International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 04 | Apr 2020 www.irjet.net p-ISSN: 2395-0072

Efficient Regular Language Search for Secure Cloud Storage

Harsha Vardhan[1], Suvin[2], D. Punitha[3]

[1,2]*Computer Science and Engineering, SRM Institution of Science and Technology, Chennai, India*
[3]*Assistant Professor, Computer Science and Engineering, SRM Institution of Science and Technology, Chennai, India*

ABSTRACT: *Cloud computing provides flexible data management and ever-present data access. However, the repository services provided by cloud server is not trusted by customers. The data's provided by cloud server can be easily stolen by intruders. Searchable encryption could provide the functions of confidentiality protection and privacy-preserving data retrieval, which is an important tool for secure storage. In this paper, we propose an efficient large universe regular language search scheme for the cloud storage, which privacy is preserving and secure against the off-line keyword guessing attack (KGA). A notable highlight of the proposal over other existing schemes is that it supports the regular language encryption and deterministic finite automata (DFA) based data retrieval. The large universe construction ensures the extendibility of the system, in which the symbol set does not need to be predefined. Multiple users are supported in the system, and the user could generate a DFA token using his own private key without interacting with the key generation centre. Furthermore, the concrete scheme is efficient and formally proved secure in standard model. Extensive comparison and simulation show that this scheme has function and performance superior than otherschemes*

KEYWORDS: *Cloud, Deterministic finite automata, Efficient Regular language, Secure*

1.INTRODUCTION

1.1 Purpose of system

The main aim of this project is to provide integrity of an organization data's which is in public cloud and retrieving data's in encrypted form and retrieving information

1.2 Project scope

In this work, cloud storage of secured data's of various fields and retrieving those data's whenever necessary. All those data's will be in different formats like text, image, video etc and all these will be encrypted and stored in database DFA search technique is used to read the data's in the form of encrypted data and convert the data's into regular language and display those information to authorized user whenever they need. In server side operations they can also manipulate all the data's which they uploaded

1.3 Product Perspective

Cloud delivers convenience to the customers and at the same time arouses many security and privacy problems. 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. 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 cipher text.

1.4 System features

Different User's will create their account with specific user name and password and with that information only cloud can retrieve their account and from that account user can access and search various kinds of data by proving the particular name of the data or any related names of that data's for the cloud to read and identify the required user information and the data's and fetched from the cloud database backend and the information is displayed in user's account so that the user can read and download the data's.

2. LITERATURE SURVEY

2.1 Two-Factor Data Security Protection Mechanism for Cloud Storage System

This paper studies to improve data security protection mechanism for cloud using two components. In this system sender sends an encrypted message to a receiver with the help of cloud system. The sender requires to know identity of receiver but no need of other information such as certificate or public key. To decrypt the cipher text, receiver needs two parts. The first thing is a unique personal security device or some hardware device connected to the computer system. Second one is private key or secrete key stored in the computer. Without having these two things cipher text never decrypted. It provides no risk of data Storage maintenance tasks, such as acquiring additional storage capacity, can be unloaded to the responsibility of a service provider. But if the security device lost or stolen, then cipher text cannot be decrypted and hardware device is revoked or cancelled to decrypt cipher text.

2.2 Public-Key Encryption with Fuzzy Keyword Search

: A Provably Secure Scheme under Keyword Guessing Attack This paper studies to avoid keyword guessing attack. So they propose a novel concept called public-key encryption with fuzzy keyword search (PEFKS), by which the un-trusted server only obtains the fuzzy search trapdoor instead of the exact search trapdoor. It is efficient under the practical condition that the size of the keyword space is not more than the polynomial level. Disadvantage is their searches take time linear in the number of cipher texts.

2.3 A Multi-Parameter Analysis of Hard Problems on Deterministic Finite Automata

This paper proposes a hard problems on deterministic finite automata (DFAs): the problem of finding a short synchronizing word, and that of finding a DFA on few states consistent with a given sample of the intended language and its complement. Natural parameterizations and classify them with the tools provided by Parameterized Complexity. But synchronizing the word is the major problem.

2.4 Learning Deterministic Finite Automata with a Smart State Labelling Evolutionary Algorithm

In this paper, we studies a novel evolutionary method for learning DFA that evolves only the transition matrix and uses a simple deterministic procedure to optimally assign state labels. This compares its performance with the Evidence Driven State Merging (EDSM) algorithm, one of the most powerful known DFA learning algorithms. They present results on random DFA induction problems of varying target size and training set density. Also study the effects of noisy training data on the evolutionary approach and on EDSM. Disadvantage is optimally assigning state labels for the DFA .

2.5 Key-Aggregate Searchable Encryption (KASE) for Group Data Sharing via Cloud Storage

This paper proposes a key aggregate searchable encryption (KASE) and instantiating the concept through a concrete KASE scheme, in which a data owner only needs to distribute a single key to a user for sharing a large number of documents, and the user only needs to submit a single trapdoor to the cloud for querying the shared documents. Large number of keys must not only be distributed to users via secure channels, but also be securely stored and managed by the users in their devices. But they cannot produce the solution for federated cloud.

2.6 Practical Techniques for Searches on Encrypted Data Practical Techniques for Searches on Encrypted Data

In this paper, we studies cryptographic schemes for the problem of searching on encrypted data and provide proof of security for the resulting crypto systems. Our techniques have a number of crucial advantages. They are provably secure. They provide provable secrecy for encryption, in the sense that the untrusted server cannot learn anything about the plaintext when only given the cipher text. But Sequential scan may not be efficient enough when the data size is large.

2.7 Searchable Attribute-Based Mechanism with Efficient Data Sharing for Secure Cloud Storage

This paper, proposes a searchable attribute-based proxy re-encryption system. Primitive supports both abilities and provides flexible keyword update service. Specifically, the system enables a data owner to efficiently share his data to a specified group of users matching a sharing policy and meanwhile, the data will maintain its searchable property but also the corresponding search keyword(s) can be updated after the data sharing. Disadvantage is needs to reduce the size of search token and needs to provide more expressive keyword search.

2.8 Towards Differential Query Services in Cost-Efficient Clouds

In this paper, we studies query services by the user demand. Mask Matrix is used to filter out as what user really wants matched data before recurring to the Aggregation and Distribution Layer (ADL).this system reduces the communication cost

2.9 Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data

In this paper, they define and solve the problem of secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by enabling search result relevance ranking instead of sending undifferentiated results, and further ensures the file retrieval accuracy. Specifically, we explore the statistical measure approach, i.e. relevance score, from information retrieval to build a secure searchable index, and develop a one-to-many order preserving mapping technique to properly protect those sensitive score information. The resulting design is able to facilitate efficient server-side ranking without losing keyword privacy.

2.10 Enabling Fine-grained Multi-keyword Search Supporting Classified Sub-dictionaries over Encrypted Cloud Data

In this paper, they address by building up the fine-grained multi-watchword hunt plans over scrambled cloud information. Unique commitments are three-fold. They present the significance scores and inclination elements upon watchwords which empower the exact catchphrase seek and customized client experience. Second, they build up a handy and exceptionally effective multi catch phrase inquiry plan. The proposed plan can backing entangled rationale seek the blended "AND", "OR" and "NO" operations of catchphrases. Third, we further utilize the ordered sub-lexicons procedure to accomplish better proficiency on list building, trapdoor producing and question. Finally, we examine the security of the proposed plans as far as secrecy of reports, security assurance of file and trapdoor, and unlink ability of trapdoor.

3. SYSTEM ANALYSIS

3.1Existing System

In existing system the data's which are uploaded by the server side are all Stored in the database with all formats in the regular language so that the security of data's are very less and there is a possibility of information theft and information loss possibility is there, so in future only encrypted data's alone should be stored in cloud.

3.2 Proposed System

In this work, cloud storage of secured data's of various fields and retrieving those data's whenever necessary. All those data's will be in different formats like text, image, video etc. and all these will be encrypted and stored in database DFA search technique is used to read the data's in the form of encrypted data and convert the data's into regular language and display those information to authorized user whenever they need. In server side operations they can also manipulate all the data's which they uploaded.

4. SYSTEM DESIGN

4.1 System Architecture

System architecture is the conceptual model that defines the structure, behaviour, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviours of the system. Figure 1 shows the architecture.

ARCHITECTURE

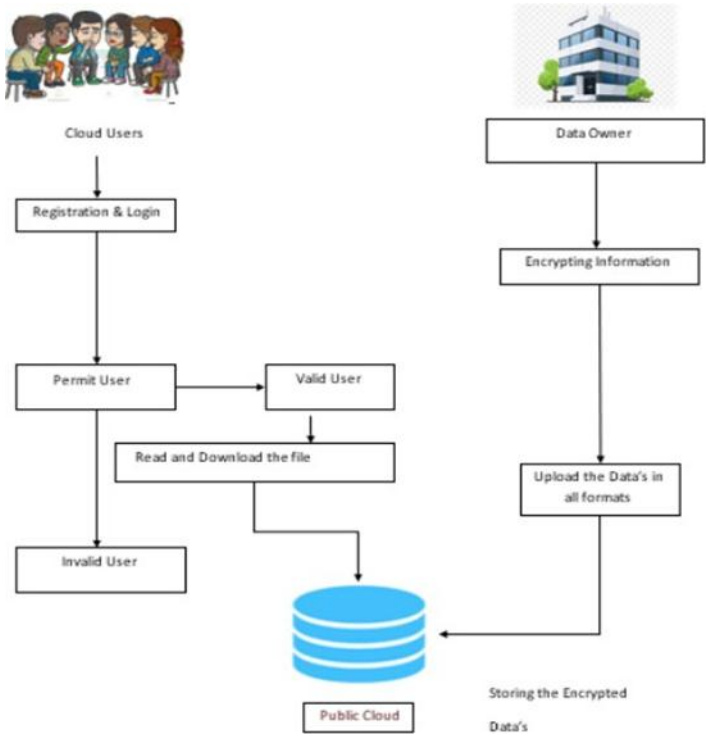

FIG.1 ARCHITECTURE

5. MODULE

The proposed system consists of four main modules. They are:

- *Data owner and user authentication*
- *Safe page file upload*
- *Authorized user's page to access cloud*
- *Server side data manipulation*

5.1 Data owner and user authentication

User has to fill the registration form with required mandatory input fields to be filled completely and enrol themselves to access the cloud information and downloading the documents from the data's retrieved from the cloud. Finally a user id and password is generated and can be used as a key to access the information from the cloud. .

5.2 Safe page file upload

In this module Server side will upload the required data's for various fields according to the basic requirements of the user and all these information are gathered and then encrypted and finally stored inside the public cloud. All the formats of data's and file like text, image, video, pdf etc. can be stored in the cloud in encrypted form.

5.3 Authorized user's page to access cloud

Different User's will create their account with specific user name and password and with that information only cloud can retrieve their account and from that account user can access and search various kinds of data by proving the particular name of the data or any related names of that data's for the cloud to read and identify the required user information and the data's and fetched from the cloud database backend and the information is displayed in user's account so that the user can read and download the data's.

5.4 Server side data manipulation

Server has the rights not only to upload the data's in cloud; he also manipulates the data's according to user needs. Some data's are hidden based on age restrictions and server can kill the information which are hazardous to society and user's so always manipulation rights will be there in server side computing and after manipulation or editing the data's all these data's once again encrypted and stored back inside the public cloud

6. Conclusion

The various units were extensively tested for all possible defects using all possible parameters applicable in the context. Integration had been done successfully after unit testing, and the whole system is tested again by integration system testing before implementation. In this system all those data's will be in different formats like text message, image, videos, ppt etc. All these will be encrypted and stored in database .And these data's will be recovered by using DFA search technique which convert into regular language and displays those information to authorized user.

REFERENCES

1. *[Erl T, Cope R, Naserpour A. Cloud computing design patterns[M]. Prentice Hall Press, 2015.*
2. *Li Z, Dai Y, Chen G, et al. Toward network-level efficiency for cloud storage services[M]//Content Distribution for Mobile Internet: A Cloud-based Approach. Springer Singapore, 2016: 167-196.*
3. *Sookhak M, Gani A, KhanMK, et al. Dynamic remote data auditing for securing big data storage in cloud computing[J]. Information Sciences, 2017, 380: 101-116.*
4. *Zhang Q, Yang L T, Chen Z, Li P. Privacy-preserving double projection deep computation model with crowd sourcing on cloud for big data feature learning[J]. IEEE Internet of Things Journal, 2017, DOI: 10.1109/JIOT.2017.2732735.*
5. *Zhang Q, Yang L T, Chen Z, Li P. PPHOPCM: Privacy-preserving High-order Possibilistic c-Means Algorithm for Big Data Clustering with Cloud Computing[J]. IEEE Transactions on Big Data, 2017, DOI: 10.1109/TBDATA.2017.2701816.*
6. *Liu J K, Liang K, Susilo W, et al. Two-factor data security protection mechanism for cloud storage system[J]. IEEE Transactions on Computers, 2016, 65(6): 1992-2004.*
7. *Boneh D, Waters B. Conjunctive, subset, and range queries on encrypted data[C]//Theory of Cryptography Conference. Springer Berlin Heidelberg, 2007: 535-554.*
8. *Q. Zheng, S. Xu, and G. Ateniese. VABKS: verifiable attribute-based keyword search over outsourced encrypted data. In INFOCOM, pp. 522C530. IEEE, 2014.*
9. *Liang K, Huang X, Guo F, et al. Privacy-Preserving and Regular Language Search Over Encrypted Cloud Data[J]. IEEE Transactions on Information Forensics and Security, 2016, 11(10): 2365-2376.*
10. *Chang V, Ramachandran M. Towards achieving data security with the cloud computing adoption framework [J]. IEEE Transactions on Services Computing, 2016, 9(1): 138-151.*