Event record-based evaluation of business scenarios in the logistics domain using process mining

Denise den Hartog
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
d.denhartog@student.utwente.nl

ABSTRACT

Process mining is an upcoming field in data science using event logs. Current methods of evaluation of business scenarios in the logistics domain focus on an ex-post evaluation in set time frames, making it not applicable in real-life scenarios. This study proposes an event record-based model of evaluation of business scenarios in the logistics domain, and in addition, proposes how to select a better scenario if the current one is subpar. We do this by first executing a literature review to classify what different attributes can be present in a log, and which KPI's could be used to quantify the success of a scenario, to finally propose a model that can be used for logistics use cases. With this information, a way to set qualifying criteria for a business scenario is described. The final model will be validated by a case study. This research proposes a method to link attributes to KPI's, and the case study proves the proposed model suitable for the event record-based evaluation of business scenarios, taking a step towards filling the gap in research in the managerial point of view in process mining.

Keywords

Process mining, Business scenarios, Event logs, KPI's, Logistics domain, Case study

1. INTRODUCTION

In today's digital society, we encounter an increasing number of embedded systems in our personal and professional life. All these devices leave a digital trace with every action happening. These traces can be stored in a event log [15]. This log can be seen as an exact description of actions executed by a particular process.

For many businesses it is important to constantly evaluate and adapt their business strategy to the changing world around them. Nowadays this is still often done manually by making a process model by hand. This usually is a model of how the company wants a process to be executed, instead of how it is actually executed. This means that the created models are often an idealized version of the real world [7]. By combining these two areas, process modelling and event logs, we can create process models based on the actual events taking place, hence making genuine process models. This is the realm of process mining, a quickly emerging field in data science [14].

Businesses in the logistics domain have to deal with a lot of uncertainty, a high degree of complexity, and a very dynamic environment [3]. Because of this, these systems have to quickly and continuously adapt to these external factors [2]. These requirements make the logistics domain well suited for process mining. New trends or circumstances can be discovered immediately, and responded to quickly. Process mining is already broadly used in other areas, such as healthcare and finance [7]. Limiting factors there are little standardization [13], and slow response time to changes [10, 12]. Terms and formats can differ greatly between different institutions, and strict regulations greatly slows down changes. This study focuses on the logistics area, because the above mentioned limiting factors are less prevalent in this domain. There are less restrictions when it comes to changing strategy, making the new scenarios directly executable. Next to this, some standards are in place for event logs in this domain [13].

Previous studies often focus on the computer science point of view of process mining, while it can also be used from a managerial point of view, for example in business decision making. This is highlighted by the result of the study of Zerbino et al., showing a research gap in this managerial focus in process mining [18].

We will further research in this area by creating a model of evaluation for business scenarios. We will build upon previous work of Van Midden [16], in which he evaluates business scenarios at set time frames. There are however some downsides to using time frames. As mentioned before, in order to keep up with today's rapid changing environment and conditions, businesses need to constantly evaluate their performance, and quickly adapt if it is not up to par. These changes can also happen inside of a time-frame, causing a delay in detection and therefore response to a disruption.

Therefore the research question we will be answering is:

RQ: How to evaluate business scenarios event record-based in the logistics domain by their logs?

This will be answered by answering the following sub questions:

1. What attributes exist in event logs, and how do these relate to key performance indicators?
2. How to evaluate a business scenario based on these attributes and key performance indicators?
3. How to evaluate a log event record-based?

We will use the knowledge gained from these subquestions to propose a model of evaluation in order to answer the main research question, which will be validated by a case study.

36th Twente Student Conference on IT Febr. 4th, 2022, Enschede, The Netherlands.

The remainder of this paper will first present some background knowledge (Section 2). Then, we will answer the above mentioned subquestions in Section 3. Following this, we will propose the model answering the main research question (Section 3.4), which we will validate with a case study (Section 4). After discussing the results of this case study (Section 5), we will draw conclusions from the previous chapters, and touch upon some future research directions (Section 6).

2. BACKGROUND

Process mining The huge amount of data that is continuously produced in the form of process logs is a useful tool for the creation, evaluation and improvement of business models. Process logs can be seen as the timeline of multiple events. Each entry in this log, also called an event, consists of at least a timestamp, an activity, and a specific case. In addition to this, custom attributes can be added. By using these event logs, a 'true' model can be created on how the process actually takes place, instead of the earlier mentioned idealized model [7].

This is the field of process mining. It "aims to discover, monitor, and improve real processes by extracting knowledge from event logs readily available in today's information systems." [1].

Business scenarios Each process in a business is executed according to some set of rules. Following Van Midden [16], we will call these sets of rules 'business scenarios'. In logistics context this could be rules like: when transporting packages from A to B, first transport the package that has been waiting the longest. More examples of business rules can be found here [4, 8, 9]. In order to determine the success of these scenario's, we need a way to quantify them. This can be done by Key Performance Indicators. KPI's can be taken directly from, or calculated from the attributes present in the event log. Business rules can be based on KPI's. If there is an attribute that states the quality of a product in the log, the quality at the end of the process could be a KPI. A business rule using this could be to discard all products that have a quality of under 80%, or to first transport the product with the highest current quality.

Current method The paper of Van Midden proposes different strategies to evaluate multiple business scenarios. It does this by comparing a KPI at set intervals. Then, the worst performing scenarios are disregarded. The advantage of this way of selecting different scenarios is that you do not have to have every simulation carried out completely in order to already disregard some scenarios that will most likely not perform good. This way of selecting which scenarios to discontinue on set time frames could however run into problems when there are different rates of new log entries per unit of time. A solution for this would be to not have set time frames on which you evaluate the success of a scenario, but to do this directly after every completed trace. A downside of this is that is does cost some computational capacity. However, by evaluating at every new log entry instead of in timeframes, you get a much better representation of reality, instead of just the image at the points in time at each time frame. When evaluating the log in this way we are able to proactively respond to fluctuations, and immediately change a business scenario, instead of waiting for the time frame to be over and evaluate ex-post. Especially in a highly evolving area such as logistics, reacting immediately to fluctuations can give a competitive advantage.

Another aspect of the method of Van Midden is that multiple scenarios are evaluated at the same time, which is not always realistic or possible in practice. For example, if you have one physical store or factory, you may only be able to test one scenario at the same time. Next to this, there are multiple KPI's that can have an influence on the success of a scenario. This paper proposes a method to evaluate multiple KPI's, evaluating directly as new event records are added.

3. RESEARCH QUESTIONS

First we will research which attributes can be present in logs. After this, we will classify the possible KPIs in the logistics domain, and how the previously found attributes can be linked to those. Now that we have a good overview on what is present in an event log, we use this to quantify the success of a business scenario. Finally, we will use all this information to propose a method of evaluation of business scenarios using event logs.

3.1 Event log characteristics

In this section we will research how we can link attributes present in an event log to Performance Indicators (PI's). By linking these attributes to PI's, we will know which attributes should be optimized in order to select a more successful business scenario, while at the same time generalizing our proposed method in order for it to be applicable to logs of different standards and formats. In the next chapter these PI's will be used as a metric to quantify the success of a scenario.

3.1.1 Attributes

The XES standard, eXtensible Event Stream, is used to describe entries in a log. It is an improvement on the previously widely used MXML. The XES standard has been chosen as a basis for the model proposed in this paper, because of its usable format, which is easy to translate to PI's, it's promising future, and the existence of XESame, a tool that can convert any ODBC database to an XES log [17]. This makes this standard widely usable. We will not go into detail on all aspects of the XES standard, but will only focus on the standard extensions. Interested readers are referred to [17]. The exact meaning of an attribute in a log in this format is defined by its extension. There are seven standard extensions, and a user can also create its own custom extensions. The standard extensions are Concept, Lifecycle, Organizational, Time, Semantic, ID, and Cost [17]. The main attributes influencing the success of the business model of the log can be retrieved from the Lifecycle, Time, and Cost attributes.

Lifecycle specifies a transition in the lifecycle of a process. States like *Cancelled, Completed* + {*Failed, Success*}, and *NotRunning* can indicate the progress of an event, which in turn can be used to derive quantitative conclusions about the process. These could for example be used to calculate the ratio of successful traces, or the time that traces are left idle. The Time attribute is important to calculate the total throughput time, and the time it takes to go from one specific state to another. The Cost attribute stores the cost that is connected to a trace or event. This can be used to calculate the cost of one step, or the total of the entire trace. Next to these attributes, a user can add custom attributes. Since these are not standardized, the custom attributes will have to be connected to PI's via other means (e.g., manually).

3.1.2 Key performance indicators

Now that we have seen some attributes that influence the success of a business scenario, we take a look at the PI's that can quantify this success.

The study by Krauth et al. provides a framework on how to measure performance in the logistics domain. The result of the research is a table containing common PI's for logistic service providers. All PI's should be either minimized or maximized (indicated by the arrow behind the PI) in order to improve the success of a business strategy [11]. Since this research focuses on event logs in the logistics domain, not all points of view from the table are represented in these logs. The goal is to quantify the success of a business scenario by looking at just the event log, therefore only PI's that can be (directly) derived from this log will be used.

As mentioned in the previous paragraph, there are three extensions that influence the success of the business scenario, the Lifecycle, Time, and Cost attributes. In Table 1 all PI's that are not able to be calculated or deduced by just these three attributes combined with other characteristics of a log, such as the total number of entries, are left out. The remaining subset of PI's are linked to one of the three extensions that are related to it. The abbreviations are L – Lifecycle, T – Time, and C – Cost. These PI's are directly related one of these attributes. This means that it is possible to evaluate these PI's by their linked attributes. The complete table from [11] also contains the PI's that cannot be directly linked with just the standard attributes, but custom attributes can be used to link to these PI's.

Table 1: Relation attributes - Performance Indicators

Internal perspective - Management point of view	
Effectiveness	
C	Revenue ↑
C	Profit margins ↑
L	Capacity utilization ↑
L	Number of deliveries ↑
L	Perfect order fulfilment ↑
L	Total number of orders ↑
Efficiency	
C	Total distribution cost ↓
L	% of failed orders↓
T	Item/Product/Grade changeover time ↓

3.2 Evaluating business scenarios

In the previous chapter the different attributes present in a log were linked to PI's. The arrows behind the PI's in Table 1 indicate whether the value of the PI should be as high or as low as possible in order to maximize success. This again indicates whether the value being above or below the threshold means disregarding or keeping the scenario. When custom attributes are used in the log, the user should determine this threshold himself.

A distinction will be made between PI's and KPI's. PI's are all performance indicators. KPI's are the selected PI's in order to evaluate the scenario.

Comparison methods There are multiple ways to compare scenarios, depending on the goal of the comparison. In order to get the single best or top scenarios from a closed set of scenarios, a threshold can be narrowed while evaluating all scenarios parallel, until only a small amount of, or only one scenario, remains.

Another applicable scenario would be to evaluate the KPI of a scenario, and if this does not satisfy the threshold, select another scenario to compare. This can also be done for multiple KPI's in parallel, each having their own threshold that needs to be kept in order to prevent exclusion. By evaluating one scenario at a time, the model is also suitable for real world scenarios, since it is possible to run only one scenario at a time in a physical store.

When evaluating multiple KPI's in parallel, assumptions can be made on which business scenario to examine next, depending on which threshold is broken. For example, when two KPI's are evaluated that are 'cost' and 'throughput time'. When the 'cost' threshold is broken, a scenario that will produce lower costs should be selected, while when the 'throughput time' is broken, a scenario the is faster than the current one.

It should be kept in mind that the first few entries in a log might not give true values on the KPI. It takes some time before different variables (such as queues), are at normal levels as if the simulation did not just start. Therefore there should be a warm-up period in which the KPI's stabilize before excluding scenarios.

Threshold There are multiple ways in which you can determine a threshold of exclusions of a KPI. When the user already has a threshold in mind, it is easy to set this as the threshold. If for example the cost needs to be below a set amount of money, or there is a maximum time something can take, the threshold can be set to this value.

If this is not the case, a suitable threshold could be determined manually. The threshold should not be too easy to achieve, which would lead to no scenarios being excluded. When the threshold is too hard to achieve, no scenario will be labeled as suitable. A way to manually set this threshold is to first determine the value of the KPI that is evaluated from a few business scenario's. When the warm up time of the logs of the different scenarios has passed, the mean of the value of the KPI('s) for these scenarios together can be calculated. The threshold can now be determined by taking this mean.

Threshold flexibility There could be one scenario in which a specific trace has poor results, while other traces have superb results. When some of the less optimal traces occur in a row, the scenario could be excluded even though it is the highest scoring one on other traces. This could be solved by working with a X strikes system. The scenario can get X amount of strikes before it is excluded.

3.3 Event record-based evaluation

There are different options on how to select the beginning of the time frame over which the evaluation will take place. To mention some:

- Add all new incoming events, and evaluate all.
- Select a bar of width X event logs, and evaluate only the event logs inside of this bar. This bar will roll over the log as new events are added.
- Select a bar of width X time, and evaluate only the event logs inside of this bar. This bar will roll over the log as time progresses.
- The previous option, but with a bar of width X percentage.

Since the world around us is constantly changing, it would not always make sense to keep on using the event log entries of months or even years ago to evaluate, discarding the first option. The same problem arises when working with a set percentage. As new entries are added, after some time the amount of entries to be evaluated grows rapidly large. Because it is this large, it takes some time before a threshold is breached, even if the last few new entries were all below this threshold. One of the benefits of event record-based evaluation is the possibility to respond

Figure 1: Flow from attributes in event log to the model.

to changes immediately, which would be diminished by using too many old log entries in the evaluation.

As stated before, the rate of entries added in a log can differ from time to time. Because of this, it makes more sense to look at a number of entries, instead of a timeframe. This leaves option two. After the warm up period, the bar will grow with each added entry until the desired length has been reached.

3.4 Proposing the model

From the previous research questions, we now know how we can quantify the success of a business scenario by its log. Now we bring the above gained knowledge together to form a model for evaluation of these scenarios.

Figure 1 illustrates the flow of the process. First we link the attributes present in the log to their corresponding KPI's. Then we can use these KPI's in the model we propose next.

The model uses event record-based evaluation. This means that every time a new log trace is completed, the scenario is reevaluated. It does so by looping through the model in Figure 2.

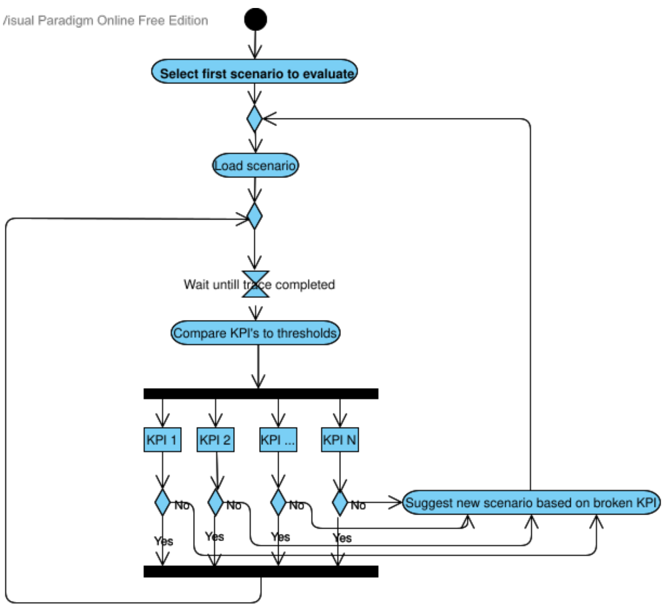

Figure 2: Proposed model of evaluation.

First, we select which business scenario to start with. Then we go through the log as new entries are added. After the warm up period has passed, at each completed trace, it checks the KPIs against the set thresholds. If the threshold is met, the model waits for the next trace to be completed. If the threshold is not met, the model gives a suggestion on how to select the next business scenario, based on which KPI threshold is not met.

4. CASE STUDY

After this theoretical part, we will do a case study in order to test our proposed method of evaluating. The tool used for this case study will be our model created in Python. The data we will be using is the dataset that is created for the paper by Bemthuis et al. We will only list the details that are relevant for our case study, the complete information can be found in [5].

This corresponding dataset [6] consists of event logs from the simulation of a factory where three kinds of vehicles are driving by different rules, which need to transport products with different characteristics from A to B. The three types of vehicles have different properties (such as speed and capacity), and select which product to transport according to different business rules, which are grouped together in scenarios.

These scenarios consist of two kinds of rules. The product-initiated rules are Random, Lowest utilization, and Shortest travel distance [5]. The vehicle-initiated rules are Random, Lowest-, and Highest quality decay. Next to this, there is a quality decay threshold value, which is set to 60% for all scenarios. This means that products that have a quality lower than 60% (before being picked up) are discarded. Next to the different dispatching rules per scenario, it also differs how many of which vehicle are driving. In total there are 27 scenarios, each having its own set of dispatching rules and amount of types of vehicles [5].

As defined in the paper from the dataset [5], the warm up period is set to two hours. This means only the traces of the log that are started after two hours are used. Different attributes are present in the log. From the XES standard there are Time (timeStamp) and Lifecycle (event) attributes directly present in the log. Next to this, there is the custom attribute *currentDecayLevel*. From the information at [5] the cost PI's could also be determined, but this is out of scope for the purpose of this research. Looking at Table 1, this means only the following PI's are used:

Lifecycle:

- Capacity utilization ↑
- Number of deliveries ↑
- Perfect order fulfilment ↑
- Total number of orders ↑
- % of failed orders↓

Time:

- Item/Product/Grade changeover time ↓

For the custom attribute, a custom PI 'Average quality' will be created, which should be maximized (↑). Due to time restrictions, it is not possible to test all possible KPIs to see which scenario is the best. We will restrict this case study by only looking at the changeover time (throughput time), and the average quality of the end product as our KPI's.

Since the data for this case study concerns results obtained from a simulation study, all event log entries are already known. Nevertheless, this case can still be used to test the proposed model for real life situations as well (e.g., in an real-time manner), since the logs are regarded as if the entries are added on their timestamp. By using the (synthetic) log files in this way, it is possible to use it to test our model for a real-life scenario.

Predefined thresholds are set for the throughput time and the product quality. As can be seen in Table 1, the KPI: Throughput time (Item changeover time) should be minimized. Therefore the threshold is the maximum that this

KPI should be. This means that the scenario should have an average throughput time less than this threshold. The custom defined KPI: Quality should be maximized. This means that this threshold is the minimum. The following starting thresholds will be used.

- Throughput time = maximum 650.000ms
- Quality = minimum 60%

Before starting the evaluation, all scenarios that do not meet these thresholds after the warm up time are disregarded. This has been done for practical modeling purposes. After this initial filtering, we take the first scenario, and evaluate this until either we are at the end of the log and there are no more new entries to evaluate, or one of the thresholds is not met. In the first case, we randomly pick another not-disregarded scenario to continue the evaluation. If one of the thresholds is not met, the current scenario is disregarded. The next scenario that is chosen is based on the previously broken threshold. In the meantime, the thresholds are gradually risen or lowered, according to Table 1.

In order to test our proposed model, we will execute three experiments. First, we will use the proposed model to create a ranking of the scenarios from the dataset using both the quality and the throughput time KPI's. This is the intended use for the model.

Second, to determine the influence of evaluating multiple KPI's at the same time, we will run the model while evaluating just a single KPI. We will be comparing these rankings to that of Van Midden [16].

Third, we will run the model with the amount of strikes set to zero, to test the influence of the strike system, as proposed in Subsection 3.2.

5. RESULTS

From the above described case study we get the following results.

The results of experiment one and two, running both KPI's at the same time, and individual, can be found in Appendix A and Appendix B. The scenarios that did not meet the starting thresholds are scenarios 22, 23, and 26.

From Appendix A we see that evaluating one or multiple KPI's can have a huge difference on the final ranking of that scenario. In the top five scenarios we see similarities between the Time KPI and both KPI's, but this is less present in the remainder of the ranking. This might indicate that the quality was not the limiting factor when the thresholds rose sufficiently high. Another interesting finding from this table is that the scenarios that score high on one KPI, usually score mediocre on the other KPI. Scenario 3 for example scores highest when we only evaluate the Quality KPI, but ends up halfway on the list when we also take the throughput time into account. This can be seen in multiple scenarios. When the two KPI's separate both score mediocre, the combined ranking usually is higher (see scenario 9, 7, 18). These two findings could indicate that there is a trade-off between the KPIs.

In Appendix B we compare our results to those of Van Midden [16]. One important note is that Van Midden only uses the Quality KPI to create the ranking, and that this research uses set timeframes. We can see that the bottom half of the ranking looks quite similar. The top 10 however differs quite. This is in line with the previously mentioned finding that the quality is not the limiting factor when using high thresholds.

Next, we tested the result of the strike system, as proposed in Subsection 3.2. We did this by also running the model with the amount of strikes allowed set to 0. The main difference here is that all scenarios were excluded based on their throughput time. This indicates that there is high fluctuation in this KPI. The ranking list if we do not use strikes can be found in Appendix B.

6. CONCLUSION

This research aimed to propose a model to evaluate business scenarios in the logistics domain event record-based by their event logs. By proposing a means to link attributes to performance indicators, and using this to detect which KPI's could be used for the evaluation of the event log, we can more broadly apply the proposed model of evaluation. The case proved the usability of the model, the influence of evaluating multiple KPI's, and that it can produce a ranking of the possible scenarios.

We now discuss some further improvements on the above described work. The case study shows that the proposed model can be of use in the selection of business scenarios. However, there are many areas that still need research in order to validate the model. For example:

- In the case study, setting the threshold is done by using predetermined values, since there were no scenarios that met the average of the throughput time of all scenarios together. A method can be developed that automatically determines a suitable value for this.
- More research should be done in order to determine the relation different KPI's have on the success of a scenario.
- We randomly selected a new scenario when no thresholds are broken, but we run out of log entries for this scenario. A different way of selecting another scenario could be tested.
- When a threshold is broken, we select the next scenario based on the scenario with the highest value of this broken KPI. As mentioned in the conclusion, this is not always the most successful scenario when we combine multiple KPI's.

Altogether, this paper contributed to filling the research gap currently present in process mining from the managerial point of view. However, this upcoming area needs more research before it is fully applicable in practice.

7. REFERENCES

[1] W. Aalst, van der. *Process mining : discovery, conformance and enhancement of business processes.* Springer, Germany, 2011.

[2] M. Barad and D. Even Sapir. Flexibility in logistic systems—modeling and performance evaluation. *International Journal of Production Economics*, 85(2):155–170, 2003.

[3] T. Becker and W. Intoyoad. Context aware process mining in logistics. *Procedia CIRP*, 63:557–562, 2017.

[4] R. Bemthuis. Business logic for resilient supply chain logistics. In *2019 IEEE 23rd International Enterprise Distributed Object Computing Workshop (EDOCW)*, pages 190–195. IEEE, 2019.

[5] R. Bemthuis, M. Mes, M.-E. Iacob, and P. Havinga. Using agent-based simulation for emergent behavior detection in cyber-physical systems. *2020 Winter Simulation Conference (WSC)*, 2020.

[6] R. Bemthuis, M. Mes, M.-E. Iacob, and P. Havinga. Data underlying the paper: Using agent-based simulation for emergent behavior detection in cyber-physical systems. 2021. https://data.4tu.nl/articles/dataset/Data_underlying_the_paper_Using_agent-based_simulation_for_emergent_behavior_detection_in_cyber-physical_systems/14743263.

[7] C. d. Garcia, A. Meincheim, E. R. Faria Junior, M. R. Dallagassa, D. M. Sato, D. R. Carvalho, E. A. Santos, and E. E. Scalabrin. Process mining techniques and applications – a systematic mapping study. *Expert Systems with Applications*, 133:260–295, 2019.

[8] I. Graham. *Business rules management and service oriented architecture: a pattern language.* John wiley & sons, 2007.

[9] B. N. Grosof, Y. Labrou, and H. Y. Chan. A declarative approach to business rules in contracts: courteous logic programs in xml. In *Proceedings of the 1st ACM Conference on Electronic Commerce*, pages 68–77, 1999.

[10] J. Haycock, A. Stanley, N. Edwards, and R. Nicholls. The hospital of the future: Changing hospitals. *BMJ*, 319(7219):1262–1264, 1999.

[11] E. Krauth, H. Moonen, V. Popova, and M. Schut. Performance measurement and control in logistics service providing. pages 239–247, 01 2005.

[12] N. Lightner and T. Bagian. The benefits of standardization in healthcare systems. *Advances in Human Factors and Ergonomics in Healthcare and Medical Devices*, page 51–57, 2021.

[13] J. P. S. Piest, J. A. Cutinha, R. H. Bemthuis, and F. A. Bukhsh. Evaluating the use of the open trip model for process mining: An informal conceptual mapping study in logistics. In *Proceedings of the 23rd International Conference on Enterprise Information Systems*, pages 290–296, 2021.

[14] W. Van Der Aalst. Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems (TMIS)*, 3(2):1–17, 2012.

[15] W. M. Van Der Aalst, H. A. Reijers, A. J. Weijters, B. F. van Dongen, A. A. De Medeiros, M. Song, and H. Verbeek. Business process mining: An industrial application. *Information Systems*, 32(5):713–732, 2007.

[16] Y. Van Midden. Using process mining and event log analysis for better business strategy decision-making. 2021.

[17] H. M. Verbeek, J. C. Buijs, B. F. van Dongen, and W. M. van der Aalst. Xes, xesame, and prom 6. *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, page 60–75, 2011.

[18] S. A. Zerbino, P. and D. Aloini. Process science in action: A literature review on process mining in business management. *Technological Forecasting and Social Change*, (172), 2021.

APPENDIX

A. RANKING OF SCENARIOS

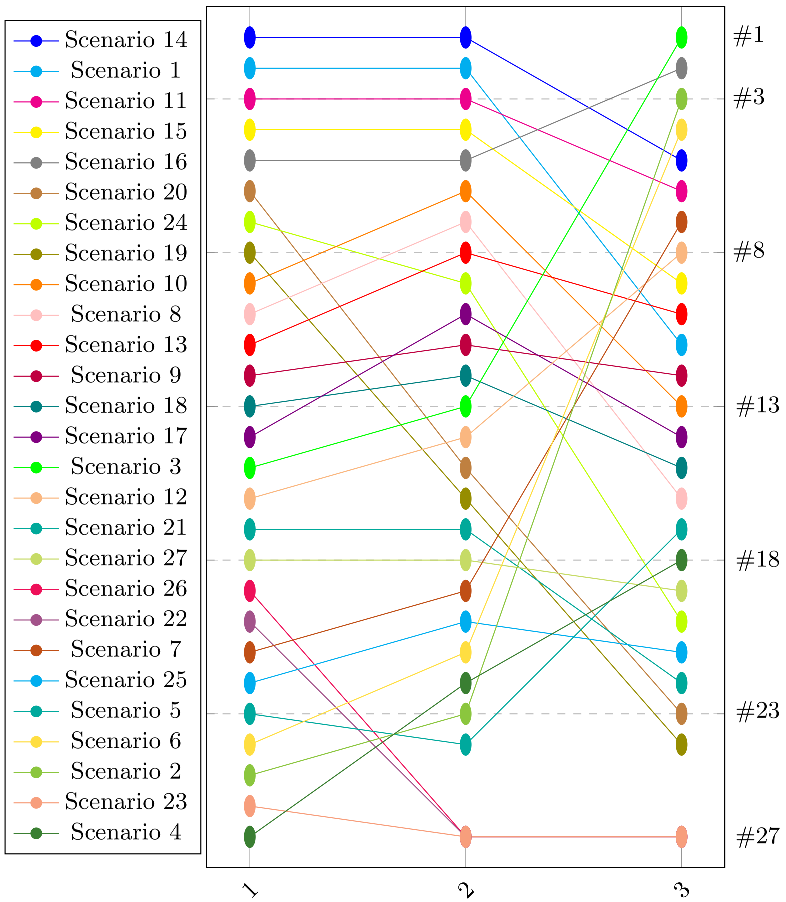

Figure 3: Ranking of scenarios based on: 1. One KPI: Throughput time, 2. Two KPI's: Quality & Throughput time, 3. One KPI: Quality

B. RANKING DIFFERENT EXPERIMENTS

Rank	*Ex. 1: Proposed model*	*Ex. 2: Time*	*Ex. 2: Quality*	*Van Midden, Ex 1, Quality*	*Van Midden, Ex 2, Quality*	**Ex. 3: No strikes**
#1	14	14	3	12	12	14
#2	1	1	16	16	16	11
#3	11	11	2	15	14	1
#4	15	15	6	14	13	13
#5	16	16	14	13	11	10
#6	10	20	11	11	15	17
#7	8	24	7	3	7	15
#8	13	19	12	2	6	24
#9	24	10	15	7	3	12
#10	17	8	13	6	2	16
#11	9	13	1	1	5	18
#12	18	9	9	9	1	7
#13	3	18	10	5	9	25
#14	12	17	17	18	18	20
#15	20	3	18	10	10	5
#16	19	12	8	17	17	6
#17	21	21	5	4	4	3
#18	27	27	4	8	8	4
#19	7	26	27	25	25	27
#20	25	22	24	27	27	21
#21	6	7	25	24	24	9
#22	4	25	21	21	21	19
#23	2	5	20	20	20	8
#24	5	6	19	19	19	2
#25		2		22	22	
#26		23		23	23	
#27		4		26	26	
Not placed	{22, 23, 26}		{22, 23, 26}			{22, 23, 26}

Table 2: Ranking of different experiments.