Clinical Processes and Its Data, What Can We Do with Them?

Eric Rojas, Michael Arias and Marcos Sepúlveda
Computer Science Department, School of Engineering, Pontificia Universidad Católica de Chile, Macul, Santiago, Chile

Keywords: Process Mining, Healthcare Processes, Process Discovery, Healthcare Data, Process Mining Methodology.

Abstract: Global healthcare services have evolved over time, and nowadays they are expected to follow high-quality optimized standards. Analyzing healthcare processes has become a relevant field of study, and different techniques and tools have been developed to promote improvements in the efficiency and effectiveness of these processes. There is a research field called process mining that can be used to extract knowledge from the event data stored in the hospital information systems. With the help of this, it is possible to discover the real executed process, examine its performance and analyze the resource interaction during its execution. The goal of this article is to provide a bibliographic survey about the use of process mining algorithms, techniques, and tools in the analysis of healthcare processes, providing a general overview about the main approaches previously used and the information required to apply them in the medical field. We provide important insights about data, algorithms, techniques and methodologies that are required to help answer medical expert questions about their processes, motivating and inspiring a broader usage. So, if we have the information and it is possible to analyze and understand the healthcare processes, why are we not doing it?

1 INTRODUCTION

Medical centers around the world perform a large variety of processes that are relevant because, depending on their proper execution, they have a direct impact on people's health. Therefore, there is a clear need to ensure that the implementation of these medical processes are carried out in the most effective and efficient manner possible. The information currently stored in healthcare information systems might provide valuable insights about how these processes are being performed helping to propose potential improvement that could enhance their performance.

Process mining has emerged as a research discipline that can discover, monitor and improve real processes by extracting knowledge from event logs readily available in today's information systems (Van der Aalst et al., 2012). This discipline, has a great potential for application in the healthcare domain because medical processes are considered complex, large and with a lot of variability in time (Homayounfar, 2012).

The objective of this work is to present a systematic overview of the different approaches that have been used to analyze healthcare processes using process mining, including a compilation of types of data needed (Kaymak & Mans, 2012; RS Mans & Aalst, 2013), questions that can be answered (RS Mans & Aalst, 2013), methodologies that have been applied (Ferreira, 2012; Van der Aalst et al., 2012), and techniques/algorithms available to perform the analysis (e.g. (Günther & van der Aalst, 2007; Weijters & van der Aalst, 2003)). It also covers a detailed analysis of the geographical location of the main case studies that have been done in this domain and the leading medical fields where it has been applied. Gathering all this information should help professionals, in both process mining and healthcare domains, to identify new research opportunities to improve healthcare services using process mining techniques.

The outline of the paper is as follows. In section 2, we introduce the basics of process mining in healthcare. In section 3, we give a list of the main approaches applied. Main challenges are explained in Section 4 and the conclusions in Section 5.

2 RELATED WORK

2.1 Process Mining

Process mining is a relatively young research discipline. It focuses on extracting knowledge from data generated and stored in databases of (corporate)

Rojas E., Arias M. and Sepúlveda M..
Clinical Processes and Its Data, What Can We Do with Them?.
DOI: 10.5220/0005287206420647
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2015), pages 642-647
ISBN: 978-989-758-068-0

information systems. Process-Aware Information Systems (PAISs) (Dumas, Van der Aalst, & Ter Hofstede, 2005) are systems that are readily able to produce event logs. Specific examples of such applications are ERP systems (e.g. SAP) and Customer Relationship Management systems (CRM). Event log data is not limited only to the data from these applications, because many other systems also provide useful data. Moreover, the data about a complex process might not come from a single source of information.

According to (Van der Aalst et al., 2007), it is possible to store event information where: (i) each event refers to an activity, (ii) each event refers to a case, (iii) each event can have a performer, also referred to as originator (the person executing or initiating the activity), and (iv) each events have a timestamp and are totally ordered.

There are three main types of process mining: process discovery, conformance checking and enhancement. In (Van der Aalst et al., 2012), it is described that automatic process discovery allows extracting process models from an event log; conformance checking to monitor deviations by comparing a given model and the event log, and enhancement allows extending or improving an existing process model using information about the actual process recorded in the event log. It is possible to extend the analysis through organizational mining, the automatic construction of simulation models, extension models, predicting cases and other approaches.

2.2 Process Mining in the Health Domain

Healthcare processes are seen as a domain that has complex models and vary a lot in time (Homayounfar, 2012). Being able to use techniques to discover workflow models and analyze the performance of them, are great opportunities to examine the information stored in the events logs of hospital information systems.

Using process mining techniques in healthcare processes, allows not only to understand what is really happening with them, but also, can generate benefits associated with process efficiency, improving the quality of services provided, as well as a positive impact on the management of medical centers. For this, it is relevant to take advantage of the capabilities offered by processes mining to collaborate in the construction of event logs and analyze healthcare processes. Also, incorporated medical knowledge can generate results that can provide data to improve medical practices in hospitals. The process mining research area has been used in the healthcare processes field to discover process models from event logs (Bose & Aalst, 2011; Mans, Reijers, van Genuchten, & Wismeijer, 2012; RS Mans & Schonenberg, 2009), do conformance checking (Kirchner, Herzberg, Rogge-solti, & Weske, 2012; Zhou, 2009) and evaluate social networks analysis (Bose & Aalst, 2011; Lang, Bürkle, Laumann, & Prokosch, 2008; Mans et al., 2012).

3 MAIN APPROACHES APPLIED

In this section, the type of processes, the type of data, the frequently posed questions, the methods/algorithms, and the used methodologies, are described. Besides, the case studies are characterized according to their medical field and their geographical location.

3.1 Types of Processes

It is critical to understand the different types of processes that are executed in the hospital environments, so the algorithms/techniques and process mining tools are correctly applied. According to (Dumas et al., 2005; Ferreira, 2012) and (Kaymak & Mans, 2012), there are two types of processes in the healthcare domain: the medical treatment processes and the organizational processes.

Medical Treatment Processes: These are the clinical processes of managing a patient. These processes include from the diagnostic actions based on symptoms, to the execution of a series of actions, to relief the patient.

Organizational Processes: These are the ones focused on the organizational knowledge of the process, capturing the collaborative information of the healthcare professionals and their organizational units.

3.2 Types of Data

Based on the type of process and the type of analysis that is needed, it is necessary certain data to do it. In this section, 2 classifications are described. The first classification is the one included in (Kaymak & Mans, 2012), based on the data used on their process mining case study. The second one is presented on

(RS Mans & Aalst, 2013), and it is based on the data source and its level of abstraction, accuracy, granularity, directness, and correctness. To allow a correct execution of the process mining techniques, and give the expert the knowledge of its processes it is necessary to have clarity of the type of data that it is available and how it should be managed.

3.3 Frequently Asked Questions

The process mining techniques gives the medical specialists the ability to answer some of the questions related to their processes, acknowledging them the improvement opportunities. According to the questions that are raised by the specialists, data and information must be gathered from different sources. Four typical questions that want to be answered by medical process specialists (RS Mans & Aalst, 2013), are:

- *What are the most followed paths and what exceptional paths are followed?*
- *Are there any differences in care path followed by different patient groups?*
- *Do we comply with internal and external guidelines?*
- *Where are the bottlenecks in the process?*

Besides these interrogations, we propose a fifth question to include the organizational collaboration process between the specialists:

- *What are the roles and social relationships between the medical staff?*

These questions guide the use and application of the process mining techniques, to provide correct and accurate answers to the real needs of the specialists.

3.4 Used Methods and Algorithms

Through the tools available for process mining there are several techniques implemented to execute the correct and desired analysis. Some of the main techniques that have been applied are: Trace Clustering (Caron et al., 2014; RS Mans & Schonenberg, 2009), Performance Sequence Analyzer (Butler-Henderson, 2012a; Caron et al., 2014), Fuzzy Miner (Kim, Kim, Song, Kim, & Yoo, 2013; RS Mans & Schonenberg, 2009), Alpha Miner (Lang et al., 2008), Genetic Miner (Fei & Meskens, 2010; Lang et al., 2008), Heuristic Miner (Caron, Vanthienen, & Baesens, 2013; Kaymak & Mans, 2012; Kim et al., 2013) and Conformance Checker (Dewandono, Fauzan, Sarno, & Sidiq, 2013; Zhou, 2009).

3.5 Used Methodologies

Two main methodologies have been used: a methodology including clustering techniques (Caron et al., 2013; Doremalen, 2012), and a methodology following the L* life-cycle model (Binder et al., 2012; Van der Aalst et al., 2012).

3.6 Implementation Strategies

A classification of strategies used to implement process mining is described. The first one is the basic direct strategy, which involves the direct application of the process mining tools to a set of data extracted directly from a data source and used to build an event log (Caron et al., 2014; Mans et al., 2012; RS Mans & Schonenberg, 2009; Rebuge, Lapao, Freitas, & Cruz-Correia, 2013). This strategy has two basic challenges: data extraction and event log construction. The second strategy is the semi-automated, where the data and event log construction is done through a specific solution. These solutions connect to one or several data sources and extract the correct data to build the event log, but still need the knowledge of the process mining tools to execute it (Helmering, Harrison, Iyer, Kabra, & Van Slette, 2008). This strategy has the disadvantage that is a local solution.
The third strategy to implement process mining is to do it in a specific suite, where you can connect to the data sources, extract the data, build an event log and execute the process mining techniques. This has the advantage that the person that will use the suite does not need to know in detail how to connect to the data sources, and how to use the process mining tools. The disadvantage is that the suites are developed for a specific environment and its data sources. Examples of this implementation include the Medtrix Process Mining Studio (Ferreira, 2012) and the Emotiva Tool (Fernández-Llatas *et al*, 2013).

3.7 Types of Case Studies

There are several ways to build or execute a case study using process mining techniques/algorithms. They all include the use of data and the creation of an event log, but can vary according to the process mining tool they use. The tool they use determines the type of case study, it can be the typical tools, a new development made for the specific case study or the use of the existing techniques with additional techniques from outside the process mining field. Following, is a proposed classification. The first type defined is the basic case study, which takes the

data from one or several data sources and builds the event log and executes the process mining techniques using the tools available. There is no new implementation done in this type of case study and the main objective is to give knowledge of the healthcare process. Several case studies have been done of this type (Mans, R., & Schonenberg, H., 2008; RS Mans & Aalst, 2013; RS Mans & Schonenberg, 2009). The second one is the case study where a new technique or algorithm has been developed to complement the actual available tools. Some of these case studies are (Bose & Aalst, 2011; Butler-Henderson, 2012b; Gupta, 2007). The third type is the one that uses the process mining tools and techniques available, but also incorporates techniques from other fields like statistical analysis (Butler-Henderson, 2012b), extended data analysis (Bose & Aalst, 2011), data mining and CRISP-DM (Zhou, 2009) and DECLARE (Grando, Aalst, & Mans, 2011; Grando, Schonenberg, & Aalst, 2011). These are important due to the mixture of techniques.

3.8 Process Mining Tools

A group of software applications is available to do analysis through process mining. The most used tool is ProM, a tool that has a large number of algorithms/techniques implemented and it is considered as the "pluggable" environment for process mining (van Dongen, *et al*, 2005). In healthcare domain, ProM has been used in several case studies, for instance: (Bose & Aalst, 2011; Lang et al., 2008; RS Mans & Schonenberg, 2009; Zhou, 2009). DECLARE is a flexible constraint-based workflow management suite, used to model medical guidelines in (Grando, Aalst, et al., 2011; Grando, Schonenberg, et al., 2011). It is interesting that the popular process mining toolkit DISCO is not mentioned as much as expected in these case studies. Only in (Perimal-lewis, Vries, & Thompson, 2014) is reported as a main tool. An emerging process mining tool called PmLab, has not been used so far.

3.9 Geographical Analysis and Medical Fields

The use of process mining has been growing in the last couple of years, becoming an important tool to analyze the medical processes and generate improvements opportunities. A quantitative count and geographical classification was performed based on the case studies available. The highest concentration of case studies is in Europe, existing only a few in North America, Asia and Australia. No case studies have been done in Africa or South America. In more detail, Netherlands is the country that has published most studies, followed by Belgium and Germany.

Besides the geographical analysis, the case studies have been divided and classified into medical domains. Some of these test cases include Cardiology data, Caregiving Processes data, Dentistry data, Diabetes data, Intensive Care Unit data, Medication data, Oncology data, Radiotherapy data and surgical data. Oncology and surgery (nine and five case studies) are the medical fields where most case studies have been done. In average two or fewer case studies have been done per field.

4 CHALLENGES

The main challenges and limitations that experts have found are the following: (1) Satisfying medical protocols and guidelines, (2) Including medical knowledge, (3) Including the physical information and conditions of the patients, (4) Identifying and accessing data sources, (5) Data integration from different sources, (6) Quality of data (incorrect and incomplete data), (7) Granularity and preprocessing of the data, (8) Using real event logs and data (and its complexity of cases), and not only synthetic data, and (9) Building the correct and complete event log. These limitations relate to data sources, event log construction and how to include semantic knowledge from the medical experts.

5 CONCLUSIONS

In process mining is possible to discover a process model, perform conformance analysis and propose opportunities to enhance any process. Healthcare experts can benefit from these tasks, allowing them to find process improvement opportunities. We provide a useful bibliographic survey, which gives an overview of main approaches needed to apply process mining in the healthcare domain. Some challenges and limitations were identified, related with data sources, event log construction, and adding semantic knowledge from medical specialists. Overall, this article aims to serve as a motivational guide that collects some case studies previously conducted in the healthcare field, and provides an outline about what must be taken into consideration to carry out a process mining project. These outlines

involve the approaches explained and the different kinds of algorithms, techniques and process mining tools. We are able to gain significant insight as a result of performing process mining in healthcare, taking advantage of the data stored in the medical information systems is what makes this possible.

REFERENCES

Binder, M., Dorda, W., Duftschmid, G., Dunkl, R., Fr, K. A., Hronsky, M., & Rinderle-ma, S. (2012). On Analyzing Process Compliance in Skin Cancer Treatment : An Experience Report from the Evidence-Based Medical Compliance Cluster (EBMC2). In *CAiSE 2012* (pp. 398–413).

Bose, R., & Aalst, W. van der. (2011). Analysis of Patient Treatment Procedures. *Business Process Management*.

Butler-Henderson. (2012a). *Health Informatics and Knowledge Management 2012* (Vol. 129).

Butler-Henderson. (2012b). *Health Informatics and Knowledge Management 2012* (Vol. 129).

Caron, F., Vanthienen, J., & Baesens, B. (2013). Healthcare Analytics: Examining the Diagnosis–treatment Cycle. *Procedia Technology*, *9*, 996–1004.

Caron, F., Vanthienen, J., Vanhaecht, K., Limbergen, E. Van, De Weerdt, J., & Baesens, B. (2014). Monitoring care processes in the gynecologic oncology department. *Computers in Biology and Medicine*, *44*, 88–96.

Dewandono, R. D., Fauzan, R., Sarno, R., & Sidiq, M. (2013). Ontology and Process Mining For Diabetic Medical Treatment Sequencing. In *7th International Conference on Information & Communication Technology and Systems (ICTS)* (pp. 171–178).

Doremalen, B. Van. (2012). *Process Mining in Healthcare Systems : An Evaluation and Refinement of a Methodology*.

Dumas, M., Van der Aalst, W. M. P., & Ter Hofstede, A. H. M. (2005). *Process-aware information systems: bridging people and software through process technology*. John Wiley & Sons.

Fei, H., & Meskens, N. (2010). Discovering Patient Care Process Models From Event Logs. *8th International Conference of Modeling.*

Fernández-Llatas, C., Benedi, J.-M., García-Gómez, J. M., & Traver, V. (2013). Process mining for individualized behavior modeling using wireless tracking in nursing homes. *Sensors (Basel, Switzerland)*, *13*(11), 15434–51.

Ferreira, D. R. (2012). Business process analysis in healthcare environments : A methodology based on process mining, *37*, 99–116.

Grando, M. A., Aalst, W. M. P. Van Der, & Mans, R. S. (2011). Reusing a Declarative Specification to Check the Conformance of Different CIGs. In *BPM 2011 Workshops* (pp. 188–199).

Grando, M. A., Schonenberg, M. H., & Aalst, W. Van Der. (2011). Semantic-Based Conformance Checking of Computer Interpretable Medical Guidelines. In *BIOSTEC 2011* (pp. 285–300).

Günther, C. W., & van der Aalst, W. M. P. (2007). Fuzzy Mining – Adaptive Process Simplification Based on Multi-Perspective Metrics. *In Business Process Management*, 328–343.

Gupta, S. (2007). Workflow and process mining in healthcare. *Master's Thesis, Technische Universiteit Eindhoven*, (May).

Helmering, P., Harrison, P., Iyer, V., Kabra, A., & Van Slette, J. (2008). *Process Mining of Clinical Workflows for Quality and Process Improvement*. *Mercy Health System* (pp. 1–7).

Homayounfar, P. (2012). Process mining challenges in hospital information systems. *In Computer Science and Information Systems (FedCSIS), 2012 Federated Conference on*, 1135–1140.

Kaymak, U., & Mans, R. (2012). On process mining in health care. *... Conference on*, 1859–1864.

Kim, E., Kim, S., Song, M., Kim, S., & Yoo, D. (2013). Discovery of Outpatient Care Process of a Tertiary University Hospital Using Process Mining, *19*(1), 42–49.

Kirchner, K., Herzberg, N., Rogge-solti, A., & Weske, M. (2012). Embedding Conformance Checking in a Process Intelligence System in Hospital Environments. In *2012 BPM workshops* (pp. 126–139).

Lang, M., Bürkle, T., Laumann, S., & Prokosch, H. (2008). Process Mining for Clinical Workflows : Challenges and Current Limitations. *eHealth Beyond the Horizon – Get IT There S.K. Andersen et Al. (Eds.) IOS Press, 2008 © 2008 Organizing Committee of MIE 2008. All Rights Reserved.*, 229–234.

Mans, R., & Aalst, W. van der. (2013). Process mining in healthcare: Data challenges when answering frequently posed questions. *... in Health Care*, 107–120.

Mans, R., Reijers, H., van Genuchten, M., & Wismeijer, D. (2012). Mining processes in dentistry. *Proceedings of the 2nd ACM SIGHIT Symposium on International Health Informatics - IHI '12*, 379.

Mans, R., & Schonenberg, H. (2008). Process mining techniques: an application to stroke care. *Studies in Health …*, *136*, 573–578.

Mans, R., & Schonenberg, M. (2009). Application of process mining in healthcare–a case study in a dutch hospital. *Biomedical Engineering*, 425–438.

Perimal-lewis, L. U. A., Vries, D. D. E., & Thompson, C. H. (2014). Health intelligence : Discovering the process model using process mining by constructing Start-to-End patient journeys, (Hikm), 59–67.

Rebuge, A., Lapao, L. V., Freitas, A., & Cruz-Correia, R. (2013). A process mining analysis on a Virtual Electronic Patient Record system. *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, 554–555.

Van der Aalst, W. M. P., Adriansyah, A., Karla, A., Medeiros, A. De, Arcieri, F., Baier, T., … Song, M. (2012). Process Mining Manifesto. *In Business Process Management Workshops*, 169–194.

Van der Aalst, W. M. P., Reijers, H. a., Weijters, a. J. M. M., van Dongen, B. F., Alves de Medeiros, a. K., Song, M., & Verbeek, H. M. W. (2007). Business process mining: An industrial application. *Information Systems*, *32*(5), 713–732.

Van Dongen, B. F., Medeiros, A. K. A. De, Verbeek, H. M. W., & Weijters, A. J. M. M. (2005). The ProM Framework: A New Era in Process Mining Tool Support. *In Applications and Theory of Petri Nets 2005*, 444–454.

Weijters, A. J., & van der Aalst, W. M. P. (2003). Rediscovering Workflow Models from Event-Based Data using Little Thumb. *Integrated Computer-Aided Engineering*, *10*, 151–162.

Zhou, W. (2009). Process mining: acquiring objective process information for healthcare process management with the CRISP-DM framework.