Use of DNA Computing for Three-Layer Privacy Preserving Cloud Storage Scheme

Puransingh Chauhan[1], Dr. Sachin Chaudhari[2], Ms. Rana Syeda[3]
[1]MTech. [2]Department of Computer Science & Engineering,
Jhulelal Institute of Technology, Nagpur

Abstract : *With the rapid development of network bandwidth, the volume of user's data is rising geometrically. User's requirement cannot be satisfied by the capacity of local machine any more. Therefore, people try to find new methods to store their data. Pursuing more powerful storage capacity, a growing number of users select cloud storage. Storing data on a public cloud server is a trend in the future and the cloud storage technology will become wide spread in a few years. Cloud storage is a cloud computing system which provides data storage and management service. With a cluster of applications, network technology and distributed file system technology, cloud storage makes a large number of different storage devices work together coordinately. Nowadays there are a lot of companies providing a variety of cloud storage services, such as Dropbox, Google Drive, iCloud, Amazon Web Services, etc. These companies provide large capacity of storage and various services related to other popular applications, which in turn leads to their success in attracting humorous subscribers. However, cloud storage service still exists a lot of security problems. The privacy problem is particularly significant among those security issues. In history, there were some famous cloud storage privacy leakage events. For example, in 2018 despite a robust legislation on data protection by UIDAI, Aadhaar numbers and bank details of over 134,000 beneficiaries on Andhra Pradesh Housing Corporation's website have been leaked. Apples iCloud leakage event in 2014, numerous Hollywood actresses private photos stored in the clouds were stolen. This event caused an uproar, which was responsible for the users' anxiety about the privacy of their data stored in cloud server.*

Keywords: *UIDAI (Unique Identification Authority of India), Cloud Computing, Cloud Storage, Anonymity*

I. Introduction

Storing and exchange of data in cloud computing become the necessity of modern working pattern in IT industry. Cloud computing provides multitudinous benefits to both service provider and customer. However, the security of cloud computing has been a challenging one. To increase security and confidentiality of data in cloud environment, the DNA sequences are used with Morse code and zigzag pattern, for encoding scheme. Use of Morse code and Zigzag pattern makes the intruder much harder to steal original data. Recent years witness the development of cloud computing technology .With the explosive growth of unstructured data, cloud storage technology gets more attention and better development. However, in current storage schema, user's data is totally stored in cloud servers. In other words, users lose their right of control on data and face privacy leakage risk. Traditional privacy protection schemes are usually based on encryption technology, but these kinds of methods cannot effectively resist attack from the inside of cloud server. In order to solve this problem, we propose a three-layer storage framework based on fog computing. The proposed framework can both take full advantage of cloud storage and protect the privacy of data. Besides, Hash Solomon code algorithm is designed to divide data into different parts. Then, we can put a small part of data in local machine and fog server in order to protect the privacy. Moreover, based on computational intelligence, this algorithm can compute the distribution proportion stored in cloud, fog and local machine, respectively. Through the theoretical safety analysis and experimental evaluation, the feasibility of our scheme has been validated, which is really a powerful supplement to existing cloud storage scheme.

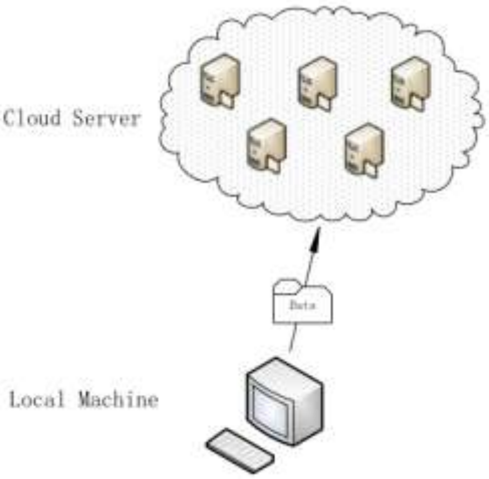


Fig. Traditional Cloud Storage Structure

In our scheme, we split user's data into three parts and separately save them in the cloud server, the fog server and the user's local machine. Besides, depending on the property of the Hash-Solomon code, the scheme can ensure the original data cannot be recovered by partial data. On another hand, using Hash-Solomon code will produce a portion of redundant data blocks which will be used in decoding procedure. Increasing the number of redundant blocks can increase the reliability of the storage, but it also results in additional data storage. By reasonable allocation of the data, our scheme can really protect the privacy of user' data. The Hash-Solomon code needs complex calculation, which can be assisted with the Computational Intelligence (CI). Paradigms of CI have been successfully used in recent years to address various challenges, for example, the problems in Wireless sensor networks (WSNs) field. CI provides adaptive mechanisms that exhibit intelligent behavior in complex and dynamic environments like WSNs. Thus in our paper, we take advantage of CI to do some calculating works in the fog layer. Compared with traditional methods, our scheme can provide a higher privacy protection from interior, especially from the CSPs.

1.1 Three-Layer Privacy Preserving Cloud Storage Scheme

In order to protect user's privacy, we propose a TLS framework based on fog computing model. The TSL framework can give user a certain power of management and effectively protect user's privacy. As mentioned, the interior attack is difficult to resist. Traditional approaches work well in solving outside attack, but when CSP itself has problems, traditional ways are all invalid. Different from the traditional approaches, in our scheme, user's data is divided into three different-size parts with encoding technology. Each of them will lack a part of key information for confidentiality. Combining with the fog computing model, the three parts of data will be stored in the cloud server, the fog server and user's local machine according to the order from large too small. By this method, the attacker cannot recover the user's original data even if he gets all the data from a certain server. As for the CSP, they also cannot get any useful information without the data stored in the fog server and local machine because both of the fog server and local machine are controlled by users.

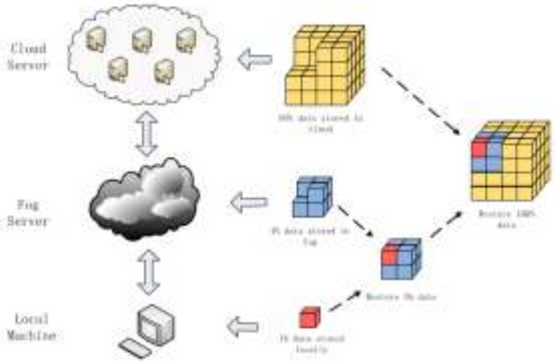


Fig. Illustration of Three-Layer storage framework based on fog computing

As shown in Fig, the TLS framework makes full use of fog server's storage and data processing capability. The architecture includes three layers, the cloud server, the fog server and the local machine. Each server saves a certain part of data, the storage proportion is determined by users' allocation strategy. Firstly, user's data will be encoded on user's local machine. Then, for example, let 1% encoded data be stored in the machine. Then upload the remainder 99% data to the server. Secondly, on the server, we do similar operations to the data which comes from user's machine. There will be about 4% data stored in the fog server and then upload the remainder data to the cloud server. The above operations are based on Hash-Solomon code. Hash-Solomon code is a kind of coding methods based on Reed Solomon code. After being encoded by Hash-Solomon code, the data will be divided into k parts and generates m redundant data. Hash-Solomon code has such property, in these k+m parts of data, if someone has at least k parts, he can recover the complete data. In other word, nobody can recover the complete data with less than k parts of data. According to this property of Hash-Solomon code, in our scheme, we let no more than k-1 parts of data be stored in higher server which has larger storage capacity and let the remainder be stored in the lower server. In this way, the stealer cannot recover the complete data even if one of the three layers' data was stolen. Thus we can ensure the privacy of user's data. Then we consider the value of k and m. assuming that we want to save r% data on the fog server.

1.2 DNA Computing

The Deoxyribonucleic Acid is a molecule in all the living organisms. The genetic instructions about all organisms are carried in DNA molecule. The Nucleotides are the basics of DNA molecule which is used to store biological information. DNA computing was introduced in the year 1994 by Adleman. The direct Hamiltonian path problem (HPP) was successfully solved by using DNA computing. From the research of DNA computing, the DNA cryptography has evolved. DNA consists of DNA strands which are polymer chains. The polymer chains are composed of four nucleotides. They are, Adenine (A), Guanine (G), Cytosine(c) and Thymine (T). The structure of double stranded DNA with sequence is shown in the following Fig.

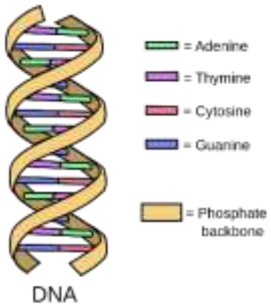


Fig. Double Stranded DNA

These nucleotides are basic building blocks of DNA strands. The DNA strands bond with each other based on Watson-crick complementary base pair rule which is shown in the following Fig.

5'T A C A G A C T C A 3'
3'A T G T C T G A G T 5'

Binary sequence		Nucleotide
0	0	A
1	1	T
1	0	C
1	1	G

Fig. Base pair (b) Nucleotide sequence

The DNA sequence are represented by binary numbers 0’s and 1’s refer Fig. The binary sequences are machine understandable sequences which are used to increase the efficiency and speed of the process.

The DNA cryptography is used to enable secure data communication for users. In DNA cryptography, the DNA sequences are used to create unbreakable encryption and decryption technology.

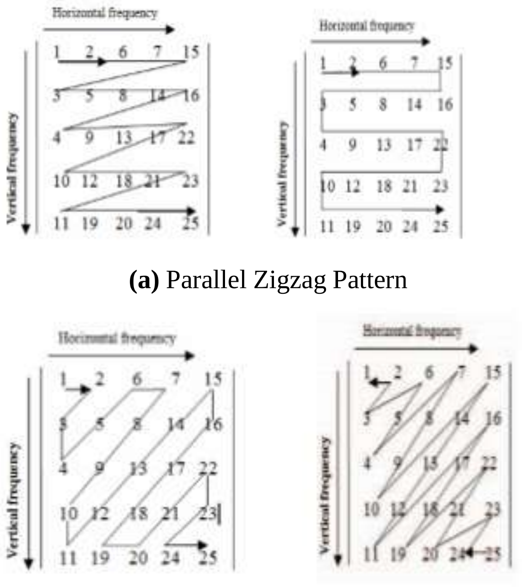
1.3 DNA Based Cryptography

DNA cryptographic techniques are based on DNA sequence. The strong encryption schemes are achieved because of randomness and uniqueness.

The cost of the process of DNA computing is high and time consuming. These can be overcome by using modern cryptographic techniques which is based on computer based algorithms. Complex problems can be solved by DNA cryptographic algorithms.

1.4 Zigzag Encryption Scheme

The Zigzag pattern is matrix structure of the N2 integers which would increase sequentially along the arrays anti diagonals. The two kinds of zigzag patterns are shown in the Fig.5. A motive wave that travels on parallel trend is known as parallel zigzag pattern refer Fig. (A). the wave that almost never travels within a parallel trend is known as diagonal zigzag pattern refer Fig. (B).



(a) Parallel Zigzag Pattern

(b) Diagonal Zigzag Pattern

Fig. Various Zigzag patterns

In this proposed method, Encryption of data can be performed using diagonal zigzag encryption pattern refer Fig (a).

II. Implementation Detail of Workflow

2.1 Stored Procedure

When user wants to store his file to the cloud server, the procedure is shown as Fig. First of all, user’s file will be encoded with Hash-Solomon code. And then, the file will be divided into several data blocks and the system will also feedback encoding information simultaneously. Assuming that 1% data blocks and the encoding information will be stored locally. The remainder 99% data blocks will be uploaded to the fog server. Secondly, after receiving the 99% data blocks from user’s machine, these data blocks will be encoded with Hash Solomon again. These data blocks will be divided into smaller data blocks and generates new encoding information. Similarly, assuming that 4% data blocks and encoding information will be stored in the fog server. The remainder 95% data blocks will be uploaded to the cloud server. Thirdly, after cloud server received the data blocks form fog side, these data blocks will be distributed by cloud manage system. Finally, the storage procedure ends when all the related information be recorded in different servers.

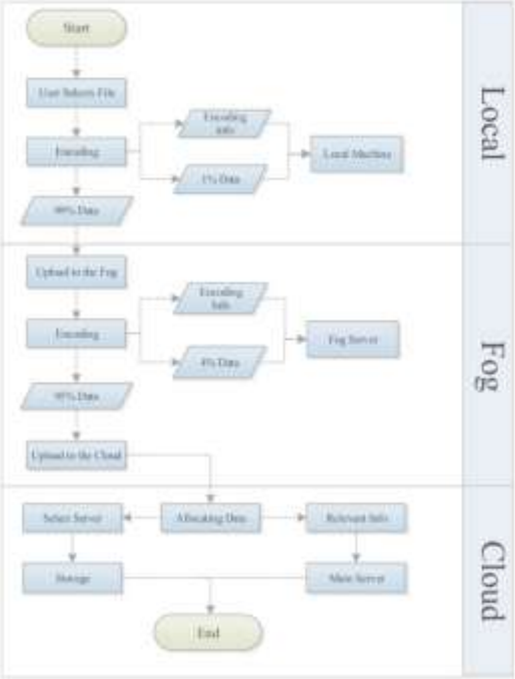


Fig. Diagram of stored procedure

2.2 Download Procedure

When user wants to download his file from the cloud server, the procedure is shown in Fig. 4. Firstly, cloud server receives user’s request and then integrates the data in different distributed servers. After integration, cloud server sends the 95% data to the fog server. Secondly, the fog server receives the data from the cloud server. Combining with the 4% data blocks of fog server and the encoding information, we can recover 99% data. Then the fog server returns the 99% data to the user. Thirdly, the user receives the data from fog server. User can get the complete data by repeating the above steps.

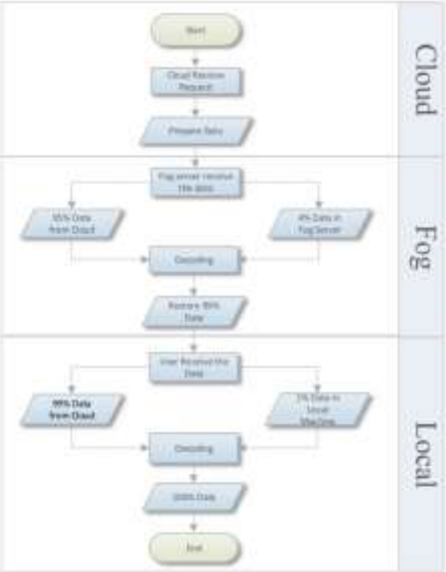


Fig. Diagram of download procedure

III. Conclusions

The development of cloud computing brings us a lot of benefits. Cloud storage is a convenient technology which helps users to expand their storage capacity. However, cloud storage also causes a series of secure problems. When using cloud storage, users do not actually control the physical storage of their data and it results in these parathion of ownership and management of data. In order to solve the problem of privacy protection in cloud storage, we propose a TLS framework based on DNA computing model and design a Hash-Solomon algorithm. Through the theoretical safety analysis, the scheme is proved to be feasible. By allocating the ratio of data blocks stored in different servers reasonably, we can ensure the privacy of data in each server. On another hand, cracking the encoding matrix is impossible theoretically. Besides, using hash transformation can protect the fragmentary information. Through the experiment test, this scheme can efficiently complete encoding and decoding without influence of the cloud storage efficiency.

Cloud Computing allows companies to use available resources at any time without a limit. The sensitive information is also transmitted and stored in the cloud at lower cost. However, the prevalence of cloud is suffered by its security challenges.

To improve the security of cloud computing the new model has been proposed. The security model is based on DNA sequences. So finding original data is harder with the existing encryption model and now the Zigzag pattern is added to improve security.

References

[1]. A.Murugan and R.Thilagavathy," Securing Cloud Data using DNA and Morse code: A Triple Encryption Scheme", International Journal of Control Theory and Applications (IJCTA), vol.10, pp.31-18, Nov 2017.

[2]. Jacob Grasha, and A. Murugan, "A Hybrid Encryption Scheme using DNA Technology" The International Journal of Computer Science and Communication Security (IJCSCS), vol.3 (2), pp.61-65, Feb 2013.

[3]. G. Feng, "A data privacy protection scheme of cloud storage," vol. 14, no. 12, pp. 174–176, 2015