Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 4, April -2017

Document Retrieve On Cloud With Multikeyword Searching

Searching document on Cloud

Shreya Thite[1],Monali Yenpure[2], Surabhi Mohite[3]

[1]*Computer Engineering, NESGOI Of Naigaon Pune*
[2]*Computer Engineering, NESGOI Of Naigaon Pune*
[3]*Computer Engineering, NESGOI Of Naigaon Pune*

Abstract —*Cloud storage contain large amount of outsourced data. For large data require to provide security to store the data on cloud storage. large data cannot be handledby the user for searching document on cloud. The cloud provides data storage and sharing services for users, and has ample storage space. The third party auditor is able to verify the integrity of shared data based on requests from users, without downloading the entire data. To find the specific data on cloud is critical task . Finding data in large data is the big challenge in this case. So in this paper, we introduce to 'Multikeyword Searching' strategy . In multikeyword searching we can find data from cloud using date, Location or keyword if mention in the document. It is easiest technique of finding data on cloud. Also this technique increases the overall performance of the system. It require less time to find document on cloud. We provide integrity and confidentiality to the data while it is stored on cloud.*

Keywords-*Cloud computing, Multikeyword searching, Privacy preserving, Public auditing, Third party auditor etc.*

I. INTRODUCTION

The cloud can be formidable and costly for the users .The overhead of using cloud storage should be minimized as much as possible such that a user does not need to perform so many operations to their outsourced data. Cloud stores large size of data. So user cannot directly retrieve data . Therefore it have chances to reduce performance of the system. So we have come up with a solution that this processing will be done on cloud. Despite of significant improvement in cloud storage ' computing capabilities, still computing requirements of cloud users, especially enterprise users, is not achieved.

Users no longer have physical possession of the possibly large size of outsourced data makes the data integrity protection in Cloud Computing a very challenging and potentially formidable task, especially for users with constrained computing resources and capabilities. Thus, enabling public auditability for cloud data storage security is of critical importance so that users can resort to an external audit party to check the integrity of outsourced data when needed. Instead of directly applying the old public auditing scheme to the multi-server setting. On cloud storage, it is also difficult to search the data as per user require. So in this paper we introduce the 'multikeyword searching' technique for searching various type of files. The files can be search by the keyword, location or date , which are mention in data. This technique is helpful to increase the system performance and efficiency.

II. LITERATURE SURVEY

A literature review or study forms an essential part for studing and searching on previouse topic:

1. **Jian Liu, Kun Huang, Hong Rong, Huimei Wang, Ming Xian, *'Privacy-Preserving Public Auditing for Regenerating-Code-Based Cloud Storage '*, IEEE Network , 2015.**

In this paper, authors propose a public auditing scheme for the regenerating-code-based cloud storage. To solve the regeneration problem of failed authenticators in the absence of data owners, they introduce a proxy, which is privileged to regenerate the authenticators, into the traditional public auditing system model.

2. **Priyanka Dehariya, Prof. Shweta Shrivastava, Dr. Vineet Richaraya, *'Surveying Cloud Storage Correctness using TPA with BLS '*, International Journal of Engineering Research and General Science Volume 3, Issue 1, January-February, 2015.**

Cloud Computing is emerging technology and gaining remarkable popularity in the recent years for its benefits in terms of flexibility, scalability, reliability, effciency and cost efectiveness. Despite of all these benefits, Cloud Computing has one problem: Security, they analyze the problems of data security in cloud storage, which is a distributed storage system. An effective and flexible scheme is proposed in our paper to ensure the integrity and correctness of the data stored in cloud server.

3. Surendra Singh, Rathod Anand Rajawat, *'The Research on Cloud Server Storage Security Using TPA '*, Volume 5, Issue 7, July 2015.

In this paper, they propose a privacy-preserving public auditing system for data storage and security in Cloud Computing. they utilize the homomorphism authenticator and random masking to assurance that TPA would not learn any information about the data content stored on the cloud server during the effcient auditing process.

4. CongWang, QianWang, Kui Ren, Wenjing Lou, *'Privacy-Preserving Public Auditing for Secure Cloud Storage '*, IEEE 2014.

In this paper, authors propose a privacy-preserving public auditing system for data storage security in Cloud Computing. They utilize the homomorphic linear authenticator and random masking to guarantee that the TPA would not learn any knowledge about the data content stored on the cloud server during the efficient auditing process, which not only eliminates the burden of cloud user from the tedious and possibly expensive auditing task, but also alleviates the users fear of their outsourced data leakage.

III. PROPOSED SYSTEM

Every coin has two side, and cloud computing is no exception. There is criticism about privacy in cloud model, because of the fact that administrator have access to data stored in the cloud. They can unintentionally or intentionally access the client data. Traditional security or protection techniques need a reconsideration for cloud. Except for private cloud where organization does not have control over the equipment, the progress of cloud is seems little slow, because organizations think instead of compromising on the security of the data, they are still willing to invest in buying private equipment to setup their own infrastructure. Security issues which are of concern to the client can be classified into sensitive data access, data segregation, bug exploitation, recovery, accountability, malicious insiders and account control issues. Like different disease have different medicines, different cloud security issues have different solutions, like cryptography, use of more than one cloud provider, strong service level agreement between client and cloud service provider. Heavy investment is needed to secure the compromising data in cloud. Cloud can grow only if it is possible to build a trust in client, and which can be built only if security concerns are being addressed. Generally the cloud services are browser based, therefore any browser enabled device such as for instance laptop, desktop, smartphone, tablets can used to gain access to these services, the services at providers end may be hosted on any platform, from windows, Linux, etc. which are accessible via internet.

IV. SYSTEM ARCHITECTURE

In this paper, we consider data storage and sharing services in the cloud with three entities: the cloud, the third party auditor (TPA), and users who participate as a group (as shown in Fig. 1). Users in a group include one original user and a number of group users. The original user is the original owner of data, and shares data in the cloud with other users. Based on access control policies, other users in the group are able to access, download and modify shared data. The clouprovides data storage and sharing services for users, and has ample storage space. The third party auditor is able to verifythe integrity of shared data based on requests from users, without downloading the entire data. When a user (either the original user or a group user) wishes to check the integrity of shared data, she first sends an auditing request to the TPA. After receiving the auditing request, the TPA generates an auditing message to the cloud, and retrieves an auditing proof shared data from the cloud. Then the TPA verifies the correctness of the auditing proof. Finally, the TPA sends an auditing report to the user based on the result of the verification.

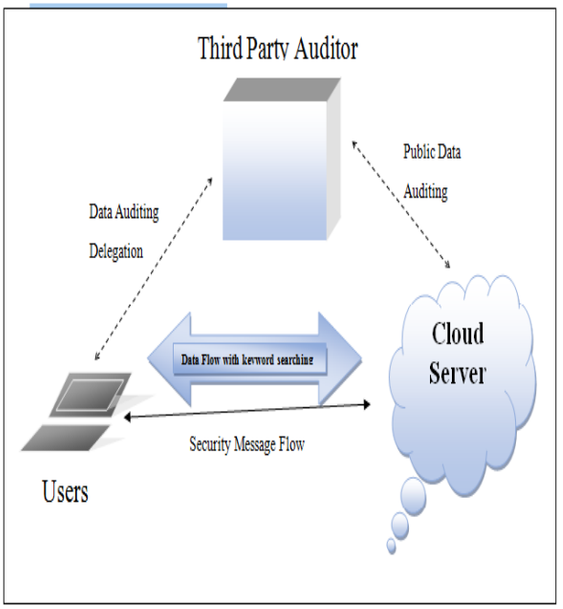

Fig 1: System Architecture

In this architecture , user can use multikeyword for searching file on cloud. User can use lacation name, date and keyword to find file on cloud. It increases the system performance and efficiency. User can register on cloud to get authority to preserve security on cloud. Once user can register on cloud then it can login in cloud using it's own account for store or download data. Users have two type in this architecture. First is author which can upload the data on cloud to store it, and second is user's which can download the data from cloud using multikeyword strategy. Using this technique user require less time to gets it's result. In this paper, we used MS-SQL database for store the data. When user register on cloud then information about user is store on MS-SQL database.

IV. ALGORITHM

In this paper **,** we use binary vector generation algorithm for view the data without extraction of data.Algorithm is given below:

Input : Each client C_m has an binary vector $b_m \in Z_2^n$, $1 \le m \le M$.

Output: $b = T_t(b_1,..., b_M)$.

1: Each C_m selects M random share vectors $b_{m,l} \in Z_{M+1}^n$, $1 \le l \le M$, such that $\sum_{l=1}^{M} b_{m,l} = b_m \bmod(M+1)$.

2: Each C_m sends $b_{m,l}$ for all $1 \le l \ne m \le M$.

3: Each C_l computes $s_l = (s_l(1),\ldots,s_l(n)) = \sum_{m=1}^{M} b_{m,l} \bmod(M+1)$.

4: Clients C_l, $2 \le l \le M-1$, send s_l to C_1.

5: C_1 computes $s = (s_l(1),\ldots,s_n(n)) = \sum_{l=1}^{M} s_l \bmod(M+1)$.

6: **for** $i = 1,\ldots,n$ **do**

7: If $(s(i) + s_M(i) \bmod(M+1)) < t$ set $b(i) = 0$ Otherwise set $b(i) = 1$.

8: **end for**

9: Output $b = (b(1),..., b(n))$.

VI. PERFORMANCE ANALYSIS

We focus on evaluating performance of our document retrieve on cloud with multikeyword searching setup, audit and repair procedure in our experiment all codes are written in .net in Visual Studio 2.10 platform on windows. All entities in our prototype in intel core i3 2450 2.5 GHz, 4 GB RAM, 7200 RPM Hitachi 5000 SATA Drive.

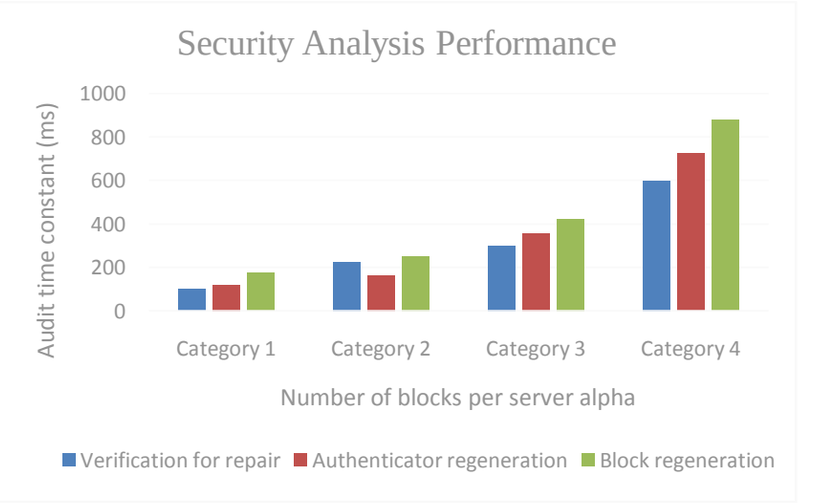

Fig 2: Performance analysis

VII. RESULT

This paper , is based on multikeyword searching and also we provide security to this system. We compare our system with previous system and calculate computational time. Our system require less computational time as compaire to existing system. Proposed system uses keyword to retrieve data , therefore its require less time than the exiisting system which can shown in bellow:

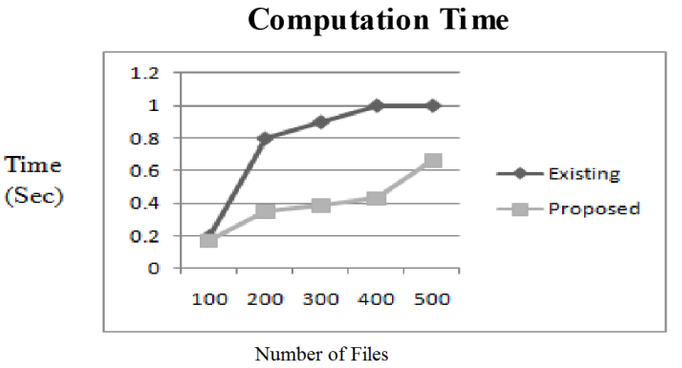

Fig 3: Result of computation time

VIII. ADVANTAGES

The advantages f proposed system is given below:

- **Usability :**Allows users to drag and drop files between Cloud storage and their local storage.
- **Bandwidth :**You can avoid emailing files to individuals and instead send a web link to recipients through your email.
- **Accessibility :**Storage files can be access from anywhere via internet connections.
- **Computation Time:**It require less computational time as compare to previous system. Because proposed system uses multikeyword searching technique.
- **Cost Reduce :**Business and organizations can often reduce an annual operating costs by using cloud storage; cloud storage costs about 3 cents per gigabyte to store data internally. Users can see additional cost savings because it does not require power to store information remotely.
-

IX. DISADVANTAGES

- **Network :** If you have no internet connection you've no access to your data.
- Several cloud storage have specific bandwidth allowance. If an organization surpasses the given allowance, the additional charges could be significant. However, some providers allows unlimited bandwidth. This is the factor that companies should consider when looking at cloud storage provider.

X. APPLICATIONS

1. Use in big data where daily millions of documents generated.
2. Use in hospitals.
3. Use in Government document search.

XI. CONCLUSION

The homomorphic linear authenticator and random number generation utilize to guarantee that the TPA would not learn any knowledge about the data content store on the cloud server during efficient auditing process,which not only eliminates the burden of cloud user from the tedious and a possibally expensive auditing task. Multikeyword searching technique is used in system. It improves the overall performance of the system and reduces the computation time as compare to existing system. It is highly efficient system. Also , this system provide seqcurity to the data to preserve from unauthorised person.

XII. FUTURE SCOPE

To enhance the security more, a mechanism to secure the keys in security cloud canbe an area of research. Also to reduce the overhead of network traffic can be another area of research.

XIII. ACKNOWLEDGEMET

This acknowledgment is intended to be thanks giving gesture to all thosepeople who have been involved directly or indirectly with our dissertation work. Firstand foremost, we express our special thanks with gratitude and great respect to ourvaluable Head of Department Prof. P. V. Mahadik, and our Project guide Prof. S. Autade for their keen interest, fruitful suggestions and valuable guidancewith motivation and constant encouragement. We are also thankful to their greatpatience, constructive criticism & useful suggestions apart from valuable guidance.We owe a great debt to our principal prof. R. J. Patil, who guided us through themaze of details associated with the seminar. Lastly we are deeply grateful to everyonewho has been associated with this seminar.

REFERENCES

[1] Jing-Jang Hwang, Hung-Kai Chuang,Yi-Chang Hsu, Chien-Hsing Wu, "A Business Model for Cloud Computing Based on a Separate Encryption and Decryption Service," Proceedings of the 2011 International Conference on Information Science and Application, April 2011.

[2] Qian Wang ,Cong Wang ,Kui Ren ,Wenjing Lou ,Jin Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 22, NO. 5, MAY 2011.

[3] Balachandra Reddy Kandukuri, Ramakrishna Paturi V, Dr. AtanuRakshit, "Cloud Security Issues", IEEE International Conference on Services Computing, pp. 517-520, September 2009.

[4] Peter Mell, Timothy Grance, "The NIST Denition of Cloud Computing", NISTSpecial Publication 800-145.

[5] Ling Li, Lin Xu, Jing Li, Changchun Zhang,"Study on the Third-party Audit in Cloud Storage Service", 2011 International Conference on Cloud and Service Computing.

[6] L. M. Vaquero,L. Rodero-Merino,J. Caceres, and M. Lindner, "A break in the clouds: towards a cloud de nition", ACM SIGCOMM Computer CommunicationReview, vol. 39, no. 1, pp. 50-55, January 2009.

[7] A. Parakh and S. Kak, "Online data storage using implicit security", InformationSciences, vol. 179, issue 19, pp. 3323-3333 ,September 2009.

[8] C. Weinhardt, A. Anandasivam, B. Blau, N. Borissov, T. Meinl, W.Michalk, andJ. Stober, "Cloud computing a classication, business models, and research di rections", Business & Information Systems Engineering (BISE), vol. 1, no. 5, pp.391-399, 2009.

[9] L. Lamport, "Password authentication with insecure communication", Communications of the ACM, vol. 24, no. 11, pp. 770-772, 1981.

[10] "Announcing the ADVANCED ENCRYPTION STANDARD (AES)", Federal Information Processing Standards Publication 197. United States National Institute of Standards and Technology (NIST). November 26, 2001. Retrieved October 2, 2012.

[11] William Stallings, "Cryptography and Network Security", 2009.

[12] Salesforce.com, Inc., "Force.com platform", Retrieved Dec. 2009, fromhttp://www.salesforce.com/tw/.

[13] http://www.cs.utsa.edu/ wagner/laws/AESintro.html.

[14] C. Verplaetse, Inertial Proprioceptive devices: Self-motion-sensing toys and tools, IBM SYSTEMS JOURNAL, VOL35, NOS 3&4, 1996.

[15] I.-K. Park, 1 –H. Kim, and K. –S. Hong, “An Implementation of an FPGA-Based Embedded Gesture Recognizer Using a Data Glove” Conference on Ubiquitous Information Management and Communications Proceeding of the second International Conference on Ubiquitous information management and communication 2008, Suwon, Korea, January 31 – February 01, 2008, pp.496-500.

[16] A. M. Khan and T. –s. Kim, Accelerometer Signal-Based Human Activity Recognization using Aumented Autoregressive Model coefficients and Artificial Neural Nets, “IEEE EMBC 2008, pp. 5172-5175.

[1] Oliver J. Woodman, An Introduction to Inertial Navigation”, Technical Report, University of Cambridge, 2007.

[2] Walid Abdel-Hamid, “Accuracy Enhancement of Integrated MEMS-IMU/ GPS Systems for Land Vehicular Navigation applications”, University of CAGARY, January 2005.

[3] TilakshanKanesalingam, Motion Tracking Glove for Human Machine Interaction : Inertial Guidance, Mc Master University, Hamilton, Ontario, Canada.

[4] Doug Vargha, Motion Processing Technology Driving New Innovations in customer Products, InvenSense.

[5] Tu X Y, Terzopoulos D. Artificial Fish :Phisics, locomotion, perception, behavior. In :Proceedingd of ACM SIGGRAPH 1994. Orlando, USA : ACM Press, 1994 : 43-50.

[6] Terzopoulos D, Rabie T F. Animat vision: Active vision in artificial animals. Journoul of Computer Vision Research, 1997, 1(1):2-19.\

