Designing a Secure Cloud-Based EHR System using Ciphertext-Policy Attribute-Based Encryption

Suhair Alshehri, Stanisław Radziszowski, and Rajendra K. Raj
Golisano College of Computing & Information Sciences
Rochester Institute of Technology
Rochester, New York 14623, USA
sxa3788@rit.edu, spr@cs.rit.edu, rkr@cs.rit.edu

ABSTRACT

As more and more healthcare organizations adopt electronic health records (EHRs), the case for cloud data storage becomes compelling for deploying EHR systems: not only is it inexpensive but it also provides the flexible, wide-area mobile access increasingly needed in the modern world. However, before cloud-based EHR systems can become a reality, issues of data security, patient privacy, and overall performance must be addressed. As standard encryption (including symmetric key and public-key) techniques for EHR encryption/decryption cause increased access control and performance overhead, the paper proposes the use of Ciphertext-Policy Attribute-Based Encryption (CP-ABE) to encrypt EHRs based on healthcare providers' attributes or credentials; to decrypt EHRs, they must possess the set of attributes needed for proper access. This paper motivates and presents the design and usage of a cloud-based EHR system based on CP-ABE, along with preliminary experiments and analyses to investigate the flexibility and scalability of the proposed approach.

Categories and Subject Descriptors

E.3 [**Data Encryption**]: Standards; H.2 [**Database Management**]: Systems; J.3 [**Life and Medical Sciences**]: Health, Medical Information Systems

General Terms

Algorithms, Design, Security, Theory

Keywords

Cloud Computing, Electronic Health Records, Attribute-Based Encryption, Ciphertext Policy, Security, Privacy

1. INTRODUCTION

The Health Information Technology for Economic and Clinical Health (HITECH) Act [26] provides federal incentives to encourage U.S. healthcare providers to adopt and use EHR systems meaningfully to improve healthcare quality. Cloud computing has been viewed as an appropriate platform to deploy electronic health records (EHRs) systems for its cost-effective services (including data management and storage, and computational resources) and features (portability, reliability, scalability, and elasticity) delivered by cloud service providers [5]. Security and privacy issues, however, have raised difficulties for the adoption of cloud-based EHR systems [10].

In the United States, compliance to HIPAA (Health Insurance Portability and Accountability Act) [25] is often cited as the requirement to preserve the confidentiality of medical records including EHRs. For completeness, it should be noted that HIPAA-compliance is not sufficient as HIPAA does not provide strong guidance to address current security issues [10]; for example, HIPAA does not mandate the use of encryption mechanisms, as stated in associated regulations such as 45 CFR Part 164.306 and 164.312 [27]. As cloud service providers are not trusted to store EHRs unencrypted, even when access controls are in place [5], EHR encryption must be required in cloud-based EHR systems.

Standard encryption techniques are not well suited for EHR systems, especially in cloud-based settings:

- **Symmetric-Key Encryption (SKE).** These techniques, e.g., AES, are usually efficient but introduce complexity in EHR systems as additional mechanisms are required to apply access control. In particular, all healthcare providers use one shared key for encryption and decryption; thus, if the shared key is compromised, all EHRs are compromised.

- **Public-Key Encryption (PKE).** These techniques, e.g., RSA, provide a secure solution but are not practical for secure EHR storage due to the requirement for an expensive public-key infrastructure (PKI) to be maintained for distributing and managing public keys and digital certificates for all healthcare providers.

These inadequacies with standard encryption techniques in supporting cloud-based EHR systems motivate the investigation of other approaches. This paper builds on Ciphertext-Policy Attribute-Based Encryption (CP-ABE) [2], a novel cryptographic approach that utilizes user roles for secure handling of data.

The two main contributions of this short paper are the proposed design for a secure cloud-based EHR system using CP-ABE, and a preliminary investigation of performance issues with secure cloud-based EHR systems.

The rest of this section reviews the necessary background in cryptography. Section 2 presents the design of the proposed cloud-based EHR system based on CP-ABE. Preliminary experiments and analyses of the proposed design are presented in Section 3. Related work in cryptography and EHR systems is discussed in Section 4.

1.1 Background

We review the needed background in cryptography to pave the way to the discussion of CP-ABE in the next section.

- **Elliptic Curve Cryptography** (ECC) is a type of public-key cryptography (PKC) that is based on the algebraic structure of elliptic curves over finite fields. The security of ECC is based on the hardness of the elliptic curve discrete logarithm problem, and achieves RSA-equivalent security with a much smaller elliptic curve group; for example, a 163-bit key in ECC is considered to be as secure as 1024-bit key in RSA [6]. ECC implementations use less memory and processing power, which allows them to be used on compact platforms such as smart phones and smart cards.

  Definition 1 (Elliptic Curves over Finite Field): *Let F_p be a finite field where $p > 3$ is a prime, and $a, b \in F_p$ such that*

  $$4a^3 + 27b^2 \neq 0 \ (mod\ p) \in F_p.$$

  An elliptic curve $E[F_p]$ is the set of solutions (x, y) to the equation

  $$y^2 = x^3 + ax + b \ (mod\ p) \in F_p[x],$$

  together with the point at infinity 0.

  Addition of points on ECC is defined to form the so-called elliptic curve group, for $P \in E(F)$, nP denotes $P + P + \ldots + P$, n times.

  Definition 2 (Elliptic Curve Discrete Logarithm Problem): *If E is an elliptic curve over a field F, then the elliptic curve discrete logarithm to base $Q \in E(F)$ is the problem of finding an $n \in Z$ such that $P = nQ$ for a given $P \in E(F)$.*

- **Bilinear Maps** construct a relationship between two cryptographic groups leading to new schemes.
  Definition 3 (Bilinear Maps): *Let G1 and G2 be cyclic groups of prime order p; g a generator of G1. e is a bilinear map, e: $G1 \times G1 \rightarrow G2$, where $|G1| = |G2| = p$. The bilinear map e has three properties:*

  - *Bilinearity: $\forall\ P,Q \in G1, \forall\ a,b \in \mathbb{Z}^*_p$, $e(aP, bQ) = e(P, Q)^{ab}$,*
  - *Non-Degeneracy: $P \neq 0 \Rightarrow e(P, P) \neq 1$,*
  - *Computability: e is efficiently computable.*

- **Attribute-Based Encryption** (ABE) extends Identity-Based Encryption (IBE), originally proposed by Adi Shamir [24], by using a public key as an arbitrary string to identify a user. Boneh and Franklin's pairing-based encryption scheme [3] was the first to use fully functional IBE that is based on a novel solution of a pairing on groups of elliptic curves over finite fields.

  Sahai and Waters subsequently introduced a new type of IDE called Fuzzy Identity-Based Encryption (FIBE) [23]. In FIBE, a private key is associated with a set of attributes, ω, and able to decrypt ciphertexts encrypted with a set of attributes, ω', if and only if at least k attributes overlap between ω' and ω. FIBE's motivation was to design an error-tolerant IBE that uses biometric identities as public keys. IBE and FIBE have limited applications, as they not allow for a scalable and fine-grained access control to ciphertexts.

  Bethencourt, Sahai, and Waters [2] first constructed CP-ABE in which private keys are labeled with sets of attributes and ciphertexts are associated with access structures consisting of AND and OR gates. Current implementations of CP-ABE are typically based on the construction of a bilinear mapping between two elliptic curve groups [2, 9, 11].

2. A CLOUD-BASED EHR SYSTEM

This section explores a proposed design—based on CP-ABE—for a cloud-based EHR system that provides secure EHR storage, with flexible, fine-grained access control.

2.1 The CP-ABE Scheme

Following Bethencourt, Sahai, and Waters [2], in our CP-ABE scheme, healthcare providers share one public key for encryption, thus avoiding PKI; however, each healthcare provider has a distinct secret key for decryption. CP-ABE supports complex policies to specify which secret keys can decrypt which ciphertexts: each healthcare provider's secret key is labeled with a set of attributes, and ciphertexts are associated with access policies. The secret key of a healthcare provider can decrypt a particular ciphertext only if the attribute set of the healthcare provider's key satisfies the access policy associated with that ciphertext, as illustrated in Figure 1. Here, the nurse practitioner with the ABCD Medical Group can access EHRs that are only allowed to physician assistants *or* nurse practitioners, *and* who work in the ABCD Medical Group; and the physician assistant with the WXYZ Medical Group is not allowed access.

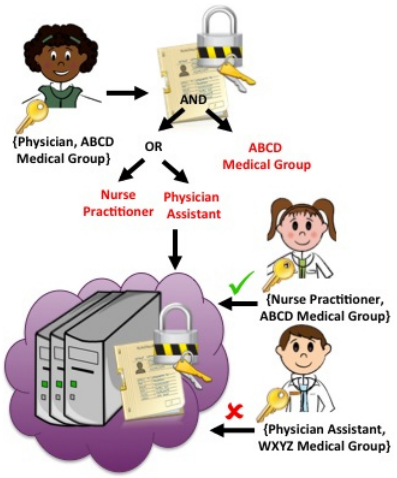

Figure 1: Using CP-ABE in a cloud-based EHR System

A CP-ABE scheme consists of four fundamental algorithms: Setup, Encrypt, Key Generation, and Decrypt, and one optional algorithm, Delegate.

- **Setup:** the setup algorithm takes no input other than the implicit security parameter. It outputs the public parameters PK and a master key MK.
- **Key Generation(MK, S):** The key generation algorithm uses the master key MK and a set of attributes S that describe the key, and outputs a private key SK.
- **Encrypt(PK, M, $\mathbb{A}$):** The encryption algorithm takes as input the public parameters PK, a message M, and an access structure $\mathbb{A}$ over a set of attributes. It will encrypt M and produce a ciphertext CT such that only a user who possesses the set of attributes satisfying the access structure will be able to decrypt CT.
- **Decrypt(PK, CT, SK):** The decryption algorithm takes as input PK, a ciphertext CT, which was obtained for an access policy $\mathbb{A}$, and a private key SK for a set S of attributes. If the set S of attributes satisfies the access structure $\mathbb{A}$, then the algorithm will decrypt the ciphertext and return a message M.
- **Delegate(SK, S'):** The delegate algorithm takes as input a secret key SK for some set of attributes S and a set $S' \subseteq S$. It outputs a secret key SK' for the set of attributes S'.

CP-ABE thus supports flexible and fine-grained access control with healthcare providers being able to access only relevant EHRs encrypted with access policies that satisfy their keys' attributes. Also, if a secret key is compromised, only EHRs that can be decrypted with that key will be compromised; other EHRs are still protected.

2.2 System Architecture

In our scheme, EHRs are stored in the cloud, and can be assessed through a web portal by multiple owners and users. Owners, who create EHRs, are responsible for generating access policies based on the attributes of authorized healthcare providers, encrypting EHRs based on the generated policies and uploading encrypted EHRs into the cloud. EHRs are organized into a labeled hierarchical data structure [1], which makes it possible to share different parts of the EHR, thus making the scheme more flexible.

Figure 2 shows the architecture for the proposed cloud-based EHR system, which consists of three main components: the cloud-based EHR system, Healthcare Providers (owners and users), and the Attribute Authority (AA). The system uses two fundamental cloud services: data storage and computing resources. The first service is for storing encrypted EHRs that are accessible only to healthcare providers through authentication mechanisms, and access policies based on complete attributes of healthcare providers. The second service is for hosting the web portal, generating access policies, and performing other needed computing tasks.

Once healthcare providers obtain their private keys from the AA, they log in to the system using their username and password; on first login, they will need to download and install lightweight software for encrypting and decrypting EHRs locally. When a healthcare provider requests access to an encrypted record, she will first locate and download it, and then use her key and the lightweight software to decrypt it. To upload a new record, she will first request the desired attributes and generate the access policy using the Access Policy Engine; encrypt the record using the lightweight software; and finally upload the encrypted record.

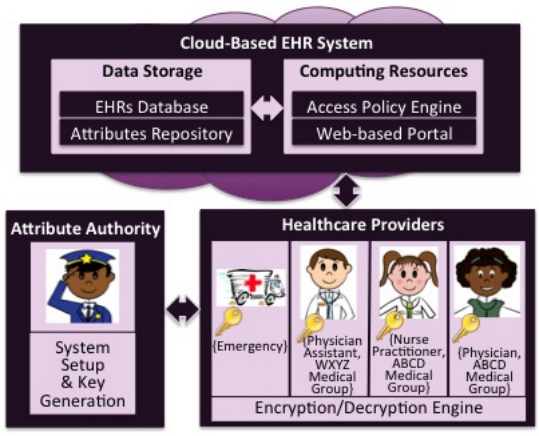

Figure 2: The Cloud-Based EHR System Architecture

2.3 Key Management

Key management must be cost-effective and and its components carefully implemented in a cloud-based EHR system.

- **Generation and Distribution:** The AA generates the public key and the master private key using the Setup algorithm. When a new healthcare provider joins the system, the AA derives a distinct secret key associated with her attributes by running the Key Generation algorithm off-line. Healthcare providers must store their secret keys securely, and regenerate keys after a predefined expiration date. Secret key regeneration is performed without the need to refresh the system parameters, public key, and master private key. Healthcare providers can only seek their secret keys through their healthcare organizations. Our design does not require backward secrecy: when a new healthcare provider joins the system, she should be able to access all previously encrypted EHRs if and only if the attributes associated with her key satisfy the access policies associated with the encrypted EHRs.
- **Revocation:** The cloud-based EHR requires forward secrecy such that when a healthcare provider's access is revoked, she should not be allowed to access EHRs that she was able to access before being revoked. Owners of encrypted EHRs have the option to add an expiration date to access policies used for encryption, or to re-encrypt them with updated access policies to prevent access by revoked healthcare providers. The problem of re-distributing secret keys is thus avoided.
- **Escrow:** In the cloud-based EHR system, the AA can regenerate secret keys for healthcare providers to access EHRs during emergencies.

In summary, key management is handled appropriately for cloud-based settings in our system.

3. PRELIMINARY ANALYSIS

To evaluate the feasibility of our proposed cloud-based EHR system based on CP-ABE, we conducted several preliminary experiments to measure overhead in terms of time

and storage. The experiments were run on a virtual machine running Fedora 14 with 1GB of RAM, and hosted on 2 x Intel Xeon E5520/2.26GHz. The CP-ABE implementation uses a constructed 160-bit elliptic curve group on the curve $y^2 = x^3 + x$ over a 512-bit finite field [8]. It also uses the Pairing Based Cryptography library to perform the pairing-based cryptosystems [15].

3.1 Time Overhead

To measure the efficiency of the encryption and decryption algorithms, five image files (1MB, 10MB, 20MB, 30MB, 40MB, and 50MB) were encrypted with six different numbers of attributes in the access structure, and then decrypted with a secret key that is associated with ten attributes. Encryption time increases almost linearly with the number of attributes in the access structures as access trees need to be created. However, decryption time tends to be independent of the number of attributes as we are using a single key. Each experiments was repeated five times, and the elapsed time was averaged to yield the efficiency measure.

In the context of EHR systems, records less than 1MB are encrypted in less than 0.2 seconds depending on the number of attributes in the access structures, and decrypted in less than 0.1 seconds, as shown in figures 3 and 4. On the other hand, medical images that are 30MB, are likely to be encrypted in less than 0.5 seconds for the number of attributes in the access structures, and decrypted in less than 0.4 seconds.

These preliminary results indicate that the time performance of CP-ABE is feasible for EHR systems.

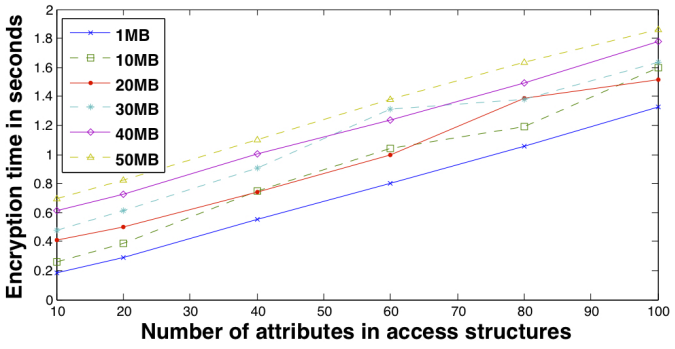

Figure 3: Encryption time for varying number of attributes in the access structure (for five image sizes)

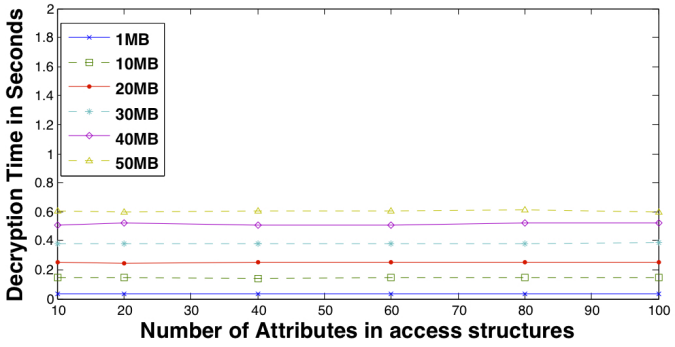

Figure 4: Decryption time for varying number of attributes in the access structure (for five image sizes)

3.2 Storage Overhead

To determine storage overhead, we measured the difference between encrypted records and decrypted records across various numbers of attributes in the access structures. The difference is almost identical for all the files and increases linearly with respect to the number of attributes in the access structures. The difference is less than 4KB for simpler access structures and less than 30KB for complex access structures.

The results suggest that storage overhead is negligible in the context of cloud-based EHR systems because cloud service providers currently (and will continue to) provide large amounts of storage at reasonable prices.

4. RELATED WORK

In this paper, the terms electronic medical record (EMR), electronic health record (EHR), and personal health record (PHR) systems follow the definitions provided by the National Alliance for Health Information Technology [20]. Numerous EMR, EHR, and PHR designs and systems [13, 16, 18, 22] have been proposed by both the academic community and the IT industry to deal with securely managing EHRs.

EMR systems manage electronic health records that can be created, gathered, and managed by authorized healthcare providers *within one health care organization*. Benaloh et al. [1] propose a patient controlled encryption (PCE) EMR system that enables patients to mediate access control decisions over their medical records. EHRs are partitioned into a hierarchical structure in which each section is encrypted with a derived public key that patients are required to manage, and decrypted with a derived subkey from a master private key. This system has several issues: potential key management overhead, no support for a key escrow agent in emergencies, and likely data integrity issues as EHRs are managed by patients, not healthcare providers.

EHR systems manage electronic health records that can be created, gathered, and managed by authorized healthcare providers across *more than one* health care organization. Though several EMR and PHR systems have been proposed, there is little research in the area of cloud-based EHR systems. Commercial cloud-based EHR systems such as Practice Fusion [21] and CareCloud [4] are available; although they are HIPAA compliant, the security of EHRs is only based on access control decisions mediated by their cloud storage providers. EHRs are stored unencrypted, which has security and privacy implications.

PHR systems manage personal electronic health records that are *imported* from EMR and EHR systems by patients. Narayan et al. [19] introduce a PCE-PHR system that enables a patient to import EHRs into the cloud, and encrypt each record with SKE using different keys. Along with each encrypted record, the patient uploads a corresponding entry consisting of encrypted metadata using a broadcast ABE, unencrypted access policy, and a search-index. For healthcare providers to access a record, they must decrypt the metadata entry to find the location and name of the record and the symmetric key. They then request the encrypted record from the cloud server, and decrypt it using the symmetric key. In Narayan's scheme, only patients can be the owners of EHRs, and so this scheme cannot be used for cloud-based EHR systems. HealthVault [17], a cloud-based PHR service offered by Microsoft, and the soon-to-be dis-

continued Google Health [7] service, both guarantee encryption during EHR transmission from patient to the cloud and back, but to secure EHR storage, they seem to rely primarily on proper access control and limiting physical access rather than full encryption.

Finally, substantial attention has been paid in recent years to querying encrypted data stored externally or in the cloud, e.g., Li at al. [12] or Liu et al. [14], but our focus is on efficient and effective access control on cloud data.

5. CONCLUSIONS

This short paper presented a design for a secure cloud-based EHR system using CP-ABE that provides effective solutions to some of the issues related to standard encryption mechanisms. It also investigated the feasibility of adopting CP-ABE in terms of performance and storage overhead. The results suggests that the proposed design would provide reasonable performance and consume negligible storage, and thus it can be used as a replacement to standard encryption mechanisms in cloud-based EHR systems. A proof-of-concept cloud-based EHR is being implemented and will be used to verify the feasibility of this approach.

6. REFERENCES

[1] J. Benaloh, M. Chase, E. Horvitz, and K. Lauter. Patient controlled encryption: ensuring privacy of electronic medical records. In *Proceedings of the 2009 ACM workshop on Cloud Computing Security*, 2009.

[2] J. Bethencourt, A. Sahai, and B. Waters. Ciphertext-policy attribute-based encryption. In *Proceedings of the 2007 IEEE Symposium on Security and Privacy*, 2007.

[3] D. Boneh and M. K. Franklin. Identity-based encryption from the weil pairing. In *Proceedings of the 21st Annual International Cryptology Conference on Advances in Cryptology*, CRYPTO '01. Springer, 2001.

[4] CareCloud. Carecloud, web-based medical practice management software, 2011. `http://www.carecloud.com/`.

[5] R. Chow, P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masuoka, and J. Molina. Controlling data in the cloud: outsourcing computation without outsourcing control. In *Proceedings of the 2009 ACM workshop on Cloud Computing Security*, CCSW '09, 2009.

[6] K. D. Edoh. Elliptic curve cryptography: Java implementation. In *Proceedings of the 1st Annual Conference on Information Security Curriculum Development*, InfoSecCD '04. ACM, 2004.

[7] Google. Google Health, 2011. `https://www.google.com/health`.

[8] B. W. John Bethencourt, Amit Sahai. Advanced crypto software collection, Feb 2011. `http://acsc.cs.utexas.edu/cpabe/`.

[9] P. Junod and A. Karlov. An efficient public-key attribute-based broadcast encryption scheme allowing arbitrary access policies. In *Proceedings of the 10th Annual ACM Workshop on Digital Rights Management*, DRM '10, 2010.

[10] D. R. Levinson. Audit of information technology security included in health information technology standards, May 2011. `http://oig.hhs.gov/oas/reports/other/180930160.pdf`.

[11] A. Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters. Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption. In *Advances in Cryptology - EUROCRYPT 2010*. Springer, 2010.

[12] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou. Fuzzy keyword search over encrypted data in cloud computing. In *Proceedings of the 29th IEEE Conference on Information Communications (INFOCOM'10)*, San Diego, 2010.

[13] M. Li, S. Yu, K. Ren, and W. Lou. Securing personal health records in cloud computing: Patient-centric and fine-grained data access control in multi-owner settings. In *Security and Privacy in Communication Networks*. Springer, 2010.

[14] N. Liu, Y. Zhou, X. Niu, and Y. Yang. Querying encrypted character data in DAS model. In *Proceedings of the 2nd International Conference on Networking and Digital Society (ICNDS)*, May 2010.

[15] B. Lynn. The pairing-based cryptography library, Feb 2011. `http://crypto.stanford.edu/pbc/`.

[16] K. Mandl, W. Simons, W. Crawford, and J. Abbett. Indivo: a personally controlled health record for health information exchange and communication. *BMC Medical Informatics and Decision Making*, 2007.

[17] Microsoft. Microsoft HealthVault, 2011. `http://www.healthvault.com/personal/index.aspx`.

[18] A. Mohan, D. Bauer, D. M. Blough, M. Ahamad, R. Krishnan, L. Liu, D. Mashima, and B. Palanisamy. A patient-centric, attribute-based, source-verifiable framework for health record sharing. Technical report, Georgia Institute of Technology, 2009.

[19] S. Narayan, M. Gagné, and R. Safavi-Naini. Privacy preserving EHR system using attribute-based infrastructure. In *Proceedings of the 2010 ACM Cloud Computing Security Workshop*, 2010.

[20] National Alliance for Health Information Technology. Defining key health information technology terms, 2008. `http://healthit.hhs.gov`.

[21] Practice Fusion. Web-based electronic health records, 2011. `http://www.practicefusion.com`.

[22] Regional Health Information Organization. Rochester RHIO, 2011. `http://www.grrhio.org`.

[23] A. Sahai and B. Waters. Fuzzy identity-based encryption. In *Proceedings of Advances in Cryptology - EUROCRYPT 2005*. Springer, May 2005.

[24] A. Shamir. Identity-based cryptosystems and signature schemes. In *Proceedings of CRYPTO 84 on Advances in Cryptology*. Springer, 1984.

[25] United States Department of Health & Human Services. Health Information Privacy, 2011. `http://www.hhs.gov/ocr/privacy/index.html`.

[26] United States Department of Health & Human Services - HITECH. HITECH Act Enforcement Interim Final Rule, 2011. `http://www.hhs.gov`.

[27] United States Department of Health & Human Services - Privacy Rule. Standards for Privacy of Individually Identifiable Health Information; Final Rule, 2002. `http://www.hhs.gov`.