ISSN: 0975-766X
CODEN: IJPTFI
Available Online through *Review Article*
www.ijptonline.com

EFFICIENT FILE RETREIVEL FROM CLOUD SERVER USING MULTI-SEARCH

[1*]Thushar.V.Jacob,[2] S.Nirmal Sam
[1]PG student, CSE department, SRM University, Kattankulathur.
[2]Assistant Professor, CSE department, SRM University, Kattankulathur.
Email: thushar1@hotmail.com

Received on 05-01-2017 *Accepted on: 12-02-2017*

Abstract:

Huge number of information proprietors are moved our information into cloud servers. Cloud information proprietors like to outsource archives in an encoded shape with the end goal of protection safeguarding. Hence it is vital to create proficient and dependable cipher text seek procedures. One test is that the relationship between archives will be ordinarily hidden during the time spent encryption, which will prompt critical hunt exactness execution corruption. They are all get to the information from cloud utilized the catchphrase based inquiry. Approach bunches the records Based on the base importance edge, and after that parcels the subsequent groups into sub-groups until the imperative on the most extreme size of bunch is come to. Here we proposed the safe multi catchphrase positioned look from the encoded information from cloud. It opens operations like upgrade, erase, and addition of archives. Here utilizing tree structure and shapeless scan strategy for recover the information from cloud. These sorts of strategy used to take care of the issue of watchword speculating assault. Here we proposed the Blowfish system for the encryption procedure. Here to diminish measurable assaults, apparition terms are added to the record vector for blinding list items. The proposed plan can accomplish linear search, semantic search, K gram1 and K gram2 searches and the query item like number of record recovery additionally manages erasure and inclusion of reports adaptable.

Keywords: Cloud search, Multi – Keyword, ciphertext search, ranked search.

1. Introduction

As Cloud Computing gets to be predominant, more touchy data are being concentrated into the cloud, for example, messages, individual wellbeing records, government archives, and so forth. By putting away their information into the cloud, the information proprietors can be calmed from the weight of information stockpiling and upkeep in order to

appreciate the on-request fantastic information stockpiling administration. Nonetheless, the way that information proprietors and cloud server are not in similar trusted area may put the outsourced information at hazard, as the cloud server may never again be completely trusted. It takes after that touchy information typically ought to be scrambled preceding outsourcing for information security and battling spontaneous gets to. In any case, information encryption makes successful information usage an extremely difficult errand given that there could be a lot of outsourced information records. Also, in Cloud Computing, information proprietors may impart their outsourced information to a substantial number of clients. A standout amongst the most well known courses is to specifically recover documents through keyword based hunt as opposed to recovering all the scrambled records back which is totally unfeasible in distributed computing situations. . In spite of the fact that encryption of keywords can ensure watchword security, it assist renders the customary plaintext seek strategies pointless in this situation.

We concentrate on empowering powerful yet privacy preserving fuzzy keyword seek in Cloud Computing. To the best of our understanding, we formalize curiously the issue of feasible soft watchword look for over encoded cloud data while keeping up keyword security. Soft watchword look fundamentally enhances system usability by giving back the planning records when customers' looking for sources of info decisively organize the predefined keywords or the closest possible organizing archives based on keyword similarity semantics, when exact match fails. More specifically, we use edit distance to quantify keywords similarity and develop a novel technique, i.e., an wildcard-based technique, for the construction of fuzzy keyword sets. This technique eliminates the need for enumerating all the fuzzy keywords and the resulted size of the fuzzy keyword sets is significantly reduced. In light of the built fuzzy watchword sets, we propose an effective fuzzy keyword seek conspire. Through thorough security investigation, we demonstrate that the proposed arrangement is secure and protection safeguarding, while successfully understanding the target of fleecy keyword look for. Segment presents the framework display, risk show, our outline objective and quickly depicts some vital foundation for the strategies utilized as a part of this paper. Area IV demonstrates a direct development of fuzzy watchword seek plot. Area V gives the point by point portrayal of our proposed plans, including the proficient developments of fuzzy watchword set and fuzzy keyword seek plot.

Vast number of information proprietors are moved our information into cloud servers. Cloud information proprietors want to outsource records in an encoded shape with the end goal of protection saving. Subsequently it is vital to create effective and dependable ciphertext look systems. One test is that the relationship between reports will be ordinarily

hidden during the time spent encryption, which will prompt huge inquiry exactness execution debasement. They are all get to the information from cloud utilized the keyword based inquiry. Approach bunches the reports Based on the base significance limit, and after that parcels the subsequent groups into sub-groups until the limitation on the greatest size of group is come to. Here we proposed the safe multi watchword positioned seek from the scrambled information from cloud. It open operations like upgrade, erase, addition of records. Here utilizing tree structure and undefined hunt technique down recover the information from cloud. These sorts of procedure used to take care of the issue of watchword speculating assault. Here we proposed the AES method for the encryption procedure. Here to decrease measurable assaults, apparition terms are added to the list vector for blinding query items. The proposed plan can accomplish sub-direct inquiry time and the query item like number of record recovery additionally manages erasure and inclusion of archives adaptably.

2. Literature survey

1. A. Swaminathan,Y. Mao, G.-M. Su,H. Gou,A. Varna,S. He,M. Wu,and D. Oard in 2007, To build up a structure for confidentiality protecting rank-requested hunt in substantial scale archive accumulations. We investigate methods to safely rank-arrange the records and concentrate the most significant document(s) from a scrambled gathering in light of the encoded seek inquiries. We show useful procedures for legitimate combination of pertinence scoring strategies and cryptographic systems, for example, arrange saving encryption, to ensure information accumulations and records and give efficient and exact inquiry capacities to safely rank-arrange archives in light of a question. Exploratory results on the W3C accumulation demonstrate that these strategies have similar execution to customary scan frameworks intended for non-scrambled information regarding seek exactness.

2. Cong Wang, Ning Cao, Kui Ren, and Wenjing Lou in 2013. Describes and deals with the issue of secure situated watchword investigates encoded cloud data. Situated look for uncommonly enhances system convenience by engaging thing criticalness situating rather than sending undifferentiated results, and further ensures the record recuperation accuracy. Specifically, we research the real measure approach, i.e., significance score, from information recuperation to produce a safe searchable record, and develop a one-to-various demand sparing mapping methodology to fittingly guarantee those sensitive score information. The resulting arrangement can support profitable server-side situating without losing catchphrase security. Passed on preparing fiscally empowers the view point of information association outsourcing. Regardless, to secure information confirmation, delicate cloud information must be blended before

outsourced to the business open cloud, which makes productive information usage advantage a particularly troublesome undertaking. In any case, standard searchable encryption procedures permit clients to safely explore encoded information through catchphrases, they bolster just Boolean pursue and are not yet adequate to meet the reasonable information use require that is regularly requested by liberal number of clients and gigantic measure of information records in cloud.

3. Dawn Xiaodong Song David Wagner Adrian Perrig in 2000, Our procedures have various essential preferences. They are provably secure: they give provable mystery to encryption, as in the untrusted individual from staff serving at table can't take in regardless of which about the plaintext when just given the ciphertext. They give question seclusion to inquiries, implying that the untrusted server can't learn much else about the plaintext than the query output; they give controlled looking, so that the untrusted server can't hunt down a subjective word without the client's approval. They likewise bolster shrouded questions, so that the client may approach the untrusted server to scan for a mystery word without uncovering the word to the server. It is appealing to store in succession on data collection servers, for instance, mail servers and file servers fit as a fiddle to decrease security and assurance threats. But this usually implies that one has to sacrifice functionality for security .

4. Dan Boneh, Giovanni Di Crescenzo, Rafail Ostrovsky,Giuseppe Persiano in 2004, A mail server that stores different messages freely scrambled for Alice by others. Utilizing our component Alice can send the mail server a key that will empower the server to recognize all messages containing some specific watchword, yet learn nothing else. We define the idea of open key encryption with catchphrase hunt and give a few developments. The issue of looking on data that is mixed using an open key system. Consider customer Bob who sends email to customer Alice encoded under Alice's open key. An email entryway needs to test whether the email contains the catchphrase "squeezing" with the target that it perhaps will course the electronic message in like way. Alice, on the other hand does not wish to give the entryway the ability to unscramble each one of her messages.

5. Cong Wang, Ning Cao, Jin Li, Kui Ren , and Wenjing Lou in 2010, To define and solve the problem of effective yet secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain relevance criteria. we motivate and solve the problem of supporting efficient ranked keyword search for achieving effective utilization of remotely stored encrypted data in Cloud Computing. This approach suffers from two main drawbacks when directly applied in the context of Cloud Computing. On the one hand, users, who do not necessarily have pre-knowledge of the encrypted cloud data, have to postprocess

every retrieved file in order to find ones most matching their interest; On the other hand, invariably retrieving all files containing the queried keyword further incurs unnecessary network traffic,which is absolutely undesirable in today's pay-as-you-use cloud paradigm.

Algorithm

Blowfish Algorithm

An encryption algorithm plays an important role in securing the data in storing or transferring it. The encryption algorithms are categorized into Symmetric (secret) and Asymmetric (public) keys encryption.

In Symmetric key encryption or secret key encryption, only one key is used for both encryption and decryption of data.

Eg: Data encryption standard(DES), Triple DES, Advanced Encryption Standard(AES) and Blowfish Encryption Algorithm

In asymmetric key encryption or public key encryption uses two keys, one for encryption and other for decryption.

Eg: RSA

K gram Algorithm

We will use the k-gram index to retrieve vocabulary terms that have many k-grams in common with the query. We will argue that for reasonable definitions of many k-grams in common," the retrieval process is essentially that of a single scan through the postings for the k-grams in the query string –q. Once we retrieve such terms, we can then find the ones of least edit distance from –q.

K gram:

Enumerate all k-grams in the query term.

Example: bigram index, misspelled word boardroom.

Bigrams: bo, or, rd, dr, ro, oo, om.

Use the k-gram index to retrieve "correct" words that match query term kgrams.

Threshold by number of matching k-grams.

E.g., only vocabulary terms that differ by at most 3 k-grams.

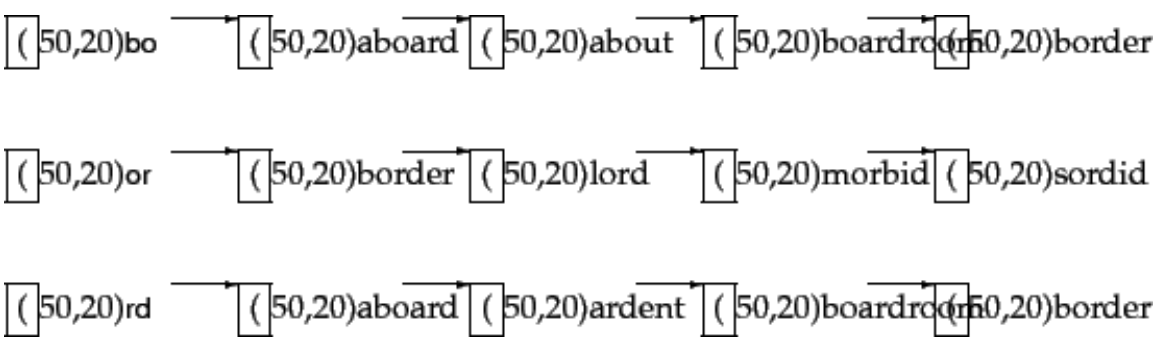
K-Gram Indices :

Heuristic:

If two words have many common kgrams, they may be similar to each other – If there are multiple candidates, find the

one with the least edit distance

Example: query “bord” – Suggest “border



Encryption Algorithm

Blowfish was designed in 1993 by Bruce Scheier as a fast, alternative to existing encryption algorithms such AES, DES and 3 DES etc.

Blowfish is a symmetric block encryption algorithm designed in consideration with,

Fast: It encrypts data on large 32-bit microprocessors at a rate of 26 clock cycles per byte.

Compact: It can run in less than 5K of memory.

Simple: It uses addition, XOR, lookup table with 32-bit operands.

Secure: The key length is variable, it can be in the range of 32~448 bits: default 128 bits key length.

It is suitable for applications where the key does not change often, like communication link or an automatic file encryption.

Unpatented and royalty-free.

Description

Blowfish symmetric block cipher algorithm encrypts block data of 64-bits at a time.

Key-expansion:

It will convert a key of at most 448 bits into several sub-key arrays totaling 4168 bytes. Blowfish uses large number of sub-keys.

These keys are generating earlier to any data encryption or decryption.

The p-array consists of 18, 32-bit subkeys:

P1,P2,…….,P18

Four 32-bit S-Boxes consists of 256 entries each:

S1,0, S1,1,……. S1,255

S2,0, S2,1,…….. S2,255

S3,0, S3,1,…….. S3,255

S4,0, S4,1,.....S4,255

Conclusion

The first occasion when we formalize and tackle the issue of supporting productive yet protection saving fuzzy hunt down accomplishing powerful usage of remotely put away, scrambled information in Cloud Computing. We plan a propelled procedure (i.e., trump card based system) to develop the capacity proficient fuzzy keyword sets by abusing a critical perception on the closeness metric of alter separation. In view of the built fuzzy watchword sets, we promote propose a productive fuzzy keyword seek plot. Through thorough security investigation, we demonstrate that our proposed arrangement is secure and protection saving, while precisely understanding the target of feathery keyword look.

Reference

1. A.Swaminathan, Y. Mao, G. M. Su, H. Gou, A. Varna, S. He, M. Wu, and D. Oard, "Confidentiality-preserving rank-ordered search," in Proc. ACM ACM Workshop Storage Security Survivability, Alexandria, VA, 2007, pp. 7–12.
2. C. Wang, N. Cao, K. Ren, and W. J. 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.
3. D. X. D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE Symp. Security Priv., BERKELEY, CA, 2000, pp. 44–55.
4. D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. EUROCRYPT, Interlaken, SWITZERLAND, 2004, pp. 506–522.
5. C. Wang, N. Cao, J. Li, K. Ren, and W. J. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. IEEE 30th Int. Conf. Distrib. Comput. Syst., Genova, ITALY, 2010, pp. 253–262.
6. M. Bellare, A. Boldyreva, and A. O'Neill, "Deterministic and efficiently searchable encryption," in Proc. 27th Annu. Int. Cryptol. Conf. Adv. Cryptol., Santa Barbara, CA, 2007, pp. 535–552.
7. D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in Proc. 4th Conf. Theory

Cryptography, Amsterdam, NETHERLANDS, 2007, pp. 535–554.

8. E.-J. Goh, Secure Indexes, IACR Cryptology ePrint Archive, vol. 2003, pp. 216. 2003.
9. C. Wang, N. Cao, K. Ren, and W. J. 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.
10. A. Swaminathan, Y. Mao, G. M. Su, H. Gou, A. Varna, S. He, M. Wu, and D. Oard, "Confidentiality-preserving rank-ordered search," in Proc. ACM ACM Workshop Storage Security Survivability, Alexandria, VA, 2007, pp. 7–12.
11. S. Zerr, D. Olmedilla, W. Nejdl, and W. Siberski, "Zerber+R: Topk retrieval from a confidential index," in Proc. 12th Int. Conf. Extending Database Technol.: Adv. Database Technol., Saint Petersburg, Russia, 2009, pp. 439–449.
12. C. Wang, N. Cao, J. Li, K. Ren, and W. J. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. IEEE 30th Int. Conf. Distrib. Comput. Syst., enova, ITALY, 2010, pp. 253–262.
13. P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Proc. Proc. 2nd Int. Conf. Appl. Cryptography Netw. Security, Yellow Mt, China, 2004, pp. 31–45.
14. L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in Proc. 7th Int. Conf. Inform. Commun. Security, Beijing, China, 2005, pp. 414–426.
15. R. Brinkman, "Searching in encrypted data" in University of Twente, PhD thesis, 2007.
16. Y. H. Hwang and P. J. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi-user system," in Proc. 1st Int. Conf. Pairing-Based Cryptography, Tokyo, JAPAN, 2007, pp. 2–22.
17. H. Pang, J. Shen, and R. Krishnan, "Privacy-preserving similaritybased text retrieval," ACM Trans. Internet Technol., vol. 10, no. 1, pp. 39, Feb. 2010.
18. N. Cao, C. Wang, M. Li, K. Ren, and W. J. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," in Proc. IEEE INFOCOM, Shanghai, China, 2011, pp. 829–837.
19. W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li, "Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking," in Proc. 8th ACM SIGSAC Symp. Inform., Comput. Commun. Security, Hangzhou, China, 2013, pp. 71–82.
20. F. Li, M. Hadjieleftheriou, G. Kollios, and L. Reyzin, "Dynamic authenticated index structures for outsourced databases," in Proc. ACM SIGMOD, Chicago, IL, 2006, pp. 121–132.

21. H. H. Pang and K. L. Tan, "Authenticating query results in edge computing," in Proc. 20th Int. Conf. Data Eng., Boston, MA , 2004, pp. 560–571.
22. C. Martel, G. Nuckolls, P. Devanbu, M. Gertz, A. Kwong, and S. G. Stubblebine, "A general model for authenticated data structures," Algorithmica, vol. 39, no. 1, pp. 21–41, May 2004.
23. C. M. Ralph, "Protocols for public key cryptosystems," in Proc. IEEE Symp. Security Priv, Oakland, CA, 1980, pp. 122–122.
24. R. C. Merkle, "A certified digital signature," in Proc. Adv. cryptol. , 1990, vol. 435, pp. 218–238.
25. M. Naor and K. Nissim, "Certificate revocation and certificate update," IEEE J. Sel. Areas Commun., vol. 18, no. 4, pp. 561–570, Apr. 2000.

Corresponding Author:

Thushar.V.Jacob *,

Email: *thushar1@hotmail.com*