iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35184

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

A Survey on Secure Ranked Keyword Search over Outsourced Encrypted Cloud Data

Akash Tidke[1], Abhishek Pandey[2], Rukmin Darku[3]
[1, 2, 3]*Student, PDEA's College of Engineering Pune*

Abstract: In this paper we present a survey on keyword based searching algorithms. Various searching techniques are used for retrieving the encrypted data from cloud servers. This survey work involves a comparative study of these keyword based searching algorithms. It concludes that till now multi-keyword ranked search MRSE scheme is the best methodology for searching the encrypted data.
Keywords: Cloud, keyword search, Ranked search, Outsourced data.

I. INTRODUCTION

Cloud storage services enable users to remotely access data during a cloud anytime and anywhere, during a pay-as-you-go manner using any device. Moving data into a cloud offers great convenience to users since they are doing not need to care about the massive capital investment in both the deployment and management of the hardware infrastructures. However, allowing a cloud service provider (CSP), whose purpose is especially for creating a profit, to require the custody of sensitive data, raises underlying security and privacy issues. To keep user data confidential against an untrusted CSP, a natural way is to use cryptographic approaches, by disclosing the info decryption key only to authorized users. Many applications like emails, file storage, business data, etc. are outsourced to cloud server. Only authorized user can access the info from the cloud server. Outsourcing unencrypted data to cloud by the owner isn't much secure because server may leak information to cyberpunks. Hence encryption plays a serious role before outsourcing the info into the cloud server. In spite of encrypting, retrieval of knowledge becomes an intriguing task when searching has got to be made on vast data. The best way is to use keyword based search on encrypted data for data concealing.

II. TECHNIQUES FOR KEYWORD BASED SEARCH

Many searchable techniques are proposed on the idea of keyword search. Discussion is made on the existing techniques that are been intend by many authors. This study analyses the algorithms for searching the encrypted content. Survey is formed on these algorithms supported the working principle, merits and demerits. It also compares the complexity, efficiency overhead of varied algorithm s and shows which technique is best to handle while retrieving the encrypted content. It includes working of encryption algorithm, how searching is completed on the encrypted content, advantages and disadvantages of every technique.

A. Symmetric Key cryptography

Symmetric key cryptography works by encrypting each word construction. Probabilistic searching is formed on the encrypted sequential scan and indexing methodologies, provable secrecy, in a file using two layered encryption data. Probabilistic searching deals with controllable searching, hidden queries, query isolation [1] are the four techniques which make the algorithm efficient,simple and fast. Sequential scan meets all the above techniques but it's not effective when searching is formed on huge data content. Therefore to induce effective searching pre-computed index plays a crucial role which support advanced search queries. But to form indexing technique secure [2], secure index data structure are often used which admits queries with a trapdoor. It is semantically secure and practicable in multi-user settings where indexes are updated frequently on the remotely located server.

B. Public Encryption Keyword Search

Public Encryption Keyword Search (PEKS), a searchable encryption technique which corresponds to symmetric key encryption. In this, file is encrypted using public key by the people who wants to store it within the server but the authorized users can search a file using their private key [3]. Consider user Bob sending mail to Alice encrypted under Alice's public key. An email gateway wants to check whether email contains word 'important' so on route the e-mail accordingly. Alice doesn't wish to offer the gateway the power to decrypt all her messages.

PEKS may be a mechanism that permits Alice to supply a key to gateway that permits the gateway to check whether the 'important' is in email without learning anything within the email. First, keyGen is employed to get public key and personal key pair for both server and user. Second, PEKS algorithm produces searchable encryption. Third, Trapdoor algorithm is employed to calculate trapdoor with private key and keyword. Fourth, Test is used to match the keyword and requested word. If matches then the file is shipped to the user.

C. Hidden Vector Encryption

Hidden Vector Encryption (HVE) supports continuative queries [6] whereas PEKS supports only comparison and subset queries. HVE works with four algorithms namely Setup ,Encrypt, GenToken and Query .First, Setup creates a bilinear group of elements using random primes and random elements. Second, Encrypt chooses the random element and using public key it encrypt the contents in a file. Third, GenToken will generate the token for the predicate using a secret key. Say, a student result is encrypted for security reasons, however a statistical program may want to know some information about the data e.g., it may want to count how many students scored over 90%, which is fairly innocuous goal. We don't want to allow such program to decrypt all the data and find out the identities of the students. Therefore, it is desirable to encode the "x >= 90%" predicate in a token T to allow the program to compute this information. Furthermore, if the predicate contains a conjunction P1^P2, we do not want the encryption scheme to leak which predicate satisfied the expression. Fourth, Query finds the keyword from the cipher text and if matches it return the file. Even so HVE fails for disjunctive queries because cipher text is linear to attribute.

D. Attribute based Encryption

In distributed settings with untrusted servers, like the cloud, many applications need mechanisms for complex accesscontrol over encrypted data. Sahai and Waters [13] addressed this issue by introducing the notion of attribute-based encryption (ABE). ABE is a new public key based one-to-many encryption that allows users to encrypt and decrypt data supported user attributes. ABE enables access control over encrypted data using access policies and ascribed attributes associated with private keys and cipher texts. There are two kinds of ABE schemes: key-policy ABE (KP-ABE) [14]-[19] and cipher text-policy key-policy ABE (KP-ABE) [14]-[19] and cipher text-policy every cipher text is associated with an access policy on attributes, and each user's private key's related to a group of attributes. A user is able to decrypt a cipher text only if the set of attributes related to the user's private key satisfies the access policy associated with the cipher text. In a KP-ABE scheme, the roles of an attribute set and an access policy are swapped from what we described for CP-ABE: attributes sets are used to annotate the cipher texts and access polices over these attributes are associated with users' private keys. It provides best quality for searching over encrypted data and faster in accessing.

E. Predicate Privacy Preserving in Public Key Encryption

Predicate privacy preserving search on keyword get the better of PEKS by using randomization technique. In which keywords are randomized and therefore trapdoors does no provide any meaningful keywords. For a word x the corresponding trapdoor t(x) is generated from the master secret held only by sender and is used to define predicate P. The property that t(x) reveals no information about the encoded predicate P is called Predicate privacy. In PEKS two types of attacks can occur. One is brute force guessing attack. Second is statistical guessing attack. To make tolerant of guessing attacks, two framework are introduced namely PEKSrand-BG for bruteforce guessing and PEKSrand-SG for statistical guessing [7]First idea is to randomize the original keywords. Hence, the transformed keywords used to generate trapdoors are not dictionary words any more. Also deterministic and direct one-toone mappings are avoided.

A solution to this is that, the receiver and all senders share a secret, which is concatenated with the original keywords. However such privacy protection is weak, since the protection of shared secret is difficult when the set of senders is large. Also there are scenarios when the membership of set of senders is dynamic which results in additional costs of key/secret management. To address the above issue we limit the entities that hold the secret used for randomization to only one or a few proxy servers, which are well protected and thus are more secure than normal senders.

The PEKSrand-BG scheme is built upon the first idea. Another idea is to spread out the statistical distribution of keywords by mapping a keyword to multiple trapdoors instead of one. It can weaken the statistical guessing attacks at the cost of increasing overhead storing trapdoors at the delegate.

The PEKSrand-SG scheme is developed through a combination of both ideas. In this design, besides 3 types of entities in PEKS system, a new type of entity is added called proxy server.

F. Privacy Preserving Keyword Search

It is a multi-round protocol between server and user on single keyword. Privacy preserving keyword search utilizes the notion of a keyword index, which is made by user u. The keyword index associates each keyword with its associated files. The keyword index is being created offline with a more powerful home machine, before the user wishes to access the files remotely with a mobile device. All keyword searches by user u are supported this index [8][4].Main idea is that user u uses pseudorandom bits to mask a dictionary based keyword index for every file and send it to server s in such how that later u can use the short seeds to assists recover selective parts of the index, while keeping the remaining parts pseudorandom. On the setup phase user chooses a random secret key to encrypt the file. Then the user submits index and file content to server. On the retrieval phase, when the user wants to look or retrieve file from the server, user retrieves the index file then computes keyword with the secret key. The computed key's sent to server, where server matches the file then sent to the user. This scheme fails when multiple keywords are used. The per-index file scheme using pseudorandom functions is that the better than using bloom filters. Bloom filters can induce false positives, which can cause mobile users to download extra files not containing the keyword. Privacy Preserving Keyword Search avoids this issue.

G. Secure Privacy Preserving Keyword Search

Secure Privacy Preserving Keyword Search (SPKS) grants cloud service provider to decrypt the info and return file containing keywords [9]. This technique overcomes the computation and communication overhead, provides query and data privacy for the users. A user may use his public key to encrypt an email and its keywords before sending it to the CSP, and then sends queries within the sort of encrypted keywords to retrieve the email. Since the key key's only known to the user himself, an attacker isn't conscious of the encrypted files, the encrypted keywords, and the user querying patterns. However,such an easy encryption scheme may introduce other problems:(1) It depletes an excessive amount of CPU capability and memory power of the client during the encryption and decryption; (2) The CSP cannot determine which emails contain keywords specified by a user if the encryption is not searchable, and can only return all the encrypted emails. Generally speaking, a thin client has only limited bandwidth, CPU, and memory; therefore, a simple encryption scheme cannot work well under these circumstances. We propose the SPKS scheme for cloud storage services to unravel the above problem. Its contributions are threefold: It is efficient and practical. The SPKS scheme enables CSPs to participate in the partial decipherment so as to reduce com- putational overhead on users, without leaking any information about the plaintext. It supports keyword searching on encrypted data. The SPKS scheme enables the CSP to work out whether a given email contains certain keywords specified by auser, but isn't conscious of any information about both the keywords and the email.It is a secure scheme. The flow of SPKS is illustrated in Fig 1. First, KeyGen used to generate a public/private key pair. Second, EMBEnc &KWEnc encrypts all the content within the file and keywords are encrypted respectively which then stored in the server. Third, Tcompute used in the retrieving phase where user generates a trapdoor and pass it to CSP. Fourth, KWtest checks whether the keyword contain in the encrypted data. Fifth, PDecrypt mainly for CSP to decrypt the intermediate result partly and sends the cipher text and the partial decrypted content. Sixth, Recovery runs by the user to decrypt the plain text. Therefore it provides semantic security in plain text attack.

III. CONCLUSION

In this paper analysis is made on encryption techniques which relate to search based retrieval of files from the outsourced encrypted data. Many searchable encryption schemes have been analyzed based on single keyword and multi-keyword search. Many disadvantages have been focused on these techniques since they rely on Boolean expressions. It has two major draw backs: 1) User has to decrypt every file that contains the keyword to match their file. 2) Since all the files containing keyword are retrieved. This will lead to network traffic.

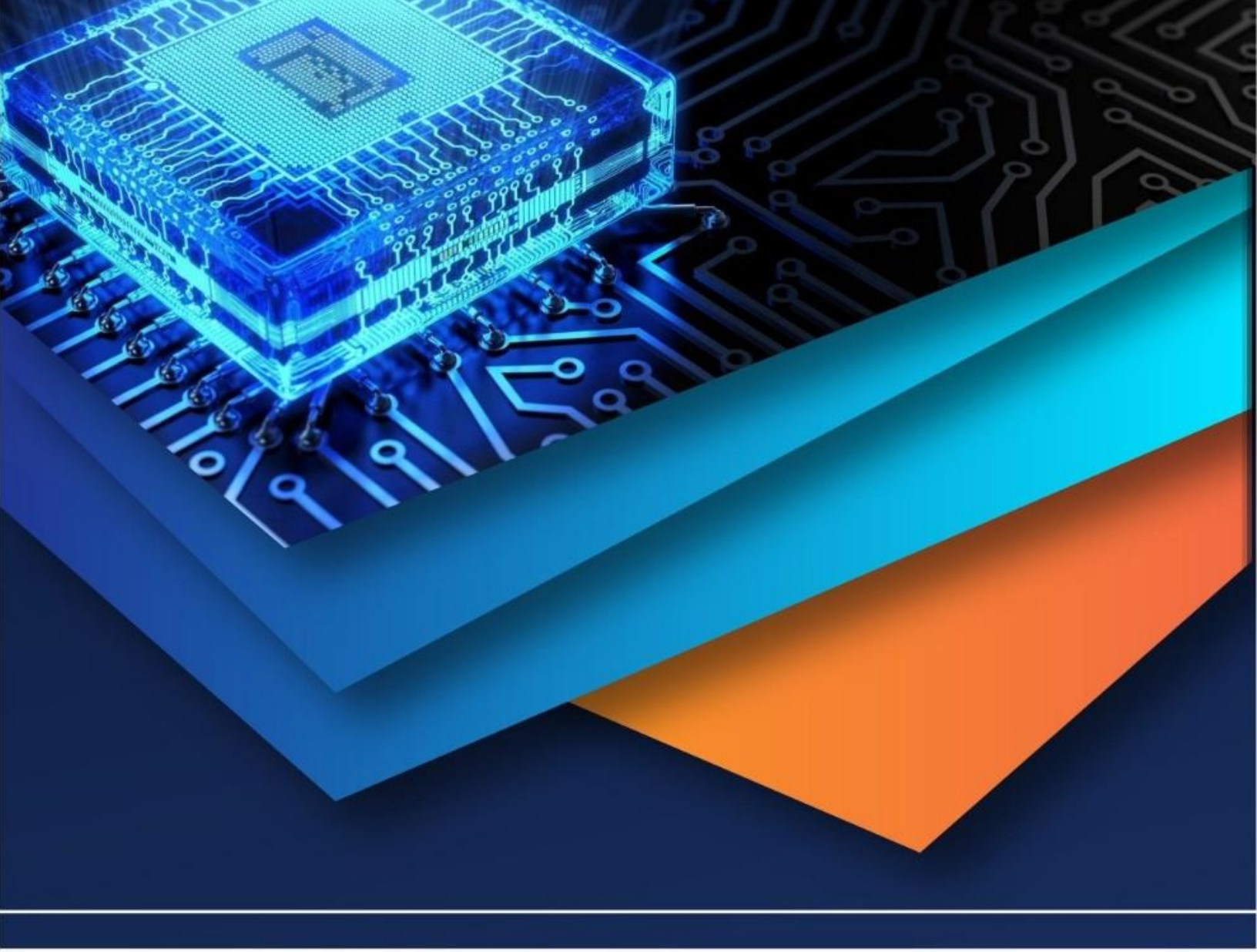
Therefore rank based retrieval of data has been discussed which proves the data security, fast search access and does not leak information to untrusted authorities. It is found that multi-keyword rank based retrieval is the most efficient for searching on encrypted data because it greatly enhances system usability by returning the matching files in a ranked order. It reduces the communication overhead and improves user searching capabilities by using multi-keyword search.

IV. ACKNOWLEDGEMENT

I would like to express my special thanks of gratitude to my guide Prof. A. B Gadewar sir, P.D.E.A. college of engineering, manjari as well as my HOD Prof. N. R. Jain Mam who gave us golden opportunity to do this wonderful project on the topic of restaurant table reservation system.

REFERENCES

[1] D.Song, D. Wagner and A. Perrig, Practical techniques for searches on encrypted data. Proceeding of the 2000 IEEE Symposium on Security and Privacy, May 14-17, Washington, DC. USA., pp: 44-55, 2000.
[2] E.J.Goh, Secure Indexes. Technical Report. . http://eprint.iacr.org/2003/216,2003.
[3] D.Boneh, G.D. Crescenzo, R. Ostrovskyand G. Persiano, Public key encryption with keyword search. Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques, May 2-6, 2004, Interlaken, Switzerland, pp: 506-522.
[4] Liu, Q., G. Wang and J. Wu., An efficient privacy preserving keyword search scheme in cloud computing. Proceedings of the International Conference on Computational Science and Engineering, Volume 2, August 29-31, 2009, Vancouver, Canada, pp: 715-720.
[5] R.Curtmola, J.A. Garay, S. Kamara and R. Ostrovsky, Searchable symmetric encryption: Improved definitions and efficient constructions. Proceedings of the 13th ACM Conference on Computer and Communications Security, October 30-November 03, 2006, Alexandria, USA., pp: 79-88.
[6] D.Boneh and B. Waters, 2007.Conjunctive, subset and range queries on encrypted data. Proceedings of the 4th Theory of Cryptography Conference, February 21-24, 2007, Amsterdam, The Netherlands, pp: 535-554.
[7] Zhu, B.and K. Ren. PEKSrand: Providing predicate privacy in public-key encryption with keyword search. Proceedings of the IEEE International Conference on Communications, June 5-9, 2011, Kyoto, Japan, pp: 1-6.
[8] Y.C.Chang, and M. Mitzenmacher, Privacy preserving keyword searches on remote encrypted data. Proceedings of the 3rd International Conference on Applied Cryptography and Network Security, June 7-10, 2005, New York, USA.,pp: 442-455.
[9] Liu, Q., G. Wang and J. Wu. Secure and privacy preserving keyword searching for cloud storage services. Journal on Network Computer Applications, 2012, Volume 35, pp: 927-933.
[10] Li, M., S. Yu, N. Cao and W. Lou. Authorized private keyword search over encrypted data in cloud computing. Proceedings of the 31st International Conference on Distributed Computing Systems, June 20-24, 2011, Minneapolis, MN., USA., pp: 383-392.
[11] Li, J., Q. Wang, C. Wang, N. Cao, K. Ren and W. Lou. Fuzzy keyword search over encrypted data in cloud computing. Proceedings of the 29th IEEE International Conference on Computer Communications, March 15-19, 2010, San Diego, CA., USA., pp: 1-5.
[12] Wang, C., N. Cao, K. Ren and W. Lou, 2012.Enabling secure and efficient ranked keyword search over outsourced cloud data. IEEE Transaction on Parallel and Distributed System, Volume 23, pp:1467-1479.
[13] V.Goyal, O. Pandey,A. Sahai, and B.Waters. Attribute-based encryption for fine-grained access control of encrypted data. Proc. ACM Conf. Computer and Communications Security, 2006, pp: 89–98.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY