Ferret Miner: A Process Mining Case Study

Martin Alvarez-Lopez*   Hardy Leung†   Carlos Hernandez‡   Divyam Sobti§

December 2022

Abstract

The present project analyzed event data logs from the travel reimbursement process at Eindhoven University of Technology (ETU/c). The objective is to identify the bottlenecks in the process of travel declaration and improve its efficiency. Using Process Mining techniques, a control-flow model and a value stream map was generated to help identify any delays in the process. Subsequently, bottlenecks were detected using a representative analysis. The implementation results recognized bottlenecks in critical tasks of the process that were optimized by assigning more resources to the area. The optimized process improved its efficiency, compared with the original, by almost 30%, significantly reducing the time processing of the travel reimbursement.

1 Introduction

Process mining consists of a group of analysis techniques that reveal core information from processes through event data logs [1]. Error and deficiencies can be detected using this methodology. Process mining and data mining are two different, but closely related data analytical systems since both belong to the business intelligence field. However, data mining concentrates on finding data relationships, whereas process mining focuses on analyzing data as well but from a systematic perspective. Currently, information systems, such as Enterprise Resource Planning (ERP), Work on Management Systems (WMS) or Customer Relationship Management (CRM), are vital tools for the execution of almost any company process [2]. The essence of these information systems relies on their capabilities of logging events. An event log is basically a record of an activity performed with a timestamp embedded [3]. Event logs facilitates process mining to discover information by tracking the chronology of the tasks and their execution time registered in each log.

1A Problem Description

This project performs a process mining analysis to detect patterns, delays and bottlenecks that obstruct the efficient flow of the process activities. The main objective is to provide a case study assessment using a dataset with real information provided from the Eindhoven University of Technology. The research data is from the BPI Conference 2020 held in Padova, Italy [4].

The appropriate supervision of a business behavior enables the opportunity to adjust any task process that is generating a problem [5]. Therefore, our approach focuses on studying the problem using process mining methodology. Process mining is an outstanding data analysis technique that generates models representing the behavior indicated by activity logs produced by business applications after the execution of a process task.

Our approach proposes a real case study analysis using process mining technique. We utilize a structured dataset based on systematic event logs. The system logs provide a detailed record of the process task, identification, resource, timestamp, and other data. These attributes provide insights of successful execution of the process, or any other problem encountered at any

---

*MS in Software Engineering, San José State University
†MS in Artificial Intelligence, San José State University
‡MS in Computer Engineering, San José State University
§MS in Artificial Intelligence, San José State University

stage of the activities process flow.

The fundamentals of our approach were introduced by Wil Van Der Aalst in his paper titled "Process Mining" [6], where process mining guidelines are detailed. In general, a process mining implementation covers three main issues: 1) Discovery, 2) Conformance, and 3) Enhancement. Our project focuses on the Discovery since it is the area that requires more effort, and it is the most deterministic for the success of the analysis.

1B Motivation

Practical application publications are scarce, and our research paper aims to add support to this section. Furthermore, another motivation to complete this investigation relies on the fact of resolving a real problem using process mining techniques implemented with python PM4PY library modules. We found python PM4PY a stable open-source library available for everyone with great capabilities to resolve our problem. The goal is to demonstrate the practicality of process mining analysis in practice. Our paper illustrates a case study in the processing of travel reimbursement for a university in Germany. The travel reimbursement issue involves several business processes for which analysis of event logs registration is worthwhile. The case being developed involves s number of students who submitted their application, added their support documentation and a series of tasks that need to be validated before the disbursement approval.

1C Approach Summary

The document is structured as follows: Section 1 provides details about process mining and its fundamentals; Section 2 presents an overview of our approach and related literature review; Section 3 analyzes the outcomes produced from our primary studies and research problem. Section 4 refers to discussion. Finally, Section 5 concludes the work.

2 Survey

The literature review for process mining related topics concentrates on the development of innovative techniques and algorithms from a control-flow discovery perspective and its applications.

2A Process Mining

The position of process mining considering other disciplines is just between data mining and process modeling [6]. Such interaction is represented in figure one.
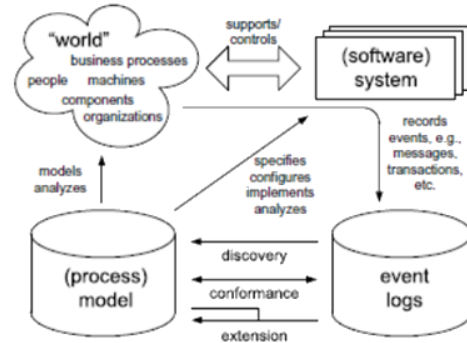

Figure 1 *General outline of process mining approach.*

Usually, information system databases do not have clearly defined structures which means that a process mining analysis must search and separate the information directly from the system. Under these terms, process mining is considered a technique where the event logs possess the right information to determine the behavior of the system. Event logs, then, are guidelines that dictate the applicability of a process mining implementation [7]. As a result, process mining approach embraces three main actions (Figure 2): 1) Discovery, which creates models by analyzing event logs 2) Conformance checking, which basically enables to compare a newly created model versus a predefined model to gain knowledge 3) Enhancement, which deals with the implementation of solutions to overcome previously detected problems.
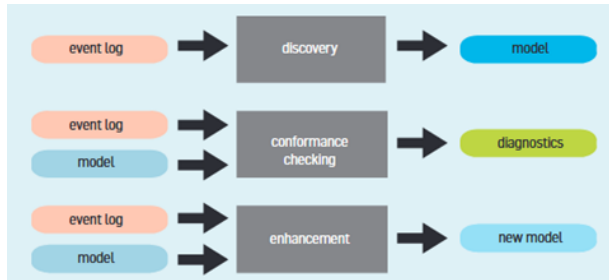
General outline of process mining approach

Figure 2 *Process mining layout.*

2B Process Models

Distinct data analysis approaches have been developed to discover processes from event logs. However, Petri Nets technique has taken relevance over the others [8]. Another approach utilizes Unified Modeling Language (UML) diagrams [9]. Furthermore, A novel design named "alpha algorithm" is capable of depicting event data logs into work-flow diagrams, which is a derivation from Petri nets. Nowadays, a popular model is "Heuristic Miner." One main characteristic of this method is to be highly permissive to distinct statistical levels and thresholds adjusted by the user [10]. In order to deal with the massive number of events that actual information systems generate, Burattin et al. [11] developed a Heuristic Miner adaptation based on Lossy Counting, as well as a sliding window that effectively analyzes streams of real datasets.

2C Applications

Most work-related publications are dedicated mainly to the enhancement of novel methodologies and algorithms with significant emphasis on the control-flow discovery area [7]. However, the public services area has been common for process mining development research. For instance, a Genetic Miner approach was tested by Alves de Madeiros et al. [12] in Dutch municipalities. Additionally, van del Aalst et al. [13] applied an analysis of organizations in another municipality process. Rozinat et al. [14] evaluated process models in two different case studies on the public sector. Nonetheless, all these research applications are focused to validate and assess rather than enhance the process. Implementation of process mining

approaches have been also developed in the private domain. Distinct discovery approaches were deployed by Goedertier et al. [15] in the telecom field. Also, Mans et al. [16] and Rebuge and Ferreira [17] created a convoluted process mining method to resolve the tracking of patients within the healthcare industry.

The aforementioned literature review provides the fundamentals of process mining from a general perspective. Each approach makes its own significant contribution to the process mining study field, and in this paper, we focus on its application.

3 Technical Approach

using proper mathematical notation and formatting

4 Experimental Methodology

description of your data set, the results you found (formatted as either tables or plots with proper labels and captions), and a discussion of your results. If you did a theoretical project, you should have an expanded technical approach instead of this section

4A Dataset Description

4B Results (include tables and Plots)

4C Discussion of Results

5 Conclusion

A brief conclusion

References

[1] "What is process mining? All you need to know." https://quixy.com/blog/all-about-process-mining/.

[2] "Tutorial:introduction | ProM tools." https://www.promtools.org/doku.php?id=tutorial:introduction.

[3] "Event log – the process mining glossary | appian." https://appian.com/process-mining/event-log.html.

[4] "BPI challenge 2020." https://icpmconference.org/2020/bpi-challenge/.

[5] F. Leymann and W. Altenhuber, "Managing business processes as an information resource," *IBM systems journal*, vol. 33, no. 2, pp. 326–348, 1994.

[6] W. Van der Aalst, T. Weijters, and L. Maruster, "Workflow mining: Discovering process models from event logs," *IEEE transactions on knowledge and data engineering*, vol. 16, no. 9, pp. 1128–1142, 2004.

[7] W. Van Der Aalst, "Process mining," *Communications of the ACM*, vol. 55, no. 8, pp. 76–83, 2012.

[8] J. E. Cook and A. L. Wolf, "Discovering models of software processes from event-based data," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 7, no. 3, pp. 215–249, 1998.

[9] S. Saito, "Understanding key business processes for business process outsourcing transition," in *2019 ACM/IEEE 14th international conference on global software engineering (ICGSE)*, 2019, pp. 35–39.

[10] A. Weijters, W. M. van Der Aalst, and A. A. De Medeiros, "Process mining with the heuristics miner-algorithm," *Technische Universiteit Eindhoven, Tech. Rep. WP*, vol. 166, no. July 2017, pp. 1–34, 2006.

[11] A. Burattin, A. Sperduti, and W. M. van der Aalst, "Control-flow discovery from event streams," in *2014 IEEE congress on evolutionary computation (CEC)*, 2014, pp. 2420–2427.

[12] A. K. A. de Medeiros, A. J. Weijters, and W. M. van der Aalst, "Genetic process mining: An experimental evaluation," *Data mining and knowledge discovery*, vol. 14, no. 2, pp. 245–304, 2007.

[13] A. Rozinat, I. S. M. de Jong, C. W. GÜnther, and W. M. P. van der Aalst, "Process mining applied to the test process of wafer scanners in ASML," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 39, no. 4, pp. 474–479, 2009, doi: 10.1109/TSMCC.2009.2014169.

[14] A. Rozinat, R. S. Mans, M. Song, and W. M. van der Aalst, "Discovering simulation models," *Information systems*, vol. 34, no. 3, pp. 305–327, 2009.

[15] S. Goedertier, J. De Weerdt, D. Martens, J. Vanthienen, and B. Baesens, "Process discovery in event logs: An application in the telecom industry," *Applied Soft Computing*, vol. 11, no. 2, pp. 1697–1710, 2011.

[16] R. S. Mans, M. Schonenberg, M. Song, W. M. van der Aalst, and P. J. Bakker, "Application of process mining in healthcare–a case study in a dutch hospital," in *International joint conference on biomedical engineering systems and technologies*, 2008, pp. 425–438.

[17] Á. Rebuge and D. R. Ferreira, "Business process analysis in healthcare environments: A methodology based on process mining," *Information systems*, vol. 37, no. 2, pp. 99–116, 2012.

Appendix: Team Contributions

In this section we have to describe what/where each team member work on (according to project guidelines on CANVAS)