# MEASURING HUMAN RESOURCES PERFORMANCE USING PROCESS MINING

Robert Kelemen[1]

**DOI: 10.24989/ocg.v335.41**

***Abstract***
*Performance of human resources is an important issue for managers in profit sector and public sector as well. Performance evaluation can be measured by numerous factors. Traditional approaches are often subjective, and based on descriptive indicators which are hard to measure. Since modern organizations use the information systems to record information about business processes and activities of human resources, it is possible to use this information utilizing process mining techniques to acquire objective information about employees' performance.*

*This paper reviews the literature and investigates the state of the art trends in human resource performance measuring using process mining, the indicators which are measured, methods, frameworks, main directions of development, and suggested future works.*

## 1. Introduction

In day-to-day work, employees are involved in executing business processes and play a key role in creating value for the organization. Organizational managers need to evaluate the effectiveness of human resources, identify best practices, identify opportunities for improvement and make decisions [10]. They often have to make decisions based on subjective estimates, which can lead to wrong conclusions about success, ineffective development strategies, and missed opportunities for improvement. The basic idea of this research has been to explore the state-of-the-art trends in the human resource performance measuring using process mining, to find out what the methods, frameworks and main directions of development are as well as suggest further research.

Process mining is a relatively new research discipline that comes between computational intelligence and data mining on the one hand, and process modelling and analysis on the other hand [1]. The idea of process mining is to discover, monitor and improve real processes by extracting knowledge from event logs readily available in today's (information) systems. Process mining includes (automated) process discovery, conformance checking, social network/organizational mining, automated construction of simulation models, model extension, model repair, case prediction, and history-based recommendations [1]. Process mining provides an important bridge between data mining and business process modelling and analysis [1]. A starting point for process mining is an event log. All process mining techniques assume that it is possible to sequentially record events so that each event refers to an activity and is related to a particular case [2]. Event logs may store additional information such as the resource executing or initiating an activity, the timestamp of an event, or data elements recorded with an event [2].

[1] Varazdin County, Franjevacki trg 7, 42000 Varazdin, Croatia, robert.kelemen@vzz.hr

Resource behavior refers to the actions or reactions of a resource in relation to business process execution [4]. Resource behavior measure is highly relevant for the performance of processes, suggesting that comprehensive support for this is needed. Several important issues are related to resource behavior measure [4]. The problem of resource behavior measure can be approached in two ways [4]. According to one approach, it may use subjective estimates by observing resource behavior in business process execution [4]. According to the other approach, the use of process mining to measure resource behavior from event logs is possible [4]. Both methods may perform equally well [4]. Process outcomes depend on human resource behavior, which is considered to be the "most important element that can affect project success" [13].

CONTRIBUTION

This paper provides a literature review on process mining techniques adopted in order to capture the performance of human resources. To achieve this goal, the following research questions are formulated:

A. What aspects of human resources behaviour can be measured using process mining?

B. What are the main opportunities for future research and identified research areas?

The rest of the paper is structured as follows: the research methodology is described in Section 2, the main explanation of research results is presented in Section 3 and finally, concluding remarks in Section 4.

## 2. Research methodology

To gain an insight on process mining implemented in the security domain a literature review has been conducted according to the general systematic review steps proposed by Kitchenham [5]. The most common reasons for undertaking a systematic review are [5]: 1) To summarize the existing evidence concerning treatment or technology e.g. to summarize the empirical evidence of the benefits and limitations of a specific agile method; 2) to identify any gaps in current research in order to suggest areas for further investigation; 3) to provide a framework/background in order to appropriately position new research activities.

### 2.1. Literature search

The first step in literature research was to define the keywords which will be used in database search. The first keyword was "process mining", the second "performance", and the third "human resource" or "HR" as the general term which includes all the essential components. Consequently, the query is as follows: "process mining" AND performance and "human resource" or HR. For the purpose of this survey, only Google Scholar was used within the time span from 2010 to 2019.

The inclusion and exclusion criteria for article selection are presented in Table 1. Based on these Criteria all the articles have been selected. The whole research has been conducted in three phases. The first phase objective has been to rise query in Google Scholar database with defined keywords and titles. The articles matching the criteria have been selected. The second phase has taken into consideration the abstracts and duplicates. If the paper has matched criteria, the third phase has been performed – the full text analysis. The papers that have not matched the criteria have also been excluded from the research.

| | PHASE 1 | PHASE 2 | PHASE 3 |
|---|---|---|---|
| **INCLUSION CRITERIA** | Title indicates that the paper is about the process mining, performance measuring and human resource<br><br>Keywords indicate that the paper is about the process mining performance measuring and human resource | Abstract of articles indicates that the paper is about the process mining, performance measuring and human resource | Topics on the process mining usage in performance measuring of human resources,<br>examples of implementation<br><br>The main indicators which are measured, methods, frameworks, main directions of development, and suggested future research |
| **EXCLUSION CRITERIA** | Title indicates that the paper is about another topic, it can include process mining but exclude human resource or performance measuring<br>Keywords indicate that the paper is about another topic<br>Book Chapter<br>PhD or Master Thesis<br>Patent | Abstract of articles indicates that the paper is not related to the topic<br><br>No abstract available<br><br>Duplicates excluded | There are no process mining techniques related to the performance measuring of human resources<br><br>There are no performance measuring of human resources |

**Table 1. Paper inclusion and exclusion criteria**

An initial search was performed during December 2018 and the initial number of 254 papers were discovered.

The first phase was performed by analysing titles and keywords, with exclusion criteria presented in Table 1 and 151 articles were selected.

The second selection was based on the analysis of the abstracts and removing duplicates. After the second phase 22 articles remained.

The third selection was performed by full text analysis and the final number of papers ready for analysis was 9.

At first glance, the final number of articles is small, but when research is put into the context that the field of research is just beginning to develop, then it is understandable that the number of articles is rather small.

## 3. Research Results

The analysis of 9 identified articles has revealed that from 2010 to 2018 only 6 articles were published in conference proceedings and 3 in scientific journals. Identified articles are the following:

- Analyzing Resource Behavior Using Process Mining [9],
- Resource behavior measure and application in business process management [4],
- An Extensible Framework for Analysing Resource Behaviour Using Event Logs [11],

- Implementing Closeness Centrality Measurements on Workflow-Supported Enterprise Social Networks [3],
- Retrieving Resource Availability Insights from Event Logs [7],
- Context-Aware Recommendation of Task Allocations in Service Systems [12],
- Mining resource profiles from event logs [10],
- A Comparative Evaluation of Log-Based Process Performance Analysis Techniques [8],
- Dynamic human resource selection for business process exceptions [6].

For the purpose of the analysis the main findings of the articles are presented in chronological order.

Nakatumba and van der Aalst put focus in their paper on the phenomenon of workload-dependent processing speeds and the analysis of that phenomenon [9]. They have presented an approach to quantify the relationship between workload and processing speed. The workload of a resource or a group of resources can be defined as either: (a) the number of work items waiting at the start of execution of an activity, i.e., the amount of work that has been scheduled for a given user or (b) the number of activities that have been executed over a particular period [9]. Authors have focused on the second option, to define "how busy" the resource has been [9]. This approach is based on regression analysis and is implemented as a new plug-in in ProM. The authors have identified aspects of resource behavior that are not captured in today's simulation tools. They point out that people do not work at constant speeds and their processing speeds are usually influenced by their workload [9]. The future research will aim at more powerful analysis techniques and a tight coupling between simulation and operational decision making. They want to make simulation more realistic by adequately modeling resources based on empirical data. Besides workload-dependent process times, it is necessary to take into account the facts that people are involved in multiple processes, they are available only part-time or they work in batches [9].

Huang et al. [4] have pointed out that the problem of resource behavior measure can be approached in two ways. According to one approach, it may use subjective estimates by observing resource behavior in business process execution. According to the other approach, it is possible to use of process mining to measure resource behavior from event. Both methods may perform equally well. Their approach of measuring resource behavior is based on four important perspectives, i.e., preference, availability, competence and cooperation, based on process mining the event logs as a basis on which they measure and apply resource behaviors in BPM context, i.e., what the typical behaviors of a resource are and how they are applied in business process management [4]. The authors address four applications of resource behavior measure: (1) resource similarity measure based on resource reference and competence measure, (2) resource network analysis based on resource cooperation measure, (3) resource dependent constraint detection based on resource cooperation measure and (4) work distribution recommendation based on resource preference, competence and availability measure [4]. Opportunities for future work are: the further research of discovery and extension issues, e.g., how to distribute work to appropriate resources with the performance consideration in business process execution, and how to handle the situation in which a group of resources work for an activity simultaneously. In these cases, resource behaviors should

be defined more precisely and their applications may remarkably improve the business process execution performance [4].

Pika et al [11] in 2014 proposed an extensible framework for extracting knowledge from event logs about the behavior of a human resource and for analyzing the dynamics of this behavior over time. The framework is fully automated and implements a predefined set of behavioral indicators for human resources research aims to make use of information recorded in event logs to extract knowledge about the behavior of a resource over time. For various dimensions of resource behavior, such as skills, utilization, preferences, productivity and collaboration, authors define a set of Resource Behavior Indicators (RBIs) which are discoverable from event logs. The framework enables the definition of new RBIs as necessary via Structured Query Language (SQL) statements. An extension of this work will be to combine several RBIs to a single performance measure. For this purpose, Data Envelopment Analysis – a non-parametric efficiency measurement method developed by Charnes et al. can be used [11]. Another possible extension of this work is using the results of the analysis to enrich event logs with knowledge about resource [11] .

Ahn et al. [3] have proposed and implemented the algorithm so as to be realized as a closeness centrality knowledge analyzer, and illustrate its run-time screen-shots with an operational example. The goal of the algorithm was to measure and calculate theoretically the degree of work-intimacy among workflow-performers involved in a workflow model or a workflow package (a group of inter-relevant workflow models) on a workflow-supported enterprise or organization. The authors have been trying to develop several sophisticated and diversified analysis techniques, such as centrality, prestige, disconnectedness, correspondence, structural equivalence, and affiliation, to be applied to workflow-supported enterprise social networks. Possibilities for further research are directly related with a converged issue of social networks analysis issue and its visualization issue [3].

Martin et al. [7] have focused on solving the problem of resources as a critical component of a business process as they execute the activities. These resources, especially human resources, are not permanently available and tend to be involved in multiple processes [7]. The goal of the paper was to retrieve resource availability in-sights from an event log. All information embedded in the event log is systematically used to obtain daily availability records, which express the resource's availability for the process under consideration [7]. The mined daily availability records are innovative as they are the first to take into account (i) the temporal dimension of availability, i.e. at which time of a day a resource is available, and (ii) intermediate availability interruptions caused by, e.g., a break or the allocation of source to another process [7]. Depending on the time horizon of the analysis period, a large number of daily availability records might be obtained. To describe resource availability on a more aggregated level, an extensible set of resource availability metrics is introduced: Total available time, Total unavailable time, Total idle time, Working day length, and Active time [7].

The goal of the Sindhgatta et al. [12] paper has been to present a context-aware collaborative filtering recommender system that predicts a worker's suitability for a task, in different contexts or situations. The context-aware recommender uses information on the performance of similar resources in similar contexts to predict a resource's suitability for a task [12]. Collaborative Filtering (CF) is a technology that has been widely used in e-commerce applications to produce personalized recommendations for users [12]. Functionally, CF builds a database of preferences or ratings done by distinct users on specific items [12]. A CF algorithm provides recommendations in following ways: (i) Prediction, a numeric value, expressing the predicted preference or rating of an

item for a user. This predicted value is within the same scale as the rating values provided by user. (ii) Recommendation, a list of items, that a user will like [12]. The recommended list is on items that have not been already rated by the user. The key in CF is to locate other users with similar profiles to that of the user for which the recommendations need to be provided (or the active user). These similar users are commonly referred to as 'neighbors' [12].

Pika et al. [10] have presented a framework for analyzing and valuating resource behavior through mining event logs. The framework provides (1) a method for extracting descriptive information about resource skills, utilization, preferences, productivity, and collaboration patterns; (2) a method for analyzing relationships between different resource behaviors and outcomes; and (3) a method for evaluating the overall resource productivity, tracking its changes over time, and comparing it to the productivity of other resources [10]. The framework consists of three modules. The goal of the first module, called Analyzing Resource Behavior, is to discover what resources have been doing - that is, to gain objective information about their skills, utilization, working preferences, productivity, and collaboration patterns [10]. The goal of the second module, called Quantifying the Outcome of Resource Behavior, is to help managers better understand the effects of resource behaviors on different process outcomes [10]. Finally, the goal of the Evaluating Resource Productivity module is to provide a method for evaluating the overall productivity of a resource by comparing it to the productivity of other resources and tracking its evolution over time [10].

Milani and Maggi [8] explore the effect of workload on service times based on historic data and by using regression analysis. They are introducing a framework for categorizing and selecting performance analysis approaches based on existing research of other authors such Pika et al. [11] or Huang et al. [4].

Lee et al. [6] propose a systematic approach that analyzes event logs to select suitable substitutes if the initial human resources become unavailable. The approach uses process mining and social network analysis (SNA) to derive a metric called degree of substitution, which measures how much the work experiences of the human resources overlap, from the two perspectives: task execution and transfer of work [6]. Along with the metric, suitable substitutes are also identified [6].

## 4. Conclusion

This research has discovered 9 articles dealing with measuring human resources behavior using process mining in the time span from 2010 to 2018. This is a rather small number of articles taking into consideration the initial number of 254 articles. One of the reasons for such a small number of articles is the fact that the process mining is a relatively new discipline and the human resource performance measuring within PM is even newer. However, this is a "hot" topic in the area, which develops fast. It can be expected that the number of research in this area will increase.

In this research two research questions have been formulated:

A. What aspects of human resources behaviour can be measured using process mining?

B. What are the main opportunities for future research and identified research areas?

The answer to the first research question is given in each of 9 articles. Nakumba and van der Aalst [9] quantify the relationship between workload and processing speed. The authors focus on the

definition of workload as the number of work items that have been executed over a particular period, but there are also workload definitions that can be used and can be explored.

Huang et al. [4] approach is based on measuring resource behavior from four important perspectives, i.e., preference, availability, competence and cooperation. Process mining techniques are used to mine the event logs as a basis on which resource behaviors in BPM context are measured and applied.

Pika et al [11] have proposed an extensible framework for extracting knowledge from event logs about the behavior of a human resource and for analyzing the dynamics of this behavior over time. For various dimensions of resource behavior, such as skills, utilization, preferences, productivity and collaboration, a set of Resource Behaviour Indicators (RBIs) which are discoverable from event logs are defined. The framework enables the definition of new RBIs using Structured Query Language (SQL) statements.

Ahn et al. [3] have been trying to develop several sophisticated and diversified analysis techniques, such as centrality, prestige, disconnectedness, correspondence, structural equivalence, and affiliation, to be applied to workflow-supported enterprise social networks. The goal of the algorithm was to measure and calculate theoretically the degree of work-intimacy among workflow-performers involved in a workflow model or a workflow package on a workflow-supported enterprise or organization.

Martin et al. [7] suggest a procedure to retrieve daily availability records from an event log The daily availability records, potentially after post-processing, allow managers to gain insight in resource allocation to a particular process, the variation of this availability over several days and weeks, etc. An extensible set of resource availability metrics is introduced: Total available time, Total unavailable time, Total idle time, Working day length, Active time.

Sindhgatta et al. [12] present a context-aware collaborative filtering recommender system that predicts a worker's suitability for a task, in different contexts or situations. CF builds a database of preferences or ratings done by distinct users on specific items [12].

Pika et al. [10] present a framework for analyzing and valuating resource behavior through mining event logs. The framework consists of three modules: Analyzing Resource Behavior (gain objective information about skills, utilization, working preferences, productivity, and collaboration patterns); Quantifying the Outcome of Resource Behavior (to help managers better understand the effects of resource behaviors on different process outcomes); Evaluating Resource Productivity (to provide a method for evaluating the overall productivity of a resource by comparing it to the productivity of other resources and tracking its evolution over time) [10].

Lee et al. [6] propose a systematic approach that analyzes event logs to select suitable substitutes if the initial human resources become unavailable. The approach uses process mining and social network analysis (SNA) to derive a metric called degree of substitution, which measures how much the work experiences of the human resources overlap, from the two perspectives: task execution and transfer of work [6]. Along with the metric, suitable substitutes are also identified [6].

B. What are the main opportunities for future research and identified research areas?

Nakumba and van der Aalst [9] next research will aim at more powerful analysis techniques and a tight coupling between simulation and operational decision-making. They want to make simulation more realistic by adequately modelling resources based on empirical data. Besides workload-dependent process times, they also take into account the fact that people are involved in multiple processes, are available only part-time and work in batches.

Huang et al. [4] will further look into how to distribute work to appropriate resources with the performance consideration in business process execution, and how to handle the situation in which a group of resources work for an activity simultaneously.

Pika et al [11] have predicted an extension of this work to combine several RBIs to a single performance measure. For this purpose, Data Envelopment Analysis – a non-parametric efficiency measurement method developed by Charnes et al. can be used. Another possible extension of this work is using the results of the analysis to enrich event logs with knowledge about resource behaviour.

Ahn et al. [3] suggest future works directly related with a converged issue of social networks analysis issue and its visualization issue.

Lee et al. [6] suggests that other dimensions such as time, quality, or cost should also be considered to assess substitutability between human resources

The author's idea is to investigate the opportunity of further development of suggested Pika et al. [10] framework in order to extend the evaluation of resource productivity which currently uses DEA technique with Balanced Scorecard. It can be justified by the fact that every organization, which wants to be successful, must achieve its goals. The same principle is valid for human resource behavior performance. Connecting the Balanced Scorecard method and DEA method could enable not only the measurement of effectiveness but also the measurement of the efficiency. However, it is the topic to be additionally investigated in further research.

## 5. References

[1] VAN DER AALST, W. and ADRIANSY, A. H., A, de Medeiros AKA, Arcieri F, Baier T, Blickle T, et al. Process Mining Manifesto [Internet]. In: Daniel F, Barkaoui K, Dustdar S, editors. Business Process Management Workshops. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012 [cited 2017 Jan 29]. p. 169–194.Available from: http://link.springer.com/10.1007/978-3-642-28108-2_19

[2] VAN DER AALST, WMP. and DUSTDAR, S., Process Mining Put into Context. IEEE Internet Comput. 2012 Jan;16(1): p. 82–86.

[3] AHN, H., PARK, M., KIM, H. and KIM, KP., Implementing closeness centrality measurements on workflow-supported enterprise social networks. In: 2015 17th International Conference on Advanced Communication Technology (ICACT). 2015. p. 763–767.

[4] HUANG, Z., LU, X. and DUAN, H., Resource behavior measure and application in business process management. Expert Syst. Appl. 2012 Jun 1;39(7): p. 6458–6468.

[5] KITCHENHAM, B., Procedures for performing systematic reviews. Keele UK Keele Univ. 2004;33:2004.

[6] LEE, J., LEE, S., KIM, J. and CHOI, I., Dynamic human resource selection for business process exceptions. Knowl. Process Manag. 2019;26(1): p. 23–31.

[7] MARTIN, N., BAX, F., DEPAIRE, B. and CARIS, A., Retrieving Resource Availability Insights from Event Logs. In: 2016 IEEE 20th International Enterprise Distributed Object Computing Conference (EDOC). 2016. p. 1–10.

[8] MILANI, F. and Maggi, FM., A Comparative Evaluation of Log-Based Process Performance Analysis Techniques. In: Abramowicz W, Paschke A, editors. Business Information Systems. Springer International Publishing; 2018. p. 371–383.

[9] NAKATUMBA, J. and VAN DER AALST, WMP., Analyzing Resource Behavior Using Process Mining. In: Rinderle-Ma S, Sadiq S, Leymann F, editors. Business Process Management Workshops. Springer Berlin Heidelberg; 2010. p. 69–80.

[10] PIKA, A., LEYER, M., WYNN, M. T. and FIDGE, CJ., Hofstede AHMT, Aalst WMPVD. Mining Resource Profiles from Event Logs. ACM Trans. Manag. Inf. Syst. 2017 Mar 23;8(1): p. 1–30.

[11] PIKA, A., WYNN, MT., FIDGE, CJ., TER HOFSTEDE, AHM., LEYER, M. and VAN DER AALST, WMP., An Extensible Framework for Analysing Resource Behaviour Using Event Logs [Internet]. In: Jarke M, Mylopoulos J, Quix C, Rolland C, Manolopoulos Y, Mouratidis H, et al., editors. Advanced Information Systems Engineering. Cham: Springer International Publishing; 2014 [cited 2017 Nov 12]. p. 564–579.Available from: http://link.springer.com/10.1007/978-3-319-07881-6_38

[12] SINDHGATTA, R., GHOSE, A. and DAM, HK., Context-Aware Recommendation of Task Allocations in Service Systems. In: Sheng QZ, Stroulia E, Tata S, Bhiri S, editors. Service-Oriented Computing. Springer International Publishing; 2016. p. 402–416.

[13] THEVENDRAN, V. and MAWDESLEY, MJ., Perception of human risk factors in construction projects: An exploratory study. Int. J. Proj. Manag. 2004 Feb;22(2): p.131–137.