Process Enhancement in Process Mining: A Literature Review

Fitri Almira Yasmin[1], Faiza Allah Bukhsh[1] and Patricio de Alencar Silva[2]

[1] Department of Computer Science University of Twente, Enschede, The Netherland
fitrialmirayasmin@student.utwente.nl, f.a.bukhsh@utwente.nl

[2] Department of Computer Science Federal University of the Semi-arid Region, Rio Grande do Norte, Brazil
patricio.alencar@ufersa.edu.br

Abstract. The need for continuous improvement of business process drives the adoption of analytics tools and techniques that provide insight into business. One of the techniques that can explore an entity's business process is Process mining. Process mining approach consists of numerous alternatives depending on their tools, types and perspectives. To better describe the topic of process mining and what benefits it has to offer, we provide an overview of one of the prominent dimensions of process mining: process enhancement. In this paper, we provide an overview of practices in process enhancement within 43 relevant articles. We highlight the trends of the tools, techniques, and perspectives that are used in the studies. We also describe the application of process enhancement in various industry. Our concluding findings emphasise the importance of considering objectives and constraint of the enhancement process in the selection of process enhancement approach.

Keywords: Process Mining, Process Enhancement, Process Model Repair, Process Model Extension

1 Introduction

The idea of process mining was introduced by Aalst in 2004 [50]. Process mining is data analytics techniques whose goals is to extract process-related information [46] and specifically focused on analysing historical data of process executions in the form of event logs [47]. Numerous process mining techniques, tools and plug-ins are available; they provide fact-based insights and support process improvements. Through this study, we contribute to give an overview of the state of the art of process enhancement practices within the process mining field that are scientifically recorded.

In order to achieve the stated objectives, the study structured as follows: Section 2 describes the research question of the study and research methodology. Following that, Section 3 presents the findings that answer our research questions. Then, in Section 4 we will further discuss the implication of our findings with some threads to validity of our analysis. Lastly, Section 5 concludes this study.

2 Research Method

The goal of this study is to provide an overview of the topic of process enhancement in process mining for both practitioners and process owners. This goal is realized by defining the following research questions:

1. How process enhancement defined? What is the most cited definition of process enhancement within the process mining field? (Definition)
2. What are the practice trends of process enhancement topics in process mining field that are scientifically reported? Which type of perspectives and tools used for process enhancement? (Practice)
3. How process mining approaches influence process enhancement? (Application)

The literature search started by querying keywords through several online databases. We acquired 46 articles in ACM Library by querying searching articles containing at least of these phrases "*Process Enhancement*", "*Enhancement*", "*Perspective*" and "*Repair*"; and limiting the results that with the keyword "*Process Mining*". We found another 48 Articles from IEEE Xplore and 226 from Scopus by applying the same query. Lastly, we search for articles with the phrase "*Process Enhancement*" and the keywords "*Process Mining*" in ScienceDirect, returning 45 articles. The aforementioned queries are chosen with the consideration to limit our search to process mining spectrum. After utilising the automatic duplication identifier in EndNote, the total 346 articles identified. These articles then are selected with the criteria of:

Include only the literature that is written in English.
Include only the latest version of the article.
Include only literature that discusses process enhancement that falls within the process mining field.
Include only the articles that implicitly indicate the utilisation of process mining techniques, conception or tools in the scope of process enhancement.
Exclude duplicating article. For instance, in some cases, there are identical articles that available on different digital libraries. In that case, we chose one of the articles.

The initial relevance determined by the title and the keywords of the literature. This phase aims to eliminate the articles that do not belong within the scope of process mining and to eliminate the duplications that were not identified automatically. After initial screening, 241 articles chosen. Next step is to narrow the scope to the topic of process enhancement within the process mining field by reading the article's abstract and conclusion which reduced the number of articles to 95. After the full-text reading, the final total of 43 articles was included in this study.

3 Findings

In this section, we elaborate on current practices of process enhancement within the process mining field that `are scientifically recorded, based on our literature review.

3.1 Process Enhancement Definition

Based on our literature analysis, "*the extension or improvement of an existing process model using information about the actual process recorded in some event log* [46]" is the most cited definition of process enhancement within the process mining field. Out of 43 articles included in our study, 10 mentioned the concept of process enhancement stated in [46]. Six among them [1] [7] [9] [11] [16] [36] adopted Aalst definition of process enhancement, and the other three [20] [21] [25] described the conception of process extension instead of the process enhancement itself. Process extension is "a form of process enhancement where apriori model is extended with a new aspect or perspective" also defined by Aalst [46].

3.2 Process Enhancement Practices

After discovering the process model, the commonalities and discrepancies between the modeled behavior and observed behavior can be found through conformance checking [46]; i.e., conformance checking process describes whether the process model reflects the actual process in the recorded event log. If the process model does not reflect the reality, then either repair or extend the model so they can replay most of the event log.

Table 1. Process Enhancement Types and Perspectives

Type	**Perspective**	**Articles**
Repair	Control-flow	[1] [4] [5] [7] [12] [22] [23] [24] [36]
Extension	Organisation	[2] [3] [8] [9] [10] [11] [13] [14] [15] [16] [17] [18] [19] [20] [21] [25] [26] [27] [28] [29] [30] [31] [32] [33] [35] [37] [38] [40] [41] [42] [43]
	Time	[6] [10] [15] [16] [18] [19] [26] [28] [29] [32]
	Case	[10] [15] [18] [19] [26] [28] [32] [34] [39]

As shown in Table 1, there is a total of 9 article that discusses process model repair and 34 article that discusses process model extension. There are three perspectives that can be added to the process model. The organizational perspective focuses on information about resources hidden in the log, the time perspective is concerned with the timing and frequency of events, and the case perspective focuses on the properties of cases [46]. As shown in the table, the mostly studied perspective is the organizational perspective. The total of 29 out of 43 articles utilizes or consider the organizational perspective in their approach.

Table 2. Tools in Process Enhancement

Tools	**Articles**
ProM	[1] [2] [4] [6] [8] [9] [10] [11] [12] [13] [15] [16] [17] [18] [19] [20] [21] [22] [23] [26] [27] [34] [35] [36] [38] [40] [41]
Disco Fluxicon	[14] [15] [35] [40]
Other	SQL [15] [28] [43], Weka [10], MP Declare [29], R [30], Proview [25], BPMN Oracle [11], Apromore [5], DpiL Miner [27], SPADA [31]
Unavailable	[3] [7] [24] [37] [38]

Tools. Another attribute to be highlighted is the tools used in the studies. As shown in Table 2, most studies in the process enhancement use ProM as their tools, followed by Disco. Besides those, some of the studies use various tools that better suited to their needs such as SQL, Weka, DpiL Miner, etc. The rest six articles do not use any tool in their approach, or they do not explicitly mention which tools they used in their study.

3.3 Process Enhancement Application

After examining the datasets used in the relevant articles, we found that process mining has been applied in various domain. The most studied field within process enhancement is Medical, followed by Governmental and Financial. Another interesting finding that we acquired while examining the process enhancement application throughout different industries is the contribution of BPI Community, in this case, to the availability of event logs. 17 articles examined in this study uses the data provided by the BPI Community. For the list of the articles and its datasets information, please refer to Appendix.

4 Threat to Validity

The main threat to validity lies in the study search and study selection procedure. Even with the defined set of the search string, there are chances of relevant studies to be left out. We have mitigated this limitation but iteratively executing the search strings and by going through the reference list of selected studies. The possible biases in study selection procedure are mitigated by applying the study selection criteria. Also, in this version, we exclude the articles that the full-text version is not available via the University of Twente's access portal. We are waiting for the full-text version from corresponding authors, and we will incorporate them in the extended version of this study. Lastly, we also want to acknowledge the challenge in term of comprehensiveness of our study. The growth in the process enhancement and also process mining field means that the knowledge within them, including this study, needs constant refinement.

5 Conclusion

Literature within process enhancement topics populated by articles that proposed a framework and/or an approach with different possible improvement and advantages for process owners, however, it's important to note that there is no one-size-fits-all solution. Therefore, it's important to highlight what is the context in the offered approaches. We suggest two variables for choosing process enhancement approach: Objectives and Constraint. Objectives refer to the incentives or goals that the approach provide; this varies from identifying resource community, repairing the discovered business model, discovers bottleneck in the process and many more. The next variable to be considered for selecting an enhancement approach is the constraint. Constraint refers to the boundary that limits the applicability of the approach. The combination of these two contexts is an important determinant for the approach selection process. Both practitioners and process owners are suggested to take these two variables in selecting process enhancement approaches.

6 Appendix

Due to the limitation of the page of this article, the authors will provide the more detailed findings of the study in the author's ResearchGate page. For interested readers, please visit this link.

References

1. Adriansyah, A. and J. C. A. M. Buijs (2013). Mining process performance from event logs. Lecture Notes in Business Information Processing. 132 LNBIP: 217-218.
2. Anuwatvisit, S., et al. (2012). Bottleneck mining and Petri net simulation in education situations. International Conference on ICT and Knowledge Engineering.
3. Appice, A., et al. (2016). Discovering and tracking organizational structures in event logs. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 9607: 46-60
4. Appice, A. and D. Malerba (2016). A Co-Training Strategy for Multiple View Clustering in Process Mining. IEEE Transactions on Services Computing 9(6): 832-845.
5. Armas Cervantes, A., et al. (2017). Interactive and incremental business process model repair. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 10573 LNCS: 53-74.
6. Ballambettu, N. P., et al. (2017). Analyzing process variants to understand differences in key performance indices. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 10253 LNCS: 298-313.
7. Basile, F., et al. (2015). Model repair of Time Petri nets with temporal anomalies. IFAC-PapersOnLine 28(7): 85-90.
8. Bozkaya, M., et al. (2009). Process diagnostics: A method based on process mining. Proceedings International Conference on Information, Process, and Knowledge Management, eKNOW 2009.
9. Burattin, A., et al. (2013). Business models enhancement through discovery of roles. 2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM).
10. De Leoni, M., et al. (2016). A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs. Information Systems 56: 235-257.
11. Djedovic, A., et al. (2016). Model business process improvement by statistical analysis of the users' conduct in the process. 2016 International Multidisciplinary Conference on Computer and Energy Science, SpliTech 2016.
12. Fahland, D. and W. M. P. Van Der Aalst (2015). Model repair - Aligning process models to reality. Information Systems 47: 220-243.
13. Ferreira, D. R. and C. Alves (2012). Discovering user communities in large event logs. Lecture Notes in Business Information Processing. 99
14. Gupta, M. (2014). Nirikshan: process mining software repositories to identify inefficiencies, imperfections, and enhance existing process capabilities. Companion Proceedings of the 36th International Conference on Software Engineering. Hyderabad, India, ACM: 658-661.
15. Gupta, M. and A. Sureka (2014). Process Cube for Software Defect Resolution. 2014 21st Asia-Pacific Software Engineering Conference.
16. Hidayat, B. N. A., et al. (2016). Process model extension using heuristics miner: (Case study: Incident management of Volvo IT Belgium). 2016 International Conference on Computational Intelligence and Cybernetics.
17. Kim, K., et al. (2012). Discovery of information diffusion process in social networks. IEICE Transactions on Information and Systems E95-D(5): 1539-1542.
18. Mannhardt, F., et al. (2015). Extending process logs with events from supplementary sources. Lecture Notes in Business Information Processing. 202: 235-247.
19. Mannhardt, F., et al. (2016). Measuring the precision of multi-perspective process models. Lecture Notes in Business Information Processing. 256: 113-125.
20. Mans, R. S., et al. (2012). Mining processes in dentistry. IHI'12 - Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium
21. Mans, R. S., et al. (2008). Process mining in healthcare - A case study. HEALTHINF 2008 - 1st International Conference on Health Informatics, Proceedings.
22. Mitsyuk, A. A., et al. (2017). Process model repair by detecting unfitting fragments? CEUR Workshop Proceedings.

23. Polyvyanyy, A., et al. (2016). Impact-driven process model repair. ACM Transactions on Software Engineering and Methodology 25(4).
24. Rogge-Solti, A., et al. (2013). Repairing event logs using timed process models. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 8186 LNCS: 705-708.
25. Rouibah, K., et al. (2007). Combining workflow and PDM based on the workflow management coalition and STEP standards: The case of axalant. International Journal of Computer Integrated Manufacturing 20(8): 811-827.
26. Rozinat, A., et al. (2009). Discovering simulation models. Information Systems
27. Schönig, S., et al. (2015). Mining the organisational perspective in agile business processes. Lecture Notes in Business Information Processing. 214: 37-52.
28. Schönig, S., et al. (2016). Discovery of multi-perspective declarative process models. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 9936 LNCS: 87-103.
29. Sturm, C., et al. (2017). Distributed multi-perspective declare discovery. CEUR Workshop Proceedings.
30. Swennen, M., et al. (2016). Capturing resource behaviour from event logs. CEUR Workshop Proceedings.
31. Turi, A., et al. (2008). Distributed discovery of multi-level approximate process patterns. SEBD 2008 - Proceedings of the 16th Italian Symposium on Advanced Database Systems.
32. van der Aalst, W. M. P., et al. (2007). Business process mining: An industrial application.
33. Wang, J., et al. (2016). Mining organizational behaviors in collaborative logistics chain: An empirical study in a port. 2016 International Conference on Logistics, Informatics and Service Sciences (LISS).
34. Werner, M. (2013). Colored Petri nets for integrating the data perspective in process audits. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 8217 LNCS: 387-394.
35. Wongvigran, S. and W. Premchaiswadi (2015). Analysis of call-center operational data using role hierarchy miner. International Conference on ICT and Knowledge Engineering.
36. Zhang, X., et al. (2018). An Approach for Repairing Process Models Based on Logic Petri Nets. IEEE Access 6: 29926-29939.
37. Zhao, W., et al. (2016). An entropy-based clustering ensemble method to support resource allocation in business process management. Knowledge and Information Systems
38. Stroinski, A., et al. (2017). A distributed discovery of communicating resource systems models. IEEE Transactions on Services Computing: 1-1.
39. van Eck, M. L., et al. (2016). Discovering and exploring state-based models for multi-perspective processes. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).
40. Chamorro, M. and S. Maturana (2017). Method for Applying Process Mining to the Distribution of Non-alcoholic Beverages. Proceedings - International Conference of the Chilean Computer Science Society, SCCC.
41. Schönig, S., et al. (2016). A framework for efficiently mining the organisational perspective of business processes. CEUR Workshop Proceedings.
42. Ikeda, M., et al. (2014). Formal concept analysis for process enhancement based on a pair of perspectives. CEUR Workshop Proceedings.
43. Schonig, S., et al. (2016). Efficient and customisable declarative process mining with SQL. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 9694: 290-305.
44. Earl, M. J. (1994). The new and the old of business process redesign. The Journal of Strategic Information Systems,3(1), 5-22. doi:10.1016/0963-8687(94)90003-5

45. Zairi, M. (1997). Business process management: A boundaryless approach to modern competitiveness. Business Process Management Journal, 3(1), 64-80. doi:10.1108/14637159710161585
46. W., Van der Aalst. (2016). Process Mining Data Science in Action. Berlin: Springer Berlin. doi:10.1007/978-3-662-49851-4
47. Adriansyah, A. and J. C. A. M. Buijs (2013). Mining process performance from event logs. Lecture Notes in Business Information Processing. 132 LNBIP: 217-218.
48. Xue, Liang, & Boulton. (2008). Information Technology Governance in Information Technology Investment Decision Processes: The Impact of Investment Characteristics, External Environment, and Internal Context. MIS Quarterly,32(1), 67. doi:10.2307/25148829
49. Arksey, H., & O'malley, L. (2005). Scoping studies: Towards a methodological framework. International Journal of Social Research Methodology, 8(1), 19-32. doi:10.1080/1364557032000119616
50. Aalst, W. V., & Weijters, A. (2004). Process mining: A research agenda. Computers in Industry,53(3), 231-244. doi:10.1016/j.compind. 2003.10.001
51. Process Mining Group. (2014, August 14). ProM Tools. Retrieved from http://www.promtools.org/doku.php