Enhancing the Cost Dimension in Process Mining through its Application to the Mining Industry

Ignacio Velásquez[1]

[1]Department of Computer Science, Pontificia Universidad Católica de Chile

Abstract

Process Mining is a discipline for the discovery, monitoring, and improvement of processes by extracting knowledge from event logs. In the discovery stage, models are often used to analyze the process performance. The time dimension is often analyzed, but there has been little research on other dimensions. For the cost dimension, the focus has been on costs annotation in event logs, but not on strategies for their analysis. Opportunities can be observed by considering costs jointly with other dimensions, such as time and resources, allowing insights like cost-aware resource allocation or identifying trade-offs between costs and time. This research aims to enhance the analysis of costs through Process Mining, by devising methods for their analysis and operational support jointly with other dimensions. To achieve this, a partnership with companies from the mining industry is proposed, as they possess the records required for analyzing the cost dimension. Moreover, this industry is appealing as there is potential in analyzing its processes through Process Mining. Following a design science methodology, methods for jointly analyzing costs with other dimensions are devised, which will be validated by applying them to instances of processes from the mining industry. The devised methods are expected to facilitate costs-oriented decision-making in the mining industry, by addressing existing design problems in it.

Keywords

Process Mining, Process Discovery, Performance Analysis, Operational Support, Cost Dimension

1. Background and Motivation

Process Mining (PM) is a discipline that allows the discovery, monitoring, and improvement of processes by extracting knowledge from event logs containing records of actual process executions [1, 2]. When used for discovery, PM aims to construct a process model based on the behavior observed in the event log [3]. Two mainstream approaches are then followed for analyzing these models: verification, which is concerned with process correctness, and performance analysis, which focuses on analyzing key performance indicators based on the dimension of interest [2]. Performance analysis research has generally focused on the time dimension, and there has been scarce research on others, such as costs [4, 5].

The first work focusing on the cost dimension is [6], which proposes the creation and utilization of cost models from management accounting to annotate event logs with cost information. This work is later complemented with a strategy for cost prediction [7] and a transition system decorated with costs [8]. Related to this research line, cost-annotated event logs are also used for business process improvement [9].

BPM 2023 - Doctoral Consortium, September 11–15, 2023, Utrecht, The Netherlands
iavelasquez@uc.cl (I. Velásquez)

Later research maintains the idea of annotating costs in event logs through cost models. [4] proposes a framework for including cost and quality information in event logs from the manufacturing industry, and their visualization. A series of conference papers [10, 11, 12, 13, 14, 15] propose improvements over [6] by defining a costs-extended metamodel and the utilization of classification algorithms to improve the definition of cost models. More recent publications use PM to leverage information for the calculation of costs in costing strategies [16, 17].

In summary, research has focused on obtaining and annotating cost information in event logs [6], as well as visualizing them on process models, with a similar approach to the time dimension [4, 8, 10]. However, current research has not focused on designing strategies for the analysis of costs, other than designing simple visualizations that replicate those for the time dimension. There is a need for further researching the analysis of costs and not only their visualization. In this regard, research opportunities can be observed by considering the analysis of the cost dimension jointly with other PM elements, such as the time dimension and resource utilization. Insights that could not be observable by analyzing the above dimensions individually, can emerge when jointly analyzing them. For example, resources could be allocated while considering the costs associated with these allocations [18, 19, 20], or trade-offs between costs and time when executing processes could be identified. Moreover, by using the insights obtained from the analyses, operational support could be provided for improving the operational aspects of processes. Operational support refers to the utilization of PM techniques on running cases of a process [2]. Possible operational support tasks of interest are prediction (anticipating the outcomes of a case [21], e.g., its expected total cost) and recommendation (given the partial execution of a case, suggest what to do next [22], e.g., to minimize its total cost).

As costs are part of the sensitive data of companies [23], they are not readily available in public event logs, This implies the necessity to partner up with companies that possess this kind of data to research the topic. In this context, the appeal of partnering with companies from the mining industry is observed. This industry revolves around the extraction, processing, and transportation of minerals from mining sites to the marketplace [24]. Mining operations are capital-intensive and often undertaken in geographically remote and isolated areas [25].

A recent review indicates that, until 2019, PM has been seldom used within the mining industry [26]. The same research group has focused on generating and analyzing event logs based on sensor data from mining machinery [27, 28, 29, 30, 31, 32]. This shows the utility of PM for analyzing the functioning of machines in mining operations, but only two publications have researched the application of PM over actual processes of the industry. Concretely, [33] uses PM to improve an emergency rescue process in Chinese coal mines, and [34] simulates event logs for the LHD (Load, Haul, Dump) loaders maintenance process to determine bottlenecks. This shows a latent potential for applying PM to analyze the actual processes of this industry.

In addition to the above, existing research highlights that the mining industry has commonly made use of costing strategies. These strategies allow addressing the underlying premise that companies face limited information about true cost behavior, by defining cost models that provide answers to common costing questions [35]. Two costing strategies have been researched within the mining industry: Activity-Based Costing (ABC), which considers activities as the driver for determining costs)[36, 37], and Life Cycle Costing (LCC), which calculates costs over the life cycle of products and services [38]. ABC has been utilized to cost underground coal mining systems [39, 40], for continuous improvement in copper mines [41], and for cost modeling of the product mix in aggregate mining [42, 43]. LCC has been applied for the selection of equipment and technology [44], and for equipment performance optimization [45, 46]. These costing strategies facilitate decision-making by modeling expected costs. However, the utilization of PM would allow for analyzing the actual costs of executing processes. These two approaches are complementary, as costing strategies can be used for defining the cost models to annotate event logs [6], whereas PM would allow evaluating the predictions from the costing strategies by comparing them to the actual process executions.

As Chile is the leading copper producer, the mining industry has been the most important driver of its economy [47]. This leverages the possibility of partnering up with several mining companies in the country.

Based on the above, the objective of this research is to enhance the state-of-the-art regarding the analysis of costs through PM, by devising methods that allow the analysis of this dimension jointly with other PM elements, such as time and resources, and to apply these methods in event logs of real processes from the mining industry, to address their needs.

The remainder of this work is structured as follows: Section 2 elaborates on the addressed research questions and the followed research methodology. Work currently in progress is presented in Section 3, coupled with the expected contribution of this research.

2. Research Methodology

This research follows the Design Science (DS) methodology [48]. DS is a paradigm for conducting and communicating applied research whose goal is to produce prescriptive knowledge in a discipline and to share empirical insights from the application of these prescriptions [49]. This is done by iterating over two activities: designing an artifact that improves something for stakeholders and empirically investigating the performance of the artifact in a context [48].

As an applied methodology, DS works with two kinds of research problems [48]: (i) Design Problems (DP), which correspond to the need of designing artifacts for contributing to the achievement of stakeholder goals; and (ii) knowledge questions (KQ), which refer to the obtainable knowledge from researching the application context of the artifacts. In this research, the partners from the mining industry are the stakeholders, whereas the artifacts are the PM methods that will be devised throughout the research. These methods will be detailed steps on how to undertake the analysis of the cost dimension jointly with other dimensions, and how to perform operational support considering these dimensions. To facilitate these analyses, it is expected that it will be necessary to generate code compatible with PM libraries, like pm4py or bupaR, and/or to outline the steps for their realization in existing PM tools, like Celonis. The development of simple independent tools that provide ad-hoc solutions for the proposed methods is also considered. All the above correspond to artifacts that will be made publicly available. Nevertheless, the conceptual definition of the methods will be platform-agnostic.

The KQ addressed in this work seek to identify the benefits of analyzing the cost dimension jointly with other dimensions and ascertaining whether this joint analysis improves decision-making, and how it can be adapted for operational support. Similarly, DP contemplate improving cost analysis and operational support of mining processes, by extracting historical data of their executions and defining PM methods for analyzing the cost dimension jointly with other dimensions, to facilitate the identification of improvement opportunities and decision-making.

DS research is performed through a design cycle [48], which considers three activities: (i) researching the problem at hand, (ii) designing one or more artifacts for treating the problem, and (iii) validating the capability of these artifacts.

Following the above, Figure 1 shows the outline of this research. On one hand, conventional PM methods will be applied to processes in the mining industry. These methods are process discovery and performance analysis of individual perspectives. This will allow getting to know the specifics of the industry, the limitations of current methods, and opportunities for jointly analyzing costs with other dimensions. On the other hand, the realization of a Systematic Literature Review (SLR) focusing on ascertaining the state-of-the-art regarding the cost dimension is considered. This SLR provides knowledge regarding existing methods used for analyzing costs, as well as the identification of research gaps that should be addressed. Subsequently, using the acquired practical experience and academic knowledge, several PM methods for the joint analysis and operational support of costs with other dimensions will be devised. The feasibility of these methods will be verified through their application on simulated data, and their utility will then be validated by applying them to historical executions of processes from the mining industry.

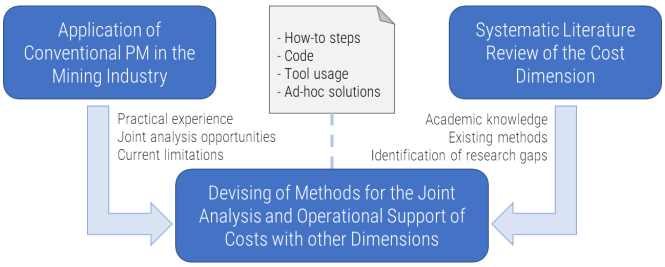

Figure 1: Outline of the research.

The limitations of this research can be described based on the validity of its results. First, as the devised methods will be validated in a specific industry, this affects their external validity in other industries. Second, the internal validity will be dependent on the data provided by the companies from the mining industry and the processes that are of interest to the stakeholders. It will also be necessary to assert that historical data is representative of current process executions. Finally, it must be mentioned that the devised methods will be validated in an academic context. Their evaluation for implementation in real-world settings will be necessary.

3. Work in Progress and Expected Contribution

Work in progress can be grouped by the three stages shown in the outline of Figure 1. Regarding the application of conventional PM, this stage has seen delay due to the need of signing non-disclosure agreements with the mining companies. So far, this has been the major problem threatening the realization of this research. However, conversations with the mining companies are ongoing, and there has been contact with interested companies from other industries, which will help to address the external validity of the methods.

As the acquisition of practical experience has been on hold, the research has focused on obtaining the necessary academic knowledge. This has been achieved through a SLR that researched the existence of methods for analyzing the cost dimension and dimensions that have been considered jointly with it, within the context of PM and other disciplines where a process perspective is relevant, such as Business Process Management and Operations Research.

Using the academic knowledge from the SLR, the following methods, which address some of the identified research gaps, have been devised so far:

- Analysis of process variants based on average costs and their standard deviation. Demo available at https://bit.ly/variant_cost_stdev.
- Analysis of the Devil's Quadrangle dimensions (costs, time, quality, flexibility) of process instances through the filtering capabilities of PM. The Devil's Quadrangle is a framework that describes the inherent trade-offs of performance dimensions [50]. Demo available at https://bit.ly/dq_dashboard.
- Visualization of all process variants through a single Directed Rooted Tree, whose leaves are the end states of variants. The tree is decorated with cost information for quickly comparing the cost of different variants. Tool available at https://bit.ly/drt-variant-costs.

The feasibility of the methods was verified by applying them over simulated executions of the blasting process of a Chilean copper mine [51]. Blasting is one of the main methods of the mining industry to fragment hard rock minerals [52]. Two versions of the simulated log are available at https://bit.ly/blasting_with_rework_log and https://bit.ly/blasting_with_incomplete_cases_log. Subsequently, the utility of the methods must be validated by applying them to real data.

Looking forward, the realization of the SLR and the currently devised methods have provided some insights regarding the roadmap for the remainder of the research. Initially, the envisioned planning of the research consisted of a linear approach where every stage focused on jointly analyzing costs and another specific dimension (i.e., time or resources), and a final stage where all dimensions will be combined for process analysis and operational support. However, based on the state-of-the-art and the characteristics of the currently devised methods, the need for a more flexible approach, which addresses current PM limitations and research gaps, was observed. Thus, based on stakeholder needs, methods will be devised for the joint analysis and/or operational support of costs and any number of other dimensions.

Following the above roadmap, it is expected that the devised methods will help, from an academic perspective, towards expanding the current state-of-the-art in PM, by allowing an enriched analysis of the cost dimension through its combination with other dimensions. Moreover, the methods will also be of utility for stakeholders of the mining industry, as their application will allow supporting cost-oriented decision-making in distinct mining processes.

In addition to the above, it is also expected that, through further research, it will be possible to validate the utility of the devised methods within contexts other than the mining industry.

Acknowledgments

This work was funded by the Agencia Nacional de Investigación y Desarrollo de Chile. Grant number ANID-Subdirección de Capital Humano/Doctorado Nacional/2021-21210022.

References

[1] H. M. Marin-Castro, E. Tello-Leal, Event log preprocessing for process mining: a review, Applied Sciences 11 (2021) 10556.

[2] W. Van der Aalst, Process mining: data science in action, Springer, 2016.

[3] M. Faizan, M. F. Zuhairi, S. binti Ismail, R. Ahmed, Challenges and use cases of process discovery in process mining, International Journal of Advanced Trends in Computer Science and Engineering 9 (2021).

[4] T. T. B. Hong, Process Mining-driven Performance Analysis in Manufacturing Process: Cost and Quality Perspective, Thesis, Ulsan National Institute of Science and Technology, 2016.

[5] T. T. B. Hong, M. Song, Analysis and prediction cost of manufacturing process based on process mining, in: 2016 international conference on industrial engineering, management science and application (ICIMSA), IEEE, 2016, pp. 1–5.

[6] W. Nauta, Towards cost-awareness in process mining, Thesis, Eindhoven University of Technology, 2011.

[7] M. Wynn, W. Z. Low, W. Nauta, A framework for cost-aware process management: generation of accurate and timely management accounting cost reports, in: Conceptual Modelling 2013: Proceedings of the Ninth Asia-Pacific Conference on Conceptual Modelling, Australian Computer Society, 2013, pp. 79–88.

[8] M. Wynn, W. Z. Low, A. ter Hofstede, W. Nauta, A framework for cost-aware process management: cost reporting and cost prediction, Journal of Universal Computer Science 20 (2014) 406–430.

[9] W. Z. Low, Towards cost model-driven log-based business process improvement, Thesis, Queensland University of Technology, 2016.

[10] D. Thabet, S. A. Ghannouchi, H. H. Ben Ghezala, A general solution for business process model extension with cost perspective based on process mining, International Conference on Software Engineering Advances 2016 (2016) 251.

[11] D. Thabet, S. A. Ghannouchi, H. H. Ben Ghézala, Towards business process model extension with cost perspective based on process mining, in: Proceedings of the 16th International Conference on Enterprise Information Systems-Volume 3, 2014, pp. 335–342.

[12] D. Thabet, S. A. Ghannouchi, H. H. Ben Ghézala, Petri net model cost extension based on process mining, in: Proceedings of the 17th International Conference on Enterprise Information Systems-Volume 3, 2015, pp. 268–275.

[13] D. Thabet, S. A. Ghannouchi, H. H. Ben Ghezala, A process mining-based solution for business process model extension with cost perspective context-based cost data analysis and case study, in: IFIP International Conference on Computer Information Systems and Industrial Management, Springer, 2018, pp. 434–446.

[14] D. Thabet, N. Ganouni, S. A. Ghannouchi, H. H. Ben Ghezala, Towards context-aware business process cost data analysis including the control-flow perspective, in: International Conference on Intelligent Systems Design and Applications, Springer, 2019, pp. 193–204.

[15] D. Thabet, S. A. Ghannouchi, H. H. Ben Ghezala, Towards business cost mining, in: International Conference On Systems Engineering, Springer, 2021, pp. 127–137.

[16] M. Halaška, R. Šperka, TDABC and Estimation of Time Drivers Using Process Mining,

Springer, 2021, pp. 489–499.
[17] M. Riesener, C. Dölle, A. Menges, G. Schuh, Time-driven activity-based costing using process data mining, in: 2021 IEEE Technology & Engineering Management Conference-Europe (TEMSCON-EUR), IEEE, 2021, pp. 1–6.
[18] M. Arias, J. Munoz-Gama, M. Sepúlveda, Towards a taxonomy of human resource allocation criteria, in: Int. Conference on Business Process Management, Springer, 2017, pp. 475–483.
[19] M. Arias, R. Saavedra, M. R. Marques, J. Munoz-Gama, M. Sepúlveda, Human resource allocation in business process management and process mining: A systematic mapping study, Management Decision (2018).
[20] G. Havur, C. Cabanillas, A. Polleres, Benchmarking answer set programming systems for resource allocation in business processes, Expert Systems with Applications (2022).
[21] M. Polato, A. Sperduti, A. Burattin, M. d. Leoni, Time and activity sequence prediction of business process instances, Computing 100 (2018) 1005–1031.
[22] R. Conforti, M. de Leoni, M. La Rosa, W. M. van der Aalst, A. H. ter Hofstede, A recommendation system for predicting risks across multiple business process instances, Decision Support Systems 69 (2015) 1–19.
[23] P. I. Okochi, S. A. Okolie, J. N. Odii, An improved data leakage detection system in a cloud computing environment, World Journal of Advanced Research and Reviews 11 (2021) 321–328.
[24] K. Muduli, K. Govindan, A. Barve, Y. Geng, Barriers to green supply chain management in indian mining industries: a graph theoretic approach, Journal of Cleaner Production 47 (2013) 335–344.
[25] H. Hall, Exploring innovation in northern canada with insights from the mining innovation system in greater sudbury, ontario, Northern Review (2017) 33–56–33–56.
[26] E. Brzychczy, An overview of data mining and process mining applications in underground mining, Inżynieria Mineralna 21 (2019).
[27] E. Brzychczy, P. Gackowiec, M. Liebetrau, Data analytic approaches for mining process improvement—machinery utilization use case, Resources 9 (2020) 17.
[28] E. Brzychczy, A. Trzcionkowska, New possibilities for process analysis in an underground mine, Zeszyty Naukowe. Organizacja i Zarządzanie/Politechnika Śląska (2017).
[29] E. Brzychczy, A. Trzcionkowska, Creation of an event log from a low-level machinery monitoring system for process mining purposes, in: International Conference on Intelligent Data Engineering and Automated Learning, Springer, 2018, pp. 54–63.
[30] E. Brzychczy, A. Trzcionkowska, Process-oriented approach for analysis of sensor data from longwall monitoring system, in: International Conference on Intelligent Systems in Production Engineering and Maintenance, Springer, 2018, pp. 611–621.
[31] M. Szpyrka, E. Brzychczy, A. Napieraj, J. Korski, G. J. Nalepa, Conformance checking of a longwall shearer operation based on low-level events, Energies 13 (2020) 6630.
[32] A. Trzcionkowska, E. Brzychczy, Practical aspects of event logs creation for industrial process modelling, Multidisciplinary Aspects of Production Engineering 1 (2018).
[33] Z. He, Q. Wu, L. Wen, G. Fu, A process mining approach to improve emergency rescue processes of fatal gas explosion accidents in chinese coal mines, Safety science 111 (2019) 154–166.
[34] N. F. A. Velásquez Aedo, Aplicación de minería de procesos y optimización del proceso de

mantención de equipos LHD, Thesis, Pontificia Universidad Católica de Chile, 2021.
[35] E. Labro, Costing systems, Foundations and Trends in Accounting 13 (2019) 267–404.
[36] P. Quesado, R. Silva, Activity-based costing (abc) and its implication for open innovation, Journal of Open Innovation: Technology, Market, and Complexity 7 (2021) 41.
[37] Z. Cidav, D. Mandell, J. Pyne, R. Beidas, G. Curran, S. Marcus, A pragmatic method for costing implementation strategies using time-driven activity-based costing, Implementation Science 15 (2020) 1–15.
[38] M. L. Kambanou, Life cycle costing: understanding how it is practised and its relationship to life cycle management—a case study, Sustainability 12 (2020) 3252.
[39] G. Lind, Activity based costing: challenging the way we cost underground coal mining systems, Journal of the Southern African Institute of Mining and Metallurgy 101 (2001) 77–82.
[40] S. Junsong, Exploring on application thought of activity-based costing method in coal mining enterprises, Journal of North China Institute of Science and Technology (2012) 93–96.
[41] J. A. Botín, M. A. Vergara, A cost management model for economic sustainability and continuos improvement of mining operations, Resources Policy 46 (2015) 212–218.
[42] D. T. Cremonese, Mineral project design using ABC methodology and dynamic simulation, Thesis, Universidade de São Paulo, 2019.
[43] D. T. Cremonese, G. D. Tomi, M. R. Neves, Cost modelling of the product mix from mining operations using the activity-based costing approach, Rem: Revista Escola de Minas 69 (2016) 97–103.
[44] T. Markeset, U. Kumar, Application of LCC techniques in selection of mining equipment and technology, Routledge, 2018, pp. 635–640.
[45] B. Balaba, M. Y. Ibrahim, Industrial informatics applications: Optimising life cycle costs of an underground mining longwall conveyor, in: 2011 IEEE International Symposium on Industrial Electronics, IEEE, 2011, pp. 1218–1223.
[46] A. Bozorgebrahimi, R. Hall, M. Morin, Equipment size effects on open pit mining performance, Int. Journal of Surface Mining, Reclamation and Environment 19 (2005) 41–56.
[47] S. Santoro, H. Estay, A. H. Avci, L. Pugliese, R. Ruby-Figueroa, A. Garcia, M. Aquino, S. Nasirov, S. Straface, E. Curcio, Membrane technology for a sustainable copper mining industry: The chilean paradigm, Cleaner Engineering and Technology 2 (2021).
[48] R. J. Wieringa, Design science methodology for information systems and software engineering, Springer, 2014.
[49] E. Engström, M.-A. Storey, P. Runeson, M. Höst, M. T. Baldassarre, How software engineering research aligns with design science: a review, Empirical Software Engineering 25 (2020) 2630–2660.
[50] N. Brand, H. Van der Kolk, Workflow analysis and design, 1995.
[51] J. E. Acuña Maldonado, Modelo de Gestión de Operaciones de la mina Rajo Abierto de la División Andina de CODELCO, Thesis, Pontificia Universidad Católica de Valparaíso, 2008.
[52] J. M. Akande, A. E. Aladejare, A. I. Lawal, Evaluation of the environmental impacts of blasting in okorusu fluorspar mine, namibia, International Journal of Engineering and Technology 4 (2014) 101–108.