Accessing Dynamic Health Records using K-Vertex Search Scheme Model towards Hierarchical Users mod Obscure Servers

K. Ketzial Jebaseeli, V.G Rani

Abstract: ***Improvement appertaining to automations paved a major way to the security systems which in turn revamps the quality and instigates services. In order to alleviate those challenges, attribute based encryption schemes have also ideated solutions to control contrivances. This proposed system accords the description of k vertex leading in the access control using encoded data over scrutinizing encryption model at the equivalent time stage. The K vertex is incapable of precise data retrieval. Orderly to rig out these issues, k-vertex search scheme has been proffered by building up a cipher text-policy hierarchical encryption (CP-HE) swank the R score counting from descending order to enumerate the file views. The data user retrieves hold, allow along with convinced strategical provisions. The previous works define the ensemble Signature paradigm that has been taken advantage of mulct signer formats. It is ascertained to take the edge off the quantum attacks. Though the data owner involves the reclamation of data, it necessitates the demand for a group of signatures. Generation of aggregate key and the arrangement in hierarchy comprehends to tortuous ingressions which are swamped down by the R score to the corresponding k values with illustrious security in the current paper with entailment of K-vertex trapdoor (TD=(K vertex(TF) & K vertex(IDF))). The results are gauged out by TF (U) and IDF (U) besides k-vertex (IDF).***

Index Terms: ***Trap door, K-vertex, R scores, Hierarchical users.***

I. INTRODUCTION

Cloud computing has become an exceeding contrapt in a multitude realms. The upper hand in the cloud computing concerns on sharing and communicating even on a heap of organizational lines of demarcation. The high-speed initiation along with progression of cloud computing on both the personages and the enterprises concede inveigle in virtue of subcontract they dossier into the cloud. Regardless of an assortment of pros in obscure services, outsourcing susceptible information conveys privacy alarms. This can cause a broad-spectrum approach to defend the data discretion to encrypt the data prior to data outsourcing. Nevertheless, on the other side, this grounds to an enormous cost in terms of data applicability. As in passable concentration on the exceeding crisis, searchable encryption (SE) proposal have equipped specific assistance .This facilitate the consumer to stockpile the encrypted data to the cloud and accomplish keyword investigation over cipher text province.

Revised Manuscript Received on September 15, 2019

K.Ketzial Jebaseeli, Asst. Professor, IT Department, holder of the BCA degree from Bharathiar University

Dr.V.G.Rani, Associate Professor, CS Department. She received her Ph.D in Computer Science from Bharathiar University

In this research, the k vertex associated with R score seek out and contact control over the hierarchical encrypted data using tree-based search has been proposed which wires multifarious-paternoster echeloned search and hyped-up maneuver on the document assortment. It builds up the cipher text-policy hierarchical encryption by authorizing the category through multifarious-paternoster search in vector space model and TF*IDF representation. Outstanding to the unique formation, the index edifice, query vector construction and automatic trap door generation can be accomplished to perk up the repossession effectiveness on self-instigated and renewed operations. A symmetric AES algorithm has been applied for data encryption [4]. The data repossess stratagem has been laid down on the basis of hierarchy of the data user and a stretchy revocation strategy has been provided to the data owners.

The memento of this paper is structured as follows. Related work is discussed in Section 2, and Section 3 gives a brief introduction to the system model and describes it in trivia. Section 4 presents the experiments and performance analysis on different metrics. As a final point, Section 5 provides the conclusion and future work.

II. RELATED WORK

A. Swaminathan, Y. Mao (2007) [1] discussed a jointed information in their paper stating the rank based search in an ordered way for the preservance of attaining extensive pragmatic techniques in cryptography. Investigational emanations on the W3C assortment evince that these knacks have analogous recital (Information Systems) Erudition frisking and Retrieval to conformist prequisition systems deliberated for non-encrypted data in requisites of rummage accurateness. The proponed methods also form the initial ambulates to fetch mutually sophisticated information repossession and protected inquisitional competencies for an extensive assortment of enactments administrant.

S. Yu, K. Ren, and W. Lou (2011) [2] confabulated the use of WSN's that put forward numerous services to users in the vicinity or athwart the Internet, and travel connecting multiple WSNs. Conversely, users should only have access to a limited detachment of overhaul. Existing research has interrogated access control in WSNs, other than habitually only reflects on authentication, with frontier consideration to authorization and policy execution. The involvement of this paper repose of three parts: (1) a substantiation protocol to make certain the legitimacy and discretion of end user service request on the basis of K lead, (2) an endorsement scaffold using role based admittance control to insure only accredited users can admittance to services, and (3) a user administration indulgence.

An accomplishment and assessment of this communications on the Contiki operating system and LooCI middleware exhibits the legality of advent.

Z. Wan, J. Liu, (2012) [3] comprised of SDN (Software Defined Network) and an emanating network technology. A hierarchical attribute-based access control scheme by integrating the hierarchical uniqueness based on encryption and cipher policy attribute based encryption (CP-ABE) structure. Putting together the hierarchical structure and distinctive inherited from CP-ABE, the proposed scheme adds on not only gullibility, but also litheness and fine-gained access control.

J. F. Wang, X. F. Chen (2015) [4] a collaborative multicloud data integrity along with the audition scheme, is conversed in this paper which is based on BLS (Boneh-Lynn-Shacham) signature and homomorphic tags. In consonance with the proposed scheme, patrons can audit their redistributed data in a one-round challenge-antiphon interaction with low recital operating cost. In turn to put off data by malicious threats, the user is acknowledged of the sensor area network (C-RAN cloud radio access network). In toting up, cloud storage facilitates the worldwide data access in any consigns. On the other hand, users lose the objective control of their outsourced data. This research outcomes the integrity of their tenuously outsourced data concern for users opting for cloud storing services.

C. Guo, X. Chen (2017) [5] discussed that users only can accumulate and administer their data on the cloud, and they can impart or boot up the data everywhere they are when there is a necessitate to compute the data. Due to the candidness of cloud, illicit organizations may be able to get hold of receptive information from the cloud servers, as emails and delicate health records, and encryption of the data is the most effectual way to defend the owners' confidentiality. This methodology of enumerating keyword search for the retrieval of data by the users randomly can be established for acquiring the encrypted data.

K.Ketzial Jebaseeli, Dr.V.G.Rani, (2019)[14] declared the use of aggregate key generation using attribute based encryption schemes. The intimation of ensemble signature scheme sets aside users to decrypt various classes of file using single key that can be collective among the users in dynamically twisted group. The projected model includes aggregate key invention for admittance of files assembled in hierarchy, hashing of key to amplify the sanctuary of keys, ring signature has been recognized to vigorously forming group for data access using cryptographic schemes Attribute based encryption has been utilized to endow with security and access control on both inside and outside assail to the outsourced data as it became well-liked on synchronization as it is comprised of Key-Policy Attribute-Based Encryption (KP-ABE),Cipher-text-Policy-Attribute-Based Encryption (CP-ABE) Hierarchical Attribute based Encryption Multi-Authority ABE to sort out key management issues that emerge at the phase of data sharing scheme on cloud services. The proposed model condenses the decryption time, recollection deployment and collision confrontation against any advances storage phase.

III. RESEARCH METHODOLOGY

The proposed system depicts the use of k-vertex in cloud sharing maneuvers. It furnishes dynamic amend on document compilation in conjunction with multi-keyword uncertainty and precise result foresights. The proposed keyword search in the user module described below gives the picture of a framework of the proposed model.

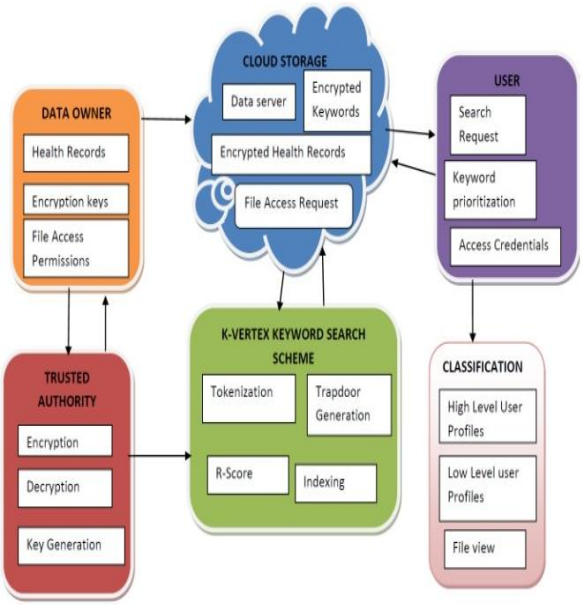

Fig 1. Proposed Framework

From the depicted figure above the data owner devolves the files into the cloud storage followed by the search request by the user. The data owner paves to trusted authority encompassing cryptography and generation of keys for the keyword search which in turn assembles R score for the engendering of trapdoor in the K-vertex proposal. The user then gets access to files in the cloud storage succeeding the acceptance of file access request.

A. Data owner

The entity has a collection of documents and wants to outsource to the cloud server in the encrypted form while still keeping the capability to search on them for effective utilization. The data owner first builds a Index T from document collection F, and then generates an encrypted document collection C for F[6]. Data owner outsources the encrypted collection C and the T index to the cloud server along the automatically generated trap door using TF and IDF calculations. Besides, the data owner is responsible for the update operation of his documents stored in the cloud server. While updating, the data owner generates the update information locally and sends it to the server [7].

B.DataUser

The given Figure 2 portrays the accession of data owner infiltrating the files through tokenization followed by arrangement of R-Score by picking up all the values in the K-Vertex. Finally the keyword using K2 is engendered through encryption leading to view of files by the data owner.

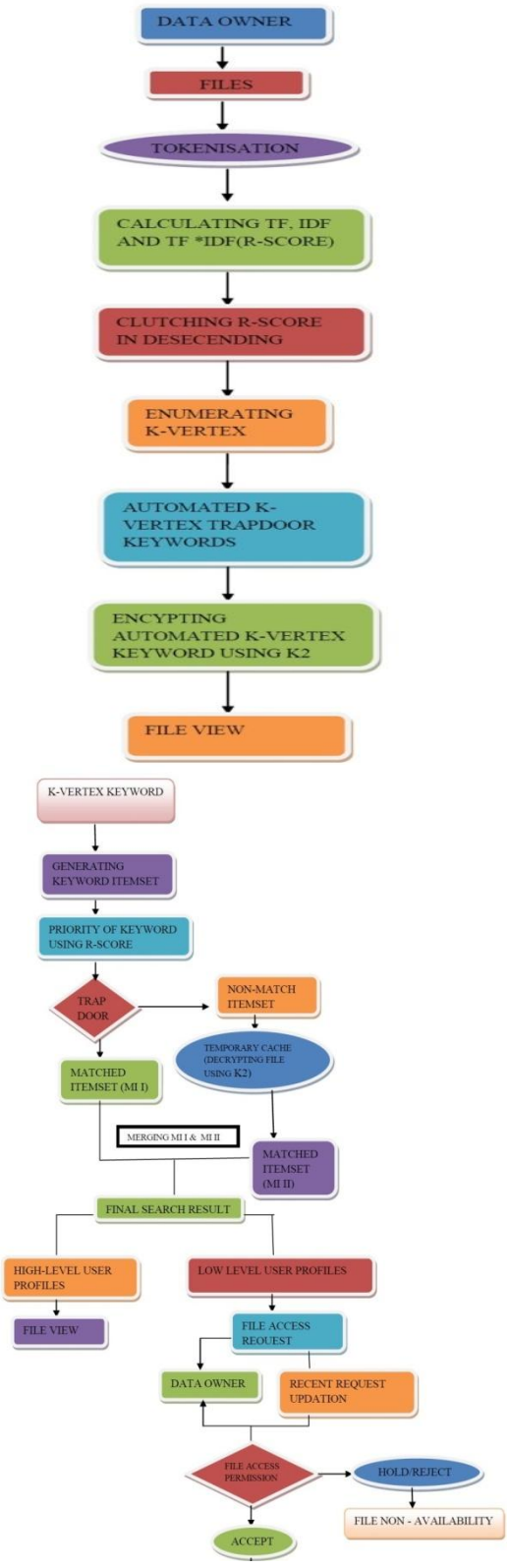

Fig 3 Access of files by user

From the Figure 3 above, the user make use of the emerging of keyword item set through k-vertex keyword. The authority by R score demonstrates k-vertex search scheme as matched and non matched

C. Cloud Server

The cloud server stores the encrypted document collection C, K-vertex k, and Index I of the data owners. The cloud server computes the retrieval operation upon receiving the query vector to trap door to return the corresponding collection of K-vertex ranked encrypted documents [8]. The figure 1 represents the architecture of the proposed mechanism.

D. Trusted Authority

The files that are from the data owner are therefore encompassed by means of trusted authority where the key generation for the two chief processes takes place. Encryption is made subsequent to the key generation which further consents to the decryption process leading to undertake of state to block transformations [9].

E. Cloud Storage

Cloud storage provides enormous storage competencies for the accession of files in a secured manner.[10] Outstanding to the unique formation, the index edifies query vector construction and k-vertex generation to be accomplished and to perk up the repossession effectiveness on self instigated and renewed operations.

F. Classification

The files are roughly classified on the basis of considerable levels of profiles as high level and low level profiles which are then directed to the file views.

IV. K-VERTEX SEARCH SCHEME:

The portrayal of the scheme in searching using K-vertex is partitioned as follows.

ALGORITHM:

The k-vertex search scheme model for the data files is generated involuntarily by taking up generate trap door function. It influences the Tf * IDF functionalities. The function is as follows

STEP I: Generate Trapdoor (W)
STEP II: Keyword set from buildindextree ()
STEP III: Initialize Keyword set =U
STEP IV: Compute TF (U) & Compute IDF (U)
STEP V: Calculating K-vertex Trapdoor
Trapdoor TD= (K vertex (TF) & K vertex (IDF))

Fig 4. Algorithm for Keyword search scheme

The above algorithm demonstrates that the search scheme on K-vertex generation involves ascertain progressions from setting keyword search using buildindextree () to computing K-vertex using TD= (K vertex (TF) &K vertex (IDF))

A. Term frequency and inverse document frequency

Term frequency is the number of times a given term. Appears within a document, and the inverse document frequency is obtained through dividing the cardinality of document collection by the number of documents containing the keyword [11]. It is designed to make rare words more important than common words Term Frequency is given by

$$Tf_{u,wi} = \frac{TF'}{\sum wi \xi W(TF')}$$

Where TF' is the TF mileage epithetical wi in document f. W is preeminent set out wherefrom keywords [12].

Inverse Document frequency is given by

$$idf_i = \log \frac{N}{n_i}$$

Where N is the number appertaining recited documents n_i is whoso number that contain word i

Filename	Keyword	TermFreq	IDF	TF_IDF
Patient1HealthRecord.txt	miscellaneous Z	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	Reason	0.006250	5.000000	3.465736
Patient1HealthRecord.txt	Inspector DoD	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	PT	0.001563	1.230000	6.238005
Patient1HealthRecord.txt	said	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	receiving	0.004688	5.000000	3.753311
Patient1HealthRecord.txt	SKELETAL	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	RESULTS	0.006250	5.000000	3.465736
Patient1HealthRecord.txt	home	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	Recalled	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	vitamin	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	download	0.003125	5.000000	4.158883
Patient1HealthRecord.txt	W	0.059375	5.000000	1.214444
Patient1HealthRecord.txt	AMLODIPINE	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	Heart	0.003125	5.000000	4.158883
Patient1HealthRecord.txt	TESTS	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	Periods Service	0.006250	5.000000	3.465736
Patient1HealthRecord.txt	panel	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	Ginkgo	0.001563	5.000000	4.851710
Patient1HealthRecord.txt	Cleared	0.001563	5.000000	4.851710

Fig.5 Values for TF-IDF

The above depicted screenshots in Fig 5, displays the values and keys that are generated and calculated in Index Document Frequency (IDF) and Term Frequency (TF).

B. Inverted index

Inverted index is a amply accustomed indexing structure that stores a list of mappings from keywords to the corresponding set of files that contain this keyword in the index vector, allowing full search on the multiple keywords through query vector [13].

C. Vector space model

Vector space model is widely used in plaintext acquaintance apocalypse, whichever dexterously affends ranked multi-keyword search on way of epitomizing notarization cross the squabbles. In this apiece chronicle is spasmodic in twain a warrant predominance tabulation conjunction represented as a vector. Queries can be represented as vectors in the tantamount thoroughfare as documents. Relevance score calculation and cosine similarity is most similarity measure. Cosine measure calculates the angle between the vectors d represents the document vector d' exemplifies the query vector. Cosine Similarity between the vectors is given by

$$\frac{d \times d'}{|d||d'|}$$

Where d represents the document vector d' betokens the query vector

D. Decryption Process

The decryption process undergoes State to block transformation on the following form from matrix to block transformation from which block to byte transformation occurs as presented below

- Decryption process ()
- Inv Sub byte ()
- *Data Encryption using AES:* The data encryption using symmetric key algorithm[12] named as Advanced Encryption Standard (AES) undergoes following steps to generate the cipher text to the document collection F. Sub byte is that a Byte is transformed into hexadecimal digits
- *Montage:* The data do minus dissevers the rabble and esoteric parameters of the system by executing KeyGen, and pre-processes the data file collection C by using Build Index to generate the Index from the unique words extracted from C

Word = Collection of Words
Block = Collection of Bytes
State = Matrix representation of Blocks

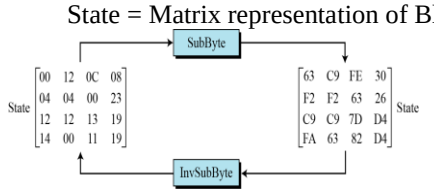

Fig 6. Employ Block to State transformation ()[13]

The figure depicts Shift rows in matrix (). It employs the 1 byte shift |2 byte shift |3 byte. Shift based on no of rows in the matrix on right side.

- *Byte transformation:* Acknowledged by Shift rows in matrix (). It employs the 1 byte shift |2 byte shift |3 byte. Shift based on no of rows in the matrix on right side
- *Mix column of the matrix ():* It is an inter byte transformation that changes the bits inside a byte, based on the bits inside the neighboring bytes.
- *Add round key ():* Add Round Key adds a round key word with each state column matrix

E. Hierarchical Data Access Strategies

The data access strategies consist of two hierarchies which high level and low level. In this higher level undergoes decryption with the decryption key whereas lower level undergoes decryption on request of decryption key from the data owner.

V. EXPERIMENTAL RESULTS

The investigational results have been calculated by the performance in comparing the corresponding keyword search and the performance evaluation between the numbers of keywords generated. we analyse the security of the proposed model against the various data size is been computed and described in terms of performance charts and tables for various security performance measure like decryption time , collusion resistance and memory usage.

A. EXPERIMENTAL SETUP

The probative upshots are wrangled por an Intel Core I3 aedile with 2620 Processors (2.0 GHz) and 4 GB RAM and 500 GB Hard disk using Dot net programming. The database server and file server has been established in the file system which acts as Cloud infrastructure. The different file sizes are used for computation of the efficiency of the model. Proxy server has been established using virtualization process as Virtual machine.

B. PERFORMANCE ANALYSIS

The reliability of the system is defined in terms of utilization of proxy server technology in data sharing systems to reduce the decryption time of the group members. It will greatly facilitate data owner delegating the access rights to other members.

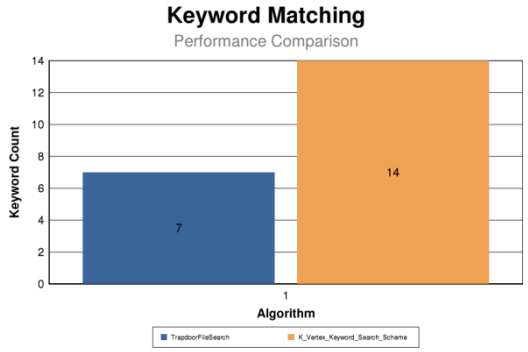

Fig 7: Performance evaluation of keyword matching between keywords generated

The keyword matching utilizes the time in terms of index vector analysis and vector splitting process analysis. The figure 7 and figure 8 represents the performance of the keyword matching on file and trap door in this work.

The contrasting matches of keywords search is given in Figure 4. The keyword based file search is lessened than the matching of multi keyword ranked search. This is attained by the above Figure 4 from the evaluation of keyword matching comparison.

Table-I: Performance Evaluation of Keyword

Searching Count	Trapdoor file search	K-vertex keyword search scheme
1	7	14

The searching of files through efficiency is not comparatively superior to the keyword search scheme according to the counts that are taken for the search format. The techniques of k-vertex searching of keywords are precisely evaluated on cipher hierarchical encryption.

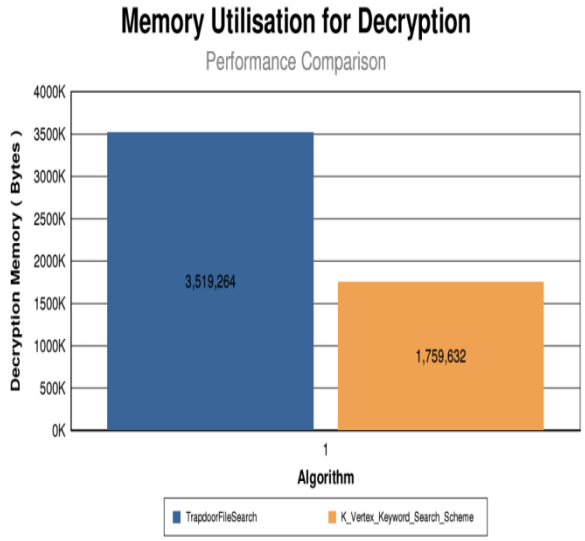

Fig 8: Performance Evaluation of the Memory utilization on Decryption process on proposed model and existing models to various file sizes

The search precision of scheme is affected by the imitated Keywords. If keyword standard deviation is used set random variable, it is supposed to obtain higher precision. The figure 6 provides the precision, recall and f measures values for the proposed and existing models on various data sizes of the health records.

Table-II: Performance Evaluation of Memory Utilization on Decryption

Records	Trapdoor file search	K-Vertex Keyword Search Scheme
1	3,519,264	1,759,632

The precision is given by the k-vertex documents or documents that are repossessed from the k-vertex articles. The real rank number is ordained by the recall of rank numbers form the retrieved k-vertex documents and measured by means of F measure from harmonic mean of recall and precision

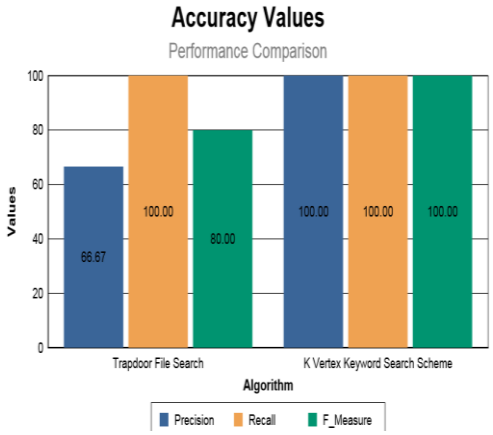

Fig 9: Performance Evaluation of Accuracy Values on proposed model with respect to keyword search scheme

The above depicted figure demonstrates the assessment of values ranging from 0 evaluating trapdoor file search in precision, recall and F-measure and K vertex searching keywords towards the same precision, recall and F-measure.

Precision = number of real k-vertex documents/ retrieved k documents.

Recall = putrid caseload of connote in the recouped k-vertex documents/ real rank number.

F measure = harmonic Mean of precision and Recall

Table III: Performance Evaluation of Accuracy values

Algorithm	Precision	Recall	F-Measure
Trapdoor File Search	66.67	100.00	80.00
K Vertex Keyword Search Scheme	100.00	100.00	100.00

Efficiency is commensurate to the amplitude of footnote whilst the dossier agglomeration is fixed as expunsion of a document takes nearly logarithmic time with the size of docket muster. The techniques of k-vertex searching of keywords are precisely evaluated on cipher hierarchical encryption. The memory utilization and accuracy are estimated by means of megabytes with regard to ensemble signature scheme.

VI. CONCLUSION

This proposed system is designed and implemented on a K-vertex keyword search on hierarchical data concerning dynamic updates of the intermittent data. It reinforces not only the multiple keyword searches but as well as the dynamic updates of the outsourced data and the flexible revocation of the data abuser. The automated trap door generation is ascertained using TF*IDF model. The competent data index has been achieved to acquire enhanced repossession efficiency to user queries. Adding up to this, query vector construction is established in order to perk up the query efficiency. Moreover, affinity summation estimation entre enciphered indicia and catechize bearings endow with high accurate results. As a final point, data owner has been imparted with flexible revocation provision. Experimental results make obvious the efficiency and accuracy of our proposed scheme on an assortment of measures.

REFERENCES

1. A. Swaminathan, Y. Mao, G.-M. Su, H. Gou, A. L. Varna, S. He, M. Wu, and D. W. Oard, "Confidentiality-preserving rank-ordered search," in Proc. ACM Workshop Storage Security Survivability, 2007, pp. 7–12.
2. S. Yu, K. Ren, and W. Lou, "FDAC: Toward fine-grained distributed data access control in wireless sensor networks," IEEE Transactions on Parallel and Distributed Systems, vol. 22, no. 4, pp. 673–686, 2011.
3. Z. Wan, J. Liu, and R. H. Deng, "HASBE: A hierarchical attribute-based solution for flexible and scalable access control in cloud computing," IEEE Trans. Inf. Forensics Security, vol. 7, no. 2, pp. 743–754, Apr. 2012.
4. J. F. Wang, X. F. Chen, J. Li, J. L. Zhao, and J. Shen, "Towards achieving flexible and verifiable search for outsourced database in cloud computing a¸ˆı," in Future Generation Computer Systems, vol. 67, pp. 266-275, 2016.
5. C. Guo, X. Chen , Y. Jie , F. Zhang, M. Li, B. Feng, Dynamic Multiphrase Ranked Search over Encrypted Data with Symmetric Searchable Encryption, IEEE Transactions on Services Computing, vol. 30, no. 10, pp. 1-12, 2017.
6. Q. Chai, G. Gong, Verifiable symmetric searchable encryption for semihonest- but-curious cloud servers, in: Communications (ICC), 2012 IEEE International Conference on, IEEE, 2012, pp. 917–922.
7. C. Wang, N. Cao, K. Ren, and W. J. Lou, "Toward Efficient Multi-Keyword Fuzzy Search Over Encrypted Outsourced Data With Accuracy Improvement," in IEEE Transactions on Parallel and Distributed Systems, vol. 23, No. 8, pp. 1467-1479, 2012.
8. B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multikeyword fuzzy search over encrypted data in the cloud," in Proc. IEEE INFOCOM, 2014, pp. 2112–2120.
9. Y. C. Chang, M. Mitzenmacher, Privacy Preserving Keyword Searches on Remote Encrypted Data, in: International Conference on Applied Cryptography and Network Security, Springer, 2005, pp. 442-455.
10. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient constructions," in Proc. 13th ACM Conf. Comput. Commun. Secur., 2006, pp. 79–88.
11. Q. Zheng, S. Xu, G. Ateniese, VABKS: verifiable attribute-based keyword search over outsourced encrypted data, in: INFOCOM, 2014 Proceedings IEEE, 2014, pp. 522–530.
12. J. F. Wang, X. F. Chen, and J. Li, "Verifiable Search for Dynamic Outsourced Database in Cloud Computing," presented at the International Conference on Broadband and Wireless Computing pp. 568-571, 2015.
13. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced Cloud data," IEEE Trans. Parallel Distribute. Syst., vol. 23, no. 8, pp. 1467–1479, Aug 2012
14. K.Ketzial Jebaseeli, Dr.V.G.Rani, "A Lightweight Data Preserving model using Ensemble Signature Scheme to the Outsourced health files in cloud", Journal of Advanced Research in Dynamical and Control Systems,Vol.11,02-Special Issue,2019.

AUTHORS PROFILE

K.KETZIAL JEBASEELI BCA., MCA., Asst.Professor, IT Department, holder of the BCA degree from Bharathiar University in 2008 and received MCA degree from Anna University in 2011. She is trailing her PhD in computer science at Sri Ramakrishna College for women. She has a practice of experience in her teaching for about 5 years. She has published papers in International journals and also presented papers in various International National and National level conferences.

DR.V.G.RANI MCA., Ph.D., Associate Professor, CS Department. She received her Ph.D in Computer Science from Bharathiar University for the period of 2013. She is embraced with her master's degree in MCA and Bachelor degree in computer Science from Bharathiar University. Her scholastic in Sri Ramakrishna CAS for women has attained for about 16 years. Ad Hoc network, Security solution, Cloud computing and IoT are her research areas. She has showed her guidance towards 10 M.Phil scholars and at present 2 PhD scholars are pursuing under her guidance. She has brought out further 10 papers in national and international conferences. Her publication of journal includes Scopus and IEEE. She has systematized more than 50 seminars and has enthusiastically participated in various FDP, Workshop and Seminars carried out by various colleges and universities.