WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Optimizing Quality-Aware Big Data Applications in the Cloud

E. Yashini [#1] ***, A. Kalaivani***[#2]

[#1].M.Sc, Department of Computer Science, Kamban College of Arts and Science for Women, Tiruvannamalai-606603

[#2].Assistant professor, Department of Computer Science, Kamban College of Arts and Science for Women, Tiruvannamalai-606603

ABSTRACT:

Privacy of data in subjects of cloud computing or big data is one of the most principal issues. The privacy methods studied in previous research showed that privacy infringement for cloud computing or big data happened because multi risks on data by external or internal attackers. An important risk to take into consideration when speaking of the privacy of the stored transactions is represented by the transactions' information which is not in the owner's control. Such a case is represented by the cloud servers that are administered by cloud providers which cannot be wholly trusted by the users with sensitive, private data such as business plans or private information. A simple method for protecting data privacy is by applying certain privacy techniques onto transactions' data, followed by the upload of the modified data into the cloud. In this paper, we are proposing a case study that is built on levels containing three models: cloud's architecture, transaction's manager and clients. Moreover, we consider that our case study is based on the premise of zero trust among the three models, therefore all the transactions take place with third-parties and the data movements are realized going through various levels of security..

KEYWORDS: Multi Key , Cipher TExt ,AES

1.INTRODUCTION:

Loud storage [1] is an emerging model of storage to provide scalable, elastic and pay-as-you-use service to cloud computing users. For individual usage, the subscribers enjoy the freedom to access to their data anywhere, anytime with any device. When cloud storage [2] is utilized by a group of users, it allows team members to synchronize and manage all shared documents. Moreover, it also saves the user a lot of capital investment of expensive storage equipments [3]. Cloud delivers convenience to the customers and at the same time arouses many security and privacy problems [4], [5]. Since the data are physically stored on the multiple servers of the cloud service provider, the customers cannot fully in charge of their data. They worry about the privacy of the stored documents since the server may be intruded by hacker or the data could be misused by the internal staff for commercial purpose [6]. The customers prefer to adopt the encryption technology to protect the data confidentiality, which meanwhile arouses another problem: how to execute data retrieval on the large volume of ciphertext. It is almost Y. Yang is with College of Mathematics and Computer Science, Fuzhou University, Fuzhou, China; Fujian Provincial Key Laboratory of Network Computing and Intelligent Information Processing, Fuzhou University, China; Key Laboratory of Spatial Data Mining & Information Sharing, Ministry of Education, Fuzhou, China; University Key Laboratory of Information Security of Network Systems (Fuzhou University), Fujian Province, China; Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University), Fuzhou, China. E-mail: yang.yang.research@gmail.com. X. Zheng and W. Guo are with College of Mathematics and Computer Science, Fuzhou University, Fuzhou, China; Fujian Provincial Key Laboratory of Network Computing and Intelligent Information Processing, Fuzhou University, China; Key Laboratory of Spatial Data Mining & Information Sharing, Ministry of Education, Fuzhou, China. E-mail: xianghan.zheng@fzu.edu.cn, guowenzhong@fzu.edu.cn. C. Rong is with Department of Electronic Engineering and Computer Science, University of Stavanger, Norway. E-mail: chunming.rong@uis.no. Corresponding authors: XianghanZheng, WenzhongGuo. unimaginable to ask the cloud subscriber to download all of their stored information and then decrypt and search on the recovered plaintext documents. No customer could tolerate the huge transmission overhead and the waiting time for the data retrieval result. Searchable encryption technology [7], [8], [9] not only exerts encryption protection of the data, but also supports efficient search function without undermining the data privacy. The data user generates a token of the content that he wants to search using his private key. Receiving the token, the cloud server searches on the encrypted data without decrypting the ciphertext. The most important point is that the server learns nothing about the plaintext of the encrypted data nor the searched content during the data retrieval procedure. However, most of the available searchable encryption schemes only support some basic search patterns, such as single keyword search, conjunctive keyword search and boolean search. Since the cloud computing is a fierce competition industry, it is of vital importance to

provide good user experience. It is urgent to design novel searchable encryption schemes with expressive search pattern for cloud storage.

While it is exciting to have convenient PHR services for everyone, there are many security and privacy risks which could impede its wide adoption. The main concern is about whether the patients could actually control the sharing of their sensitive personal health information (PHI), especially when they are stored on a third-party server which people may not fully trust.

In particular, the majority professional users are managed distributive by attribute authorities in the former, while each owner only needs to manage the keys of a small number of users in her personal domain. In this way, our framework can simultaneously handle different types of PHR sharing applications' requirements, while incurring minimal key management overhead for both owners and users in the system. In the public domain, we use multi authority ABE (MA-ABE) to improve the security and avoid eyes crow problem. Each attribute authority (AA) in it governs a disjoint subset of user role attributes, while none of them alone is able to control the security of the whole system.

We provide a thorough analysis of the complexity and scalability of our proposed secure PHR sharing solution, in terms of multiple metrics in computation, communication, storage, and key management. We also compare our scheme to several previous ones in complexity, scalability and security. Furthermore, we demonstrate the efficiency of our scheme by implementing it on a modern workstation and performing experiments/simulations.

2. LITERATURE SURVEY

[1] 2016**, Joseph K. Liu, Kaitai Liang, Willy Susilo,** In this paper, we propose a two-factor data security protection mechanism with factor revocability for cloud storage system. Our system allows a sender to send an encrypted message to a receiver through a cloud storage server. The sender only needs to know the identity of the receiver but no other information (such as its public key or its certificate). The receiver needs to possess two things in order to decrypt the ciphertext. The first thing is his/her secret key stored in the computer. The second thing is a unique personal security device which connects to the computer. It is impossible to decrypt the ciphertext without either piece. More importantly, once the security device is stolen or lost, this device is revoked. It cannot be used to decrypt any ciphertext. This can be done by the cloud server which will immediately execute some algorithms to change the existing ciphertext to be un-decryptable by this device. This process is completely transparent to the sender. Furthermore, the cloud server cannot decrypt any ciphertext at any time. The security and efficiency analysis show that our system is not only secure but also practical.

[2] 2017, **Qingchen Zhang, Laurence T. Yang, Zhikui Chen, and Peng Li,** As one important technique of fuzzy clustering in data mining and pattern recognition, the possibilistic c-means algorithm (PCM) has been widely used in image analysis and knowledge discovery. However, it is difficult for PCM to produce a good result for clustering big data, especially for heterogenous data, since it is initially designed for only small structured dataset. To tackle this problem, the paper proposes a high-order PCM algorithm (HOPCM) for big data clustering by optimizing the objective function in the tensor space. Further, we design a distributed HOPCM method based on MapReduce for very large amounts of heterogeneous data. Finally, we devise a privacy-preserving HOPCM algorithm (PPHOPCM) to protect the private data on cloud by applying the BGV encryption scheme to HOPCM, In PPHOPCM, the functions for updating the membership matrix and clustering centers are approximated as polynomial functions to support the secure computing of the BGV scheme. Experimental results indicate that PPHOPCM can effectively cluster a large number of heterogeneous data using cloud computing without disclosure of private data.

[3] 2014, **QingjiZheng† ShouhuaiXu† Giuseppe Ateniese .**It is common nowadays for data owners to outsource their data to the cloud. Since the cloud cannot be fully trusted, the outsourced data should be encrypted. This however brings a range of problems, such as: How should a data owner grant search capabilities to the data users? How can the authorized data users search over a data owner's outsourced encrypted data? How can the data users be assured that the cloud faithfully executed the search operations on their behalf? Motivated by these questions, we propose a novel cryptographic solution, called verifiable attribute-based keyword search (VABKS). The solution allows a data user, whose credentials satisfy a data owner's access control policy, to (i) search over the data owner's outsourced encrypted data, (ii) outsource the tedious search operations to the cloud, and (iii) verify whether the cloud has faithfully executed the search operations. We formally define the security requirements of VABKS and describe a construction that satisfies them.

[4] 2016,**Kaitai Liang, Xinyi Huang, FuchunGuo,** we should guarantee the privacy of search contents, i.e. what a user wants to search, and return results, i.e. what a server returns to the user. Furthermore, we also need to guarantee privacy for the outsourced data, and bring no additional local search burden to user. In this paper, we design a novel privacy-preserving functional encryption based search mechanism over encrypted cloud data. A major advantage of our new primitive compared to the existing public key based search systems is that it supports an extreme expressive search mode, regular language search. Our security and performance analysis show that the proposed system is provably secure and more efficient than some searchable systems with high expressiveness.

[5] 2014**, G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner**, Remote data integrity checking is of crucial importance in cloud storage. It can make the clients verify whether their outsourced data is kept intact without downloading the whole data. In some application scenarios, the clients have to store their data on multi-cloud servers. At the same time, the integrity checking protocol must be efficient in order to save the verifier's cost. From the two points, we propose a novel remote data integrity checking model: ID-DPDP (identity-based distributed provable data possession) in multi-cloud storage. The formal system model and security model are given.

3. PROPOSED SYSTEM

In this project we have to secure the file is the main motivation. In this, there is two parts are there one is user side and another one is admin side.In user side ,only they will upload the data in the form of file.After that in an admin side,there are user wants the file they needs acknowledgements of the other then only they will use the file else they are not accepting the file in the user side they used to download file by using the key .The main motive is that,if the first user wants the file the other than acknowledgement system is very important then the process initisalised by using the fragmentation part.

4. METHODOLOGY

4.1 **USER INTERFACE DESIGN:**

This is the first module of our project. Theimportant role for the user is to move login window to user window. This module has created for the security purpose. In this login page we have to enter login user id and password. It will check username and password is match or not (valid user id and valid password). If we enter any invalid username or password we can't enter into login window to user window it will shows error message. So we are preventing from unauthorized user entering into the login window to user window. It will provide a good security for our project. So server contain user id and password server also check the authentication of the user. It well improves the security and preventing from unauthorized user enters into the network.

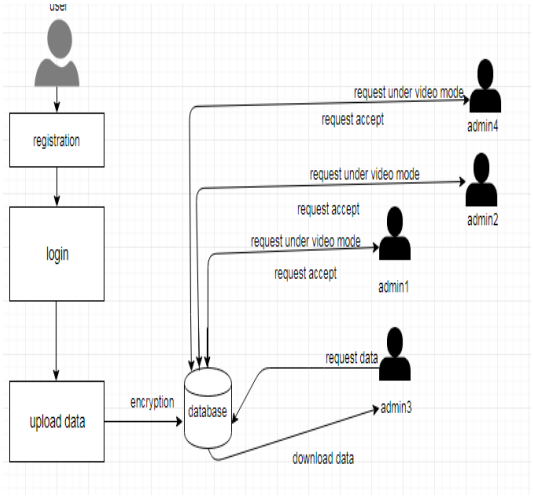
4.2 **ADMIN MONITORING THE FILE:**

In this part admins will maintain the file, after that the admin will monitor the files in the way of video mode.If anyone of the admin from the admins team are going to request a file,the request will go by video mode.

4.3 **VIEW/READ FILE:**

For reading each file which have been uploaded and split into 4 parts we should be owner of the file otherwise we should know the four different key which have been combined by random algorithm after reading the file you can also download the file otherwise with wrong key you can't open content.

5. SYSTEM ARCHITECTURE:



6. CONCLUSION

In this paper, we introduce a large universe searchable encryption scheme to protect the security of cloud storage system, which realizes regular language encryption The cloud service provider could test whether the encrypted regular language in the encrypted cipher text is acceptable by the DFA embedded in the submitted search token. In the test procedure, no plaintext of the regular language or will be leaked to the cloud server. We also put forth a concrete construction with lightweight encryption and token generation algorithms. An example is given to show how the system works. The proposed scheme is privacy-preserving and indistinguishable against, fragmentation which are proved in standard model. The comparison and experiment result confirm the low transmission and computation overhead of the scheme.

The framework addresses the unique challenges brought by multiple PHR owners and users, in that we greatly reduce the complexity of key management while enhance the privacy guarantees compared with previous works.

FUTURE ENHANCEMENT:

An accumulation is often needed to gather the partial results from these parallel executions in different servers. The runtime system captures new events and run corresponding actions to analyze the page and store more URLs into the URL set to generate new events.

7. RESULTS

FIG 1: DATA TRANSFERING

FIG 2 : PORT VALIDATION

FIG 3: FILE UPLOAD

FIG 4: ADMIN UPLOAD

REFERNCES:

1. {Ard12} D Ardagna, E Di Nitto, et al. MODAClouds: A model-driven approach for the design and execution of applications on multiple Clouds, Proceedings of MiSE 2012, 50--56.
2. {Ber12} S. Bernardi, J. Merseguer, D. C. Petriu. Dependability modeling and analysis of software systems specified with UML. ACM Computing Surveys, 45(1), p. 2, 2012.
3. {Deb11} P. Debois. Devops: A software revolution in the making?, J. Information Technology Management, 2011
4. {Men10} D. A. Menascé, J. M. Ewing, H. Gomaa, S. Malek, J. P. Sousa. A framework for utility-based service oriented design in SASSY. Proceedings of ACM/SPEC WOSP/SIPEW 2010, 27--36.
5. {Mar10} A. Martens, H. Koziolek, S. Becker, R. Reussner. Automatically improve software architecture models for performance, reliability, and cost using evolutionary algorithms. Proceedings of ACM/SPEC WOSP/SIPEW 2010, 105--116
6. {Fra13} D. Franceschelli, D. Ardagna, M. Ciavotta, E. Di Nitto. Space4Cloud: A tool for system performance and cost evaluation of cloud systems. Proceedings of
7. MultiCloud workshop, 27--34, 2013.
8. {Per13} J. F. Perez and G. Casale. Assessing SLA compliance from Palladio component models. Proceedings of the 2nd Workshop on Management of resources and services in Cloud and Sky computing (MICAS), IEEE Press, 2013.
9. {Per15} J. F. Pérez, G. Casale, and S. Pacheco-Sanchez. Estimating Computational Requirements in Multi-Threaded Applications. IEEE Transactions on Software Engineering, to appear in 2015.
10. {Zha10} H. Zhao, C. H. Xia, Z. Liu, D. F. Towsley. A unified modeling framework for distributed resource allocation of general fork and join processing networks. Proceedings of ACM SIGMETRICS 2010: 299--310.
11. {Zik11} P. Zikopoulos, C. Eaton. Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data. McGraw-Hill Osborne, 2011