Requirements of Secure Storage Systems for Healthcare Records

Ragib Hasan[1], Marianne Winslett[1], and Radu Sion[2]

[1] University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
(rhasan, winslett)@cs.uiuc.edu
[2] Network Security and Applied Cryptography Lab
Stony Brook, NY 11794, USA
sion@cs.stonybrook.edu

Abstract. Recent compliance regulations are intended to foster and restore human trust in digital information records and, more broadly, in our businesses, hospitals, and educational enterprises. In the health sector, storage and management of electronic health records have become a vital issue. Specifically, with the passing of the Health Insurance Portability and Accountability Act (HIPAA), the security of medical records has come into focus. HIPAA and other regulations in the health sector require strict compliance with specific privacy and security requirements. Unfortunately, existing storage solutions do not live up to the task of ensuring compliance with mandated legislation. In this position paper, we discuss the main characteristics of the health sector record management regulations, and present a set of requirements for secure, trustworthy storage that complies with these regulations. We also briefly analyze existing storage models, and show that they are not suitable for meeting the requirements of health-care record storage.

1 Introduction

Accurate and detailed record-keeping, along with ensuring their privacy and authorized access, are integral parts of managing medical information. With the advent of electronic computing, medical records, like many other application domains, have depended heavily on computerized storage systems for storage and archival of health information.

However, in the digital realm, the adversaries and attackers are quite different than the physical world – digitally stored information can be copied verbatim, and records may be exposed to a wide variety of adversaries. To protect privacy and security of such electronic medical records, many countries worldwide have enacted consumer protection and privacy laws. These laws have strict guidelines and requirements for regulation of medical record management.

Unfortunately, existing storage architectures are not capable of providing the strong security and privacy guarantees mandated by the laws associated with this new digital information domain. For example, several regulations require mandatory record retention (with data integrity) for periods of up to 30 years. But storing such records for a long time would require inevitable change of storage hardware and/or storage format. The resulting migration to new servers must be trustworthy, and verifiable. Similarly, if a medical record needs to be removed after the mandated retention period, the storage system must guarantee its secure deletion. Such features are not available in most of the current storage architectures for medical records.

In this paper, we look into HIPAA and several other regulations on protection of medical records, and discuss the security and privacy requirements that such regulations impose on record management. From this discussion, we derive a set of common requirements for electronic health-care record storage systems. Finally, we briefly look into several storage architectures, and show the limitations of current architectures in meeting all the requirements. The contribution of this paper is to map out the open research problems in the area, and to direct future research endeavors for secure storage of health-care records.

2 Health Care Regulations

Management of health information has become an important and regulated area in most countries. In the following, we briefly discuss several of these laws from different countries, and outline their essential common mandated features.

2.1 HIPAA

The Health Insurance Portability and Accountability Act of 1996, commonly known as HIPAA [3, 7], is an attempt to update the health sector and insurance record keeping in order to bring more accountability and better protection of consumer rights. Besides regulating the insurance industry, one of HIPAA's significant effect is to mandate the confidentiality and integrity of medical information.

HIPAA is divided into two titles. Title I regulates health insurance coverage. Title II discusses digital health care records, their security, privacy, as well as other facets of their management. The following main security and privacy requirements are mandated by HIPAA:

- **Privacy and Data Confidentiality.** The privacy rule of HIPAA requires organization to ensure that they have taken reasonable steps to ensure the confidentiality of health care records and communication with individuals. Individuals have the right to request correction of health care records.
- **Security.** Organizations outsourcing some of their record management tasks must ensure that the third-parties also comply with HIPAA. Each organization must have established *internal audit* procedures for medical records. All records must be disposed of in a trustworthy manner at the end of their *retention* period. Access to hardware and software should be limited to properly authorized individuals. Data *integrity* must be ensured by means of checksums, message authentication, or digital signatures. Each entity is responsible for ensuring that data within its systems have not been erased or tampered with.

Specifically, the General Rule (Section 164.306) requires entities to:

- Ensure the confidentiality, integrity, and availability of all electronic protected health information (EPHI) the covered entity creates, receives, maintains, or transmits;
- Protect against any reasonably anticipated threats or hazards to the security or integrity of such information;
- Protect against any reasonably anticipated uses or disclosures of such information that are not permitted or required by the Privacy Rule; and
- Ensure compliance by its workforce.

Additionally, Section 164.310 of HIPAA mandates storage media disposal, media-reuse, accountability, and data backup/storage for medical records. It requires the following:

- **Disposal.** 164.310(d)(2)(i) requires that covered entities must have policies and procedures that handle the final disposition of electronic health information records, and the media or hardware on which the records are stored.
- **Media re-use.** 164.310(d)(2)(ii) states that covered entities must implement "procedures for removal of electronic protected health information from electronic media before the media are made available for re-use."
- **Accountability.** 164.310(d)(2)(iii) states this: the covered entity must "Maintain a record of the movements of hardware and electronic media and any person responsible therefore." In other words, organizations must *log all data migration and data provenance information.*
- **Backup and Storage.** Finally, 164.310(d)(2)(iv) mandates that a covered entity "must create a retrievable, exact copy of electronic protected health information, when needed, before movement of equipment."

2.2 Occupational Safety and Health Administration regulation

In the United States, the Occupational Safety and Health Administration regulation (Standards - 29 CFR) "Access to employee exposure and medical records. - 1910.1020" [11], controls the management of medical records for employees, and all exposure records. Section 1910.1020(d)(1)(ii) requires that "Each employee exposure record shall be preserved and maintained for at least thirty (30) years". It also requires all employee medical records to be kept for at least 30 years. For businesses changing ownership, it must ensure the transfer of the records to the new owner.

2.3 EU Directives

In Europe, the Directive 95/46/EC of the European Union on the protection of personal data, provides privacy and security guarantees for personal information, including health care records [4]. In particular, Article 6 of the directive requires accuracy guarantees of personal records, and guaranteed disposal after the retention period. Article 17 requires measures for ensuring the confidentiality and availability of records. In addition, most countries in Europe have their own data protection laws. For example, in the United Kingdom, the Data Protection Act of 1998 [2] regulates, among other information, personal health-care records. It requires mandatory disposal of electronic records after retention period, accuracy of information, logging any changes, and strict confidentiality.

3 Requirements

As seen in the previous section, a set of relatively consistent, broadly mandated assurances can be found in a multitude of regulations. In the following we discuss the main requirements that storage systems would need to adhere to, for compliance purposes, including data confidentiality, records integrity and availability, as well as secure retention, deletion and migration mechanisms.

Confidentiality and Access Control. As health-care records contain sensitive information, the storage systems must ensure their confidentiality. Moreover, only authorized personnel should have access to confidential medical records. Consequently, to ensure confidentiality, storage systems must deploy strong encryption in both the actual storage and the data pathways leading to and out. Moreover, in the case of storage media re-use or disposal, the confidentiality of records previously stored in such media should be ensured.

Integrity. The storage system must ensure the integrity of medical records. In particular, it must ensure the integrity of medical records even in the case of malicious insiders. The security mechanisms must identify any tampering of information.

Availability and Performance. The health-care records must be accessible in a timely manner. Medical records are frequently expanded, and patients may also ask for correction of records. Hence, appropriate storage models should be used to allow both performance, security, and mutability.

Timely access to medical records would require indexing techniques. However, regular indexing schemes such as keyword index can breach privacy as the mere existence of a word in a document can leak information [9]. For example, if the keyword “Cancer” is present in a medical, then an adversary can assume that the patient might have Cancer. So, the index itself must be trustworthy, and confidential.

Logging, Audit Trails, and Provenance. All access to the storage system should be logged in a trustworthy manner. HIPAA mandates recording all medical record access information. Many of the regulations require extensive logging to record the movement of records between systems, and the access and modification history. Consequently, the storage system must provide verifiable audit trails and the maintenance of provenance information on the chain of records custody.

Support for Long Retention and Secure Migration. Many of the regulations require long retention periods for certain types of health-care records. The storage system must be capable of providing long term retention guarantees. Since it is conceivable that the failure of storage servers, as well as obsolescence of technology and formats will require migration of records, the storage system must provide trustworthy and verifiable migration mechanisms.

Backup. There must exist strong backup and restore operations. The backup copies should be located in a separate off-site location to ensure survival in case of fire or natural disasters.

Cost. The storage system must also be cost effective, possibly using cheap off-the-shelf hardware. Compliance with HIPAA and other regulations have significant management overhead. The cost of training personnel is also another factor. So the storage system should notbe cost-prohibitive. Media used by the storage system should be cheap.

4 Limitations of Existing Storage Models

We now discuss the suitability and limitations of existing storage models with respect to the above requirements. In particular, we look at relational databases, object-storage systems, and compliance WORM storage.

Commercial solutions for HIPAA compliant storage tends to focus on using strong encryption to provide security for electronic records [13]. Unfortunately, however, such schemes do not protect against malicious insiders. Moreover, such encryption based solutions do not account for maintaining provenance information.

Most of the early storage systems for electronic records involved relational databases. However, securing relational databases to the extent of compliance with the requirements described in section 3 is difficult. Relational databases are geared more towards performance rather than security. Specifically efficiently performing queries on encrypted data in the presence of malicious insiders as well as guaranteeing secure record retention are significant open problems to consider.

A promising alternative is IBM's Hippocratic Database Technology [6], which aims at providing regulatory compliance with data protection laws. It provides fine-grained access control by transparently rewriting user queries and enforcing various access and disclosure policies. Hippocratic databases also provide compliance auditing, in which database access information is logged for future forensic analysis in case of a privacy breach. However, without underlying security support, just defining semantics and enforcing them in a software query processor still leaves things vulnerable to insider attacks with direct disk access.

In object based storage systems, usually document content hashes are used as object IDs to locate documents [8]. This renders such mechanisms suitable for efficient storage of read-only content and read operations are efficient and optimized. Moreover, information integrity can be easily assured. However, appends and writes in the presence of malicious adversaries are difficult to achieve in object storage, and likely slow in performance.

The most promising technology for secure storage of health records is compliance WORM storage [5, 9, 10]. In such systems, records are kept in write-once, read-many times storage media. The media can be optical, or magnetic. Trustworthy indexing mechanisms [9] can ensure fast retrieval of data, as well as ensuring privacy and integrity of the index. Trustworthy migration [10] can ensure guaranteed and verifiable transfer of records among systems. Trustworthy deletion mechanisms can ensure complete removal of expired records. However, compliance WORM storage is mainly suitable for records that do not require corrections. Since medical records are expected to be corrected, and individuals have the right to request such corrections to their medical records, allowing corrections is an important feature. Currently, trustworthy WORM storage systems do not support such corrections.

Ultimately, the trade-off between security and performance makes it difficult to use existing secure storage systems. Most of the existing systems are geared towards read-only settings, optimizing read operations via smart indexing and caching. However, to support efficient and trustworthy write operations, data retention, secure deletion, migration, such systems simply do not live up to the requirements.

Moreover, an additional missing feature in all these systems is storage of provenance information [1, 12]. Since access to storage records must be recorded for later audits, it is critical to record such information. With migration of records between different systems, it is important to ensure a proper chain of custody for the ownership and transfer of records. However, current storage systems do not implement trustworthy provenance, and therefore, cannot fulfill this requirement of health-care record storage.

5 Conclusion

In this paper, we explored major health care regulation acts and discussed their impact on the requirements for associated storage support systems. We showed that unfortunately, existing systems and data models fall short of the resulting desiderata. We thus believe it is important to explore novel avenues and solutions in this area that would possibly combine existing functionality creating a hybrid model suited for trustworthy regulatory-compliant health-care record storage. Additionally, it is important to explore and consider the impact of the additional costs and overhead burdens such mechanisms would put onto their users and the healthcare system in general. Ultimately, as increasing amounts of health information are created and stored digitally, we believe compliance storage to be a vital tool in providing trust and privacy assurances.

Acknowledgments

The authors would like to thank the anonymous reviewers for their suggestions. This research was supported by the NSF under grants IIS-0331707, 0331690, CNS-0325951, and CNS-0524695. Radu Sion was in part supported by the NSF through awards CT CNS 0627554, CRI CNS 0708025 and CT CNS 0716608. Sion would also like to thank Motorola Labs, IBM Research, CEWIT and the Stony Brook VP Office for Research for their support.

References

1. U. Braun, S. Garfinkel, D. Holland, K.-K. Muniswamy-Reddy, and M. Seltzer. Issues in automatic provenance collection. In *Proceedings of the International Provenance and Annotation Workshop*, pages 171–183, 2006.
2. British Parliament. Data protection act of 1998. Online at `http://www.staffs.ac.uk/legal/privacy/dp10rules/index.php`, 1998.
3. Center for Medicare & Medicaid Services. The Health Insurance Portability and Accountability Act of 1996 (HIPAA). Online at `http://www.cms.hhs.gov/hipaa/`, 1996.
4. European Parliament. Legislative documents. Online at `http://ec.europa.eu/justice_home/fsj/privacy/law/index_en.htm`, 2006.

5. W. Hsu and S. Ong. WORM Storage is not Enough. *IBM Systems Journal*, April 2007.
6. C. Johnson and T. Grandison. Compliance with data protection laws using hippocratic database active enforcement and auditing. *IBM Systems Journal*, 46(2), April 2007.
7. N. Lawson, J. Orr, and D. Klar. The HIPAA privacy rule: An overview of compliance initiatives and requirements. *Defense Cousel Journal*, 70:127–149, 2003.
8. M. Mesnier, G. Ganger, and E. Riedel. Object-based storage: pushing more functionality into storage. *IEEE Potentials*, 24(2):31 – 34, April-May 2005.
9. S. Mitra, W. Hsu, and M. Winslett. Trustworthy keyword search for regulatory-compliant record retention. In *Proceedings of the 32nd International Conference on Very Large Data Bases*, pages 1001–1012. ACM, September 2006.
10. S. Mitra and M. Winslett. Secure deletion from inverted indexes on compliance storage. In *StorageSS '06: Proceedings of the Second ACM Workshop on Storage Security and Survivability*, pages 67–72, New York, NY, USA, 2006. ACM Press.
11. Occupational Safety and Health Administration. Access to employee exposure and medical records. - 1910.1020 regulations (standards - 29 cfr). Online at `http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10027`.
12. Y. Simmhan, B. Plale, and D. Gannon. A survey of data provenance in e-science. *SIGMOD Rec.*, 34(3):31–36, September 2005.
13. Smart Card Alliance. HIPAA compliance and smart cards: Solutions to privacy and security requirements. Online at `http://www.datakey.com/resources/HIPAA_Compliance_and_Smart_Cards_FINAL.pdf`, September 2003.