Secret Shares to Protect Health Records in Cloud-based Infrastructures

Prof William J Buchanan, Elochukwu Ukwandu
Edinburgh Napier University, Edinburgh. UK
w.buchanan@napier.ac.uk

Dr Nicole van Deursen, Dr Lu Fan, Dr Gordon Russell, Dr Owen Lo, Prof Christoph Thuemmler
Edinburgh Napier University, Edinburgh. UK

Abstract— **Increasingly health records are stored in cloud-based systems, and often protected by a private key. Unfortunately the loss of this key can cause large-scale data loss. This paper outlines a novel Cloud-based architecture (SECRET) which supports keyless encryption methods and which can be used for the storage of patient information, along with supporting failover and a break-glass policy.**

Keywords—secret shares, keyless encryption, patient records

I. INTRODUCTION

Healthcare records are one of the most sensitive areas for protection, but two-thirds of healthcare organisations experienced a security incident in 2014 [9, 11]. Patients who have had their records lost or stolen are at risk of medical and financial identity theft. These patients will feel embarrassed, may suffer from mistreatment, and experience a stressful and costly road to repair the damage. Stealing medical records is an attractive criminal business, as the data gained could be worth at least ten times the value of credit card data on the black market [10]. The number of healthcare data breaches rises at a worrying pace. Since last year, medical identity theft incidents increased 21.7% [12], and there are forecasts that healthcare breaches will keep increasing in the near future [8], due to the potential economic gain and digitization of records.

The most important causes of breaches are organizational threats such as employee negligence, and computer-related threats such as cyber-attacks and public cloud services [9, 11].

Cloud services are often advised as a counter-measure against insider threats. It is thought that removing the physical data from the premises, and thus making it more difficult to maintain personal relationships with the persons who have access to the data, will lower the risk of a data breach. However, due to many data breach reports in the media, cloud services are also mistrusted and approximately half of the respondents (52% in the US; 49% globally) of a study by BT, admit that they are ‘very or extremely anxious’ about the security implications of these services [13]. The number of cloud data breaches are highest in the entertainment industry, but health care ranks second [7] and is rising. More than 83% of hospitals and health systems are using the cloud for at least some technology, according to a recent HIMSS Analytics survey of 150 organizations. About half are using the cloud for clinical operations, and around three-quarters are using it for administration. To protect their reputation and their patients, healthcare needs to make data security a top priority.

II. MAJOR DATA BREACHES AROUND HEALTH CARE DATA

In the US, there have been a large number of data breaches including:

- In 2009, Blue Cross Blue Shield of Tennessee lost 1.02 million patients. Cause: stolen hard disks.
- In 2011, 4.9 million user records was stolen from a Science Applications International Corporation. Cause: records left in employee’s car in 2011. Also, Health Net lost 1.9 million records. Cause: Loss of unencrypted server hard drives from data center. In 2011, Nemours Foundation lost ten years of data on 1.05 million patients. Cause: theft of backup tapes.
- In 2013, Advocate Health and Hospitals Corporation lost 4.03 million individuals. Cause: stolen computers.
- In 2014, Complete Health Systems, based in Tennessee lost 4.5 million people. Cause: Network hack. Also Montana Department of Public Health and Human Services lost 1.06 million people. Cause: Network hack.

We analysed two publicly available sources of original healthcare data breach information from the UK and the U.S. in 2014 to find out which factors contribute to incidents. In the UK, the website of the ICO [14] lists the number of breaches that are reported, and details the most important ones that lead to prosecution or fines. From this list, we found that healthcare is the sector with the highest number of reported data breaches.

We analysed the undertakings related to healthcare in 2014 and created a list of actions contributing to incidents in healthcare (Figure 1). Health Care services could benefit greatly from the usage of Cloud Computing, but many risks still exist, including security and performance. In a bid to provide a solution to this, [3] presented a novel architecture and its implementation for inter-organizational data sharing, which provides a high level of security and privacy for patient data in semi-trusted cloud computing environments. This architecture features attribute-based encryption for selective access authorization and cryptographic secret sharing in order to disperse data across multiple clouds, reducing the

978-1-4673-8325-7/15/$31.00 ©2015 IEEE

adversarial capabilities of curious cloud providers. An implementation and evaluation by several experiments demonstrate the practical feasibility and good performance.

Rabin's IDA has been deployed in sharing health data alongside Shamir's Secret sharing scheme in a multi-cloud environment, the result shows that the adoption of Rabin's IDA would create a low overhead and more feasible that Shamir's Secret sharing scheme, as Shamir's reconstruction phase had a huge marginal increase when the document size or the threshold grows despite the optimisation by Lagrange interpolation [12].

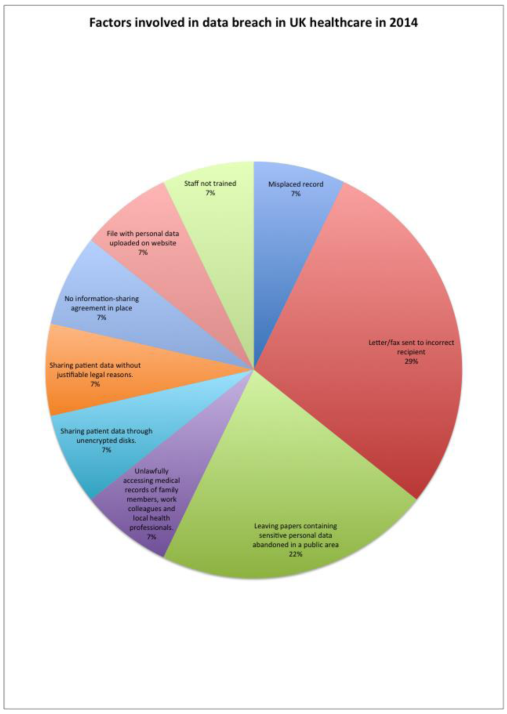


Figure 1: Factors involved in data breach

III. Secure Data Storage of Health Care Records

Health care records are one of the most sensitive documents, and for them to be stored in a cloud-based system they must be properly controlled for access and rights to data. A major risk is around the insider threat, especially anyone with System Administrator rights, as they can often gain access to encrypted content. Many existing architectures for cloud-based systems have often just scaled from legacy infrastructure, with the additional of encryption. The main weakness is often through a loss of the key used to encrypt the data. New types of Cloud-based architectures have been proposed for health care, but these still often rely on the usage of encryption key-pairs (with a public and a private key) to protect the encryption key. A breach of the encryption key used, such as from an insider, can cause a major data leak.

Overall the important attributes for the support of health care records in cloud-based systems we have the following:

- **Keyless encryption method**. While the methods employed in modern encryption are often water-tight, it is the actual implementation in real-life systems which is often flawed. Most systems still rely on digital certificates holding a key-pair, and these certificates are often easily stolen or compromised through brute force. The Superfish compromise should how sloppy some developers can be, where a digital certificate with both the public and private key was distributed with the software, and which has a default password with the name of the company who developed the software [6].
- **Self-destruct.** The secret sharing method supports a self-destructing data system, where all the information and their copies, as well as decryption keys become destructed after a user-specified time, without any user involvement. SeDas [1] causes sensitive information, such as account numbers, passwords and notes to irreversibly self-destruct, without any action on the user's part. It is applicable in a multiple server-based systems with operational overhead.
- **Break-glass data recovery.** Many systems have strict access control policies for data, but in some circumstances, such as for a life-threatening situation, it is important to have access to data in emergencies. This is typical in health and social care. The loss of an encryption key can also cause major problems in data recovery.
- **In-built failover protection.** The usage of secret shares allows the striping of data across multiple cloud-based storage systems using an any k-from-n sharing policy, such as with storage across three Cloud storage systems. For example a policy of 2-from-3 will support one of the Cloud storage systems failing, and the original data will still be able to be recovered.
- **Both public and private Cloud storage should be protected.** With the increasing risk of insider threats, there should be no differentiating in security between internal and external storage of data in Cloud-based system within health care.
- **Every data object should be protected.** There is an increasing risk of large scale data loss if a single key is used for data, thus every objective should be protected with its own security protection.

IV. Sharing Methods

1) *Perfect Secret sharing scheme (PSS)*

Shamir provides a good example of a perfect secret sharing [15]. Its perfectness is measured by two parameters: if, and only if, t-1 shares, provide absolutely no information regarding the hidden secret, and when the ratio of the length of the secret to the length of each of the shares known as information rate, is equivalent to 1.

Shamir's PSS relies on the idea that it takes two points to define a straight line, three points to fully define a quadratic, four points to define a cubic curve, … and T points to define a polynomial of degree T-1. Hence, a method for T-out-of-N secret sharing is to create a polynomial of degree T-1 with the secret as the first coefficient and the remaining coefficients picked at random. Next, find N points on the curve and give one to each of the players. As a result, when at least T out of the N players reveal their points, there is sufficient information to fit a (T-1)th degree polynomial to them, in which the first coefficient being the secret.

2) *Information Dispersal Algorithm (IDA)*

Rabin suggested splitting a secret S into n pieces such that a person can obtain the secret only if $k < n$ of these pieces are available, where k is the threshold. Here, each secret S_i, $i \leq n$, is of size |S|/k, where |S| is the size of the secret. The total sizes of all the secrets are:

$$(n/k) \times |S|$$

Thus, with Rabin's Information Dispersal Algorithm, the storage complexity of a secret sharing system can be significantly reduced in comparison to Shamir perfect secret sharing (PSS) scheme. But, the security flaw in this method is that, if the data exhibits some pattern frequently, and that the attacker gets hold of $m < k$ slices, there are great possibilities for him gaining the secret S.

3) *Krawczyk's Computational Secret Sharing*

Hugo Krawczyk [4] proposed the Computational Secret Sharing (CSS) technique (a.k.a. *secret sharing made short*), which combines Rabin's IDA with Shamir's PSS. Data is first encrypted with a randomly generated key, using a symmetric encryption algorithm. Next this data is split into n fragments using Rabin's IDA with a threshold t configured. In this case, the scheme is t times more efficient than Shamir's PSS.

4) *Publicly Verifiable Secret Sharing (PVSS)*

Some applications of a secret sharing scheme, such as e-auction, e-voting and multiparty computation, require public verifiability. Namely, it must be publicly verifiable without revealing the secret or any of its shares that all of the n shares are consistently generated from a unique share-generating polynomial such that any t of them can reconstruct the same secret. This capability is crucial to overcome the problem of dishonest share dealers.

Tal Rabin [16, 5] devised a publicly verifiable secret sharing system that allows players to detect dishonesty on the part of the dealer or on part of up to one-third of the threshold number of players, even if those players are co-ordinated by an adaptive attacker who can change strategies in real-time depending on what information has been revealed. The PVSS is essentially a combination of secret sharing and publicly verifiable encryption.

V. SECRET Architecture

The secret sharing methods show potential in being used in Cloud-based infrastructures, as they inherently preserve the data without the requirement for private keys. Figure 2 outlines a novel architecture known as SECRET which supports a secret sharing scheme in a multi-cloud environment. It has five main elements:

- **Application Platform.** Its function is to: determines the access structure; encodes secrets; sends secrets to the main multi-cloud proxy server for distribution to multi-cloud service providers, as well as keeping the secret shares when recovered.
- **Main Multi-Cloud Proxy Server with Router.** This splits and distributes encoded shares to multi-cloud based on pre-determined access structure and manages the fail-over protection of shares.
- **Metadata server**. This includes the functionality of: User management; Server management; Session management; and File metadata management [14].
- **Multi-cloud Proxy server**. This Gathers shares and reconstruct secret as well manages break-glass data recovery.
- **Sub-Routers.** This creates a path between Cloud Service Provider (CSP) 1 (considered here as front-end) with other Cloud Service providers (considered here as the Back-ends) thereby creating a quick and alternative recovery path for all the shares. For example: R4 connects with R3+R2+R1, so R4 is a path for CSP1+R3+ R2+R1+S1, and so on.

The process entails that when the document/key/secret to be shared is called up at the Application platform with the determination of user's access level and authentication. Initially the document is first encoded in relation to the pre-determined access structure of the scheme to be used, such as for Shamir's Secret Sharing scheme. In Figure 2 we have a 3-out-of-5 access structure. The encoded secret is sent to the multi-cloud proxy server for onward dissemination to the CSPs. Next the Application platform maintains a close interaction between the Proxy servers and Metadata server for effective operation of the system.

The design incorporates unique features in implementing a secret sharing scheme in a multi-cloud environment and implements keyless encryption method. This is done by breaking the secret/key into chunks called (k-out-of-n) threshold in such a manner that less than k shares cannot recover the secret. The incorporation of a Self-Destructive system solves the problem of cloud user's privacy as there is no way the user's data can be accessed, copied, cached or used without the data owner's consent within a pre-determined time-frame as all data and their copies become destructed or unreadable after a user-specified time, without any user intervention [16]. The self-destructive system defines two modules: a self-destruct method object; and survival time parameter for each secret key part. In this case, a secret sharing algorithm is used to implement share distribution in an object storage system so as to ensure safe destruct with equally divided shares. Based on active storage framework, object-based storage interface will be used to store and manage the equal divided shares.

The design also incorporates **Break-Glass data recovery** and which is implemented using one of the Proxy Servers. An access to Multi-Cloud proxy Server II entails an access to CSPs1, 3 and 5 and this in turn ensures a quick recovery of shares in order to reconstruct the secret as it is a quick link to all other CSPs and moreover, following the access structure, such access ensures the possibility of reconstructing the secret in an emergency situation. This is an important feature as there could be a period of cloud outage as stated in [2], and in such situation, data recovery could be done from 3-out-of-5 Cloud service providers being used for data storage. That is to say, if 2-out-of-the-5 cloud service providers fail, data recovery is still possible in such an extreme condition.

VI. Conclusions

Within health care there is a move toward cloud-based systems, but these often still use private keys to secure the data, and where the loss of the key could cause a large-scale loss of data. With the secret share method, every piece of data is secured, and can only be recovered when the other shares are brought back together.

VII. References

[1] Rao, C., Rodi, P., Palande, A., & Bhusari, V. (2015). SEDAS: A Self-Destructing Data System Based on Shamir's Secret Sharing Algorithm. International Journal Of Scientific Research And Education, 3(03).

[2] Srinivasan, S. (2014). Building trust in cloud computing: Challenges in the midst of outages. Proceedings of Informing Science & IT Education Conference (InSITE) 2014 (pp. 305-312). Retrieved from http://Proceedings.InformingScience.org/InSITE2014/InSITE14p305-312Srinivasan0544.pdf

[3] Padsala, C., Palav, R., Shah, P., & Sonawane, S. (2015). Survey of Cloud Security Techniques. International Journal for Research in Applied Science & Engineering Technology, 3(3) pp. 47-50.

[4] H. Krawczyk, "Secret Sharing Made Short," in Proceedings of the 13th Annual International Cryptology Conference on Advances in Cryptology, 1993.

[5] K. Peng, "Critical survey of existing publicly verifiable secret sharing schemes", Information Security, vol. 6, no. 4, pp. 249-257, 2012.

[6] Buchanan W.J, Storing The Keys to your House Under a Plant Pot, https://www.linkedin.com/pulse/storing-keys-your- house-under-plant-pot-william-buchanan

[7] Elastica. 2015. Q2 2015 Shadow Data. Available from: https://www.elastica.net/q2-2015-shadow-data-report/

[8] Experian. 2015. 2015 Second Annual Data Breach Industry Forecast.

[9] HIMSS. 2015 HIMSS Cybersecurity Survey. Available from: himms.org.

[10] Humer, C., Finkle, J. "Your medical record is worth more to hackers than your credit card." Reuters.com U.S. Edition, 24 Sep 2014. Available from: http://www.reuters.com/article/2014/09/24/us-cybersecurity-hospitals-idUSKCN0HJ21I20140924?feedType=RSS&feedName=healthNews

[11] Ponemon Institute. 2015. Fifth Annual Benchmark Study on Privacy & Security of Healthcare Data.

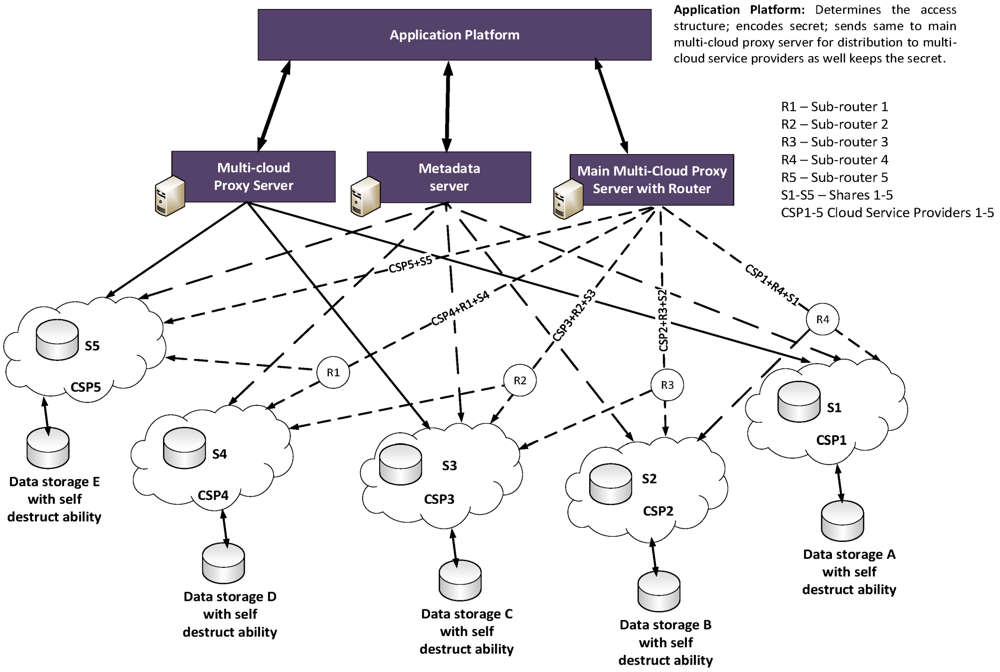
[12] Ponemon Institute. 2015. Fifth Annual Study on Medical Identity Theft.

[13] BT. 2014. Business trust in data security in the cloud at an all-time low. Available from: http://www.globalservices.bt.com/uk/en/news/business_trust_in_data_security_in_cloud_at_all_time_low

[14] ICO. 2015. Actions we have taken. https://ico.org.uk/action-weve-taken/data-breach-trends/

[15] Zeng, L., Chen, S., Wei, Q., & Feng, D. (2012). Sedas: a self-destructing data system based on active storage framework. In APMRC, 2012 Digest (pp. 1-8). IEEE.

[16] T. Rabin and M. Ben-Or, "Verifiable secret sharing and multiparty protocols with honest majority," in Proceedings of the Annual ACM Symposium on Theory of Computing, 1989.



Application Platform
Multi-cloud Proxy Server
Metadata server
Main Multi-Cloud Proxy Server with Router
CSP5+S5
CSP4+R1+S4
CSP3+R2+S3
CSP2+R3+S2
CSP1+R4+S1
R1
R2
R3
R4
S5
CSP5
S4
CSP4
S3
CSP3
S2
CSP2
S1
CSP1
Data storage E with self destruct ability
Data storage D with self destruct ability
Data storage C with self destruct ability
Data storage B with self destruct ability
Data storage A with self destruct ability
R1 – Sub-router 1
R2 – Sub-router 2
R3 – Sub-router 3
R4 – Sub-router 4
R5 – Sub-router 5
S1-S5 – Shares 1-5
CSP1-5 Cloud Service Providers 1-5

Application Platform: Determines the access structure; encodes secret; sends same to main multi-cloud proxy server for distribution to multi-cloud service providers as well keeps the secret.

Figure 2: Proposed architecture of a Secret Sharing Scheme in a Multi-cloud environment