Turkish Journal of Computer and Mathematics Education Vol.12 No.2 (2021), 2845-2853

Research Article

Enhanced Effective and Privacy Preserving Multi Keyword Search over Encrypted Data in Cloud Storage Using Blowfish Algorithm

Ch. Chakradhara Rao[1]Dr. Tryambak Hiwarkar[2] ,Dr. B. Santhosh Kumar[3]

[1]Research Scholar, CSE Department, Sri Satya Sai University of Technology and Medical Sciences, Sehore, MP
[2]Associate Professor, CSE Department, Sri Satya Sai University of Technology and Medical Sciences, Sehore, MP
[3]Sr.Assistant Professor, CSE Department, GMR Institute of Technology, Rajam, AP
b.santhoshkumar@gmail.com

Article History: Received: 11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

Abstract: The data owner administers their data to the public cloud due to regulatory. Even this data encoded once it easily transmitted to the cloud. In order to ensure the privacy and security, cloud subscribers need a very different type of online data. This specific information must be germane to the recipient's query. The user to inspect the query too in the cloud with umpteen keywords can accomplish this. With the intention to defend the privacy of data, confidential data encoded before subcontracting by the metadata owner, attempting to make the traditional and efficient plaintext keyword search tactic pointless. Therefore, it is crucial to explore a secure data search service for encrypted user data. Due to the growing of massive number of digital users and legal documents in the cloud, umpteen keywords forced in the search request and documents returned in the order of their relevance to these keywords.Cloud assistance end-user demands cloud information by various paternoster inquiry, whichlabelled as umpteenkeywordstratified exploration over scrambled information. In the cloud server, all the client inquiries exchanged. Server looks through the significant matter by utilizing the harmonic equivalency and sends the applicable outcomes to the client. Information got from cloud server is in the scrambled configuration. To get to control to the client,Information proprietor furnishes with key for unscrambling of information. Message Authentication Code(MAC) calculation is utilized to check and confirm trustworthiness of information. Thusly this paper depicts the method of giving security to redistribute information on cloud and checking the trustworthiness of information using blowfish algorithm. Comparison results shows the improvement in efficiency of Data Retrieval and Time efficiency.

Keywords: Cloud Computing; Metadata; Encryption; Message Authentication Code; Privacy; Query.

1. Introduction

In view of sharing processing assets as opposed to having neighborhood servers or individual gadgets to deal with applications, cloud computing is characterized as a sort of figure out.Public Cloud, Private Cloud, hybrid Cloud. Open Cloud gives benefits that are open for open use[1].Accessible-free functionality is the primary goal of the Open cloud administrations in which the single association and oversaw outlook are utilized for private cloud framework.Distinctive elements nevertheless, bound together forms the hybrid cloud.In light of the fact that the specialist co-op can get to the information that is on the cloud whenever the cloud specialists worried about security. Hardly sometimes, it could coincidentally or purposely modify or even erase data.Many cloud suppliers can impart data to outsiders. Before they begin utilizing, the cloud administration clients need to concur in their security approaches that allowed [2].To handle with to relish brilliant applications and administrations on interest from a common arrangement of configurable figuring assets.

It's extraordinary adaptability and economicvitality propelling the individuals and organizations to redistribute their intricate provincial information framework in the cloud. To secure the protection of information and to contradict spontaneous access in the cloud and past, information proprietors must scramble secret information. For instance, messages, individual wellbeing records, photograph collections, charge reports, and so forth before subcontracting them to the business open cloud [3].“Online data as a service” is one of the use of distributed computing which provides the authority of the data on the cloud [4]. We strongly recommended to ensure the information protection against assaults from the cloud server classified information must be scrambled before being transferred to the cloud server. Mainly the data provider encrypt and upload the data to the central storage server, whichshared to the distributed online users, the key management authority will provide the key for the providers and subscribers. The overall layout of the sharing scheme depicted in figure 1.

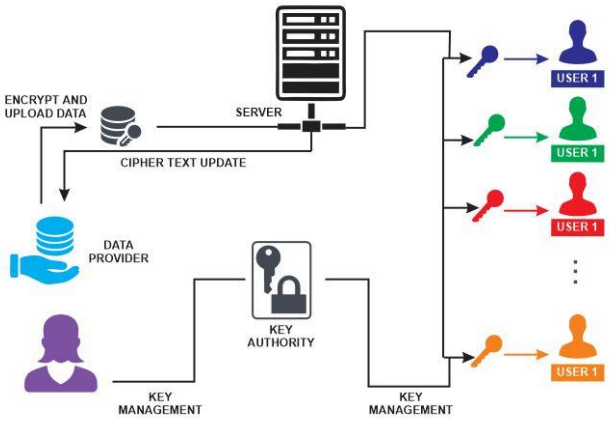

Fig 1: Layout of sharing scheme

A productive ordering strategy intended in this venture to assist quicker inquiry assessment than the trifling straight inspection methodology. Rather than returning uniform and constant outcomes, to accomplish the powerful information recovery process, the huge measure of archives request the cloud server to perform result pertinence positioning. Previously mentioned positioned inquiry framework, empowers information clients to locate the most important data rapidly [5]. As just the most significant information, which is very alluring in the compensation as-you-use cloud worldview, Positioned pursuit can likewise decrease superfluous system traffic by sending aftereffect of inquiry. When security insurance taken into consideration this positioning often called as ranking task, notwithstanding, ought not to release any password related data. Then again, to enhance the query output precision just as to improve the client looking background, it is additionally important for such positioning framework to help different keyword look, as single catchphrase seek frequently yields extremely coarse outcomes [4].

To refine the result significance the systemize matching,[6] i.e., whatever number matches as could reasonably be expected, which is considered as a proficient closeness measure among such multi keyword semanticsand availed in the plaintext data recovery society.The rest of this paper sorted out as pursues: segment II gives the related work regard to the data encryption and hunt on information put away out in the open cloud. Area III depicts design for the framework and modules to execute the framework. Additionally depicts proposed framework calculation, Blowfish calculation for data encryption, pursued by Section IV, which portrays test results of execution. Area V presents finish and future enhancements of the proposed work.

2. Related Works

Uponsteering of cloud security, heterogeneous techniques put forward in accomplishing studies on providingmulti-keyword search. Some of the approaches summarized in the following phase.

To execute information protection and security in distributed computing, information protection what's more and get to control, a few existing plans that had been endeavored. Information is put away as open or private consequently distinctive looking tactics are accessible for the twain sorts of information. Hence, to access the secret information those are put away in the cloud in scrambled arrangement so just the verified individuals who realize the key can get to the information. In providing, the encryption techniques for the remote documents the authors pointed the issue of productively recover a portion of the scrambled documents containing explicit catchphrases, keeping the catchphrases themselves mystery and not to imperil the security of the remotely put away documents [2]. In answers for this issue under chiseled characterized security, necessities advertised. With the implementation of this approach,the encoding techniques are effective in view of the fact that no open key cryptography utilized. While taking Deterministic and Efficiently Searchable Encryption (DESE) [7]into consideration where the encryption calculation deterministic. On the other hand, when the re-appropriated databasestaken the information sent to a remote server. In that case, the database server is untrusted. The information in each field in the database scrambled independently under the open key of a beneficiary, who should have the capacity to question the server to recover the scrambled records containing specific information. The researchers provide a conceivable answer for the issue. To resolve the issues faced in the DESE, [8] presents the blurkeyword seek over encoded cloud information and furthermore keeping up the keyword security in which the umpteen keywords are encoded and decoded accordingly. The study [9] highlights the test of secure positioned keyword look over encoded cloud information, additionally Ranked pursuit using KNN calculation is

forecasted whichincredibly upgrades the ergonomics by empowering query output significance positioning as opposed to sending constant outcomes, and further guarantees the record recovery exactness.

3. Proposed System

In this segment, we present framework design for proposed system illustrating the process flow. Fig 2 portrays design of proposed framework for detachment protectedumpteen-catchphrase look over encoded cloud.

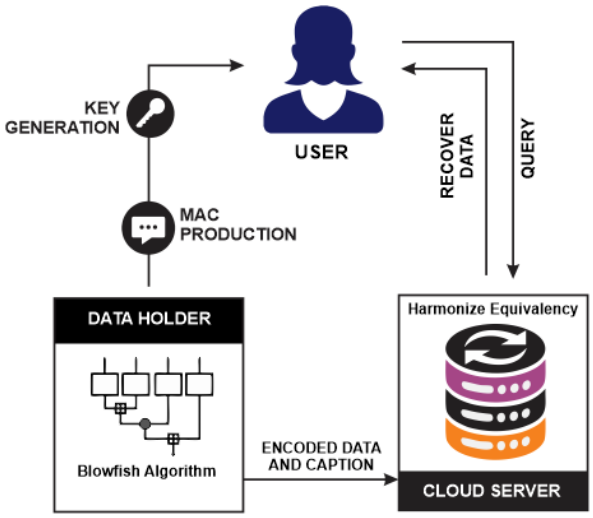

Fig 2: Proposed Layout

The framework ordered into three stuffs

(i) Data Owner (ii) Cloud Server and (iii) User or Customer. The information proprietor contains gathering of information records redistribute on cloud server in encoded frame. Scrambled caption for report constructed before forecasted to the user. Then the gathered information along with twain file and scrambled archive gathering redistributed to the Cloud Server. Cloud Server contains scrambled reports and corresponding captions redistributed by the Data holder. Followed by the data gathering the customer triggers inquiry for looking record. Message confirmation Code Algorithmapplied for the periodic checking of Information trustworthiness.

The framework partitioned into following components:

1. Binary data origination
2. Ciphering the Data
3. Enumerating the access control
4. Data holder inquiry

(i) Binary data origination

In Binary data origination sector data holderpicks the information and the corresponding bit vector is conceived for that information. The binary information generated by capturing the bit vector of the information. A BLOB is a gathering of parallel information put away as a solitary element in a database the board framework[10]. The pictures sound or other sight and sound items are commonly comes under the multitudes of BLOB.

The twofold information arrangement of the encoded information restricts the end clients to get to this paired information.The twofold information is the list for the information in the information proprietor, which are largely considered. The bit vector is the bytes type of the information in the data holder. The bit vector converted into the binary information. Generally, the data folded in form of chunks in its simplicity the linear text subdivided into cipher blocks.

For the information ciphering, the bit vector and the binary snippet generated followed by the production of the message validation code. The parallel vectors are considered as (j+2) dimensioned in the first step which the data holder haphazardly yields and two (j+2) × (j+2) avertible grid $\{N_i, N_{ii}\}$ in which mcommensurateto the component of word impediments that are we acquired for each data. Later the individual bit components are expanded to m+2 dimensioned to provide the security of the binary data.

(ii) Ciphering the Data

In Data encryption section, the foremost thing of the Data holder is to scramble the sole information by blowfish computation and formerly distributed it to server. Abundantly the data scramble the paired information or the captions and further transmit it to central authority. Mostly the service providers did not think about the trusted content in the data holder.Additionally, the cloud central adminneed to categorize the exploration results in order to promote the accuracy of document repossession with the effect of some cataloguing criteria (harmonizecounterpart) and have to assign some unique ID to the cloud participants to ensure the cloud data is secure enough.Generally, thetiltsemployed to allude the information in the central admin. By keeping in mind with the goal, that the aggressors cannot utilize the information it gives greater security in the server side. The main goal of the framework is to keep the central admin from adapting any furthercommunication between plaintext qualities and cipher content values excluding those gotten by earlier information. To avoid the inquiries from unveiled to server in the receiver side the encoded information kept secure to keep track of the plain-content qualities for any scrambled records.

(iii) Enumerating the access control

The client needs information from the cloud server for which the data client get the access to the control Module. The cloud subscribers have diverse options and so that the client pops the interrogation to the server or service co-op. In previous works, the service seekerenumerates the origin and startup of initiation from the data holder. To incorporate the access control rights the service seekers halts the insights regarding the person to the information proprietor when needed.Now the data holdergathers the online information in the cloud that are received from customer and accessible to address the decoding key. The access control tactics utilized to oversee decoding capacities forecasted to the data clients. Thus, a conveyed settingmaintained where the server is in risky environment that too in the remote side.The detailed description of the data access control in cloud the clear model illustrated in figure 3.

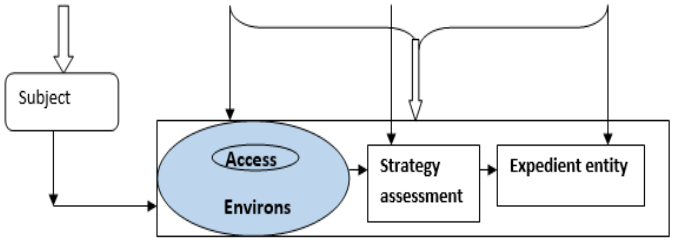

Fig 3: Access control in blowfish

Let us consider S, E, EN as subjects entity,expedient entity and environs set correspondingly, and ENT_S, ENT_E, ENT_EN are subjects entity , expedient entity and environs set, and SE_l, RE_m,EE_n are defined attributes of subject, expedient and environs. The detail expression as denoted in (1).

ENT _S={SE_1, SE_2,…….. SE_n} l< l < M

ENT _E = { EE_1, EE_2,…….. EE_m} l< m < M

ENT _E={EE_1, EE_2,…….. EE_n} l < n < M…..(1)

The three attribute sets' value realms are indicated as R(SE), R(RE), R(EE); and use a tripartite entities with a set (s, e, en) for the implementation of single user access control , including subject s, expedient e and environment en. They are described in (2) respectively:

S= se_1……se_n Є D (SE_1)X…..X D (SE_n), sεS

E= ee_1……ee_n Є D (EE_1)X…..X D (EE_n), eεE

EN= ene_1……ene_n Є D (ENE_1)X…..X D (ENE_n), enεEN ……(2)

Access policy enumerated as in (3):

Policy = {(s_1 , e_1 , en_1) , (s_2 , e_2 , en_2) , (s_n , e_n , en_n) } --------→ (3).

Only when the (s, e, en)∈ Policy, the subject s is accepted as trusted source to contact the expedient e under the environs en otherwise it halts the data transfer. In many literature retrospect lookup based method used to adopt the access control and other studies widely use algebra to fix the values.

(iv) Data holder inquiry

In Data, client inquiry sectionthe data query is processed which the cloud owners throw. Once received the query from the data seeker the data holder co-op generates the bit vector at that instant. At that point, the service

provider holds the duty of modifying the bit vector into binary data. Followed by the dataproviderdecodes the more or less it featured information from the file. Then it sends the scrambled information to the customer. By utilizing blowfish calculation at that point, the decoding of the gathered information done by the client by using the key from the data owner. Furthermore by employing the Data validation Code the reliability and integrity of the information is accomplished.

3.1 Proposed system Algorithm

Stage 1: Certification verification for data holder.
Step2: For Binary information genesis and Message Authentication code production the data proprietor selects the appropriate information.
Step 3: By blowfish calculation the data holder need to encode the startup information and send it to server. At that point it encode the binary coded data or the timestamp and forward it to the central administrator.
Stage 4: To accomplish the information decoding the Bestow emanate is to be get by the data holder to data seeker.
Stage 5: Data constituency thoroughly monitors the reliability of the information by adopting the Message confirmation code System.

3.2 Blowfish Encryption Algorithm

Blowfish is a symmetric square figure that can be viably utilized for encryption and defending of information. It takes a variable-length key, from 32 bits to 448 bits, making it perfect for anchoring information [19] as shown in figure 4. Blowfish Encryption Algorithm is encryption procedure for changing plaintext information into figure message, this calculation produce key for encryption .Input to this calculation is report chosen by the information proprietor for re-appropriating .yield is figure content of information in report.

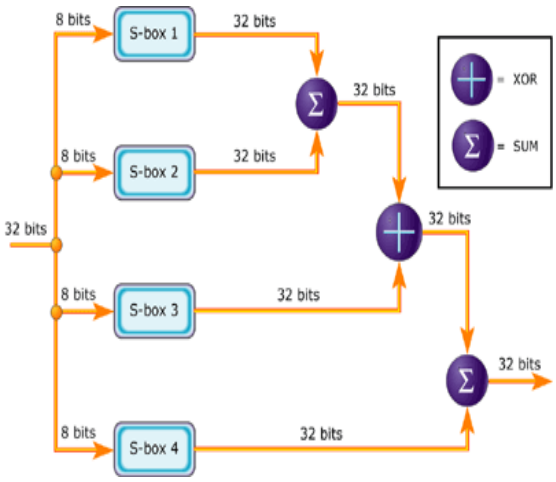

Fig 4: Schematic view of Blowfish algorithm

Key Generation:
Blowfish uses a large number of sub keys. The P array consists of 18, 32-bit sub keys: P1, P2,...,P18

Algorithm: Encoding the text

Rift the linear 64 bit value P in binal 32 bit:P1 and Pe
Fori = 1 to 16 do
Pl = Pl XOR Q[i]
Pe = Pe OR F(Pl)
Switch Pl and Pe
Untie the recent switch
Pe = Pe XOR Q[17]
Pl = Pl XOR Q[18]
Re-blend Pl and Pe into a aught 64-bit value- Blowfish subsists 16 progression.

Encryption system flow:

- Input i= 64-bit data entity.
- Split x into two 32-bit splits: iLS, iRS.
- After spliting for i = 1 to 16:
  Then set the iLS = iLS XOR Bl
  Calculate iRS = G(iLS) XOR iRS
  Swap iLS and iRS

-Do the same for the 15 rounds then after the 16^{th} round, switch iLS and iRS one more time to refresh the most recent switchover.

- Then, iRS = iRS XOR R17 and
iLS= iLS XOR R18.
- Finally, cause to combine the iLS and iRS to get the cipher text.

Decryption:

To interpret the last an incentive back to the first one, a similar technique is utilized, then again, actually the exhibit P is strolled in reverse:

Algorithm

Rift the linear 64 bit value P in binal 32 bit:P1 and Pe
Fori = 1 to 16 do
Pl = Pl XOR Q[i]
Pe = Pe OR F(Pl)
Switch Pl and Pe
Untie the recent switch
Pe = Pe XOR Q[2]
Pl = Pl XOR Q[1]
Re-blend Pl and Pe into a original 64-bit value- Blowfish subsists 16 progression

Decoding the data is as same as the procedure followed in encoding excluding that R1, R2... R18 are used in the reverse order.

With focus on the index and querysystem consists of four algorithms as follows.

- **Setup(A^{∞})**
  - ✓ Consider the refugalparameter ∞ as input
  - ✓ Yield ofdata holder is a symmetric key US.
- **Construct Caption(G, US)**
  - ✓ Upon depends on the type of the dataset D , the dataholder constructs a easily recoverablecaptionC that can be encoded by the symmetric key USwhich in turn envisioned to the cloud server.
  - ✓ After the caption building the data gathering can be separatelyscrambled and broadcasted. Mainly for theinformation security, if the cloud server halts any correspondencein thekeywordsand encoded archives from list which in turn it gain proficiency with the real subject of a report, even the substance of a short archive[20].
- **Trapdoor(ɲ)**A relevant trapdoor Tɲ is generated with qumpteen keyword of interest in ɲ as input in this algorithm.
- **Inquiry (T ɲ, m, IC)**

Upon receiving a queryrequest as (Tɲ, m), the cloud holder performs the following events: (i) ranked search on thecaption C with the help of trapdoor Tɲ,

(ii)returnsFɲ, the ranked id list of top-k documents sorted by theirsimilarity with Wɲ.

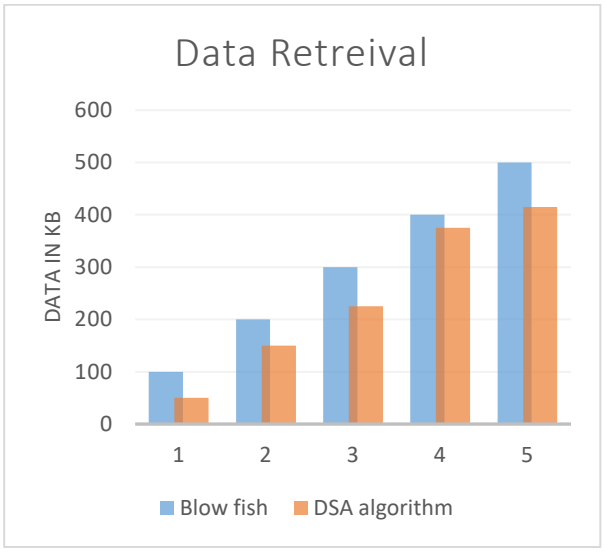

Fig 5.Efficiency of Data Retrieval

The proposed blowfish algorithm is analyzed for its time efficiency(Fig 6) and data retrieval capability(Fig 5) with the existing DSA encryption techniques [21].

The time efficiency graph is gotten by the multi-keyword search performed by both the systems for various time intervals. The graph clearly shows the blowfish technique yield the better performance. Simultaneously the data retrieval is plotted in terms of data in KB which shows the suggested framework produces better results.

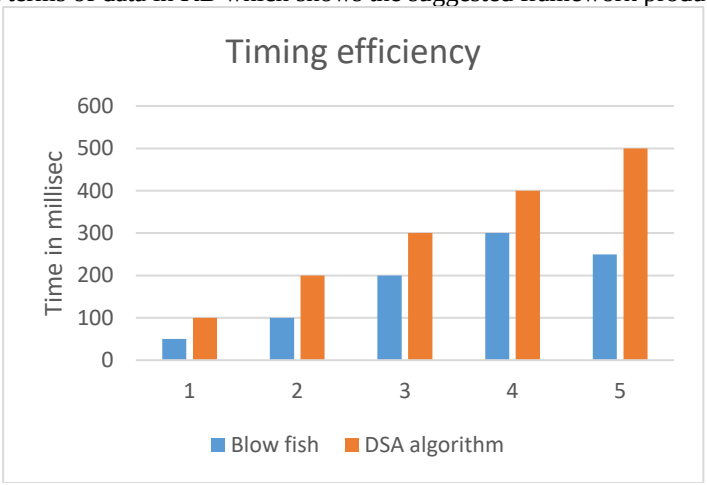

Fig 6. Timing efficiency

The DES and Blowfish algorithm are compared with various parameters such as number of revolutions, the time taken to run the search, the agility in Bytes, power backup in %

Encryption algorithm	Number of revolutions(M)	The time taken to run the search(s)	Agility in Bytes(Bytes/s)	Power backup in %
DES	20	1233.89	6815769	65%
	40	2258.23	6962539	75%
Blowfish	20	975.78	8512768	95%
	40	2787.79	8655255	78%

The multi-keyword search performed for increase in number of revolutions the different values plotted in graph and it is obvious that the DES is comparatively less with Blowfish algorithm(fig 7)

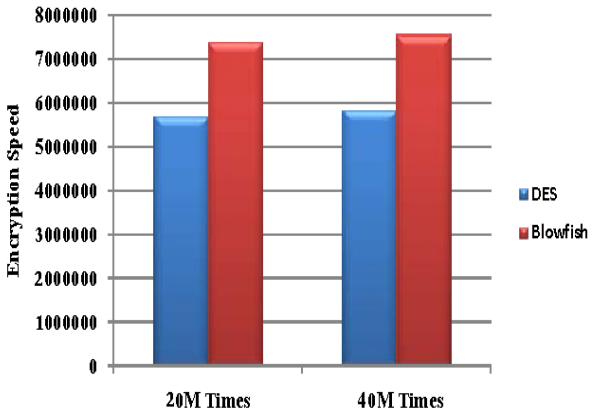

Fig 7: Agility in Bytes

The DES and blowfish compared with various values of number of rounds, power consumption the graph plotted, and it is obvious that the Blowfish shows the higher performance rate.

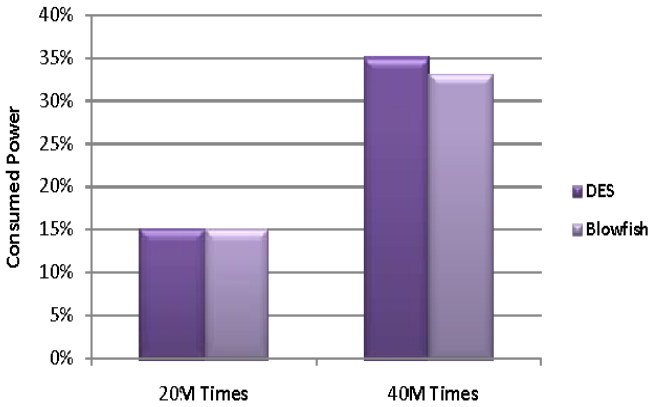

Fig 8: Battery consumption

5. Conclusion

In this paper, a productive umpteen-keyword search on encoded cloud informationsuggested. Here the Bit vector and Binary coded information created for ensuring the information security and efficient data retrieval. This venture encompasses efficacy as accepting better implication results to service requesters rather than misleading and different results .This procedure gives security as far as catchphrase protection, data reliability and caption protection,. Additionally data integrity checking performed utilizing Message Verification Code Algorithm is the vital parameter of this proposed system.Indeed numerous secure challenges are in the web when taking the multi-user scheme. Primarilyall the users have to maintain the same secure key while performing the trapdoor generation in the cloud.

References

1. Snehal M. Shewale1, Prof. Y. B. Gurav, Privacy Preserving Multi-Keyword Graded Search on Encrypted Cloud Data with Integrity Checking, International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064, Volume 4 Issue 7, July 2015.
2. a Jamdare, Prof. K. Vishal Reddy, Privacy Preserving on multi-keyword search with Lucene Indexer over Encrypted Data in Cloud, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 11, November 2017.
3. Y.-C. Chang and M. Mitzenmacher, “Privacy Preserving Keyword Searches on Remote Encrypted Data,” Proc. Third Int’l Conf. Applied Cryptography and Network Security, 2005.
4. N.Cao, C.Wang, M.Li, K Ren and Lou, “Privacy Preserving Multi-keyword Ranked Search over Encrypted Cloud Data”, IEEE Transactions on Parallel & Distributed System Vol.25, January2014.

5. N. Cao, C. Wang, M. Li, Karen, and W. Lou, “Enabling Secure and Efficient Ranked keyword Search over Outsourced Cloud Data”,IEEE Transactions on Parallel & Distributed System Vol.23 No.8, August,2012.
6. I.H. Witten, A. Moffat, and T.C. Bell, “Managing Gigabytes: Compressing and Indexing Documents and Images”. , Morgan Kaufmann Publishing, May 1999.
7. N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou,” LT Codes-Based Secure and Reliable Cloud Storage Service”, IEEE Transactions on Parallel & Distributed System 2012.
8. J. Li, Q. Wang, C. Wang and et. al,” Fuzzy Keyword Search Over Encrypted Data in Cloud Computing”, Proc. 27th Ann. Intl Cryptology Conf. Advances in Cryptology, 2007.
9. M. Premkumar and et. al, "Enhanced chaotic JAYA algorithm for parameter estimation of photovoltaic cell/modules", ISA Transactions,2021,ISSN 0019-0578,https://doi.org/10.1016/j.isatra.2021.01.045
10. Tianyue Peng and et. al, An Efficient Ranked Multi-Keyword Search for Multiple Data Owners Over Encrypted Cloud Data, Digital Object Identifier 10.1109/IEEE ACCESS.2018.2828404, Volume 6,2018.