

# IMPLEMENTATION OF PRIVACY PRESERVING AND DYNAMIC SEARCHING MECHANISM WITH BIOMETRIC AUTHENTICATION IN CLOUD STORAGE

Sarita Motghare<sup>1</sup>, Dr. Dheeraj Rane<sup>2</sup>,

<sup>1</sup>Research Scholar, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India

<sup>2</sup>Associate Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India

**Abstract:** In the recent times, cloud storage tends to be a primary storage means for external data. Cloud defense of the data against attacks is the main challenge. Private or semi-private information growth has rapidly expanded over the information network; privacy safeguards have failed to address the search mechanisms. In the field of information networks, privacy protection is an important factor in carrying out various data mining operations with encrypted data stored in different storage systems. A tolerance and protection against data corruption mechanism should be developed which is difficult to achieve. Furthermore, as there is no adequate audit mechanism, the integrity of the stored data become questionable. In addition to this, the user authentication is another challenge. The current solution provides only a remote audit mechanism. It requires data owners to always remain online so that the auditing process is manually handled, which is sometimes unworkable. In this paper, we propose a new, regenerative, public audit methodology accompanied by third-party audits. The existing data search system provides one solution that can be used to maintain the confidentiality of indexing. Documents are stored on a private server in plain word form, which compromise the protection of privacy. So that this system is improved to make the document more secure and efficient, we first store the documents in encrypted form on server, and use the Key Distribution Center (KDC). To generate keys the KDC uses the user's biometric feature. In order to improve the search experience, we also implement

TF-IDF, which provides an efficient evaluation of the results. Lastly, we carry out comprehensive data set experiments to evaluate our proposed system performance. Experimental results demonstrate that in terms of safeguarding the privacy, efficient and safe search for encrypted distributed documents the proposed system is better than existing. The methodology suggested also includes an auditing mechanism by third parties to ensure data integrity.

**Keywords—** Privacy Preserving, Third Party Auditing, Regenerating Data, Provable Secure, Dynamic Searching, Biometric Authentication

## I. INTRODUCTION

In view of its various amazing focal points as centralised storage of data, bulky calculation, low-value benefit and an appropriate approach to accessing data, cloud computing is widely undertrained by numerous associations and individuals. Virtualization is the basic idea behind cloud computing. Virtualization in cloud computing intends to make a virtual range of a gadget or asset such as a server, storage gadget, organize or work system where the structure partitions the asset into the required number of running conditions. Cloud Computing is a dominant cloud storage management system that enables data owners to store their data from cloud to cloud. Many customers prefer to store their data on the cloud. In any case, there is also a security and productivity issue in the new data management convention.

Cloud providers are called organizations providing cloud computing services. The use of Cloud computing technology is categorized into four main types: private cloud, public cloud, hybrid cloud, and

community cloud, depending on internal or external ownership and technical architecture. A private cloud is a privately owned, internally developed cloud infrastructure. Public cloud is a model of cloud computing in which large companies offer their customers cloud facilities.

Most of the features are limited to customers, and customers are free to purchase more services on a per-use basis on requirement. Azure by Microsoft and Amazon Web Services (AWS) by Amazon are some well-known examples of public cloud. Hybrid cloud is both a private and a public cloud for various activities in an organization. The Community Cloud is the implementation of cloud-based services shared by certain community organizations that have a common working field, such as security, jurisdiction, compliance, etc.

Cloud storage is a storage facility in which users can store their information on IaaS. Some of the famous and popular cloud storage providers are Google Drive, Drop box, Apple iCloud, and more. The cloud-based storage provider keeps cloud data intact and is accessible to users 24 x 7. It is also the responsibility of cloud storage providers to maintain the physical environment.

Cloud services are susceptible to faults, failures and attacks, as are other computing technologies. The major difference is that if a cloud has a fault or attack many users / customers may be affected, which a major concern for cloud security is. Cloud security is not only concerned with access and access to cloud data for its owners, but also with providing data integrity[1-5] and confidentiality to ensure that cloud-based services are implemented effectively. A service provider should ensure that user data are efficiently secured and the data is accessed only by authenticated and authorized users. Because users can store certain personal or confidential information without being aware of the place or storage, it is likely that this confidential or private information can be disclosed by Cloud storage providers.

As we know, biometric identification is becoming increasingly important as the solution to user identification proved to be precise and promising. It is considered to be a reliable and convenient way, as opposed to existing authentication methods for user

identification. The most challenge here is to develop a method for an efficient and reliable biometric identification scheme to protect cloud privacy.

Generally, an audit is done when an independent third party group is involved in finding evidence through several processes, including investigation, physical inspection, observation, confirmation, etc. Cloud implementation audit involves the monitoring of all internal and external processes implemented by any organisation, identifying compliance requirements, including legislation and regulations, Service Level Agreements (SLAs) and corporate policies.

These needs for knowledge acquisition are lacking in existing solutions. Some systems give the implementation of storage complexity a weaker guarantee: the server should save at least as large an associated amount of knowledge as the customer's knowledge but not essentially constant precise knowledge. In addition, all previous techniques require the server to access the entire file that cannot be handled once large quantities of expertise are addressed.

The next issue is the safety mechanism associated with the cloud access to data. The traditional encryption key generation mechanism cannot prevent the advance attacks of the attackers. A user-based encryption method could be an effective way in this scenario. It provides a unique key based on a user attribute that can make access to data very difficult without authentication but still contains a definite range of errors. An attribute-based mechanism coupled with user fingerprint can overcome such an event and a hybrid approach and impair the system's unauthenticated access to data.

Too much focus on the security issues and to implement the resolution of them can also result in the efficiency degrading of the legitimate user's access to information. It therefore cannot be a good idea to focus only on improving safety measures and ignore efficiency. The system should be safe, but efficiencies must be maintained in accessing and searching data.

Data proprietors can save their files in the data networks on a number of distributed servers. It provides users with services for the storage and access to data from and into server number everywhere and via any device. The efficient search for distributed

files and the privacy of owners' files is a very difficult task. In order to resolve this problem, the existing system uses a technology called privacy indexing. The main objective of PPI is a globally controlled search facility run by a third-party entity. For providers such as complete access control and privacy, the PPI architecture is appropriate.

PPI [6] is a directory service for public cloud. Control of different public cloud private servers. The data is stored over the number of private servers by distribution. Different users can locate files via distributed data. A search can be made by the corresponding file user with related PPI [6] [7] server keywords. This public server return contains a list of private servers in the network.

Afterwards, users are able to access the private server in their candidate list before searching locally for authentication. This system allows users to search the required files directly through the saving of information to private server. Data security is nevertheless essential, so the data on the proposed systems is encrypted on private servers. Therefore, after authentication, users must authenticate and access encrypted files on the private server. After acquiring encrypted files, the files are decrypted with the KDC. A key distribution center (KDC) is part of a cryptography system designed to minimize the hazards inherent in commercial keys. KDC works in frameworks in which customers can consent to the use of certain services several times and not others. KDC provides authorized users with a key to decrypt the files. If the original files are added, system then implements the TF-IDF file ranking to achieve highest results in classification format.

In this paper, we concentrate on the integrity control problem of cloud storage regeneration-based codes, in particular the functional repair strategy [12]. Bo Chen et al. [8] and H conducted similar studies. Chen et al. [9] is independent and separate. [8] extended CPOR to the regenerating code scenario for a single server version (private version [13]); [9] designed and implemented an FMSR cloud-based data integrity (DIP) system and a Thin Cloud Set-Act scheme (TCSS) [14]. However, both are for private audit, the integrity of the faulty servers can be only verified and repaired by the data owner. In view of the size of

outsourced data, and the limited resources of users, cloud auditing and repair tasks can be tremendous and costly for users [15]. The overhead of using cloud storage should be minimized so as to prevent a user from performing too many out-of-source operations (besides retrieving them) [16]. Users may not want the complexity of checking and repairing in particular. The auditing schemes in [8] [9] imply the difficulty in which users should always remain online, which in practice, particularly in long-term storage, can hamper their adoption. We propose a public audit system in which integrity checks and regenerations (of failed persons) are carried out to ensure the data integrity and to save users' computing resources and online burdens.

Instead of adjusting the existing public auditing scheme [13] directly to multi-server settings, we design a new authenticator that is more suitable for regenerating codes. Auditor and a semi-trusted proxy separately on behalf of data holders. In addition we are "encrypting" the data privacy coefficients from the auditor that is lighter than the application of evidence blind technology in [15], [16] and data blind methods in [17]. In our new system model with a proxy, there are several challenges and threats, and a security analysis shows that our system works well. In particular, the following aspects can summarize our contribution:

- We design a new homomorphic authenticator based on BLS [18], which can be generated and checked publicly by a few secret keys. The authenticators can be calculated effectively using the linear subspace of the regenerating codes. It can also be adapted for data owners who only require signatures of native blocks, equipped with low end computing devices (e.g. Tablet PC etc.).
- Our scheme is, in our best knowledge, the first to enable public audit for code-based cloud storage to be privacy-preserved. During the configuration phase the coefficients are masked with a PRF (Pseudorandom Functions). This is a lightweight method that does not introduce any cloud server or TPA computational overhead.
- In order to regenerate block and authenticators on faulty servers, we release data holders completely from online burden and it gives the privilege to a proxy for repairs.

- Optimisation measures are taken to increase our auditing system's flexibility and efficiency, thereby effectively reducing the overhead storage of servers, the overhead computing of the data owner, and overall communication during the audit phase.

## II. RELATED WORK

Biometric identification has turned out to be progressively in recent times. With the advancement of cloud computing, data owners are persuaded to redistribute the expansive size of biometric data. Identification errands to the cloud to dispose of the costly storage and calculation costs, which, nevertheless, conveys potential dangers to clients' privacy.

In this study presented by Liehuang Zhu et al. the author proposes a productive and privacy-saving biometric identification-redistributing plan [19]. In particular, use of biometric to execute a biometric identification. The database owner encodes the data and submits it to the cloud. The cloud performs identification tasks over the database and returns the outcome to the database owner. Here author accept that the biometric data has been handled to such an extent that its portrayal can be utilized to execute biometric coordinate. Without loss of simplification, the system target fingerprints and utilize Finger Codes to speak to the fingerprints. To assess the proficiency and security prerequisites, the author actualizes another encryption calculation and cloud authentication confirmation. The evaluation result and examination indicate it can oppose the potential assaults.

In this study by XialiHei et al. proposes a lightweight secure access control conspire for IMDs amid crises [20]. This plan uses patient's biometric data to forestall unauthorized access to IMDs. The plan comprises of two levels: level 1 utilizes some essential biometric data of the patient and it is lightweight; level 2 uses patients' iris data for authentication and it is exceptionally viable. In this exploration, the author additionally makes commitments to human iris check: we find that it is conceivable to perform iris confirmation by looking at halfway iris data instead of the whole iris

data. The evaluation aftereffects of the system demonstrates that the safe access control plot is exceptionally viable. It has little overhead henceforth possible for IMDs. In particular, the false acknowledgment rate (FAR) and false dismissal rate (FRR) of our safe access control conspire are near 0.000% with a reasonable edge, and the memory and calculation overheads are satisfactory.

For security, it is essential that the customer does not pick up anything on the database. The server ought not to get any data about the asked for biometry and the result of the coordinating procedure. The proposed convention by Mauro Barni et al. pursues a multi-party calculation approach and makes broad utilization of homomorphic encryption as fundamental cryptographic crude [21]. To keep the convention intricacy as low as could be expected under the circumstances, a specific portrayal of fingerprint pictures, named Finger code, is received. In spite of the fact that the past chips away at privacy-saving biometric identification center around choosing the best coordinating character in the database, this arrangement is a nonexclusive identification convention and it permits to choose and report all the enlisted personalities whose separation to the client's Finger Codes is under a given limit. Variations for basic authentication reasons for existing are given. According to the evaluation result, these conventions gain a remarkable data transfer capacity sparing (around 8 to 24%) whenever contrasted and the best past work and its computational many-sided quality is still low and appropriate for down to earth applications.

Here by Yan Huang et al. and by Yan Huang et al. introduces a productive coordinating convention that can be utilized in numerous privacy-preserving biometric identification systems in the semi-fair setting [22]. Our most broad specialized commitment is another backtracking convention that uses the result of evaluating a jumbled circuit to empower productive neglectful data recovery. We additionally present a more effective convention for computing the Euclidean separations of vectors and streamlined circuits for

finding the nearest coordinate between points held by one gathering and an arrangement of focuses held by another. For evaluation reason, usage of a down to earth privacy-safeguarding fingerprint coordinating system is been finished. The fundamental downside is that present conventions for privacy-protecting calculations are extremely costly and unfeasible for genuine scale issues. In this work, the author has demonstrated that those expenses can be significantly diminished for an expansive class of biometric coordinating applications by creating productive conventions for Euclidean separation, finding the nearest coordinate, and recovering the related record.

Data sharing turns into an outstandingly alluring administration provided by cloud computing stages in view of its accommodation and economy. As a potential procedure for acknowledging fine-grained data sharing, attribute-based encryption (ABE) has drawn wide consideration. The issue of at the same time accomplishing fine grainedness, high productivity on the data owner's side, and standard data privacy of cloud data sharing in reality still stays uncertain. This paper by Jin Li et al. addresses the testing issue by proposing another attribute-based data sharing plan reasonable for asset restricted portable clients in cloud computing. The system model for attribute based data sharing system [23]. This plan dispenses with a lion's share of the calculation undertaking by including system public parameters other than moving fractional encryption calculation disconnected. What's more, a public ciphertext test stage is performed before the decryption stage, which disposes of the vast majority of the calculation overhead because of ill-conceived ciphertexts. For data security, a Chameleon hash work is utilized to produce a prompt ciphertext, which will be blinded by the disconnected ciphertexts to get the last online ciphertexts.

Identity-Based Encryption (IBE), which rearranges the public key and endorsement administration at Public Key Infrastructure (PKI), is a critical option in contrast to public key encryption. In any case, one of the primary productivity downsides of IBE

is the overhead calculation at Private Key Generator (PKG) amid client disavowal.

The system by Jin Li et al. going for handling the basic issue of identity denial, the author brings re-appropriating calculation into IBE and presents a revocable IBE conspire in the server-supported setting [24]. This plan offloads the vast majority of the key age related activities amid key-issuing and key-refresh procedures to a Key Update Cloud Service Provider, leaving just a consistent number of basic tasks for PKG and clients to perform locally. To accomplished this objective use of a novel plot safe procedure is utilized i.e. utilizing a mixture private key for every client, in which an AND door is included to associate and bound the identity part and the time segment. According to the evaluation results, the system accomplishes consistent effectiveness for both calculation at PKG and private key size at the client. Additionally, User needs not to contact with PKG amid the key refresh, as it were, PKG is permitted to be disconnected in the wake of sending the disavowal rundown to KU-CSP. In addition, finally, no protected channel or client authentication is required amid key-refresh between client and KU-CSP.

This paper by Jin Li et al. endeavours to address the issue of accomplishing productive and solid key administration in secure deduplication [25]. The system presents a gauge approach in which every client holds a free ace key for scrambling the focalized keys and redistributing them to the cloud. Nevertheless, such a gauge key administration plot produces a huge number of keys with the expanding number of clients and expects clients to dedicatedly ensure the ace keys. To this end, the author proposes Dekey, another development in which clients do not have to deal with any keys without anyone else however rather safely appropriate the focalized key offers over various servers. Security examination exhibits that Dekey is secure as far as the definitions determined in the proposed security show.

ABE gives a protected way that enables data owner to share outsourced data on untrusted storage server rather than a confided in server with a

predefined gathering of clients. This preferred standpoint makes the strategy engaging in cloud storage that requires secure access control for an extensive number of clients having a place with diverse associations. By the by, one of the principle proficiency drawback of ABE is that the computational expense amid the decryption stage develops with the intricacy of the access recipe. Subsequently, before broadly sent, there is an expanding need to enhance the effectiveness of ABE. To address this issue, outsourced ABE, which gives an approach to redistribute escalated computing assignment amid decryption to CSP without uncovering data or private keys, was presented. Going for wiping out the overhead calculation at both the attribute authority and the client sides, we propose an outsourced ABE plot supporting outsourced decryption as well as empowering delegating key age. In this development by Jin Li et al. presents an insignificant arrangement controlled by a default attribute and utilize an AND door associating the inconsequential strategy and client's approach [26]. Amid key issuing, attribute authority can re-appropriate calculation through appointing the assignment of producing a halfway private key for client's arrangement to a key age specialist co-op (KGSP) to decrease neighbourhood overhead. In addition, the outsourced decryption is acknowledged by using key blinding. All the more accurately, the client can send the blinded private key to a decryption specialist co-op (DSP) to perform fractional decryption and do the total decryption at nearby. Following our method, steady effectiveness is accomplished at the two attributes authority and client sides.

Mysterious attribute-based encryption (unknown ABE) empowers fine-grained access control over cloud storage and jelly beneficiaries' attribute privacy by concealing attribute data in ciphertexts. Nevertheless, in existing unknown ABE work, a client knows whether attributes and a concealed arrangement coordinate or not just in the wake of rehashing decryption endeavours. What's more, every decryption as a rule requires numerous pairings and the calculation overhead develops

with the many-sided quality of the access recipe. Henceforth, existing plans endure an extreme proficiency downside and are not reasonable for portable cloud computing where clients might be asset compelled. The study of system by Yinghui Zhang et al. proposes a novel procedure called "coordinate then-unscramble", in which a coordinating stage is also presented before the decryption stage [27]. This method works by computing unique segments in ciphertexts, which are utilized to play out the test that if the attribute private key matches the shrouded access strategy in ciphertexts without decryption. For quick decryption, exceptional attribute mystery key segments are created which permit accumulation of pairings amid decryption. We propose an essential mysterious ABE development and afterward get a security-improved expansion based on emphatically existentially unforgeable one-time marks. In the proposed developments, the calculation cost of an attribute coordinating test is short of what one decryption activity, which just needs a little and consistent number of pairings. Formal security investigation and execution examinations show that the proposed arrangements at the same time guarantee attribute privacy and enhance decryption proficiency for outsourced data storage in portable cloud computing.

The study of system presented by John Bethencourt et al. for acknowledging complex access control of encoded data that we call Ciphertext-Policy Attribute-Based Encryption by utilizing this strategy, scrambled data can be kept private regardless of whether the storage server is untrusted, also, this technique counters intrigue assaults [28]. Past Attribute-based Encryption, systems utilized attributes to depict the encoded data and incorporated arrangements with client's keys; while in this system, attributes are utilized to portray a client's qualifications, and a gathering scrambling data decides an approach for who can unscramble. Pretty much this technique is adroitly nearer to customary access control strategies, for example, Role-Based Access Control (RBAC). Apart from secure key distribution and privacy preserving, one of the main challenge is to assure

the confidentiality and integrity of the data. At present, there is a significant increment in the measure of data put away in storage administrations, alongside the emotional advancement of systems administration methods. In storage administrations with gigantic data, the storage servers might need to diminish the volume of put away data, and the customers might need to screen the integrity of their data with a minimal effort since the expense of the capacities identified with data storage increment in extent to the span of the data. To accomplish these objectives, secure deduplication and integrity auditing appointment procedures have been contemplated, which can lessen the volume of data put away in storage by taking out copied duplicates and allow customers to effectively check the integrity of put away records by designating expensive activities to a confided in party, individually. The plan displayed by Taek-Young Youn et al. bolsters both secure deduplication and integrity auditing in a cloud situation [29]. Specifically, this plan gives a protected deduplication of scrambled data. The proposed plot additionally bolsters public auditing utilizing a TPA (Third Party Auditor) to help low-fuelled customers. The technique fulfils all essential security prerequisites and is more effective than the current plans that are intended to help deduplication and public auditing in the meantime.

To secure outsourced data in cloud storage against defilements, adding adaptation to non-critical failure to cloud storage together with data integrity checking and disappointment reparation winds up basic. As of late, recovering codes have picked up notoriety because of their lower repair transfer speed while giving adaptation to non-critical failure. Existing remote checking strategies for recovering coded data just give private auditing, requiring data owners to dependably remain on the web and handle auditing, and also repairing, or, in other words. The paper presented by Jian Liu et al. centers around the integrity confirmation issue in recovering code-based cloud storage, particularly with the utilitarian repair methodology [30]. The

featuring parts of this study can be outlined by the accompanying angles:

- This system utilizes a novel homomorphic authenticator based on BLS signature, which can be created by two or three mystery keys and confirmed publicly. Using the straight subspace of the recovering codes, the authenticators can be registered effectively. In addition, it very well may be adjusted for data owners furnished with low-end calculation devices (e.g. Tablet PC and so on.) in which they just need to sign the local squares.
- It is the principal plan to permit privacy-protecting public auditing for recovering code-based cloud storage. A PRF (Pseudorandom Function) amid the Setup stage to keep away from spillage of the first data veils the coefficients. This technique is lightweight and does not acquaint any computational overhead with the cloud servers or TPA. Largely, this plan totally discharges data owners from the online weight for the recovery of squares and authenticators at broken servers and it gives the benefit to a proxy for the reparation.

With data storage and sharing administrations in the cloud, clients can undoubtedly change and offer data as a gathering. To guarantee shared data integrity can be checked publicly, clients in the gathering need to register marks on every one of the squares in shared data. Distinctive squares in shared data are for the most part marked by various clients because of data alterations performed by various clients. The paper presented by Boyang Wang et al. proposes a novel public auditing component for the integrity of imparted data to productive client renouncement at the top of the priority list [31]. By using the possibility of proxy re-marks, the system enables the cloud to leave hinders in the interest of existing clients amid client repudiation with the goal that current clients don't have to download and re-sign squares without anyone else's input. What's more, a public verifier is constantly ready to review the integrity of shared data without recovering the whole data from the cloud, regardless of whether some piece of shared data has been re-marked by the cloud. Besides, our component can bolster clump auditing by

confirming different auditing errands at the same time.

Keeping in mind the end goal to address the issue of data integrity over the cloud and further accomplish a safe and reliable cloud storage benefit by Cong Wang et al. proposes an adaptable circulated storage integrity-auditing instrument, using the homomorphic token and appropriated deletion coded data [32]. The proposed configuration enables clients to review the cloud storage with exceptionally lightweight correspondence and calculation cost. The auditing result guarantees solid cloud storage accuracy ensure as well as at the same time accomplishes quick data mistake restriction, i.e., the identification of getting into mischief server. Considering the cloud data are dynamic in nature, the proposed configuration additionally underpins secure and effective unique activities on outsourced data, including square alteration, erasure, and affix.

Cloud computing can gather and redesign a gigantic measure of IT assets and obviously, the cloud servers can give more anchor, adaptable, different, financial and customized administrations contrasted and the neighbourhood servers. Likewise, to make full utilization of the data on the cloud, the data clients need to access them adaptable and efficient. Therefore, a colossal test of re-appropriating the data to the cloud is the means by which to ensure the classification of the data legitimately while keeping up their accessibility. The paper presented by Na Wang et al. has a hierarchical attribute-based encryption conspire is first intended for a record gathering [33]. An arrangement of reports can be scrambled together on the off chance that they share a coordinated access structure. Contrasted and the ciphertext-arrangement attribute-based encryption (CP-ABE) plans, both the ciphertext storage space and time expenses of encryption/decryption are spared. At that point, a file structure named attribute-based recovery highlights (ARF) tree is developed for the report gathering based on the TF-IDF show and the archives' attributes. A profundity first looks calculation for the ARF tree is intended to enhance

the pursuit effectiveness, which can be additionally enhanced by parallel computing.

The paper presented by Yuzhe Tang et al. addresses the understudied issue for the PPI methods i.e how to give separated privacy protection within the sight of multi-keyword report look [34]. The separation is vital as terms and expressions bear natural contrasts in their semantic implications. In this paper, we present e-MPPI, the principal work to furnish the conveyed report look with quantitatively separated privacy safeguarding. In the plan of e-MPPI, we recognized a suite of difficult issues and proposed novel arrangements. For one, we figured quantitative privacy calculation as an enhancement issue that strikes a harmony between privacy safeguarding and seeks productivity. We likewise tended to the testing issue of secure e-MPPI development in the multi-space data organize which needs common trusts between areas. Towards a safe e-MPPI development with satisfactory execution, we proposed to upgrade the execution of secure multi-party calculations by making a novel utilization of mystery sharing. We executed the e-MPPI development convention with a working model.

The paper by S Sharma et al. studied the issue of building a protected cloud storage benefit on best of a public cloud foundation where the specialist organization isn't totally trusted by the client [35]. We portray, at an abnormal state, a few structures that join later and non-standard cryptographic natives so as to accomplish our objective. The paper by D Boneh et al. presents the issue of searching for data that is encoded utilizing a public key system [36]. Consider client Bob who sends email to client Alice scrambled under Alice's public key. An email gateway needs to test whether the email contains the keyword "dire" with the goal that it could course the email as needs are. Alice, then again does not wish to enable the gateway to unscramble every one of her messages. We characterize and develop a component that empowers Alice to give a key to the gateway that empowers the gateway to test whether "earnest" is a keyword in the email without getting the hang of whatever else about the email. We allude to this

instrument as Public Key Encryption with Keyword Search. As another model, consider a mail server that stores different messages publicly scrambled for Alice by others. Utilizing our instrument Alice can send the mail server a key that will empower the server to recognize all messages containing some explicit keyword, yet pick up nothing else.

Here by D Boneh et al. demonstrated to make a public-key encryption plot for Alice that permits PIR searching over scrambled archives [37]. Our answer gives a hypothetical answer for an open issue presented by Boneh, et al. on "Public-key Encryption with Keyword Search", giving the principal conspire that does not uncover any incomplete data in regards to client's search in the public-key setting and with non-inconsequentially little correspondence unpredictability. The primary procedure of our answer additionally takes into account Single-Database PIR composing with sub linear correspondence unpredictability, which we consider autonomous intrigue. Here by D. X. Song et al. depicted their cryptographic plans for the issue of searching on encoded data and give verifications of security to the subsequent cryptosystems [20]. Our methods have various urgent points of interest. They are provably secure: they give provable mystery to encryption, as in the untrusted server can't pick up anything about the plaintext when just given the ciphertext; they give question separation to searches, implying that the untrusted server can't get the hang of much else about the plaintext than the search result; they give controlled searching, so that the untrusted server can't search for a discretionary word without the client's authorization; they likewise bolster concealed inquiries, so the client may approach the untrusted server to search for a mystery word without uncovering the word to the server.

The paper by Y. C. Chang et al. offer answers for this issue under very much characterized security necessities [38]. Their plans are proficient as in no public key cryptosystem is included. To be sure, our methodology is free of the encryption strategy decided for the remote documents. They are likewise gradual, in that you can submit new

records which are secure against past inquiries yet searchable against future questions. Paper presented by R. Curtmola et al. started exploring existing thoughts of security and propose new and more grounded security definitions [39]. We at that point present two developments that we demonstrate secure under our new definitions. Strangely, notwithstanding fulfilling more grounded security ensures, our developments are more productive than every single past development.

### III. MOTIVATION

It is noted that data owners lose ultimate control over the fate of their outsourced data; thus, the authenticity, availability and integrity of the data are being put at risk. On the other hand, the cloud service providers may act dishonestly, attempting to hide data loss or corruption and claiming that the files are still correctly stored in the cloud for reputation or monetary reasons. Thus it makes great sense for users to implement an efficient protocol to perform periodical verifications of their outsourced data to ensure that the cloud indeed maintains their data correctly.

User Attribute-based access control provides a high level of flexibility that promotes security and information sharing. It is an interesting alternative to standard public key encryption, which is based on simplifying key management in a certificate-based Public Key Infrastructure by using human-features or unique identities (e.g., unique name, email address, IP address, etc.) as public keys. Therefore, sender using this system does not need to look up public key and certificate, but directly encrypts message with receiver's identity. Accordingly, receiver obtaining the private key associated with the corresponding identity from Private Key Generator is able to decrypt such ciphertext.

As the sensitive data should be encrypted before outsourcing for privacy requirements, which obsoletes data utilization like keyword-based document retrieval. One challenge is that the relationship between documents will be deformed in the process of encryption, which will lead to significant search accuracy performance degradation. Also the size of data in data storages has experienced a significant

growth. This will make it even more complex to design encoded data search schemes that can provide efficient and reliable data retrieval on large size of encrypted data. Therefore, proposing a method which can maintain and utilize this relationship to speed the search phase is desirable.

The outsourced information in cloud storage is vulnerable to corruption or modification. Fault tolerance to cloud storage along with integrity verification of data gets hard. Apart from this the security of the data over cloud and privacy of the user on client side is also an important issue. While addressing the above issue system gets so complicated that the accessing capacity or efficiency of the systems

gets affected. In this study we need to focus on all the aspects of the cloud storage which includes,

- Security, privacy and integrity of the data over cloud
- Privacy and verification of authenticity of the user while accessing the data.
- Efficiency of the system while searching and accessing the data.

IV. SYSTEM MODEL AND DESIGN

The detailed system architecture is illustrated in Figure 5.1. This system is made up of three different entities: cloud server, data owner and data user.

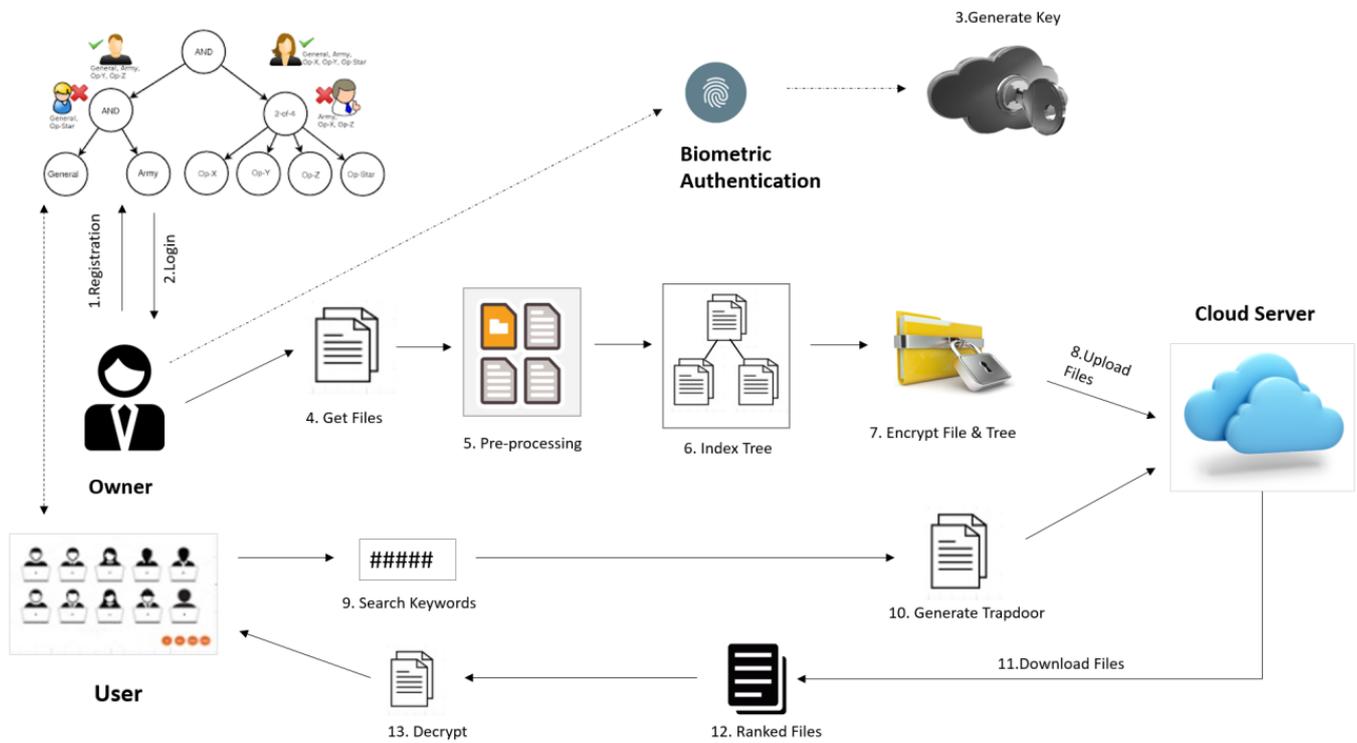


Figure 1. System Architecture

Data Owner has a list of cloud-encoded and cloud-specified files. This encrypted information can be searched by the user. The data owner creates a secure index tree on this system, and the encrypted file is created thereafter. This index tree and encrypted files are saved on the server of the cloud. Data owner shall be responsible to the authorized users for the key distribution required for decryption of files. Based on a user query request for a specific document, the cloud server searches through the index tree and lists the user

with encrypted top-k results. Use the secret key provided by the data owner to decrypt the obtained files.

Cloud server enables users to store their respected hash and encrypted file blocks. A distributed KDC is available for this file block encryption. System uses distributed KDC, since another one will be used if one KDC is busy. This distributes the load to KDC and improves performance. Using the key, the file blocks can be encrypted. The user generates and store the

hash of block files on the server before storing block files on cloud storage.

The user can request a cloud server store in TPA for file block completeness checks. The hash blocks are stored by TPA. It calls for a User to check the integrity of particular file requests. It compares the file block

hash received in its database with the hash store received. If the hash matches, then it will send the message to the user that does not corrupt the store files on the server. TPA requests a proxy to correct the file if the file is corrupted. Proxy with code for regeneration.

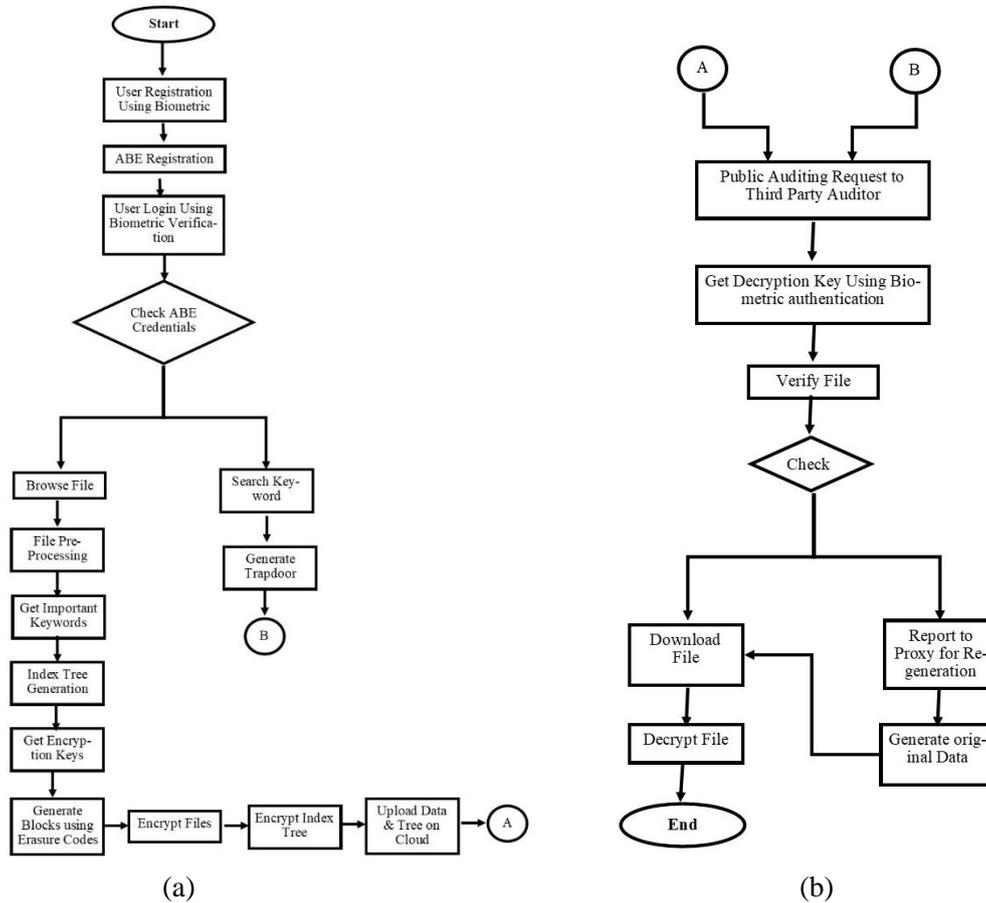


Figure 2 (a) (b) Flowchart of the System

Proxy can recover corrupted files on the server by using this regeneration code. And then TPA again checks if the file is recovered. TPA informs the user that the file has been recovered. The proposed system flows are illustrated in Figures 2 (a) and (b).

**System model for Biometric Identification Scheme**

In our scheme, we assume that the biometric data has been processed such that its representation can be used to execute biometric match. Without loss of generality, we target fingerprints and use FingerCodes to represent the fingerprints. More specifically, a FingerCode consists of n elements and each element is a 1-bit integer (typically n = 640 and l = 8). Given two FingerCodes x D [x<sub>1</sub>; x<sub>2</sub>... x<sub>n</sub>] and y D [y<sub>1</sub>; y<sub>2</sub>;...; y<sub>n</sub>], if

their Euclidean distance is below a threshold ε, they are usually considered as a good match, which means the two fingerprints are considered from the same person.

In order to achieve practicality, both security and efficiency are considered in the proposed scheme. To be more specific, design goals of the proposed scheme are described as follows:

- Efficiency: Computational costs should be as low as possible at both the database owner side and the user side. To gain high efficiency, most biometric identification operations should be executed in the cloud.

- Security: During the identification process, the privacy of biometric data should be protected. Attackers and the semi-honest cloud should learn nothing about the sensitive information.

Here, we list the main notations used in the remaining section as follows.

$b_i$  - the  $i$ -th sample FingerCode, denoted as an  $n$ -dimensional vector  $b_i \in \mathbb{R}^n$  [ $b_{i1}; b_{i2}; \dots; b_{in}$ ].

$B_i$  - the extended sample FingerCode of  $b_i$ , denoted as an  $(n+1)$ -dimensional vector  $B_i = [b_{i1}; b_{i2}; \dots; b_{i(n+1)}]$ , where  $b_{i(n+1)} = -0.5(b_{i1}^2 + b_{i2}^2 + \dots + b_{in}^2)$ .

$b_c$  - the query FingerCode, denoted as an  $n$ -dimensional vector  $b_c = [b_{c1}; b_{c2}; \dots; b_{cn}]$ .

$B_c$  - the extended query FingerCode of  $b_c$ , denoted as an  $(n+1)$ -dimensional vector  $B_c = [b_{c1}; b_{c2}; \dots; b_{c(n+1)}]$ , where  $b_{c(n+1)} \in \mathbb{R}$ .

$W$  - The secret keys collection, denoted as  $W=(M_1; M_2; M_3; H; R)$ , where  $M_1, M_2$  and  $M_3$  are  $(n+1) \times (n+1)$  invertible matrices, and  $H, R$  are  $(n+1)$ -dimensional row vectors.

$I_i$  - the searchable index associated with the  $i$ -th sample FingerCode  $b_i$ .

The Steps for implementation of the scheme is as follows:

*Step 1:* The database owner randomly generates an  $(n+1) \times (n+1)$  matrix  $A$  where  $H \times A^T \in \mathbb{R}^{(n+1)}$  and  $A_i$  is a row vector in  $A$ ,  $1 \leq i \leq (n+1)$ .

Then, the database owner generates a corresponding matrix. After that, the database owner performs the following operations:

$$C_i = M_1 \times D_i \times M_2, \quad (1)$$

$$C_h = H \times M_1^{-1}, \quad (2)$$

$$C_r = M_3^{-1} \times R^T. \quad (3)$$

Subsequently, the database owner uploads  $(C_i; C_h; C_r; I_i)$  to the cloud, where  $I_i$  is the index of  $B_i$ .

*Step 2:* After Step 1 is executed, the cloud has stored many tuples in its database  $C$ . When a user requests to identify his/her identity, he/she extends  $b_i$  and then submits the extended query  $B_i$  to the database owner.

On receiving the request from the user, the database owner generates a random  $(n+1) \times (n+1)$  matrix  $E$  such that  $E_i \times R^T = 1$ , where  $E_i$  is a row vector in matrix  $E$ .

The database owner then generates a corresponding matrix to hide the query FingerCode  $B_c$ . The Database owner then performs the following operations:

$$C_f = M_2^{-1} \times F_c \times M_3. \quad (4)$$

Then, the database owner uploads  $C_f$  to the cloud.

*Step 3:* On receiving  $C_f$ , the cloud begins to search for the best match. Specifically, the cloud computes  $P_i = C_h \times C_i \times C_f \times C_r$  for all encrypted biometric database to compare the Euclidean distances between  $b_c$  and  $b_i$ . Then, the database owner compares the Euclidean distance with the standard threshold. If the distance is less than the threshold value, the query is identified. Otherwise, the identification fails.

*Step 4:* Finally, the database owner returns the identification result to the user.

### Mathematical formulation for Encryption Process

System  $S$  is represented as  $S = \{U, CS, KU-CSP\}$

1. User  $U = \{R, L, Q, E, V\}$

Where,  $R$  = Registration Process

$L$  = Login Process

$Q$  = Key Request Process

$E$  = File Encryption Process

2.  $KU-CSP = \{PK, SK\}$

Key Generation  $PK = \{pk_1, pk_2, pk_3 \dots pk_n\}$

Where  $PK$  is the set of generate public keys.

$SK = \{sk_1, sk_2, sk_3 \dots sk_n\}$  Where  $SK$  is the set of generate private keys related to public key.

3. Cloud Server  $CS = \{U, D\}$

Where,  $U$  = Upload file

$D = \{T, F\}$

- Where  $D$  = Self-Destructive Process
- $T$  = Time Interval

- F=Number of files

*Step1:* Setup ( ): PKG run the setup algorithm. It chooses a random generator as well as a random integer and sets  $g_1 = g^x$ . Then, A random Element PKG picked by and two hash functions. Finally, output the public key  $PK = (g, g_1, g_2, H_1, H_2)$  and the master key  $MK = x$ .

*Step 2:* KeyGen (MK, ID, RL, TL, and PK): PKG firstly checks whether there quest identity ID exists in RL, for each user's private key request on identity ID, if so the key generation algorithm is terminated. Next, PKG randomly selects  $X_1, Z_q$  and sets  $x_2 = x - x_1$ . It randomly chooses, and computes. Then, PKG reads the current time period  $T_i$  from TL. Accordingly, it randomly selects and computes, where and. Finally, output  $SK_{ID} = (IK [ID], TK [ID]_{Ti})$  and  $OK_{Id} = x_2$ .

*Step 3:* Encrypt (M, ID,  $T_i$ , and PK): Assume a user needs to encrypt a message M under identity ID and time  $T_i$  period. He/She chooses a random value  $s \in \mathbb{Z}_q$  and computes,  $C_0 = Me(g_1; g_2) s$ ;  $C_1 = gs$ ;  $EID = (H_1(ID))s$  and . Finally, publish the ciphertext as  $CT = (C_0; C_1; EID; E_{Ti})$ .

*Step 4:* Decrypt (CT; SKID; PK): Assume that the ciphertext CT is encrypted under ID and  $T_i$ , and the user has a private key  $SKID = (I_k [ID]; T_k [ID] T_i)$ , where  $IK [ID] = (d_0; d_1)$  and  $TK [ID] T_i = (dT_i_0; dT_i_1)$ .

**Mathematical formulation for Searching Mechanism**

Let S be a System.  $S = \{I, P, O\}$

Where,

Input I: The input for the system is multi word query from the user.

Output O: Ranking results.

Process P:

1. Single-term publication

$$\xi = \frac{1 - \sigma_j \cdot \beta_j(t)}{1 - \beta_j + \sigma_j} \Rightarrow \beta_j = \frac{1}{[(\sigma_j^{-1} - 1) (\epsilon_j^{-1} - 1)]^{-1}}$$

Where,  $\beta_j$  is number of probability values produces by source analytical computation for term.

2. False Positive Rate:

$$FP(0,1) = \frac{F(0,1)}{F(0,1) + \sigma_0 \sigma_1}$$

Where, FP (0, 1) is the false positive values,  $\beta_0, \beta_1$  are the probability at which a non-positive owner publishes data as a positive owner.

3. Index Generation  $I = \{I_1, I_2 \dots I_n\}$

Where I is the set of all index of all private servers.

4. Merge and upload index at private cloud.  $MI = \{MI_1, MI_2 \dots MI_n\}$

Where MI is the set of all merge indexes collected from monitoring system.

5. User Query to public cloud  $Q = \{Q_1, Q_2 \dots Q_n\}$

Where, Q is the set of all queries poses to public cloud.

6. User Authentication at private server  $U = \{U_1, U_2 \dots U_n\}$

Where U is the set of all authenticated users of private server.

7. Token Generation and distribution  $T = \{T_1, T_2 \dots T_n\}$

Where T is the set of all tokens generated by private server for its authenticated users.

8. Key Generation at KDC  $K = \{K_1, K_2 \dots K_n\}$

Where K is the set of all keys stored at KDC, used for decryption of data at user side.

9. Data decryption and TF\_IDF ranking  $R = \{R_1, R_2 \dots R_n\}$

Where R is the set of all ranked results for particular input query.

### Build Index Tree

Input: the document collection  $F = \{f_1, f_2, \dots, f_n\}$  with the identifiers  $FID = \{FID—FID = 1, 2, \dots, n\}$ .

Output: the index tree T

1. for each document {FID} in F do
2. Construct a leaf node u for {FID},
3. Insert u to CurrentNodeSet;
4. end for
5. while the number of nodes in CurrentNodeSet is larger than 1 do
6. if the number of nodes in CurrentNodeSet is even, i.e. 2h then
7. for each pair of nodes  $u_1$  and  $u_2$  in CurrentNodeSet do
8. Generate a parent node u for  $u_1$  and  $u_2$ ,
9. Insert u to TempNodeSet;
10. end for
11. else
12. for each pair of nodes  $u_1$  and  $u_2$  of the former (2h - 2) nodes in CurrentNodeSet
13. do
14. Generate a parent node u for  $u_1$  and  $u_2$ ;
15. Insert u to TempNodeSet;
16. end for
17. Create a parent node  $u_1$  for the (2h - 1)-th and 2h-th node, and then create a parent
18. node u for  $u_1$  and the (2h + 1)-th node;
19. Insert u to TempNodeSet;
20. end if
21. Replace CurrentNodeSet with TempNodeSet and then clear TempNodeSet;
22. end while
23. return the only node left in CurrentNodeSet, namely, the root of index tree T ;

## V. EXPERIMENTAL ANALYSIS

### A. Experimental Setup

The system is built using Java framework version JDK 1.8 on Windows platform. The Netbeans

version 8.2 is used as a development tool. The system doesn't require any specific hardware to run, any standard machine is capable of running the application. The system analysis is carried out on datasets consisting of files.

### B. Dataset

Dataset for peer to peer network with 2500 file names and their sizes are used for proposed system. Dataset consists of 1.6 million queries and data derived from NIST's available TREC WT10g. We can create index table by uploading files on private server. For Analysis of the System we have used multiple files. The Input Files are of various sizes varying from 1 KB to 100MB.

### C. Results

By using non-grouping based approach of PPI the proposed system will going to provide better preservation of user's privacy in terms of data confidentiality through encryption and better quality of results i.e. relevant results to the queries using ranking techniques.

### Similarity Measurement

In the table 1 evaluate the similarity for the both existing and proposed system. Run the project four times and obtained the result which contain different similarity values. From the values it shows that the similarity values for the existing system is less than the proposed system.

Table 1. Similarity Table

	Existing	Proposed
D1	0.43	0.93
D2	0.71	0.95
D3	0.37	0.98
D4	0.38	0.94

Graph in Fig. 3 shows that the proposed system performs better than existing system in terms of similarity measures for 4 text file.

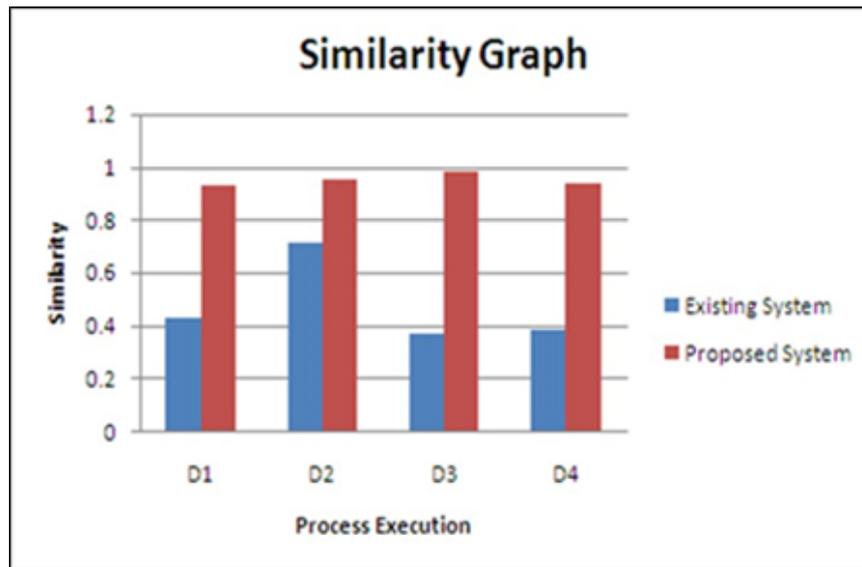


Figure 3. Time Graph for Process Execution

**Time Measurement**

In the table 2 measure the time for different process like uploading the file, query searching,

encryption time, token generation, and ranking time. Run the project twice and plot the graph

Table 2. Time Measurement Table

	File Upl	Que ry	Encrypti on Time	Tok en	Ranki ng
D	3.0	0.97	2.03	0.23	0.84
D	5.9	0.38	3.87	0.47	0.72

In the graph Fig. 4 shows the time graph for the proposed system. Fetch the value from the above table and plot the graph.

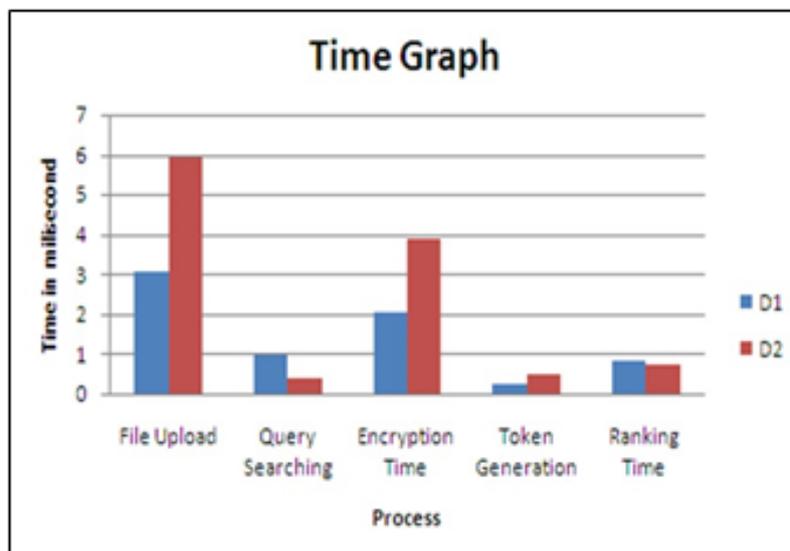


Figure 4. Time Graph

Fig. 5 represent that, the proposed system reduces the memory overhead in bytes than the proposed system, because proposed system uses distributed KDC. The X-axis shows the number of key requests and y-axis represents the memory in bytes.

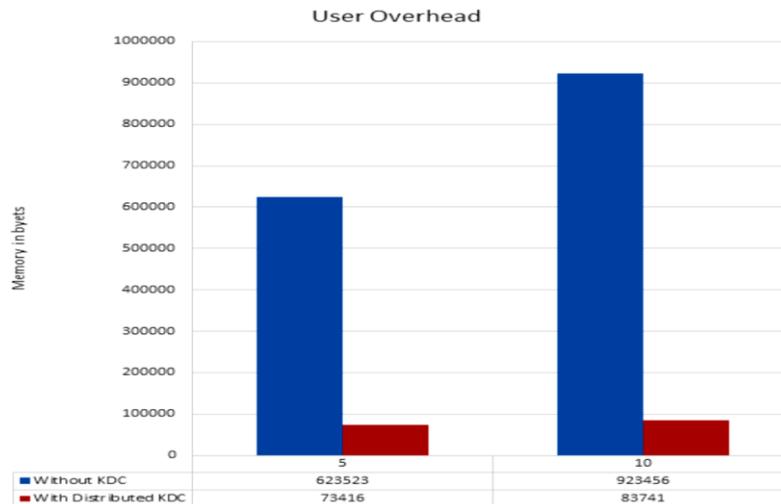


Figure 5. User overhead graph comparison

## CONCLUSIONS

In order to share data between users, the proposed system will link the local server with the cloud server. Specific data or information require some authentication. This authentication is processed by a system of encryption. By using a PPI and encryption technique, the system user can access the required data by order.

In this proposed study we propose a fully reliable system that not only tackles data security and user privacy issues with easy access to data. We propose a public audit system, based on a semi-confident third-party auditor, to maintain integrity. This TPA is accompanied by a proxy server based on regenerative code, which reconstructs the original data from the original data's hash values. Saving data fragments on different servers reduces the likelihood of information loss but this data fragment stowage on different servers expands storage space for information backup. This data block could damage the cloud server store. Our system implements a replacement coding technique at the proxy to retrieve the corrupted data blocks where any blocks are damaged or lost. The system also uses cloud servers for data storage in order to lower calculation costs, as the server has some advantages such as security, low cost, high availability etc. To

minimise load on the single KDC, the system uses distributed KDC. This allows users to request a key to another KDC if one KDC is active.

We encrypt the data and generate the encrypted index tree for the security of user data before we upload it to a cloud. When the user tries to access information, this encrypted tree traps the multi keyword search request to search the information. This mechanism helps to safeguard the confidentiality of the data while keeping the search efficient. The privacy of users is also addressed because we use a unique identity to generate encryption key. We use biometric as an identification attribute in this case. Following the extensive study we trust in the reliability and efficiency of our proposed system which covers the whole range of user cloud interaction.

Various data-set tests including the number of files to calculate the performance of our system. The size of the file is between 1 kb and 100 mb. Test results show that, in terms of storage space, costs, data availability, reduced overload at kdc and file recovery, our system is better than the previous one.

## REFERENCES

- [1] B. Chen, R. Curtmola, G. Ateniese, and R. Burns, "Remote data checking for network coding-based distributed storage systems," in

- Proceedings of the 2010 ACM workshop on Cloud computing security workshop, 2010, pp. 31–42.
- [2] H. Chen and P. Lee, “Enabling data integrity protection in regenerating coding- based cloud storage: Theory and implementation,” IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 2, pp. 407–416, 2014.
- [3] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, “A survey on network codes for distributed storage,” Proceedings of the IEEE, vol. 99, no. 3, pp. 476–489, 2011.
- [4] H. Shacham and B. Waters, “Compact proofs of Retrievability,” Advances in Cryptology-ASIACRYPT 2008. Springer, pp. 90– 107, 2008.
- [5] Y. Hu, H. C. Chen, P. P. Lee, and Y. Tang, “Nccloud: Applying network coding for the storage repair in a cloud-of-clouds,” in USENIX FAST, 2012.
- [6] Yuzhe Tang and Ling Liu , Fellow , —Privacy-Preserving Multi-Keyword Search in Information Networks I, IEEE transactions on knowledge and data engineering ,vol. 27, no. 9, Sept 2015
- [7] Tseng, Ching-Yang, ChangChun Lu, and Cheng-Fu Chou. "Efficient privacy-preserving multi-keyword ranked search utilizing document replication and partition." Consumer Communications and Networking Conference (CCNC), 2015 12th Annual IEEE. IEEE, 2015.
- [8] B. Chen, R. Curtmola, G. Ateniese, and R. Burns, “Remote data checking for network coding-based distributed storage systems,” in Proceedings of the 2010 ACM workshop on Cloud computing security workshop. ACM, 2010, pp. 31–42.
- [9] H. Chen and P. Lee, “Enabling data integrity protection in regenerating coding- based cloud storage: Theory and implementation,” Parallel and Distributed Systems, IEEE Transactions on, vol. 25, no. 2, pp. 407–416, Feb 2014.
- [10] K. Yang and X. Jia, “An efficient and secure dynamic auditing protocol for data storage in cloud computing,” Parallel and Distributed Systems, IEEE Transactions on, vol. 24, no. 9, pp. 1717–1726, 2013.
- [11] Y. Zhu, H. Hu, G.-J. Ahn, and M. Yu, “Cooperative provable data possession for integrity verification in multicloud storage,” Parallel and Distributed Systems, IEEE Transactions on, vol. 23, no. 12, pp. 2231– 2244, 2012.
- [12] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, “A survey on network codes for distributed storage,” Proceedings of the IEEE, vol. 99, no. 3, pp. 476–489, 2011.
- [13] H. Shacham and B. Waters, “Compact proofs of retrievability,” in Advances in Cryptology-ASIACRYPT 2008. Springer, 2008, pp. 90–107.
- [14] Y. Hu, H. C. Chen, P. P. Lee, and Y. Tang, “Nccloud: Applying network coding for the storage repair in a cloud-of-clouds,” in USENIX FAST, 2012.
- [15] C. Wang, Q. Wang, K. Ren, and W. Lou, “Privacy-preserving public auditing for data storage security in cloud computing,” in INFOCOM, 2010 Proceedings IEEE. IEEE, 2010, pp. 1–9.
- [16] C.Wang, S. S. Chow, Q.Wang, K. Ren, andW. Lou, “Privacy-preserving public auditing for secure cloud storage,” Computers, IEEE Transactions on, vol. 62, no. 2, pp. 362–375, 2013. IEEE Transactions On Information And Security Vol 1 No 2015
- [17] C. Wang, Q. Wang, K. Ren, and W. Lou, “Towards secure and dependable storage services in cloud computing,” Service Computing, IEEE Transactions on, vol. 5, no. 2, pp. 220–232, May 2012.
- [18] D. Boneh, B. Lynn, and H. Shacham, “Short signatures from the weil pairing,” Journal of Cryptology, vol. 17, no. 4, pp. 297–319, 2004.
- [19] Liehuang Zhu, Chuan Zhang, Chang Xu, Ximeng Liu, And Cheng Huang, “An Efficient and Privacy-Preserving Biometric Identification Scheme in Cloud Computing”, Volume 6, IEEE Access March 2018.
- [20] XialiHei, Xiaojiang Du, “Biometric-based two-level secure access control for Implantable Medical Devices during emergencies”, in 2011 Proceedings IEEE INFOCOM.
- [21] Mauro Barni, Tiziano Bianchi, Dario Catalano, Mario Di Raimondo, Ruggero Donida Labati, Pierluigi Failla, “Privacy-Preserving

- FingerCodes Authentication”, in Proceedings of the 12th ACM workshop on Multimedia and security, Pages 231-240 , September 2010.
- [22] Yan Huang, Lior Malka, David Evans, Jonathan Katz, “Efficient Privacy-Preserving Biometric Identification”, 18th Network and Distributed System Security Conference (NDSS 2011), 6-9 February 2011.
- [23] Jin Li, Yinghui Zhang, Xiaofeng Chen, Yang Xiang, “Secure attribute-based data sharing for resource-limited users in cloud computing”, in computers & security, Volume 72,p 1–12, Elsevier 2017
- [24] Jin Li, Jingwei Li, Xiaofeng Chen, Chunfu Jia, and Wenjing Lou, “Identity-Based Encryption with Outsourced Revocation in Cloud Computing”, IEEE Transactions On Computers, Vol. 64, NO. 2, FEBRUARY 2015.
- [25] Jin Li, Xiaofeng Chen, Mingqiang Li, Jingwei Li, Patrick P.C. Lee, and Wenjing Lou, “Secure Deduplication with Efficient and Reliable Convergent Key Management” ,IEEE Transactions On Parallel And Distributed Systems, Vol. 25, NO. 6, JUNE 2014.
- [26] Jin Li, Xinyi Huang, Jingwei Li, Xiaofeng Chen, and Yang Xiang, “Securely Outsourcing Attribute-Based Encryption with Checkability”, IEEE Transactions On Parallel And Distributed Systems, VOL. 25, NO. 8, AUGUST 2014.
- [27] Yinghui Zhang , Xiaofeng Chen, Jin Li, Duncan S. Wong, Hui Li, Ilsun You, “Ensuring attribute privacy protection and fast decryption for outsourced data security in mobile cloud computing”, in computers & security, Elsevier 2016.
- [28] John Bethencourt, Amit Sahai, Brent Waters, “Ciphertext-Policy Attribute-Based Encryption”, IEEE Symposium on Security and Privacy (SP '07), IEEE 2007.
- [29] Taek-Young Youn, Ku-Young Chang, Kyung Hyune Rhee, And Sang Uk Shin, “Efficient Client-Side Deduplication of Encrypted Data with Public Auditing in Cloud Storage”, IEEE Access 2018.
- [30] Jian Liu, Kun Huang, Hong Rong, Huimei Wang and Ming Xian, "Privacy-Preserving Public Auditing for Regenerating-Code-Based Cloud Storage", IEEE Transactions On Information And Security, Vol. 1 No 2015.
- [31] Boyang Wang, Baochun Li, Hui Li, “Panda: Public Auditing for Shared Data with Efficient User Revocation in the Cloud”, IEEE Transactions On Services Computing, Vol. 8, No. 1, January/February 2015.
- [32] Cong Wang, Qian Wang, KuiRen, Ning Cao, and Wenjing Lou, “Toward Secure and Dependable Storage Services in Cloud Computing”, IEEE Transactions On Services Computing, Vol. 5, No. 2, April-June 2012.
- [33] Na Wang, Junsong Fu, Bharat K. Bhargava, Jiwen Zeng, “Efficient Retrieval over Documents Encrypted by Attributes in Cloud Computing”, IEEE Transactions on Information Forensics and Security, Vol13, Issue 10, Oct 2018.
- [34] Yuzhe Tang, Ling Liu, “Privacy-Preserving Multi-Keyword Searching Information Networks”, IEEE Transactions on Knowledge and Data Engineering, VOL. 27, NO. 9, SEPTEMBER 2015.
- [35] S. Kamara and K. Lauter, “Cryptographic cloud storage,” in Financial Cryptography and Data Security. Springer, 2010, pp. 136- 149.
- [36] 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. 506-522.
- [37] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, “Public key encryption that allows piracy queries,” in Advances in Cryptology-CRYPTO 2007. Springer, 2007, pp. 50-67.
- [38] D. X. Song, D. Wagner, and A. Perrig, “Practical techniques for searches on encrypted data,” in Security and Privacy, 2000. S and P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000, pp. 44-55.
- [39] Y.-C. Chang and M. Mitzenmacher, “Privacy preserving keyword searches on remote encrypted data,” in Proceedings of the Third international conference on Applied Cryptography and Network Security. Springer-Verlag, 2005, pp. 442-455.

- [40] 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.