Self-destructing Data Method Based on Privacy Cloud

Hao Wu
Computer and Information Engineering College of
HOHAI University, HHU
Nanjing, China
whforhhu@gmail.com

Zhijian Wang
Computer and Information Engineering College of
HOHAI University, HHU
Nanjing, China
W51178@sina.com

Xiao Fu
Computer and Information Engineering College of
HOHAI University, HHU
Nanjing, China
Nhri.fuxiao@gmail.com

Yu Wang
Computer and Information Engineering College of
HOHAI University, HHU
Nanjing, China
whforhhu@gmail.com

Abstract—Cloud Data Security is gaining more and more attention from enterprises and users. Email security may be the most highlighted, since the past and archived cloud data is under risk of compromising privacy. However, the traditional cryptology method cannot meet the security demand in privacy protection. Data self-destruction is the ultimate way to protect the security and privacy of data. Our research aims at protect the privacy of past and archived cloud data, such as emails remained in cloud service provider. This article proposes a self-destructing data method to realize the self-destruction of cloud data, and apply it to the implementation of self-destructing email system. In this method, split cipher data with key are distributed in storage called privacy cloud. Consequently, the system achieves the aim of data self-destruction and has good user experience. In the article, comparative analysis and performance measurement show that the scheme is able to satisfy the application demands.

Keywords- Self-destruct; Cloud Data; Information Security; Privacy Protection; Email

I. INTRODUCTION

Personal privacy has been experiencing great challenges due to technical and legal landscape nowadays. The increasing use of Cloud services is a double-edged sword, making life more convenient, but the personal data is also at the risk of being stored, copied and archived without awareness.

Our study is aimed to establish the protection of past, archived cloud date, such as copies of emails maintained by email provider.Researchers hope to guarantee that after user-specified time, even if an attacker gets both a copy of cloud data and the user's cryptographic keys and passwords, all the copies of cloud data become unreadable. For users, they need to do nothing special.

Data self-destruction is the ultimate way to protect the security and privacy of data. In recent years, researches in the network data self-destruct technology have made some achievements. Geambasu et al. [1] proposed a novel data self-destructing program, designed and implemented Vanish system without trusted service or manual action. In this system, private historical data is encrypted using a symmetrical encryption key which will be split into pieces based on threshold secret sharing. The pieces of key are sprinkled throughout DHT (Distributed Hash Table) that has a timeout and periodically discards the shares, so that the cipher data cannot be decrypted any more. Wang.[2] improved the Vanish System by distributing part of the cipher data with key shares into the DHT to resist the cryptanalysis and brute force attack. Reference [3] proposed a deterministic data deletion method applicable for cloud storage to effectively improve the efficiency of key management. The method used key derivation tree to organize and manage keys which distributed into DHT, and deterministically deleted the keys with the similar method like Vanish. However, Reference [4] argued that the Vanish which using Vuze DHT was under threat of Sybil attack that attacks could get enough key shares before expiration to reconstruct the key, so references[1-3] all had this security issue. Reference [6] proposed the SafeVanish program that extended the length of key shares and encrypt key by RSA to avoid the Sybil hopping attack and sniffing attack respectively.

So far, the self-destruct data technology bases on the threshold secret share mechanism that cipher data cannot be decrypted any more due to the deletion of key shares. With the development of cloud computing, what researchers concerned is that cipher data does not erase itself from storage and of which the information can be recovered some day. Since researchers can split the key into pieces, why don't split the data of content instead? Different from Vanish System, the article designs a self-destruct data method applicable for cloud data.Researchers split the cipher data into pieces and save them into standalone cloud storage. When only one piece of cipher data goes missing, the cipher data is uncompleted and can never be decrypted anymore, so that to destruct it is the best way to protect it.

II. SECURITY MODEL BASED ON PRIVACY CLOUD

A. *Cloud Data*

Due to the features of cloud computing, cloud data will be copied, cached and archived in cloud. Users and enterprise have no way to fully clear the cloud data. Cloud Security Alliance (CSA) proposed the concept of life cycle of data security [7], which summarize the life cycle of data security into six processes, including produce, store, use, share, archive and destruct. The essence of security of cloud data is that users are out control of cloud data they uploaded. Attack from hackers and covetousness of cloud service providers may cause the privacy leak[8].

B. *Privacy Cloud Data*

As the name suggests, privacy cloud data is the cloud Data containing sensitive and confidential information. Privacy Cloud data has a feature of timeliness. At the end of life cycle, all copies of privacy cloud data should be clear fully.

C. *Privacy Cloud.*

Privacy Cloud is the storage service for privacy cloud data. Privacy Cloud should be a private cloud to serve enterprise business, or a trusted public cloud to serve personal demand.

D. *Data proxy*

Data proxy is not the real user data. Users can retrieve the information stored in privacy cloud via corresponding data proxy. Data proxy is independent of real user data and it cannot reveal any information from its appearance.

E. *Security Model*

As shown in Fig. 1, researchers introduce privacy cloud and data proxy mainly with the following concern:

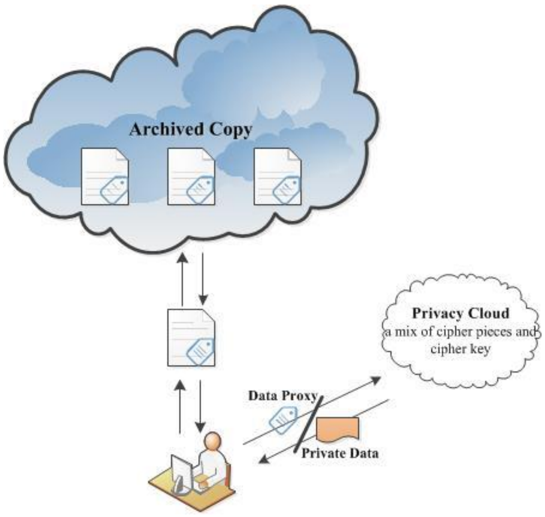


Figure 1. Security Model based on Privacy Cloud

1. Take full control of cloud data with privacy cloud to make privacy cloud data isolated from Cloud Service Provider;

2. Data proxy that is not stored in privacy cloud is independent from real data and it is the only approach for user to retrieve real data from privacy cloud.

3. The form of data stored in privacy cloud is a mix of privacy data cipher pieces and cipher key, so that they are hard to recognize in aspect of space and time for an intruder even if he can access the database.

III. SELF-DESTRUCTING DATA METHOD BASED ON EXPIRATION TIME

Researchers design a Self-destructing Data method based on Expiration Time using non-threshold secret sharing. In this paper, “self-destructing” means to become unreadable sequences automatically. This method bases on the patent [8] with some modifications.

First, encrypt the data with a key by using symmetric encryption algorithm. The key is a pseudo-random sequence k. Function *Encrypt* encrypts data d with key k and output the cipher data c. Considering capability and security of encryption algorithm, researchers choose AES[9] to do the encryption in this method.

$$c = \text{Encrypt}\ (d\ ,\ k) \tag{1}$$

Second, split the cipher data. Function *Split* splits c into pieces p_1, p_2, $p_3 \ldots p_n$ according to a specified length l. Function *Disguise* outputs p_0 by appending random bytes to k to make length of k equal to l.

$$(p_1, p_2, p_3 \ldots p_n) = \text{Split}(c,\ l) \tag{2}$$

$$p0 = \text{Disguise}\ (k,\ l) \tag{3}$$

Third, change the position of p_0. Function *GenIndex* generates i between (0, n-1], Function *Swap* exchanges the value of p_0 and p_1.

$$i = \text{GenIndex}(n)\ ,0< i \leq n \tag{4}$$

$$\text{Swap}(p_0\ ,\ p_i) \tag{5}$$

Forth, encapsulate.Researchers introduce the notion of *SDO* in this method. *SDO* (Self-destructing Data Object) encapsulates one of cipher pieces, identification of piece and user-specific *ET* (expiration timestamp). Function *Identify* generates UUID u_i (Universally Unique Identifier) for every piece pi to guarantee p_i has unique identifier. Function *Encapsulate* outputs *SDOi* which has three attributes including $u_{i,}$ p_i and ET_i and persist SDO$_i$ into distributed storage.

$$u_i = \text{Identify}\ (p_i),\ i=0,1,2\ldots \tag{6}$$

$$\text{SDO}_i = \text{Encapsulate}\ (u_i\ ,\ p_i\ ,\ \text{ET}_i\ )\ ,\ i=0,1,2\ldots \tag{7}$$

Fifth, generate data proxy. In this method, data proxy is just a string which starts with “SDD://”.Researchers call the data proxy SDD link. Function *Link* combines all the ui of SDOs as a sequence *da*, which indicate their offset address. As SDD link indicates the offset of identifier of every piece, researchers can easily retrieve the data content via SDD link by recover the original sequence of all cipher pieces.

$$da = \text{Link}\ (u_0, u_1, u_2, u_3\ldots) \tag{8}$$

In order to make data self-destruct after user-specific expiration time, researchers design a process to keep watching the ET_i of SDO_i in storage. When the ET_i of any SDO_i is up, the process will delete the SDO_i from storage. However, the massive concurrency delete operations in database cause a reduction on performance, so researchers designed a delete-on-delay mechanism: except the first SDO_i, the ET_i of the every other SDO_i would be delay randomly n minutes and n should be adjusted according to the total SDO number in storage to get the optimal performance.

IV. PROTOTYPE SYSTEM

Self-destructing Data method based on Expiration Time is the core technology.Researchers implement the prototype system based on this technology and the security model and apply it in self-destruct email. In fact, self-destruct email is a self-destructing body of content of email as researchers can't change SMTP protocol [10].

As shown in Fig. 2, the system includes client side and server side. The client side is a browser plugin based on Google Chrome, it has two main functions:

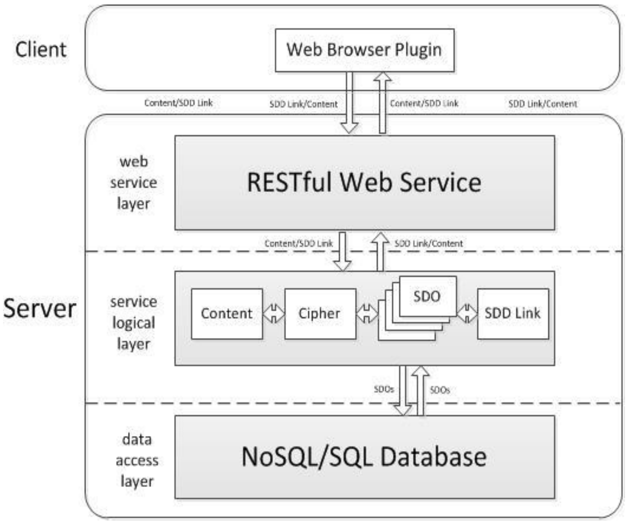
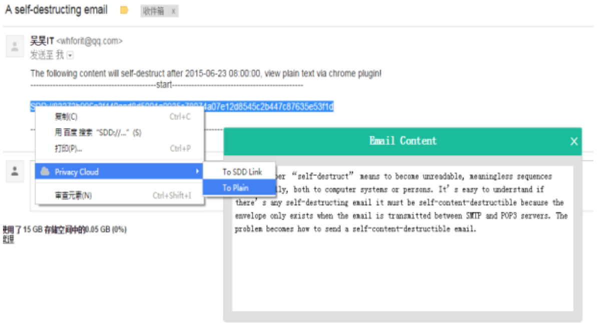


Figure 2. The Structure of Prototype System

1. Provide user with an interface to send private content to privacy cloud via SSL channel and get the data proxy back. In the self-destruct email system, users can select the finished email content in web page using left mouse button, and then click right mouse button to select "To SDD Link" menu. Before sending the email content to privacy cloud, user must set an expiration timestamp for content. Then SDD link will come back replace the selected private content in web page.

2. Provide user with an interface to retrieve the real data from privacy cloud via SSL channel and show the decrypted private information. In the self-destruct email system, user select the SDD link text and click right mouse button to select "To Plain". Then the private content will be shown in a dialog if the content hasn't expired. If the content expired, browser plugin gives user a prompt. Fig. 3&Fig. 4 shows the email content self-destructed when the expiration time is up.


A self-destructing email

吴昊IT <whforit@qq.com>
发送至 我

The following content will self-destruct after 2015-06-23 08:00:00, view plain text via chrome plugin!
--start--

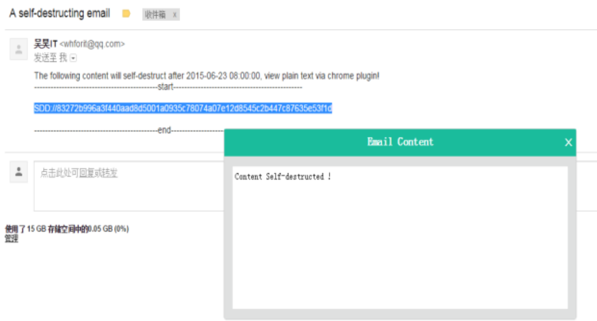
4a07e12d8545c2b447c87635e53f1d

Email Content

er "self-destruct" means to become unreadable, meaningless sequences lly, both to computer systems or persons. It' s easy to understand if there' s any self-destructing email it must be self-content-destructible because the envelope only exists when the email is transmitted between SMTP and POP3 servers. The problem becomes how to send a self-content-destructible email.

使用了 15 GB 存储空间中的0.05 GB (0%)
管理

Figure 3. The Self-destructing Email before expriration


A self-destructing email

吴昊IT <whforit@qq.com>
发送至 我

The following content will self-destruct after 2015-06-23 08:00:00, view plain text via chrome plugin!
--start--

SDD://83272b996a3f440aad8d5001a0935c78074a07e12d8545c2b447c87635e53f1d

--end--

Email Content

Content Self-destructed !

使用了 15 GB 存储空间中的0.05 GB (0%)
管理

Figure 4. Self-destructed Email Content

The server side in self-destruct email system is adopts three-layer structure: web service layer, service logical layer and data access layer. The service logical layer works based on the self-destruct data method to implement the conversion between SDD Link and SDOi. The Data Access Layer persist the SDOi into distributed storage.

V. ANALYSIS AND EXPERIMENTS

A. *Comparative Analysis.*

When user sends a self-destruct email via the system, the body of content is replaced with data proxy which is a SDD Link. SDD Link indicates how to acquire the pieces of the real data. Considering a normal content-encrypted email, even though its cipher key is missing, the cipher of the content still exists in the cloud and may be decrypted one day. Unlike the Vanish, when a self-destructing email expired, the content of email copies cannot be revealed again since the data has already broken.

Use' private data is stored as SDOs in distributed storage. Due to the delete-on-delay mechanism and the UUIDs, all the persisted SDOs have hardly collisions on database attitudes. In the other way, there is no possibility recognizing which SDOs are the components of an unexpired real data.

The time cost of a brute-force attack would be N square while N is the total number of records in the system. Benefiting from SSL protocol, the communication between client and server can be resistant to Man-in-the-Middle attack.

B. *Performance*

In aspect of performance, researchers measured the performance of the system, focusing on the times to store and retrieve SDOi from database. Considering in the self-destruct data method, the cipher data split into pieces according to a specific length causing more database operation compared to the normal situation in which the cipher data is directly saved by only one operation. Our goal was to determine what size the private data has reached, the reduction of performance will cause unbearable experience to user. Our measurements use a CPU Inter Core2 Duo 2.66GHz with 4GB of RAM, Java 1.7, MySQL 5.5 and a broadband network. AES key is 256bit and the length of piece is 64KB.

Researchers tested read/write performance in situations of using or not using self-destruct data method. The size of single data ranges from 32KB to 10MB.

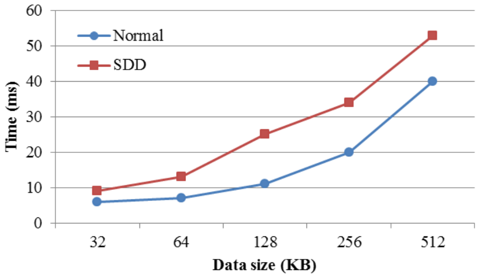


Figure 5. Write data less than 1MB

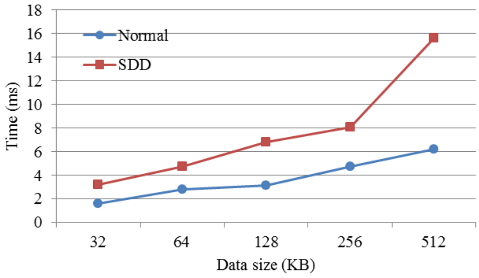


Figure 6. Read data less than 1MB

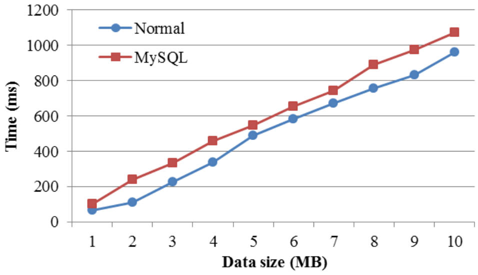


Figure 7. Write data more than 1MB

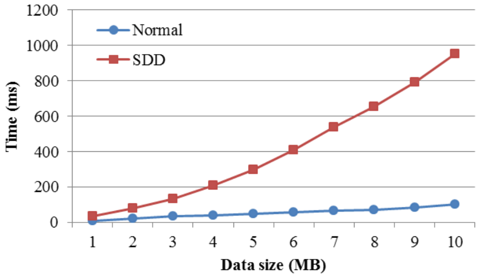


Figure 8. Read data more than 1MB

As shown in Fig. 5-8, SDD indicates the situation using self-destruct data method, whereas normal indicates not. What researchers can conclude from the graphs is that the cost of time is linearly independent with data size. When the data size is less than 1MB, the reduction of performance is fairly small which is about a few milliseconds to tens of milliseconds. When the data size is more than 1MB, because of too many pieces, more operations are needed, performance reduction is relatively obvious. In general, the size of email content is at scale of KB, so in the self-destruct email system, the reduction of user experience is very little.

VI. CONCLUSIONS

In the era of cloud computing, more attention has been paid to the protection of user privacy and the security of user data. In this article, researchers design the self-destruct data method based on privacy cloud and apply it to realize the self-destruction of e-mail content after expiration time, in this way, the confidentiality and the privacy of the emails can be protected. To store the cipher data, a stand-alone storage system is used in this system.Researchers suggest the storage system to be not only enterprise private clouds, but also the trusted public clouds. To improve the security of emails, identity authentication mechanism will be added to the system in the future, only those who have got accredits can use privacy cloud services through the browser plug-in. Privacy cloud data is certainly not confined to e-mails, e-mail is just a typical example. Researchers aimed at creating a common platform for users, on which users create, manage all sorts of data containing sensitive information and share data proxy to various web applications. To ensure the real controllability and security of user privacy, the applications can only obtain and use users' information within a certain period of time. Information security is the combination of technologies and consciousness.Researchers believe that if users' security awareness is improved, more and more attention will be paid to the research.

ACKNOWLEDGMENT

The work is supported by the Natural Science Foundation of Jiangsu Province (No. BK20130852).

REFERENCES

[1] Geambasu R,Kohno T,Levy A,and Levy M, "Vanish:Increasing data privacy with self-destructing data". Proceedings of the 18th USENIX Security Symposium. Montreal,Canada, 2009:299-315

[2] Wang G,Yue F, and Liu Q. "A secure self-destructing scheme for electronic data".Journal of Computerand System Sciences, 2013,79(2):279-290

[3] Wang Li-Na,Ren Zheng-Wei, and Yu Rong-Wei, "A data assured deletion approach adapted for cloud storage".Acta Electronica Sinica,2012,40(2):266-272

[4] Wolchok S,Hofmann O S,Heninger N, "Defeating vanish with low-cost sybil attacks against large DHTs". 17th Annual Network & Distributed System Security Conference.San Diego,USA,2010:1-15

[5] Zeng L,Shi Z,Xu S, "SafeVanish:An improved data self-destruction for protecting data privacy", IEEE Second International Conference on Cloud Computing Technology and Science.Athens,Greece,2010:521-52

[6] Cloud Security Alliance: Security Guidance for Critical Areas of F ocus in Cloud Computing V2.1 http://www.cloudsecurityalliance.o rg/guidance/csaguide-cn.v2.1.pdf

[7] X. Fu, "A network-based data self-destruction method", CN102571949A, Beijing: State Intellectual Property Office, 2012.7.11

[8] Yu, Xiaojun, and Q. Wen. "A View about Cloud Data Security from Data Life Cycle." Computational Intelligence and Software Engineering (CiSE), 2010 International Conference on IEEE, 2010:1 - 4.

[9] Pub N F. 197: Advanced encryption standard (AES) [S]. Federal Information Processing Standards Publication, 2001, 197: 441-0311.

[10] KLENSIN J. RFC 2821: Simple Mail Transfer Protocol [J]. AT&T Laboratories.2001.