Effective Audit Trails – A Taxonomy For Determination of Information Requirements

Phillip V. Asaro, MD, Robert L. Herting Jr, MD, Allan C. Roth, PhD, Mike R. Barnes, MD
Department of Health Management and Informatics,
University of Missouri-Columbia School of Medicine, Columbia, MO

Current methods of detecting confidentiality breaches in electronic medical record systems are inadequate, partially due to the lack of necessary information at the point of audit trail analysis. In order to determine the information requirements for effective audit trail analysis, we have formulated a taxonomy of confidentiality breaches. By considering scenarios in which an inappropriate access might occur, we have identified "indicators" of confidentiality breaches, which may be thought of as evidence suggesting the possibility that a confidentiality breach has occurred. The collection of facts needed to describe the indicators provides insight into the types of information needed to improve confidentiality breach detection. Much of the information needed is unlikely to be available in the patient record. Research is needed exploring means of collecting and utilizing information from sources other than the patient record for use in improving patient information security.

INTRODUCTION

In health care, the need to make patient information readily available at the point of care limits the degree to which access by authenticated users can be restricted[1,2]. This leads to the need for other means of implementing the confidentiality component of the information system security triad (availability, integrity, and confidentiality). Thorough and frequent audit trail analysis with user awareness of the audit has been shown to discourage abuse[3]. Recognizing this, the proposed Health Insurance Portability and Accountability Act (HIPAA) mandates complete and frequent audit trail analysis[4]. As practiced today audit trail analysis usually consists of entirely manual audit log review[5], although there has been some work on use-pattern deviation detection[6]. More effective tools are needed to maintain continuous surveillance of audit trail information in healthcare[6-8]. Automation of the analysis is technically feasible, but insufficiency of the information available is a limiting factor. We have developed a scenario-based taxonomy as a basis for determining the information needs in audit trail detection of confidentiality breaches.

SCENARIOS OF CONFIDENTIALITY BREACHES

A confidentiality breach scenario is an imagined situation in which a breach might occur. As an example, a user might access the medical record of a fellow employee to confirm the rumor that she is pregnant. A scenario can be characterized in part by the associated motive or motives for the imagined breach of confidentiality. In our example, the motives are (1) the presence of a diagnosis or test for pregnancy and (2) the relationship of the user to the patient, in this case a fellow employee of the health care institution. Motives relate to (1) characteristics of the patient or patient record, (2) a relationship between the user and the patient, or (3) a relationship between the user and another person represented in the patient record (such as a guarantor, spouse, or child).

Taking the guarantor as an example of an "other person" in the medical record, the patient record typically contains identifying information of this person such as name, address, telephone number, and social security number. If the patient and the guarantor are different people, the relationship between the user and the guarantor may relate to a motive behind a confidentiality breach. When an "other person" in the patient's medical record is the user (a user accessing her child's record, for example), a relationship between the user and patient is directly established.

A scenario involving false authentication may be best characterized by noting deviations in user behavior from what is expected from a particular user. For example, a login from an unexpected site suggests the possibility of misuse by someone other than the owner of the user ID.

Generalizing from a specific scenario can lead to the suggestion of numerous other scenarios. From the first example above, the pregnancy diagnosis could be generalized to any diagnosis or test result of a potentially sensitive nature. The relationship of "fellow employee" could be generalized to include other relationships such as neighbor, ex-wife, or opponent in a lawsuit. From the "login from an unexpected site" example above, the unexpected login location could be generalized to include variances in other session characteristics, such as login time-of-day, login during a shift that the user is not scheduled to work, or login failures prior to successful login. Using this approach, we considered a number of scenarios of confidentiality breaches and developed a collection of patient characteristics, user-patient relationships, and session characteristics that might be associated with a confidentiality breach.

1091—8280/99/$5.00

INDICATORS OF CONFIDENTIALITY BREACHES

We have organized a collection of indicators of confidentiality breaches in the form of a taxonomic tree (pictured in Figure 1). The indicators are described in terms of the patient characteristics, user-patient relationships, and session characteristics identified in breach scenarios as explained above. An indicator may be thought of as evidence suggesting the possibility that a confidentiality breach has occurred. The overarching rationale for the categorization scheme chosen here is evident in the first level of branching of the tree. One of the two main branches of the tree represents potential motivators of confidentiality breaches. These indicators involve information at the individual patient access level. The other main branch represents user behavioral deviations. Information about user behavioral deviation is found at the session level. This collection of indicators, though intentionally extensive, is not exhaustive and other categorization schemes might suggest other indicators.

To simplify the figure, we do not specifically include a set of indicators for "other persons represented in the medical record." This set of indicators would be very similar to the patient level indicators under the user-patient relationship heading. Substituting "patient" with "other person" would produce this set.

Patient Access Level - Motivational Indicators

At the individual patient access level, consideration is given to: (1) any relationship between the user and patient that could be a motivator to breach confidentiality, and (2) characteristics of the patient or of the patient record that could be a motivator to breach confidentiality. Patient characteristics that might motivate a breach include that of being a person with a high degree of public visibility, such as a politician. Patient record characteristics that might motivate a breach include the presence of potentially sensitive information in the record.

"Negative indicators" of confidentiality breach are also useful in audit trail analysis and we believe that combining positive and negative indicators would lead to the most effective detection of confidentiality breaches. Where "positive indicators" are potential evidence of a confidentiality breach, "negative indicators" are evidence that the user is expected to access the patient's record, typically based on an established provider role. For example, a user with a role of outpatient clinic medical assistant will at times have legitimate reason to access the records of a patient seen in that clinic. This counterbalances, to some uncertain extent, positive breach indicators which may be present. Problems associated with establishing negative indicators include determination of (1) when a particular provider role becomes associated with a particular patient to form a provider-patient relationship, and (2) the expected duration of that provider-patient relationship for access purposes.

Session Level - Behavioral Deviation Indicators

At the session level, consideration is given to characteristics of a user session. Session characteristics include: (1) session level statistics, such as total session length and number of patient records accessed, and (2) login characteristics such as failed attempts prior to successful login, time of day of login, and location from which login occurs. The indicators represent deviations from expected behavior of the user. The deviations may be defined in terms of predetermined parameters or in terms of deviation from previously established patterns of the user or a group of users.

ANALYSIS OF A POSSIBLE BREACH

If a combination of indicators associated with a confidentiality breach scenario is present in a given patient information access or user session, a breach may have occurred. If used alone, an indicator such as "accessing the record of a patient with a recent positive drug screen" would identify every user involved in the care of that patient along with any true misuse. However, if this indicator was combined with "user is related to patient as an ex-spouse" and "user and patient are currently involved in a legal action over custody of a child," there is a higher probability that this access represents a confidentiality breach.

Indicators contribute variably to the probability that a breach in confidentiality has occurred, suggesting the need for probability scoring and an indicator weighting mechanism. The probability of a breach having occurred may differ from one scenario to another. Even within one scenario, the indicators may contribute unequally to the probability of a breach.

The severity of the potential consequences associated with different confidentiality breach scenarios varies. Potential consequences could relate to the patient, individuals related to the patient in various ways, professionals involved in the care of the patient, the health care institution, or the community. Any method used to analyze patient information access will likely flag accesses that do not actually represent a confidentiality breach as well as those that do. The degree to which "false positives" can be tolerated depends upon the volume of the "false positives" and the potential severity of the consequences in the particular scenario.

A scoring system for both the probability of a confidentiality breach and for the severity of potential consequences in a particular scenario would support prioritization and appropriate allocation of resources including time spent investigating possible breaches.

INFORMATION REQUIREMENTS

We have itemized the information requirements for our collection of indicators. The information elements are categorized in the form of a taxonomic tree (Figure 2). Some of the information elements needed, such as the demographic and clinical data, are available in most medical record systems. Other information, such as marriage and divorce records, and information regarding current legal actions, is not available in the typical medical record system.

Demographic information and relationship information is needed primarily as a means to establish a motivating relationship between the user and either the patient or another person in the patient's record. These information elements are essentially the same for the user, the patient, and "others in the patient record."

By expanding our consideration beyond "currently available information" to "potentially useful information," we may discover previously unrecognized useful sources of information. Additionally, planning for future changes in the information system may be positively influenced. The security needs, and therefore appropriate indicators and information requirements, vary by institution. We have attempted to give a range of illustrative examples.

FUTURE DIRECTIONS FOR RESEARCH

There are several technical issues requiring further work. To implement analyses with some of the indicators discussed, it will be necessary to devise mechanisms to obtain the required data from various sources. For example, a system capable of taking addresses as input and returning intervening distance as output could be utilized as a means of determining "neighbor" status. It may be possible to obtain information on current legal actions and divorce records through databases increasingly available on the Internet, although this should be done with careful consideration to existing laws and policies regarding the confidentiality of such data. Effective audit trail analysis will depend upon analysis techniques which may include simple Boolean rules, rules with weighted factors, fuzzy logic methods, and use-pattern analysis. Research will be needed regarding construction and effective maintenance of systems implementing such methods.

While the technical aspects of a confidentiality breach detection system are important, the overall effectiveness of such a system will also depend upon human factors such as institutional policy and procedural issues. Therefore, research into the best use of the results of audit trail analysis will continue to be of value. Even a very effective audit trail with sophisticated analysis tools would still only act as a filter with output of "possible" confidentiality breaches. Human involvement will always be necessary in making the final judgement of whether a true breach has occurred. Policies and procedures are likely to significantly affect the extent to which a detection system becomes a deterrent to breaches of confidentiality.

The legal and ethical issues related to using private information about users and patients to enhance security are certainly complex and largely unexplored. Ongoing careful consideration of these issues will be a necessity in this line of research.

CONCLUSION

We have demonstrated the use of a taxonomy of confidentiality breaches as a means of determining information needs for effective audit trail generation and analysis. Based on scenarios of confidentiality breaches, motivational and behavioral deviation indicators can be developed which can be used to determine the information needed to generate effective audit trails. Much of the information needed to generate effective audit trails is not generally available in medical record systems. Other sources of information exist and methods of making this information available in the analysis of audit trails should be explored.

Acknowledgements

Drs. Asaro, Herting and Roth are supported by NLM Training Grant LM07089-07. Dr. Barnes is supported in part by MIAMS Grant LM05415.

References

1. Safran C, Rind D, Citroen M, Bakker AR, Slack WV, Bleich HL. Protection of confidentiality in the computer-based patient record [see comments]. MD Comput 1995;12(3):187-92.
2. Roger France FH. Control and use of health information: a doctor's perspective. Int J Biomed Comput 1996;43(1-2):19-25.
3. Bowen JW, Klimczak JC, Ruiz M, Barnes M. Design of access control methods for protecting the confidentiality of patient information in networked systems. Proc AMIA Annu Fall Symp 1997:46-50.
4. Carrington C. Keeping data safe: New HIPAA regs hit hard. Telehealth 1998;4(6):30-33.
5. Council NR. For the Record: Protecting Electronic Health Information. Washington, DC: National Academy Press, 1997.
6. White GB, Fisch EA, Pooch UW. Auditing and intrusion detection. Computer system and network security. Boca Raton: CRC Press, 1996:91-115.
7. Bauer DS, Eichelman FR, II, Herrera RM, Irgon AE. Intrusion detection: An application of expert systems to computer security. Proc 1989 Int Carnahan Conf Secur. 1989:97-100.
8. Hayam A. Security Audit Center--a suggested model for effective audit strategies in health care informatics. Int J Biomed Comput 1994;35(Suppl): 115-27.