ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 8, Issue 3, March 2020

An Efficient and Dynamic Semantic Aware Search Scheme over Encrypted Cloud Data Using Sent2Vec Model

P. Sukumar[1], V.Madhumitha[2], P.R. Malarvizhi[3], R. Dharan Gokul[4]

Department of CSE, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India[1,2,3,4]

ABSTRACT: Cloud Computing involves the process of storing encrypted data in cloud and retrieve the data from Cloud storage at any time. This platform provides data privacy and confidentiality. Doc2Vec model andWord2Vecmodel are Existing models which uses large computational resources such as data storage and time. The proposed system involves Sent2Vec model, it is used to extract the semantic information's from document. A multi-keyword top-k ranked search scheme is used over encrypted cloud data which is used for dynamic update operations in document and keys. A Group of tree-based indexes are constructed for all the document.This effective index is based on Greedy Depth- First search algorithm. It is a Keyword based document retrieval. Symmetric encryption is performed by the same secret key generated by encrypted data before outsourcing it to cloud. The secret key generated over encryption is used for retrieving the data from cloud. It reduces less computational overhead. Large size of file can be uploaded and secret key can be viewed by the data owner and exchanges the encrypted secret key for retrieving the data to user.

KEYWORDS: Cloud computing, multi-keyword ranked search, semantic-aware search, Searchable encryption.

I. INTRODUCTION

Cloud computing plays a vital role in storing an individual private data. Meanwhile cloud is not a completely trustworthy, private data can be attacked or hacked. Thus, secure encryption of data is necessary before outsourcing to cloud. Here data utilization becomes a hard issue. Many search algorithms are available but they are not much effective. Recently, several researchers have projected variety of searchable secret writing schemes like single keyword search [1], multi-keyword search [8], fuzzy keyword search [10] and conjunctive keyword search [17] etc. These searchable schemes usually have five steps: extracting dataset options, building the secure index and encrypting the dataset, generating search trapdoor, searching over the index and returning the results. C.Nagarajan *et al.* [3,6,9] has proposed this method TheWord2Vec [18] and Doc2Vec [19] are used models in Natural Language Processing.

In this paper, we propose a dynamic semantic aware search scheme over encrypted cloud data using Sent2Vec model. This model used for Natural Language Processing and encryption, decryption and search. It is also used for gathering semantic information's from the document. Multi-keyword top-k search scheme is used for effective search process and Greedy Depth First Searchis used for index generation at the time of encryption. The importance of preferring multiple keyword search than single keyword search.

A. Single Keyword Search

A single keyword searchable encryption schemes typically builds associate data searchable, its content is hidden to the server, unless its applicable trapdoors generated via secret keys. However, it solely supports single keyword search. Wherever anyone with key will write to the data kept on server, however solely approved users with private key will search. Single keyword search schemes use encrypted data for searchable index. These indexes content is going to be hidden to the server. It's not comfy enough to specific complex information desires is that the major downside of single keyword search.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 8, Issue 3, March 2020

B. Multiple Keyword Search

This scheme retrieves search results supported the existence of keywords and can't give acceptable result ranking practicality. Multi-keyword search on encrypted cloud knowledge are investigated in. It provides security and economical search by using two thread models, cipher text model and background model. A secure top-k formula is used. Economical privacy conserving search over encrypted cloud data that utilizes min hash functions to boost the exactness rate. It gives an accurate result than Single Keyword Search.

II. LITERATURE REVIEW

A. Secure Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing

With the advent of cloud computing, it becomes more and more in style for owners to source their information to public cloud servers whereas permitting users to retrieve this information. For privacy issues, secure searches over encrypted cloud information actuated many researches beneath the owner model. However, most cloud servers in observe don't simply serve owner instead, they support multiple owners to share the advantages brought by cloud servers. During this paper, we tend to propose schemes to manage secure hierarchal multi-keyword search during a multi-owner model. To change cloud servers to perform secure search while not knowing the particular information of each keywords and trapdoors, we tend to consistently construct a completely unique secure search protocol. To rank the search results and preserve the privacy of connectedness scores between keywords and files, we tend to propose a completely unique Additive Order and Privacy protective operate family. In depth experiments on real-world datasets ensure the effectualness and potency of our projected schemes. [20]

B. Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing

With the arrival of cloud computing, it's become progressivelystandard for data owners to source their information to public cloud servers whereas permitting data users to retrieve this information. For privacy considerations, secure searches over encrypted cloud information have impelled many analyses works underneath the only owner model. However, most cloud servers in follow don't simply serve one owner; instead, they support multiple house owners to share the advantages brought by cloud computing. During this paper, we have a tendency to propose schemes to cope with privacy protective graded multi-keyword search during a multi-owner model. To change cloud servers to perform secure search while not knowing the particular information of each keywords and trapdoors, we have a tendency to consistently construct a unique secure search protocol. To rank the search results and preserve the privacy of connected scores between keywords and files, we have a tendency to propose a unique additive order and privacy protective operate family. To forestall the attackers from eavesdropping secret keys and deceit to be legal information users submitting searches, we have a tendency to propose a unique dynamic secret key generation protocol and a replacement information user authentication protocol. What is more, supports economical information user revocation. In depth experiments on real-world datasets ensure the effectual and potency. [21]

C. Secure Multi- Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing

Now days with the arrival of cloud computing, it's become progressively well-liked for knowledge owners to store their information to public cloud servers whereas permitting data users to retrieve this information. Users stores their information in encrypted format on the cloud that's why unauthorized user cannot access the information. Once user desires to access the information they need to urge cryptography key from user. However, most cloud servers in observe don't simply Serve distinctive owner; instead, they support multiple house owners to share the advantages of the cloud computing. This paper, we propose 1) To stay safe the secrecy 2) Several-owners model search many keywords and hierarchical. to form attainable cloud servers to execute safe to seem omission knowing the important information of each keywords and trapdoors, 1) To keep alive the privacy of connected scores between keywords and files and rank the search result, we propose a unique Order and Privacy protective perform family.2) Dynamic hidden key creation rule and a replacement knowledge user to ascertain as real rule. [22]

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 8, Issue 3, March 2020

D. An Efficient Ranked Multi-Keyword Search for Multiple Data Owners Over Encrypted Cloud Data
With the event of cloud storage, additional information data are inclined to source their information to cloud services. For privacy issues, sensitive information ought to be encrypted before outsourcing. There are numerous searchable cryptography schemes to confirm data availableness. However, the present search schemes pay very little attention to the potency of data users' queries, particularly for the multi-owner situation. During this study, we tend to plan a tree-based hierarchic multi-keyword search theme for multiple data owners. Specifically, by considering an outsized quantity of information within the cloud, we tend to utilize the TF×IDF model to develop a multi-keyword search and come the top-k hierarchic search results. To modify the cloud servers to perform a secure search while not knowing any sensitive knowledge example keywords and trapdoors, we tend to construct a completely unique privacy-preserving search protocol supported the additive mapping. To attain AN economical search, for every knowledge owner, a tree-based index encrypted with AN additive order and privacy-preserving perform family is built. The cloud server will then merge these indexes effectively, victimization the "Depth-first Search" rule to search out the corresponding files. Finally, the rigorous security analysis proves that our theme is secure, and also the performance analysis demonstrates its efficacious and potency. [23]

E. Efficient Multi-keyword Ranked Search for Multiple Data Owner in Cloud Computing by Using AES 3DES Hybrid Approach
With the fast development of varied transmission technologies, more and more transmission information is generated and transmitted within the medical, conjointly the internet permits for wide distribution of digital media information. It becomes abundant easier to edit, modify and duplicate digital information. Besides that, digital documents are simple to repeat and distribute, so it'll be faced by several threats. Privacy protective graded Multi-keyword Search during a Multi-owner model. To change cloud servers to perform secure search while not knowing the particular information of each keywords and trapdoors, during this projected system consistently construct a unique secure search protocol. To rank the search results and preserve the privacy of connect scores between keywords and files, we have a tendency to propose a unique Additive Order and Privacy protective perform family. To stop the attackers from eavesdropping secret keys and dissimulation to be legal information users submitting searches, we have a tendency to propose a unique dynamic secret key generation protocol and a brand-new information user authentication protocol. [24]

III. PROPOSED SYSTEM

The use of Sent2Vec help in effective Encryption, Decryption and Search when compare to Word2Vec and Doc2Vec model. This model provides a Dynamic semantic search scheme and also extract the semantic information from the document. Group of indexes are constructed when encryption process is done which is handled by Greedy Depth-First Search. The base algorithm for key generation is Elliptic Curve Cryptography which is cryptographic technique used for Key generation for both data owner and data vendor. This improves the security of the document.

- The secure Greedy Depth-First Search algorithm is used to encode the index and query vectors, and ensure correct relevance score calculation between encrypted index and query vectors.
- Abundant works are projected underneath completely different threat models to realize varied search functionality.
- To resist completely different attacks in numerous threat models, the proposed systemconstruct two secure search themes the essential dynamic multi-keyword ranked search scheme within the notedcipher text model, ranked increased dynamic multi-keyword ranked search theme is understood as background model.

A. Advantages of Proposed System

- Uses Dynamic update operation of document and key
- Sent2Vec models are faster to train the sentences compared to methods like Skip Thought and Doc2Vec
- The Secret Key can be viewed by user and data owner

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 8, Issue 3, March 2020

- The uploading file size is increased considerably.

IV. SYSTEM ARCHITECTURE

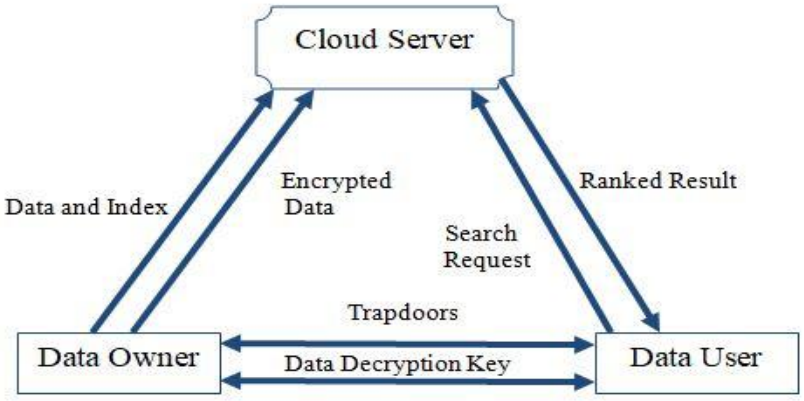

Figure.1 System Architecture

A. Description of the system

To enable an efficient and dynamic search over encrypted cloud data in this system has the following

- Sent2Vec
  Sent2Vec model is used to extract the semantic features such as sentence and keywords from the document. This model is used for the encryption, decryption and search with the base algorithm that is Elliptic Curve Cryptography for key generation. Greedy Depth-First Search Algorithm is used for index generation at the time of encryption. These keys generated separately for both data owner and data vendor for secure transaction.
- Ranked search
  The ranked search model is to return the most relevant *k* documents from the document set we measure, the relevance score for the query and a document by the cosine distance between the corresponding vectors of them. This gives two *n*-dimensional vectors.
- Multi-keyword top-k search scheme
  Based on the DMRSE, EDMRSE scheme, the proposed system has Multi-keyword top-k search scheme algorithm to enhance its security. In the known background model, Cloud Server can view the data owner's and data user's details, the statistical information of the documents and query vectors. Because the query vectors generated by the same queried keywords have the same value, cloud server can link these query requests by inferring the search result's relevance score, which stays unchanged with the same trapdoor. Thus, we introduce some phantom terms on the vectors of the documents and trapdoor to obfuscate the ranked search result.
- Search efficiency
  By exploring a special tree-based index and an effective search algorithm, the scheme aims at achievingsub-linear search efficiency.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 8, Issue 3, March 2020

- Keyword privacy
  Cloud Server should not identify the special keyword in documents, index, and trapdoors by analyzing the statistical semantic information. The keyword vector is replaced by the feature vector in our scheme. The cloud server cannot tell the difference between keyword vectors and feature vectors. Thus, the semantic feature privacy is also discussed in the analysis of keyword privacy.

V. SYSTEM DESIGN

A. Module description

- Data Owner
  In this module the owner can register their details along with login details. The owner can upload their file along with the encryption process. This ensures the files to be protected from unauthorized user. Data owner sends the document to the authorized user in encrypted form along with trapdoor. Trapdoor can be edited by the data owner. Data owners would define the access policy and compute the authorization cipher text for each document.
- Data User
  Users are approved ones to access the documents of data owner. With t query keywords, the approved user will generate a trapdoor in line with search management mechanisms to fetch k encrypted documents from cloud server. Then, the information user will decode the documents with the shared secret key.
- Cloud Server
  This module is used to observe the process between data owner and data user. Cloud server stores the encrypted document collection C and encrypted searchable tree index. When receiving the trapdoor from the user, the cloud server executes search among the index tree t, returns the corresponding assortment of top-k hierarchic encrypted documents. The server must update the index I and document collection C in line with the received information.

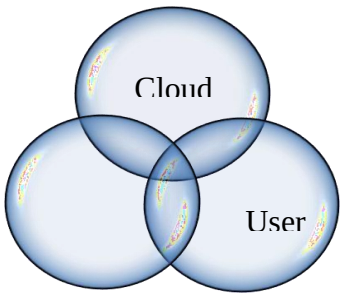

Figure.2 Module description

VI. CONCLUSION AND FUTURE WORK

The proposed work is, a secure, effective and dynamic scheme which is used over encrypted cloud data. Dynamic update of document and key can be performed. Depth First Search Algorithm is used for effective index generation at the time of encryption of the document. System can achieves better efficiency in terms of functionalities and computational overhead. In the future, Owner can add zip and rare format file to cloud server, the size of the file will be increased for supporting Big data. Computational resources such as data storage and time can bereduced and offline accessing can be included.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 8, Issue 3, March 2020

REFERENCES

[1] Y. Chang and M. Mitzenmacher, ``Privacy preserving keyword searches on remote encrypted data," in *Proc. Int. Conf. Appl. Cryptogr.Netw.Secur.*, New York, NY, USA, Jun. 2005, pp. 442_455.
[2] A. Swaminathan, Y. Mao, G.-M. Su, H. Gou, A. L. Varna, S. He, M. Wu, and D. W. Oard, ``Con_dentiality-preserving rank-ordered search," in *Proc. ACMWorkshop Storage Secur.Survivability*, Alexandria, VA, USA, Oct. 2007, pp. 7_12.
[3] C.Nagarajan, M.Muruganandam and D.Ramasubramanian – 'Analysis and Design of CLL Resonant Converter for Solar Panel - Battery systems- International Journal of Intelligent systems and Applications (IJISA), Vol.5 (1),pp.52-58, 2013.
[4] C. Wang, N. Cao, K. Ren, and W. Lou, ``Enabling secure and efficient ranked keyword search over outsourced cloud data," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 8, pp. 1467_1479, Aug. 2012.
[5] D. X. Song, D. Wagner, and A. Perrig, ``Practical techniques for searches on encrypted data," in *Proc. IEEE Symp. Secur.Privacy*, May 2000, pp. 44_55.
[6] C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- *Electric Power Components and Systems*, Vol.39 (8), pp.780-793, May 2011
[7] S. Zerr, D. Olmedilla, W. Nejdl, and W. Siberski, ``ZerberC*R*: Top-k retrieval from a congenital index," in *Proc. 12th Int. Conf. Extending Database Technol., Adv. Database Technol.*, Saint Petersburg, Russia, Mar. 2009, pp. 439_449.
[8] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, ``Privacy-preserving multikeyword ranked search over encrypted cloud data," in *Proc. IEEE INFORMATIONCOM*, Shanghai, China, Apr. 2011, pp. 829_837.
[9] K Umadevi, C Nagarajan, "High Gain Ratio Boost-Fly Back DC-DC Converter using Capacitor Coupling", 2018 Conference on Emerging Devices and Smart Systems (ICEDSS), 2nd and 3rd March 2018, organized by mahendra Engineering College, Mallasamudram, PP. 64-66,2018
[10] B. Wang, S. Yu, W. Lou, and Y. T. Hou, ``Privacy-preserving multikeyword fuzzy search over encrypted data in the cloud," in *Proc.IEEE Conf. Comput. Commun.*, Toronto, ON, Canada, Apr./May 2014,pp. 2112_2120.
[11] A. Helsinger and T. Wright, "Cougaar: A robust configurable multi agent platform," in Proc. of the IEEE Aerospace Conference,2005.
[12] J. Brunelle, P. Hurst, J. Huth, L. Kang, C. Ng, D. C. Parkes, M. Seltzer, J. Shank, and S. Youssef, "Egg: an extensible and economics-inspired open grid computing platform," in Proc.of the GECON, Singapore, May 2006.
[13] J. Norris, K. Coleman, A. Fox, and G. Candea, "Oncall: Defeating spikes with a free-market application cluster," in Proc.of the International Conference on Autonomic Computing,New York, NY, USA, May 2004.
[14] C. Pautasso, T. Heinis, and G. Alonso, "Autonomic resource provisioning for software business processes," Information and Software Technology, vol. 49, pp. 65–80, 2007.
[15] A. Dan, D. Davis, R. Kearney, A. Keller, R. King, D. Kuebler, H. Ludwig, M. Polan, M. Spreitzer, and A. Youssef, "Web services on demand: Wsla-driven automated management,"IBM Syst. J., vol. 43, no. 1, pp. 136–158, 2004.
[16] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A berkeley view of cloud computing," University of California, Berkeley, Tech. Rep. USB-EECS-2009-28, Feb 2009.
[17] W. Sun, X. Liu, W. Lou, Y. T. Hou, and H. Li, ``Catch you if you lie to me: Ef_cientveri_able conjunctive keyword search over large dynamic encrypted cloud data," in *Proc. IEEE Conf. Comput. Commun.*,Kowloon,Hong Kong, Apr. 2015, pp. 2110_2118.
[18] T. Mikolov, K. Chen, G. Corrado, and J. Dean, ``Ef_cient estimation of word representations in vector space," in *Proc. 1st Int. Conf. Learn.Represent.*, Scottsdale, AZ, USA, May 2013, pp. 1_12
[19] Q. V. Le and T. Mikolov, ``Distributed representations of sentences and documents," in *Proc. 31st Int. Conf. Mach. Learn.*, Beijing, China,Jun. 2014, pp. 1188_1196.
[20] Wei Zhang, Sheng Xiao , Yaping Lin , Ting Zhou , Siwang Zhou" Secure Ranked Multi-keyword Search for Multiple Data Owners in Cloud Computing," in 2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks.