Capturing Resource Behaviour From Event Logs

Marijke Swennen[1], Niels Martin[1], Gert Janssenswillen[1,2], Mieke Jans[1], Benoît Depaire[1], An Caris[1], and Koen Vanhoof[1]

Hasselt University, Agoralaan Bldg D, 3590 Diepenbeek, Belgium
Research Foundation Flanders (FWO), Egmontstraat 5, 1060 Brussels, Belgium
{marijke.swennen, niels.martin, gert.janssenswillen, mieke.jans, benoit.depaire, an.caris, koen.vanhoof}@uhasselt.be

Abstract. Process mining mainly focuses on the retrieval of process models from event logs. As these discovery algorithms make assumptions, performance analyses based on these models can present a biased view. In literature, algorithm-agnostic process metrics have been introduced. Given the critical importance of resources in the light of continuous process improvement, this paper extends the process metrics framework towards the resource perspective. New metrics are added and existing metrics are adapted. Using the extended framework, organisations can retrieve useful insights in resource behaviour.

Keywords: Process mining, Operational excellence, Resource behaviour, Log-based process metrics

1 Introduction

Process mining aims to retrieve valuable insights in business processes from event logs. Discovery, one of three types of process mining, aims to discover process models from event data. Related to control-flow, a plethora of algorithms have been developed [1], which all make particular assumptions and have a representational bias. Carrying out process performance analyses using such a process model entails the risk of presenting a biased view on the underlying process.

In this respect, the need for unbiased, algorithm-agnostic information retrieved from event logs is advocated in [16]. Hence, a set of metrics, related to time and structuredness of the process, is identified. These unbiased insights in process behaviour can support performance measurement [12] and continuous process improvement [11] which is related to methodologies such as lean management and Six Sigma.

A classification of resources in the domain of project management can be found in [7]. However, this paper focuses on process participants, software systems or equipment, as resources are defined in the field of BPM [4]. The importance of resources is not yet recognised in [16]. Nevertheless, resources are a source of process variability and their behaviour is essential in the light of continuous process improvement [2]. Consequently, this dimension should be taken into account to convey a more comprehensive picture on process behaviour to organisations. This is consistent with the research recommendation in [14] as it targets the resource perspective in process mining.

Given the need to include the resource perspective, this paper extends the work in [16] to include resource-related process insights. To this end, new metrics are proposed and existing metrics are redefined or complemented with additional analysis levels.

Within the context of quantifying the resource perspective using event logs, metrics that mainly focus on the relationship between resources are proposed by [15]. While the latter specifies metrics with the purpose of mining organisational models, [5] and [13] focus on defining resource behaviour measures. The current paper complements this as well as the recently introduced resource availability metrics [9]. Besides the general contribution of providing algorithm-agnostic resource insights to organisations, the extension of [16] can also support organisations in performing knowledge management, for instance when creating a knowledge map [3], or project management with applications such as resource levelling or resource allocation [8].

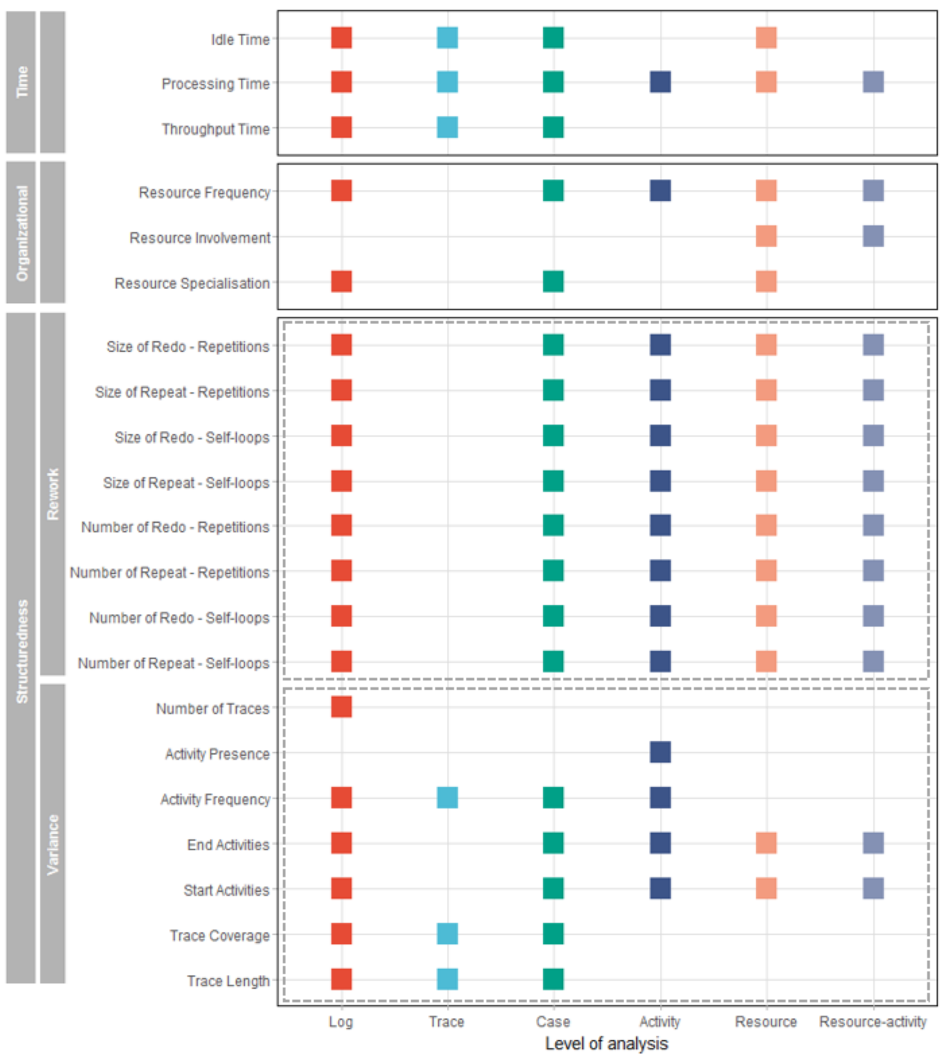

Fig. 1. Extended overview of log-based process metrics [16].

2 Resource metrics

Next to three new metrics concerning the concept of resource behaviour, adaptations to existing metrics in [16] are presented. Levels of analysis currently presented are the log level, trace level, and activity level. New levels of analysis related to resources are the resource level, representing characteristics of the resources executing the activities, and the resource-activity level in which resource-activity combinations are considered as introduced in [10]. All metrics have been implemented as functions in the R-package edeaR, which is available on the Comprehensive R Archive Network [6]. An overview of all metrics presented in [16] extended with the new work presented in the paper at hand is provided in Figure 1.

2.1 Running example

The metrics presented in this paper will be illustrated by applying them to the running example introduced in Figure 2. The latter visualises an event log for five patients undergoing five different medical activities executed by six distinct staff members of a hospital. The duration of activities is presented by their length.

2.2 New resource metrics

Resource frequency. The frequency of resources executing activities in a business process can be insightful, e.g. during company restructuring. In Figure 2, a resource is associated to on average 5.167 activity executions, with a minimum of 1 (resource R6) and a maximum of 8 (resource R4), implying that R4 is probably more active and fundamental for this process. At trace level, this metric is less informative because, even though the sequence of activities is the same, the resources executing them can differ. Other analysis levels are activities, resources and resource-activity combinations. For the latter, both the resource perspective and activity perspective can be useful for the relative number. R4 executes a surgery four times, representing 50 % of the total number of executions by R4 and 57.14 % of the total number of surgeries.

Resource involvement. The involvement of resources in cases can be of interest to, e.g., decide how "indispensable" they are. R5, for instance, only helps three patients, while R3 is involved in the process of 100 % of the patients. On the resource-activity level, the involvement of specific resource-activity combinations in cases is provided.

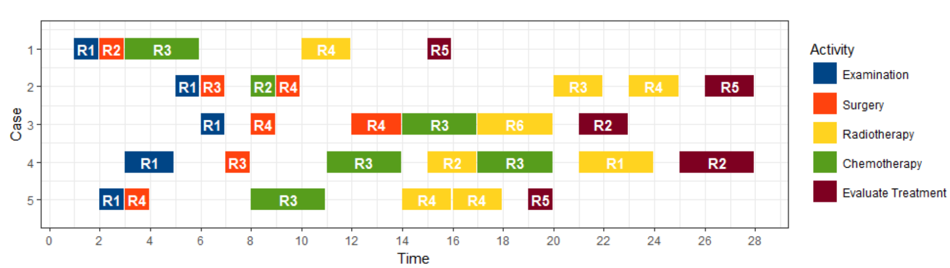

Fig. 2. Visualisation of running example event log.

Resource specialisation. Finally, the specialisation level of resources in a company demonstrates which resources are performing certain activities more than others. This information can be used to tackle challenges such as team selection or brain drain, as presented in [3]. Next to the log level, the resource level shows the absolute and relative number of distinct activities that each resource executes. R1 only executes two types of activities, examinations and radiotherapy, while he executes these activities six times in total. At the level of distinct activities, we find that examinations are always executed by the same resource, while radiotherapy is executed by everyone.

2.3 Adaptations to existing metrics presented in [16]

Looking at the **processing time** per case on the resource level provides a company an overview of the amount of time each resource spends on a case and which resources spend more time on cases than others. R3 spends on average around five times more time per patient than R5. However, comparing different resources or activities can be more insightful at the resource-activity level.

Extra levels of analysis can also be added to the structuredness metrics regarding **start and end activities** in a process. Which resources execute the first or last activity per case can be of interest for a company. Probably this person is the contact person for the patient or customer or is responsible for all communication.

Moreover, resource information was not taken into account in the metrics regarding **rework**. The metrics presenting self-loops and repetitions should take into account if the activity is redone by the same resource or by another one. Therefore, these metrics are rewritten according to four concepts, presented in Figure 3. A *repeat self-loop* is found in case 5, where radiotherapy is executed twice immediately after each other by R4. Case 2 holds a *redo self-loop*, where R4 executes radiotherapy after R3 executed it. Case 2 also holds a *redo repetition*, i.e. surgery is executed by R3 and R4 (but not immediately after each other). Finally, a *repeat repetition* is present in case 4 where R3 executes chemotherapy twice with an execution of radiotherapy in between.

Immediately following

Same Resource	Different resource
Repeat self-loop	Redo self-loop
Repeat repetition	Redo repetition

Not immediately following

Fig. 3. Dimensions of rework metrics.

3 Conclusions and Future Work

This paper presents an extension of [16] towards the resource perspective. New metrics are presented and existing metrics are adapted. This way, it anticipates on the importance of resources in the light of continuous process improvement and the need to gain algorithm-agnostic insights in their behaviour.

Future work involves applying the metrics to a real-life event log and formulating managerial recommendations based on these insights. Moreover, studying the evolution of metric values over time and examining the combined evolution of metrics can provide valuable knowledge. For instance, a decrease in processing time at resource-activity level combined with an increase in rework over the same time period might not be desirable for organisations.

References

1. van der Aalst, W.: Process mining: data science in action. Springer, Heidelberg (2016)
2. van Assen, M.F.: Position paper - Operational Excellence for Services (2011)
3. Creemers, M., Jans, M.: Social mining as a knowledge management solution. CEUR Workshop Proceedings 1612, 57–64 (2016)
4. Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: Fundamentals of business process management. Springer, Heidelberg (2013)
5. Huang, Z., Xudong, L., Huilong, D.: Resource behavior measure and application in business process management. Expert Systems with Applications 39(7), 6458–6468 (2012)
6. Janssenswillen, G., Swennen, M., Depaire, B., Jans, M.: Enabling event-data analysis in r: Demonstration. CEUR Workshop Proceedings 1527, 189–198 (2015)
7. Jugdev, K., Mathur, G.: Classifying project management resources by complexity and leverage. International Journal of Managing Projects in Business 5(1), 105–124 (2012)
8. Kerzner, H.: Project Management: A Systems Approach to Planning, Scheduling, and Controlling. John Wiley & Sons (Feb 2013)
9. Martin, N., Bax, F., Depaire, B., Caris, A.: Retrieving resource availability insights from event logs. Proceedings of the 2016 IEEE International Conference on Enterprise Distributed Object Computing pp. 69–78 (2016)
10. Martin, N., Swennen, M., Depaire, B., Jans, M., Caris, A., Vanhoof, K.: Batch processing: definition and event log identification. CEUR Workshop Proceedings 1527, 137–140 (2015)
11. Melnyk, S.A., Stewart, D.M., Swink, M.: Metrics and performance measurement in operations management: dealing with the metrics maze. Journal of Operations Management 22, 209–217 (2004)
12. Neely, A., Gregory, M., Platts, K.: Performance measurement system design: a literature review and research agenda. International Journal of Operations & Production Management 25(12), 1228–1263 (2005)
13. Pika, A., Wynn, M.T., Fidge, C.J., ter Hofstede, A.H.M., Leyer, M., van der Aalst, W.: An extensible framework for analysing resource behaviour using event logs. Lecture Notes in Computer Science 8484, 564–579 (2014)
14. Recker, J., Mendling, J.: The state of the art of business process management research as published in the BPM conference. Business & Information Systems Engineering 58(1), 55–72 (2016)
15. Song, M., van der Aalst, W.: Towards comprehensive support for organizational mining. Decision Support Systems 46(1), 300–317 (2008)
16. Swennen, M., Janssenswillen, G., Jans, M., Depaire, B., Vanhoof, K.: Capturing process behavior with log-based process metrics. Tech. rep. (2015), http://hdl.handle.net/1942/20239