8-7-2011

A Threat Tree for Health Information Security and Privacy

Jeffrey P. Landry
University of South Alabama, jlandry@usouthal.edu

J. Harold Pardue
University of South Alabama, hpardue@usouthal.edu

Tom Johnsten
University of South Alabama, TJohnsten@usouthal.edu

Matt Campbell
University of South Alabama, mattcampbell@usouthal.edu

Priya Patidar
priya.vpatidar@gmail.com

A Threat Tree for Health Information Security and Privacy

Jeff Landry
University of South Alabama
jlandry@usouthal.edu

Harold Pardue
University of South Alabama
hpardue@usouthal.edu

Tom Johnsten
University of South Alabama
TJohnsten@usouthal.edu

Matt Campbell
University of South Alabama
mattcampbell@usouthal.edu

Priya Patidar
priya.vpatidar@gmail.com

ABSTRACT

This paper begins a process of organizing knowledge of health information security threats into a comprehensive catalog. We begin by describing our risk management perspective of health information security, and then use this perspective to motivate the development of a health information threat tree. We describe examples of three threats, breaking each down into its key risk-related data attributes: threat source and action, the health information asset and its vulnerability, and potential controls. The construction of such a threat catalog is argued to be useful for risk assessment and to inform public health care policy. As no threat catalog is ever complete, guidance for extending the health information security threat tree is given.

Keywords

Health Information Privacy, Information Security, Risk Assessment, Threat Modeling.

INTRODUCTION

Technological progress brings the promise of advancement, but not without accompanying risk. Advances in electronic health records (EHR) management is no exception. EHR advances hold the promise of improving health care outcomes, reducing medical and medication errors, and reducing health care cost by making information readily available for use in any place, at any time, and in the right format. As with other industries, however, the health care industry has experienced the cost of technological advancement, as threats to patient privacy and information security are ever present. The threats of unauthorized data access, disclosure, manipulation, loss, and corruption present risks to patients, institutions, and health care outcomes. Moreover, the complex interplay of technologies, people, policies, and legislation in an increasingly wireless networked world make the health information systems environment an interesting and challenging information security and privacy landscape.

This paper builds on recent research providing frameworks and taxonomies for cataloging threats to health information security and privacy. We start with a basic set of assumptions about health information security and privacy. The first assumption is that we take a risk management view of information security. By information security, we mean the process of identifying threats to information assets and the analysis of those threats in order to select countermeasures that mitigate the risk posed by the threats. Information assets in the health care context might include data (or information, terms used interchangeably) in all of its forms, and other technologies, such as hardware, software, networks, and electronic devices, and people, procedures, and the physical infrastructure. By considering information in all its forms, we mean not only electronic records stored in a computer system, but also paper-based records, records such as X-rays, and that which can be photographed, such as patient injuries. Procedures include processes of health care institutions, including business processes and patient-care processes. People, particularly patients and their families, but also health care institutional personnel vendors and others in the health care extended enterprise (Appari and Johnson, 2010), are the most important part of any information system. Another assumption is that a threat is any action posed by a human or non-human source that threatens the confidentiality, integrity, or availability (CIA) of information, where confidentiality involves preventing unauthorized disclosure, integrity is about preventing data modification, and availability is about making data accessible for authorized use when needed (Nematzadeh and Camp, 2010). The ultimate outcomes of threats to CIA may include, but are not limited to,

exposing patients to economic harm, mental anguish, and social stigma; enabling identity theft and medical identity theft; and impacts on medical decision making, including delayed treatment and ineffective decisions (Appari and Johnson, 2010; Johnson, 2009); financial impacts on third party payers; and failures due to inaccurate data in medical research.

The current research will utilize existing taxonomies of threats, both in the health information context and in the general case of information systems security. Existing taxonomies usually use one or both of the threat-source or threat action as an organizing principle. Threat-sources either intentionally exploit or accidentally trigger vulnerabilities in information assets. Threat actions are the methods, or the manner in which threat-sources exploit/trigger vulnerabilities. One of the existing sources of health information security and privacy threats that will inform the current research include Kotz (2011), who provides a taxonomy consisting of 25 threats, organized around three main categories: identity threats, access threats, and disclosure threats. Samy, Ahmad, and Ismail (2010) organized health information systems threats around 22 categories, finding five areas to be most critical: power failure/loss, acts of human error or failure, technological obsolescence, hardware failures or errors, and software failures or errors. More general threat categories can be found (Stoneburner, Goguen, and Feringa 2002; Whitman, 2003). The NIST 800-30 provides a categorization of threat sources in six categories: human-deliberate, human-unintentional, technical, operational, environmental, and natural. Whitman identified 12 categories of threat actions. Another source of health information privacy threats is the most widely cited law, the Health Information Portability and Accountability Act (HIPAA). The privacy implications of HIPAA influenced the development of the threat tree through the health information privacy literature that we reviewed.

THE HEALTH INFORMATION THREAT TREE

Our approach to health information threat modeling involves the use of threat trees and threat catalogs. The threat tree is organized around the goal of an attacker or outcome of a threat, depending on whether the threat is intentional or not. At the top level, or root, of the threat tree, is node 1, which is labeled *threaten health information asset*. With an orientation on information (data) as the primary vulnerable asset that leads to risks to other health information assets, we then focus on two main threats: disclose health information (1.1) and manipulate health information (1.2). From a threat tree standpoint, the overall outcome of threatening a health information asset, the root node, can be accomplished in two ways, either through data disclosure or data manipulation, its direct descendant child notes. These child nodes are broken down further.

An overview of the health information threat tree, depicted in indented outline form, is shown in Figure 1.

threat id	**node type - outline number - threat action**				
1	O	1	threaten health information asset		
2		O	1.1	disclose health information	
3			O	1.1.1	on the inside
4				T	1.1.1.1 by accident
5				O	1.1.1.2 due to curiosity
15					T 1.1.1.2.1 post patient X-rays on the Internet
16					T 1.1.1.2.2 disclose photos of patient taken with a hidden camera
6				T	1.1.1.3 through subornation of insider
7			O	1.1.2	on the outside
8				T	1.1.2.1 by hacking
9				T	1.1.2.2 by unauthorized access
10				T	1.1.2.3 by trespassing
11				T	1.1.2.4 by theft
12				T	1.1.2.5 by information extortion
13				T	1.1.2.6 by impersonation
27				T	1.1.2.7 by tracking patient

26			T	1.1.2.8	by extrapolation
14	T	1.2	manipulate health information		
17		T	1.2.1	by vandalism	
18		O	1.2.2	by loss or corruption of data	
19			T	1.2.2.1	due to faulty hardware
20			T	1.2.2.2	due to faulty software
21			T	1.2.2.3	due to human error
22			T	1.2.2.4	due to malware
23			T	1.2.2.5	due to a natural disaster
24			T	1.2.2.6	due to a database attack
25			T	1.2.2.7	due to unauthorized access

Figure 1. A Health Information Security Threat Tree

Organizing Threats

For the health information threat catalog to be useful for risk assessment, it needs to document more than just what is shown in the Figure 1 threat tree. Risk assessment requires an understanding of the threat source and threat action and how that source can trigger or exploit a vulnerability in a health information asset. As such, we have defined a set of fundamental threat attributes. Besides a unique identifier used to catalog, the threat, and its tree-related attributes, node type and outline number, we document the threat source and action, the health information asset and how it is vulnerable. Finally, we document potential controls to mitigate the threat, and provide a descriptive scenario of at least one example, real or hypothetical, of a threat incident.

Threat Modeling Examples

Examples of documented threats for three nodes are provided. Tables 1 through 3 document selected threats shown in the tree above as nodes 1.1.1.2, 1.1.2.7, and 1.1.2.8, respectively. There is an insider disclosure threat and two different types of outsider disclosure threats.

Threat Attribute	**Value**
ID-Node Type-Outline No.	5 – O – 1.1.1.2
Source	Human-deliberate insider
Action	Disclose health information due to curiosity
Health Information Asset	Patients and their families, patient information
Vulnerability	Difficulty controlling information residing in mobile or networked devices
Potential Controls	Require that only facility-owned equipment be used for photos; train personnel on mobile device awareness; conduct random tests of awareness policy
Scenario(s)	Nurses at a health clinic were dismissed for posting patient X-

	rays, taken with a cell phone, on the Internet; in another case, employees of a health care facility in Florida took photos of a shark attack victim with their personal cell phones (Shepherd 2010)

Table 1 - Insider Disclosure Threat

Disclose Health Information Due to Curiosity

Disclosure of health information due to curiosity is an insider threat, motivated by curiosity, and carried out by taking advantage of insider access. Insiders can include employees of the health care institution, patients, visitors, and any other outsiders, such as vendors, that are provided authorized access to health care facilities. The mid-level node, 1.1.1.2, and its two siblings, are organized around the motivation of the insiders to disclose. Insiders may either disclose accidentally, deliberately by subornation of another insider, or deliberately due to curiosity. The 1.1.1.2 node models the personal curiosity motivation. Medical personnel or other insiders abuse their insider access out of curiosity or for their own purposes. Some do so out of concern for the well being of fellow employees or family members. Some want to know about celebrities being treated. Some may be concerned about the possibility of sexually transmitted diseases in a colleague they are dating. Actions that may be taken include accessing computer records, or taking and releasing photos of patients, their injuries, or their records, and then releasing these photos. The scenario in Table 1 describes two published incidents where insiders used their personal cell phones with built-in cameras to disclose patient information. The vulnerability is the difficulty in controlling the small and thus easily concealable, and also wireless and Internet-connected devices. The potential controls suggested include policies against the use of personal phones, as well as awareness, testing, and enforcement of policies.

Threat Attribute	**Value**
ID-Node Type-Outline No.	27 – T – 1.1.2.7
Source	Human-deliberate outsider
Action	Disclose patient location and health information by tracking patient or patient-related equipment
Health Information Asset	Patient information
Vulnerability	Difficulty in securing information transmitted wirelessly
Potential Controls	Require all wireless transmissions be encrypted. Avoid weak encryption methods such as WEP. Restrict access to tracking software to only those who need it to do their job.
Scenario(s)	Some hospital systems use RFID (both active and passive) technology to track assets within and across their facilities. For instance, intravenous (IV) pumps and blood pressure monitors can be equipped with RFID active transmitters and located wherever they travel within the health system. These transmitters often operate using existing 802.11 networks. An intruder who gains

	access to the software tracking application would be able to track a patient as he/she is transferred within and between member hospitals. The intruder could also gain access to medications delivered by the IV pump.

Table 2 - Tracking a Patient Threat

Disclose Health Information on the Outside by Tracking Patient

Whereas the insider disclosure threat modeled threats committed by inside threat-sources, there are also disclosure threats committed by outside threat sources. The 1.1.2 branch models these threats, and the next two detailed threats come from this branch. In this branch, disclosure of health information is motivated by malicious intent. An outsider to a medical information system can pose a threat to health information by accessing health information in the following ways:

- Hacking
- Unauthorized access
- Trespassing
- Theft
- Information extortion
- Impersonation/identity theft
- Patient tracking
- Extrapolation

The terms "hack" and "hacking" are used to refer to a modification of a program or device to give the user access to features that were otherwise unavailable, such as by circuit bending. This gives a hacker open access to the patient's personal and medical data which can be misused by the hacker. An unauthorized read access to the data may lead to data disclosure. Vindictive former employees, angry patients, network intruders, or others may steal information, damage systems, or disrupt operations. With the computing devices and storage devices becoming smaller yet powerful with vast storage, these devices are becoming easier to steal and easier for attackers to use to steal the patient personal and medical information. Information extortion occurs when an attacker either threatens to steal, or actually steals, information from a company, or for agreeing not to disclose the information. Identity theft is gaining access to another person's details to gain access to the patient personal and medical details. This can be done by stealing the authentication details to of a person from inside the patient care institution and accessing the databases of the institution to gather the information.

From our list of outside disclosure threats, we now describe two: patient tracking an extrapolation. An example of patient tracking involves radio frequency identification (RFID) technology. RFID technology allows healthcare providers to keep track of equipment, consumables, and even patients and staff, but this information can cause privacy concerns when it is not kept secure from those outside the organization that do not have a legitimate need to see this information. Outsiders include individuals not affiliated with the organization that are motivated by curiosity, intent to harm, or intent to gain financial or other advantage through the illicit use of data they acquire. One form of tracking involves determining the patient's current location in a facility by querying the RFID tag attached to either the patient themselves or a piece of equipment that is attached to the patient for the duration of their stay (e.g. IV pump). Another form of tracking is accessing drug and dosage information stored in an IV pump connected to a patient. The vulnerability is the difficulty in securing wireless broadcasts against interception by outsiders and securing the tracking application on the organization's network. The potential controls suggested include the use of encryptions methods and restricting access to RFID tracking applications running on the network.

The vulnerabilities exist because of the difficulties in preventing and detecting the various unauthorized disclosures. In the case of the patient tracking example in Table 2 above, the vulnerability is the difficulty in securing wireless transactions, and one suggested control is the use of encryption.

Threat Attribute	Value
ID-Node Type-Outline No.	26 – T – 1.1.2.8
Source	Human-deliberate outsider
Action	Disclose health information on the outside by extrapolation
Health Information Asset	Violation of patient anonymity
Vulnerability	Difficult to control (a) linkage of medical records to public data sources and (b) disclosure of sensitive information through data mining technology
Potential Controls	Design software to accurately identify true violations in patient anonymity, maximize the amount of released data, and maintain to the greatest extent the integrity of the released data
Scenario(s)	Medical record of William Weld, former governor of MA, was identified through linkage with the voter registration list for Cambridge, MA (Sweeney, 2002)

Table 3 – Disclosure by Extrapolation Threat

Disclose Health Information on the Outside by Extrapolation

While tracking the patient is one way for an outsider to access and ultimately disclose health information, tracking and linking patient data through extrapolation is yet another. When such extrapolation techniques are carried out successfully by attackers, the patient's anonymity is violated.

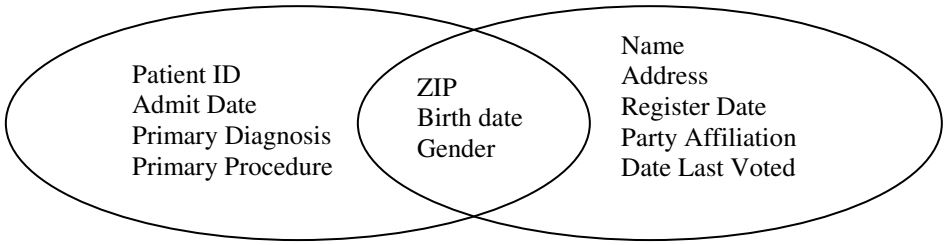

Figure 1 - A potential threat to patient privacy

Health-care organizations frequently collect and distribute person-specific data for use in medical research as well as for use in formulating local, state, and national health-care policies. Under existing guidelines, health-care organizations frequently disclose person-specific data for purposes unrelated to treatment, payment, or health-care operations based on the removal of attributes that could potentially be used to identify specific individuals. The removal of such attributes is frequently based

upon guidelines referred to as the safe harbor provision[1]. Unfortunately, in some instances data are unnecessarily removed resulting in the release of data that do not satisfy the information needs of the requesting parties (Benitez, Loukides and Malin, 2010).

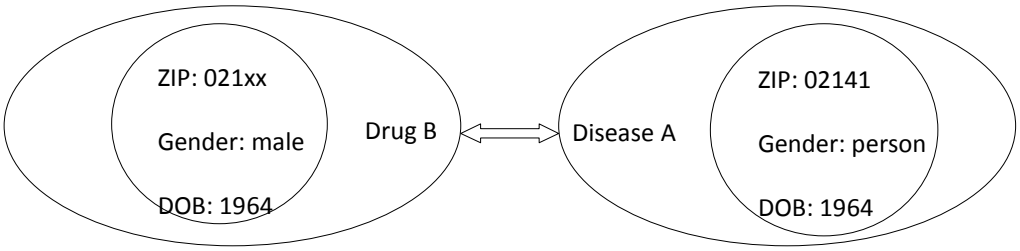

Figure 2 - A violation in patient anonymity

Violations in anonymity can occur when person-specific data records are uniquely joined with external datasets containing explicit identifiers. Figure-1 illustrates this potential threat to privacy through the joining of medical claims records with publicly available voter registration records (Sweeney 2002). In this instance a voter registration record can be linked to a medical claims record using the attribute set {ZIP, BirthDate, Gender}. Such an attribute set is referred to as a quasi-identifier. Defining a quasi-identifier only as those attributes on which two datasets can be joined is limiting and ignores cases such as the one shown in Figure-2. In this example, the attributes Drug and Disease are being used to join two datasets, by applying domain knowledge, or knowledge discovered through data mining technology, to infer that Drug B is linked to disease A. That is, anonymity holds, separately, in the two data sets prior to the introduction or discovery of the relationship between Drug B and Disease A. With knowledge of this relationship, however, the anonymity property can be violated, in part, through the linkage of the records, <ZIP:021xx, Gender:male, DOB:1964> and <ZIP:02141, Gender:person, DOB:1964> to produce the derived patient record <Zip:02141, Gender: male, DOB: 1964>. Subsequently, the derived record could be linked to publicly available datasets resulting in the violation of a patient's anonymity.

Using the Threat Tree

The systematic modeling of health information threats described in this paper will support analysts in the evaluation of health information risk. By using a common structure in which to describe risk-related data on threats, we enable a comparative assessment of risk. Such as comparison informs the risk analyst so that a threats may be analyzed, vulnerabilities prioritized, and countermeasures selected based on some efficient means, such as cost-benefit analysis. The assessment of risk may be made using either a quantitative or qualitative approach, whereby calculated risk is a function of the probability of a threat's occurrence and the severity of the consequences—or impact—should the threat occur. The implementation of risk analysis might result in probability, impact, and risk being added as attributes added to each threat. Tables 1 through 3 provide examples of threat attributes for three specific threats. Probability, impact, and risk would be added to those listed in the tables.

Because the threat catalog uses accepted categorical frameworks for organizing threats, there is the potential for creating a comprehensive catalog that would be applicable at the national level, and could inform public policy on health information security and privacy. Certainly, the catalog would be of value to risk analysts employed on behalf of health care institutions, who would use selective portions of the catalog to model the local institution's threat landscape. The catalog would also be useful for extending the breadth and depth of health information security and privacy research.

A continuous approach to health information security is urged in keeping with the idea that information security is an escalating war. As technologies change, and associated changes occur in the behavior of attackers, new threats invariably will emerge. The catalogs, extensible as they are with the tree as an organizing framework, can and must be updated.

CONCLUSION

In our effort to create a comprehensive catalog of health information security and privacy threats, we have:

- begun constructing a catalog using a risk management approach to health care information security and privacy

[1] http://www.hhs.gov/ocr/privacy/hipaa/understanding/summary/privacysummary.pdf.

- described threats documented in the catalog
- explained how the threat tree will be useful

The current threat catalog is far from complete. The most immediately obvious extension is a third branch on level two for availability threats. Such a threat might be labeled *disrupt health information services*. Additional efforts to extend the tree would include using the existing frameworks, such as the NIST 800-30 and other literature cited, to systematically account for all known threats within sub-categories of the tree. Then, additional data gathering, using the academic and practitioner literature, activity and data flow modeling of health care processes, and perhaps interviews with experts in the field, to model additional threats not appearing in the literature. Eventually, the catalog would have to be validated for correctness and completeness. Approaches to validation might include: comparison of trees to other frameworks, such as government standards, the aforementioned frameworks (Stoneburner, et al., 2002; Whitman, 2003) or threat trees in a different domain, such as election operations (Pardue, Landry, and Yasinsac, 2009); and expert review.

REFERENCES

1. Appari, A. and Johnson. M. E. (2010) Information security and privacy in healthcare: current state of research, *International Journal of Internet and Enterprise Management*, 6, 4, 279 – 314.
2. Benitez, K., Loukides,G., and Malin, B. (2010) Beyond safe harbor: Automatic discovery of health information de-identification policy alternatives, in Tiffany Veinot (Ed.) *Proceedings of the 1st ACM International Health Informatics Symposium*, November 10-12, 2010, Arlington, VA, 163-172.
3. Johnson, E. (2009) Data hemorrhages in the health-care sector, *Financial Cryptography and Data Security*, 5628, 71-89.
4. Kotz, D. (2011) A threat taxonomy for mHealth privacy, in *Workshop on Networked Healthcare Technology (NetHealth)*, January 4, 2011, Bangalore, India.
5. Nematzadeh, A. and Camp, L. J. (2010) Threat analysis of online health information system, in Fillia Makedon, Ilias Maglogiannis, and Sarantos Kapidakis (Eds.) *Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '10)*, June 23-25, 2010, Samos, Greece, 31, 1-7.
6. Pardue, H., Landry, J., and Yasinsac, A. (2009) A risk assessment model for voting systems using threat trees and Monte Carlo simulation," *First International Workshop on Requirements Engineering for E-voting Systems (RE-Vote)*, August 31, 2009, Atlanta, GA.
7. Samy, G. N., Ahmad, R., and Ismail, Z. (2010) Security threats categories in healthcare information systems, *Health Informatics Journal,* 16, 3, 201-209.
8. Shepherd, A. (2010) Negative exposure, *For the Record*, 22, 11, 10-13; retrieved from http://www.fortherecordmag.com/archives/060710p10.shtml.
9. Stoneburner, G., Goguen, A., and Feringa, A. (2002) *Risk Management Guide for Information Technology Systems, Recommendations of the National Institute of Standards and Technology*, *Special Publication 800-30*, Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.
10. Sweeney, L. (2002) k-anonymity: a model for protecting privacy, *International Journal on Uncertainty, Fuzziness, and Knowledge-based Systems*, 10, 5, 557-570.
11. Whitman, M. (2003) Enemy at the gate, *Communications of the ACM*, 46, 8, 91-96.