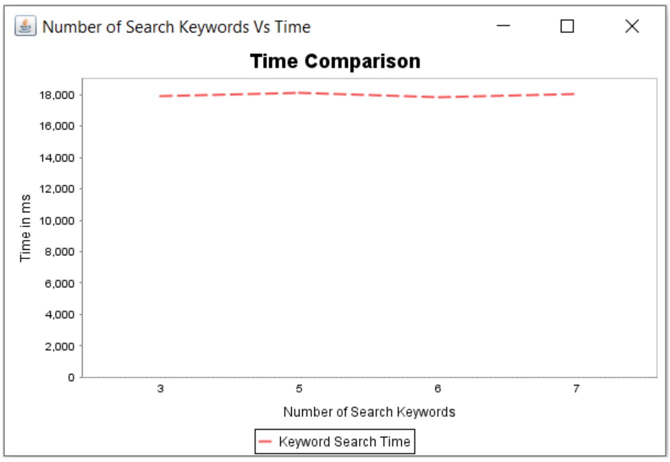

Figure 2. Time required to search documents for number of multi keyword search.

Figure 2 shows the time required to search number of documents for multi keywords search. From graph we can clearly say that with increasing number of keyword time required to search the documents over encrypted data is nearly same.

Figure 3 shows the computation memory required to search different number of keywords. In Figure 3 X-axis shows Number of Keywords while Y-axis shows memory required to search the keywords in bytes.

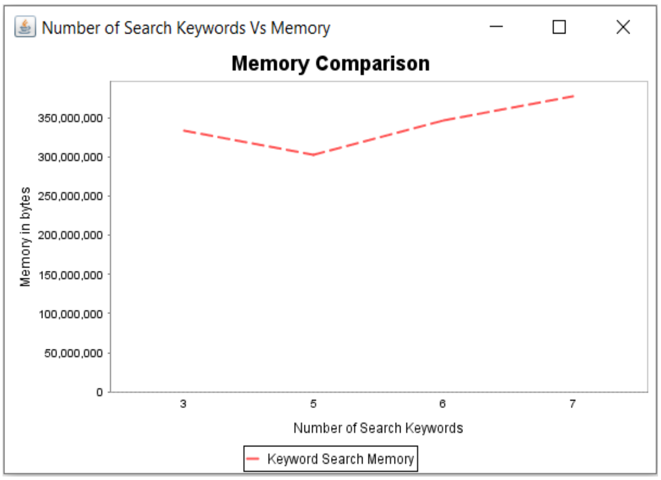

Figure 3. Computation memory required to search documents for number of multi keyword search

Figure 4 shows the server space utilization of existing and proposed system. Existing system not performed deduplication checking while proposed system performed deduplication checking of files before uploading files at cloud sever, in our experiment we got 2 duplicate files, so those another instance of files gets skip before uploading all data at cloud server which stores the server space as shown in below graph.

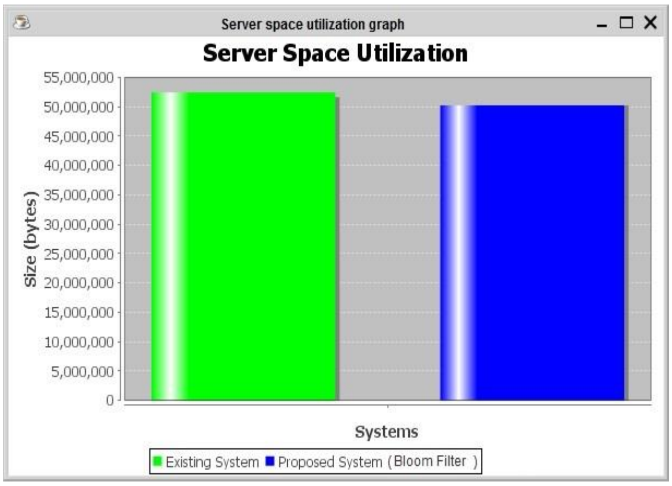

Figure 4. Server space utilisation comparison of existing and proposed system.

Table 1 shows the comparison with similar system.

Table 1. Comparison with similar system

Schemes	Multi Keyword Ranked Search	ABE	De-duplication Checking
[1]	Yes	No	No
[3]	Yes	Yes	No
Ours	Yes	Yes	Yes

V. CONCLUSION

The most important aspect that is to be considered in storing data is the security mechanisms associated with it. The proposed system presents a bloom filter and searchable Attribute Based Encryption scheme that is much more efficient than the previous systems. It provides security for appropriate users by using the user based access control attributes. In order to reduce the computation overhead of the user, the system provides modified outsourced ABE scheme which supports the outsourced key-issuing. One of the advantage of system is that is supports secure searching over encrypted data also faster de-duplication checking allows to remove duplicate files before uploading at cloud servers. Results show that our system is proficient as well as practical.

VI. FUTURE SCOPE

Cryptography has remained a core research and development area. New security techniques have made encryption strong and difficult to break. But there is always a scope for improvement and the efficiency of system could be improved by other ‘n’ number of ways. In future we can implement traceability method to trace or detect the fake user and make system more secure.

VII. REFERENCES

[1]. Dai, Xuelong & Dai, Hua & Yang, Geng & Yi, Xun & Huang, Haiping. (2019). An Efficient and Dynamic Semantic-Aware Multikeyword Ranked Search Scheme Over Encrypted Cloud Data. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2944476.

[2]. Patil, Rashmi & Gandhi, Yatin & Sarmalkar, Vinaya & Pund, Prajakta & Khetani, Vinit. (2020). RDPC: Secure Cloud Storage with Deduplication Technique. 1280-1283. 10.1109/I-SMAC49090.2020.9243442.

[3]. Shangping Wang, Duo Zhang, Yaling Zhang, And Lihua Liu, "Efficiently Revocable and Searchable Attribute-Based Encryption Scheme for Mobile Cloud Storage", IEEE Access Volume 6, June 2018.

[4]. Y. Li, F. Zhou, Y. Qin, M. Lin, and Z. Xu, “Integrity-veri_able conjunctive keyword searchable encryption in cloud storage,” Int. J. Inf. Secur., vol. 17, pp. 1_20, Nov. 2017, doi: 10.1007/s10207-017-0394-9.

[5]. D. Cash, J. Jaeger, S. Jarecki, C. Jutla, H. Krawczyk, M.-C. Rosu, and M. Steiner, Dynamic searchable encryption in very large databases: Data structures and implementation, in Proc. of NDSS, vol. 14, 2014.

[6]. J. Lai, R. Deng, C. Guan, and J. Weng, “Attribute-based Encryption with Verifiable Outsourced Decryption”, IEEE Trans. Inf. Forensics Security, vol. 8, no. 8, pp. 1343-1354, Aug. 2013.

[7]. M. Green, S. Hohenberger, and B.Waters, ``Outsourcing the decryption of ABE ciphertexts,'' in Proc. 20th USENIX Conf. Secur. (SEC). Berkeley, CA, USA: USENIX Association, 2011, p. 34.

[8]. S. Yu, C. Wang, K. Ren, and W.Lou, “Achieving Secure, Scalable, Fine-Grained Data Access Control in Cloud Computing”, in Proc. IEEE 29th INFOCOM, 2010, pp. 534-542.

[9]. L. Cheung and C. Newport, “Provably Secure Ciphertext Policy ABE”, in Proc. 14th ACM Conf. CCS, 2007, pp. 456- 465.

[10]. V. Goyal, O. Pandey, A. Sahai, and B. Waters, “Attribute- Based Encryption for Fine-Grained Access Control of Encrypted Data,” in Proc. 13th ACM Conf. Comput. Commun. Security, 2006, pp. 89-98

[11]. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, Searchable symmetric encryption: improved definitions and efficient constructions, in Proceedings of the 13th ACM conference on Computer and communications security. ACM, 2006, pp. 79-88.

[12]. A. Sahai and B. Waters, “Fuzzy Identity-Based Encryption”, in Proc. Adv. Cryptol.- EUROCRYPT, LNCS 3494, R. Cramer, Ed., Berlin, Germany, 2005, pp. 457-473, Springer- Verlag.

[13]. D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, Public key encryption with keyword search, in Advances in Cryptology- Eurocrypt 2004. Springer, 2004, pp. 506-5

Cite this article as :

Vishal Jalindar Gondil, Prof. H. A. Hingoliwala, "ABE and Bloom Filter Based Searchable Attribute-Based Encryption Scheme over Cloud Data", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 8 Issue 3, pp. 208-215, May-June 2021.

Journal URL : https://ijsrst.com/IJSRST218353