EXTENDED ABSTRACT

ANALYSING ELECTRONIC HEALTH RECORD LOGS TO DETECT PRIVACY VIOLATIONS VIA CONFORMANCE CHECKING

Chathurika Wickramage,*,[1] Colin Fidge,[2] Chun Ouyang,[2] and Tony Sahama[3]

[1]University of Ruhuna, Sri Lanka
[2]Queensland University of Technology, Australia
[3]School of Health Information Science, University of Victoria, Canada
* chathurika@dcs.ruh.ac.lk

(Published 15 October 2021)

Abstract

According to the Healthcare Data Breach Report - April 2021, more than 11 million healthcare data breaches have been reported in the year 2021. Major causes of these data breaches include accidental and deliberate information misuses in Health Information Systems (HISs) leading to privacy violations. After-the-fact reviewing for information misuses needs to be integrated with current HISs, and a prominent solution is identifying information misuses through healthcare event log analysis. However, the lack of proper log formation and log analysis imposes major challenges for auditing after-the-fact to detect information misuse. This study exploited the idea of conformance checking to detect health information privacy violations by examining healthcare logs. A case study of applying our approach to an emergency health management condition demonstrates how a violation of Health Insurance Portability and Accountability Act (HIPAA) privacy policies during clinical practices can be detected automatically.

Keywords: Health information systems, emergency care, information accountability, log analysis

1. Introduction

The availability of electronic Protected Health Information increases disclosure risks in Health Information Systems (HISs) (Cohen Mello 2018). Healthcare Data Breach Report - April 2021 which is published by the Health Insurance Portability and Accountability Act (HIPAA) reports that more than 11 million health data breaches have been reported in the year 2020 (Health IT Dashboard, 2020). Major causes of these data breaches include accidental and deliberate information misuses, such as hacking, improper disposal, loss, theft and unauthorized access or disclosure, in HISs leading to privacy violations (Health IT Dashboard, 2020).

These scenarios motivate the need for the concept of Information Accountability (IA) along with information security and privacy (Weitzner et al., 2008). As an important mechanism for IA, after-the-fact reviewing for information misuses needs to be integrated with current HISs in addition to traditional security mechanisms of blocking access to information (Kacianka, 2017). While a prominent solution for implementing IA is to identify information misuses through appropriate recording and analysis of event logs, after-the-fact reviewing via log analysis in the healthcare domain is still an open issue (Weitzner et al., 2008). Inappropriate log formation and a lack of log analysis tools dedicated to healthcare practitioners impose major challenges in auditing after-the-fact to detect information misuse (Wickramage et al., 2016).

In the area of enterprise systems management, the concept of conformance checking has been used to detect inconsistencies between a best practice workflow model and the execution log of the corresponding process (Rozinat Van der Aalst, 2008). Unusual behaviours or deviations can be identified if the corporate IT system captures records of every admittance of information related to actual process execution in the event logs. This study aims to exploit the idea of conformance checking to detect health information privacy violations by examining the event logs recorded in HISs. Although we may potentially use existing conformance checking tools, inappropriate log formation for analysis purposes is one of the main problems that obstructs the application of such an approach. Incompleteness and inconsistency of existing logging mechanisms in healthcare, the volume and the variety of logs and collecting required information from various HISs and relevant IT systems compounds the challenges.

A main contribution of this research is the development of systematic log analysis methods to detect health information privacy violations via conformance checking. To demonstrate our approach, we present a case study based in an Emergency Department in a hospital, specifically involving a Break-the-Glass situation. It shows how a violation of HIPAA privacy policies during clinical practices (Yale University, 2021) can be detected automatically. Promising results were obtained from the application of our approach which successfully distinguished unacceptable scenarios from acceptable ones using the logs recorded by a healthcare system.

2. Methodology

In enterprise systems management, conformance checking can be used to detect inconsistencies between a workflow model representing expected business operations and the corresponding event logs that record the actual operations (Rozinat Van der Aalst, 2008). Hence, our overall approach aims to apply this existing conformance checking mechanism to the healthcare context for privacy violation detection.

Figure 1 depicts an overview of our approach. The main idea is to audit healthcare practice by checking the conformance between the relevant event sequences recorded in healthcare logs against health standards and policies. On the one hand, a typical patient journey, compliant with healthcare standards and policies, can be specified in the form of a so-called clinical workflow model. On the other hand, we propose log preparation activities, to make the existing healthcare logs capable of being used with a conformance checker. Next, the log data obtained from log preparation can be checked against the clinical workflow model using an existing tool for conformance checking, and the results be recorded in log analysis reports.

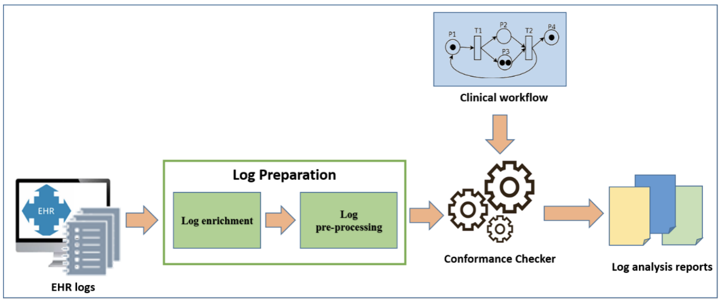


Figure 1. An overall approach of auditing healthcare logs against health standards via conformance checking.

We introduced log preparation activities which are carried out in two main phases. Firstly, to enrich healthcare log files so that they contain sufficient information for the purpose of policy

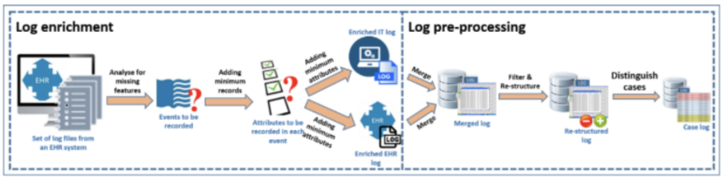


Figure 2. The log file preparation consisting of log enrichment and pre-processing phases.

violation detection, and secondly, to pre-process the log files to obtain the data required by the relevant conformance checker.

- Log enrichment. This is to analyse existing health log data to clarify the missing features and identify the additional events that must be recorded in a healthcare log and/or an IT system log to capture information relevant for privacy violation detection.
- Log pre-processing. This involves merging separate logs, filtering and restructuring the logs, and distinguishing individual 'cases' (i.e. process instances) in the logs.

3. Case Studies

The general approach presented in the above section was used for a typical emergency health management scenario in the Emergency Department (ED) in a hospital. A Break-the-Glass (BTG) procedure was chosen as an example of a complex treatment process with a potential information privacy risk. Below, we firstly describe a BTG procedure specified in HIPAA compliant healthcare standard, and then show how our approach can be used to detect privacy violations in such a context.

- Case scenario – patient data disclosure in a BTG situation. HIPAA defines a standard BTG procedure as granting immediate access to electronic Protected Health Information (ePHI) in a critical situation. A typical scenario is an emergency condition where a clinical practitioner needs access privileges for the safety of a patient beyond those normally allowed. This is managed by maintaining emergency login accounts with minimum necessary access rights as well as audit trails associated with them. A clean-up process, which often takes a selected period (such as eight hours) from the start, should also be defined to disable pre-owned emergency accounts to prevent their reuse.

3.1 Clinical Workflow Model: An ED process involving BTG

This is to specify expected behaviours (i.e. norms) using a workflow model. We selected an emergency health management example and created a workflow model to capture a medical treatment process involving a BTG situation in an emergency condition as part of our previous work (Wickramage et al., 2019). The workflow model is expressed using Business Process Modelling Notation (BPMN), a mainstream process modelling language.

3.2 Preparing EHR Logs from OpenEMR

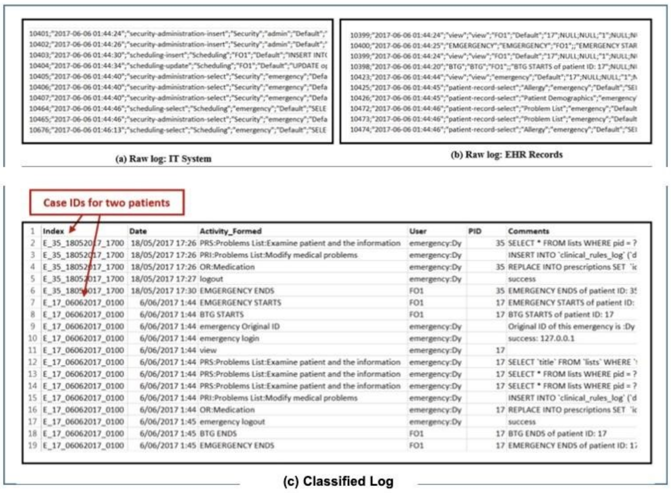
The first step for preparing conformance check-ready log files is to enrich health log files from an existing HIS. It is necessary to record an essential set of event types in healthcare log files to accomplish after-the-fact reasoning. However, the existing healthcare logs may not have all these features. To illustrate this, we implemented our proposed anatomy (Wickramage et al., 2016) of a log file for healthcare systems. We used OpenEMR, which is an open source health management system, to capture actual events in an ED process. The second step of log preparation is to pre-process the log files into an appropriate format. For our purposes we used the open-source process mining tool ProM to perform conformance checking. The enriched logs generated by OpenEMR were pre-processed such that the conformance checking plug-in module in ProM could be applied against a pre-defined workflow model. Figure 2 depicts the raw log files obtained from the OpenEMR system and the final processed log files that can be used with an existing conformance checker. The raw log files of the IT system (Figure 2 (a)) and the healthcare records (Figure 2 (b)) were merged based on the timestamps. The merged log file was then filtered to remove unnecessary information, such as administrative and financial events, and re-structured using existing log information. The filtered and restructured log file was extended to include the case IDs and led to the classified log (Figure 2 (c)).

3.3 Log Analysis via Conformance Checking

Given a process model that depicts the expected functionality of a business process and an event log that evidences its actual execution, detecting and describing the differences between the process model and the event log is known as conformance checking. The final phase in our approach was to check compliance, using a standard conformance checking tool, between the enhanced healthcare event logs and the pre-defined emergency procedure workflows. The process mining tool ProM was used to search for the alignments and the deviations of the event logs produced from OpenEMR (Figure 2 (c)) against the emergency procedure workflow. The standards of ProM were used to load the workflow model and the pre-processed event log file. Cases that deviate from the workflow model can be interpreted as an occurrence of a particular privacy violation given that we have accurately captured log-level indicators of policy violations in our workflow model. Distinct cases were analysed, to determine the exact transitions which deviate from the workflow, in each case of an unacceptable scenario.

4. Results

The generated log file from the modified OpenEMR system was analysed using ProM 6.5. Ten cases simulating an emergency health management condition were selected to be pre-processed. The cases include two normal cases, two emergency cases and six BTG cases. The BTG cases include two which may indicate a privacy violation.



```
10401;"2017-06-06 01:44:24";"security-administration-insert";"Security";"admin";"Default";"
10402;"2017-06-06 01:44:26";"security-administration-insert";"Security";"admin";"Default";"
10403;"2017-06-06 01:44:30";"scheduling-insert";"Scheduling";"FO1";"Default";"INSERT INT(
10404;"2017-06-06 01:44:34";"scheduling-update";"Scheduling";"FO1";"Default";"UPDATE o
10405;"2017-06-06 01:44:40";"security-administration-select";"Security";"emergency";"Defa
10406;"2017-06-06 01:44:40";"security-administration-select";"Security";"emergency";"Defa
10407;"2017-06-06 01:44:40";"security-administration-select";"Security";"emergency";"Defa
10464;"2017-06-06 01:44:46";"scheduling-select";"Scheduling";"emergency";"Default";"SELE
10465;"2017-06-06 01:44:46";"security-administration-select";"Security";"emergency";"Defa
10676;"2017-06-06 01:46:13";"scheduling-select";"Scheduling";"emergency";"Default";"SELE
```

(a) Raw log: IT System

```
10399;"2017-06-06 01:44:24";"view";"view";"FO1";"Default";"17";NULL;NULL;"1";NULL;N
10400;"2017-06-06 01:44:25";"EMGERGENCY";"EMGERGENCY";"FO1";;"EMERGENCY STAR
10399;"2017-06-06 01:44:24";"view";"view";"FO1";"Default";"17";NULL;NULL;"1";NULL;N
10398;"2017-06-06 01:44:20";"BTG";"BTG";"FO1";;"BTG STARTS of patient ID: 17";NULL;N
10423;"2017-06-06 01:44:44";"view";"view";"emergency";"Default";"17";NULL;NULL;"1";A
10425;"2017-06-06 01:44:45";"patient-record-select";"Allergy";"emergency";"Default";"SEI
10426;"2017-06-06 01:44:45";"patient-record-select";"Patient Demographics";"emergency
10472;"2017-06-06 01:44:46";"patient-record-select";"Problem List";"emergency";"Default
10473;"2017-06-06 01:44:46";"patient-record-select";"Problem List";"emergency";"Default
10474;"2017-06-06 01:44:46";"patient-record-select";"Allergy";"emergency";"Default";"SEI
```

(b) Raw log: EHR Records

1	Index	Date	Activity_Formed	User	PID	Comments
2	E_35_18052017_1700	18/05/2017 17:26	PRS:Problems List:Examine patient and the information	emergency:Dy	35	SELECT * FROM lists WHERE pid = ?
3	E_35_18052017_1700	18/05/2017 17:26	PRI:Problems List:Modify medical problems	emergency:Dy		INSERT INTO `clinical_rules_log` ('d
4	E_35_18052017_1700	18/05/2017 17:26	OR:Medication	emergency:Dy	35	REPLACE INTO prescriptions SET `ic
5	E_35_18052017_1700	18/05/2017 17:27	logout	emergency:Dy		success
6	E_35_18052017_1700	18/05/2017 17:30	EMGERGENCY ENDS	FO1	35	EMERGENCY ENDS of patient ID: 3!
7	E_17_06062017_0100	6/06/2017 1:44	EMGERGENCY STARTS	FO1	17	EMERGENCY STARTS of patient ID:
8	E_17_06062017_0100	6/06/2017 1:44	BTG STARTS	FO1	17	BTG STARTS of patient ID: 17
9	E_17_06062017_0100	6/06/2017 1:44	emergency Original ID	emergency:Dy		Original ID of this emergency is :Dy
10	E_17_06062017_0100	6/06/2017 1:44	emergency login	emergency:Dy		success: 127.0.0.1
11	E_17_06062017_0100	6/06/2017 1:44	view	emergency:Dy	17	
12	E_17_06062017_0100	6/06/2017 1:44	PRS:Problems List:Examine patient and the information	emergency:Dy	17	SELECT `title` FROM `lists` WHERE `
13	E_17_06062017_0100	6/06/2017 1:44	PRS:Problems List:Examine patient and the information	emergency:Dy	17	SELECT * FROM lists WHERE pid = ?
14	E_17_06062017_0100	6/06/2017 1:44	PRS:Problems List:Examine patient and the information	emergency:Dy	17	SELECT * FROM lists WHERE pid = ?
15	E_17_06062017_0100	6/06/2017 1:44	PRI:Problems List:Modify medical problems	emergency:Dy		INSERT INTO `clinical_rules_log` ('d
16	E_17_06062017_0100	6/06/2017 1:44	OR:Medication	emergency:Dy	17	REPLACE INTO prescriptions SET `ic
17	E_17_06062017_0100	6/06/2017 1:45	emergency logout	emergency:Dy		success
18	E_17_06062017_0100	6/06/2017 1:45	BTG ENDS	FO1	17	BTG ENDS of patient ID: 17
19	E_17_06062017_0100	6/06/2017 1:45	EMGERGENCY ENDS	FO1	17	EMERGENCY ENDS of patient ID: 1;

(c) Classified Log

Figure 3. Final log after the "log preparation" phase.

For the unacceptable cases we simulated unauthorised access of a physician who obtained emergency access rights in a BTG situation for a patient. This inappropriate behaviour occurred before the "clean-up" process to disable the pre-owned emergency accounts to prevent their reuse (Yale University, 2021). In the simulated scenario the physician takes advantage of the access granted to

Table 1. Conformance checking results from ProM.

Case-ID	BTG/Emergency	Acceptable/Unacce	The deviated transitions
E-18-18052017-1600	BTG	Acceptable	N/A E-35-18052017-1700
Emergency	Acceptable	N/A E-35-18052017-1900	BTG
Unacceptable	Emergency starts, BTG starts, Emergency ends and BTG ends E-17-06062017-0100	Normal	Acceptable
N/A E-18-25092017-0000	BTG	Unacceptable	Emergency starts, BTG starts, Emergency ends, and BTG ends

the patient's health record to extract data from it after the emergency has ended, possibly for financial gain by selling the data to the media in the case of a celebrity patient. The transitions that deviate from the workflow model in each case were visualised using the conformance analysis user interface. These deviated transitions indicate potential health information privacy violations. Eight cases were found to be aligned with the workflow model while two were noted as cases with deviations. Each case can be viewed in ProM to detect the exact points of deviation within a case. Table 1 summarises 5 cases out of 10 simulated cases and the transitions (events) that were automatically identified as invalid (i.e., unacceptable) are indicators of a potential privacy violation.

References

Cohen, I. G., Mello, M. M. (2018). *HIPAA and protecting health information in the 21st century. Jama,* 320(3), 231-232.

Health IT Dashboard, (2020, March). *The Office of the National Coordinator for Health Information Technology. https://dashboard.healthit.gov/index.php*

Weitzner, D. J., Abelson, H., Berners-Lee, T., Feigenbaum, J., Hendler, J., Sussman, G. J. (2008). *Information accountability. Communications of the ACM*, 51(6), 82-87.

Kacianka, S., Beckers, K., Kelbert, F., Kumari, P. (2017, November). How accountability is implemented and understood in research tools. *International Conference on Product-Focused Software Process Improvement* (pp. 199-218). Springer, Cham.

Wickramage, C., Sahama, T., Fidge, C. (2016, September). Anatomy of log files: implications for information accountability measures. *2016 IEEE 18th International Conference on e-Health IEEE.Networking, Applications and Services (Healthcom)*, (pp. 1-6).

Rozinat, A., Van der Aalst, W. M. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems, 33(1), 64-95.*

Yale University, (2021, January). *Health Insurance Portability and Accountability Act. https://hipaa.yale.edu /security/break-glass-procedure-granting-emergency-access-critical-ephi-systems*

Wickramage, C., Fidge, C., Ouyang, C., Sahama, T. (2019, January). Generating log requirements for checking conformance against healthcare standards using workflow modelling. *Proceedings of the Australasian Computer Science Week Multiconference,*(pp. 1-10).