Outsourcing Prioritization for Bottleneck Processes Using Process Mining: A Logistics Case Study

VENELINA POCHEVA, University of Twente, The Netherlands

If properly monitored, systems in the real world generate data when a process is executed. This data is a valuable resource for creating process models and analyzing performance. Process mining can be used to identify, predict, and avoid bottlenecks and inefficiencies in processes. Thus far, (most) existing research has been done regarding the discernment and subsequent understanding of bottlenecks. Inspired by the limited amount of research regarding the prediction of bottlenecks and the subsequent elimination strategies, this research proposes a method for mitigating bottlenecks using process mining and outsourcing techniques. Outsourcing is a technique in which an organization contracts a third party to complete work, manage operations, or deliver services on its behalf. Many organizations outsource so they can use their limited internal resources better. However, organizations usually face the challenge of determining exactly what to outsource. If there are several bottlenecks, it can be challenging to select one. Thus, this study aims to provide a (full-fledged) method for outsourcing prioritization of bottlenecks in various processes. Furthermore, a logistics case study demonstrates the feasibility of the proposed method. Although preliminary, the presented method is expected to enrich the scientific field of process mining and bottleneck processes.

Additional Key Words and Phrases: Process mining, Bottleneck, Process, Prioritization, Outsourcing, Logistics

1 INTRODUCTION

1.1 Overview

Contemporary enterprise applications and systems usually exhibit optimal behaviour; however, these systems have also been known to occasionally show (unexpected) performance degradation behaviour, which, in turn, has been associated with process downtime (and subsequent operational expenses) [12]. These respective behavioural changes and performance reductions have often been indications of bottlenecks in the systems. Bottleneck conditions such as system congestion and resource exhaustion have reportedly led to prolonged and intermittent system outages [17]. A step toward improving the bottleneck is to identify it. As a result, before looking for the bottleneck, it is critical first to define what a bottleneck is. There are many different definitions of a bottleneck. As mentioned by Roser et al. (2015) [19], these include: (1) a function that limits the capacity of the system, and (2) a process whose capacity is less than expected or the process that reduces throughput. Furthermore, Heo et al. (2018) [11] describe a bottleneck as (1) the overload point of systems that slows down the entire operating system and (2) the point at which the execution time of activity takes the most time.

TScIT 37, July 8, 2022, Enschede, The Netherlands

Knowledge of bottleneck processes is of great importance for improving the performance of the processes. Regardless of the causes of the problem, the challenge is to recognize performance reductions, identify possible causes and find an appropriate solution [12].

Using process mining techniques, bottleneck processes can be detected [5]. As mentioned by dos Santos Garcia et al. (2019) [8], process mining focuses on understanding processes and helps capture the more essential insights during real-world execution. Process mining techniques have been used to analyze and monitor various business processes based on event data. An activity identifier, a case identifier, and the time of execution (timestamp) are the three minimum components of the data that characterize the execution of activities [20]. According to van der Aalst et al. (2011) [1], process mining distinguishes three forms of mining: (1) process discovery, (2) conformance verification, and (3) process enhancement. The goal is to provide insights on improving the processes' efficiency. Process optimization includes detecting and minimizing bottlenecks, avoiding needless states and reprocessing loops, and analyzing the areas that are excessively time-consuming [10].

One way of minimizing bottlenecks is by outsourcing. Outsourcing is the process of shifting work, responsibility, and decision-making authority to a third party [9]. When outside providers create products or services more efficiently and effectively, a competitive advantage can be acquired [16]. The goal is to lower overhead expenses, reduce payroll, and gain better access to outside information.

Logistics has become one of the most popular domains for outsourcing [3]. Logistics is the function responsible for transporting and storing materials on their way between suppliers and customers [22]. A logistics process usually involves several parties, such as shippers, carriers, service providers, or transportation companies, and is spread across different countries [18]. This results in processes that are extremely complex and dynamic. Many companies use logistics outsourcing as a means to reduce costs, increase flexibility, improve performance and the ability to concentrate on their primary business [2, 23].

1.2 Problem definition

Mitigating inefficiencies in processes is of great importance for enhancing process performance. Several studies have highlighted and described the significance of eliminating bottlenecks [19]. Previous studies on process mining and bottleneck processes are mainly focused on identifying and detecting bottlenecks [5]. A limited amount of studies propose a technique for bottleneck mitigation.

1.3 Objective

The main objective of this study is to develop an approach for outsourcing prioritization for bottleneck processes using process mining.

1.4 Research questions

The problem definition and objective lead to the subsequent research question:

- RQ1: How can process mining be used in outsourcing processes in order to mitigate bottlenecks?

The main question will be answered with the help of the following four sub-questions:

- RQ1.1: What is state-of-the-art in the work being done on mitigating bottlenecks through process mining techniques?
- RQ1.2: How can outsourcing help to mitigate bottlenecks?
- RQ1.3: How can an approach for bottleneck mitigation strategies be designed?
- RQ1.4: How can the proposed method be applied in a case study?

1.5 Tasks

Achievement of the objective has been done by means of completing the following tasks:

(1) First, a systematic literature review has been conducted to become acquainted with state-of-the-art bottleneck mitigation strategies concerning process mining in particular.
(2) Next, an in-depth analysis has been conducted to become familiar with bottleneck process and outsourcing techniques.
(3) Based on the obtained knowledge, an approach for outsourcing prioritization for bottleneck processes using process mining has been proposed.
(4) Finally, the proposed approach has been demonstrated through a case study.

1.6 Structure

This paper has been organized as follows. Section 2 examines the related work and gives an overview of state-of-the-art methodologies used in mitigating bottlenecks. Furthermore, in the section, an explanation of how outsourcing can help to mitigate bottlenecks is given. Section 3 describes the designed bottleneck outsourcing approach. Section 4 focuses on the validation and the demonstration of the proposed approach. Section 5 evaluates and discusses the results of this study. Finally, the conclusions, limitations and future work are discussed in Section 6.

2 RELATED WORK

This section outlines the related work. A systematic literature review on the topics of process mining, bottlenecks and outsourcing had been conducted with the aim of understanding the (state-of-the-art) techniques utilized for the investigation of the above-mentioned research questions.

To begin with, Kremic et al. (2006) [15] developed a decision support model for decision-makers in public organizations, which considered typical (ad-hoc) outsourcing motives. However, this model incorporates factors beyond standard quantitative measures and did not use process mining. Furthermore, Southier et al. (2020) [7] developed a method to reduce the lead-time of business processes based on PERT/CPM (Program Evaluation and Review Technique/ Critical Path Method) techniques, which allowed for the utilization of a critical path identification framework for the identification of the relevant activities (and bottlenecks), concerning a business process. Additionally, Kinast et al. (2022) [14] analyzed the potential of hybridizing an optimization algorithm with process mining techniques and suggested improving the solution quality by assigning priority to the workstations that cause bottlenecks and focusing the improvement strategies on these workstations.

On the other hand, Iqbal et al. (2013) [13] outlined the advantages, disadvantages, and risks that should be considered when making an outsourcing decision. In addition, Gandhi et al. (2012) [9] proposed a systemic approach for prioritizing outsourcing risks, which includes an understanding of the external factors that might influence prioritization. Although the two studies discuss outsourcing techniques and risks that should be considered when an outsourcing decision is made, neither discussed bottleneck prioritization.

In recent years, outsourcing bottleneck processes, with the aim of reducing overall lead time, has become an essential corporate tool for system optimization [16]. For this purpose, the method developed within this study (see Section 3) has facilitated an exploratory approach for bottleneck mitigation by means of process mining - and outsourcing techniques.

3 METHODOLOGY

This section describes the developed method for outsourcing bottleneck processes based on their respective prioritization using process mining. For this, it has been important to note that this approach has been based on NASA's Reliability-Centered Maintenance (RCM) Guide For Facilities and Collateral Equipment [6]. This, in the sense of prioritizing bottlenecks based on their specific effects, occurrence and severity, within a well-structured framework [1]. Thus, the goal of the proposed approach has been to improve process performance by prioritizing - and subsequently outsourcing bottlenecks. For this, the developed method has been described as follows (Note that Figure 1 has provided a schematic overview of the developed bottleneck outsourcing method.):

(0) **Assumptions and boundary conditions**:
   - The generated list of bottlenecks contains only bottleneck processes that can be outsourced (i.e. the main process/activity of the company should not be included).

(1) **Developed approach**:
   - **Preparation**
     (i) First, the raw data from the selected database is cleaned by means of transformation and statistical filtering.
     (ii) Second, the cleaned data was transformed into a format suitable for process mining, where the data had been simplified and converted into an event log format. Accordingly, each event in the converted data referred to a specific case, activity, timestamp and (if necessary) resource (or other information).
     (iii) Third, the event logs were imported into the workspace, where they were converted into an XES file.

[1]Here, just as the case with the NASA RCM methodology [6], the proposed method aimed to facilitate a structured framework for delineating the highest priority bottlenecks

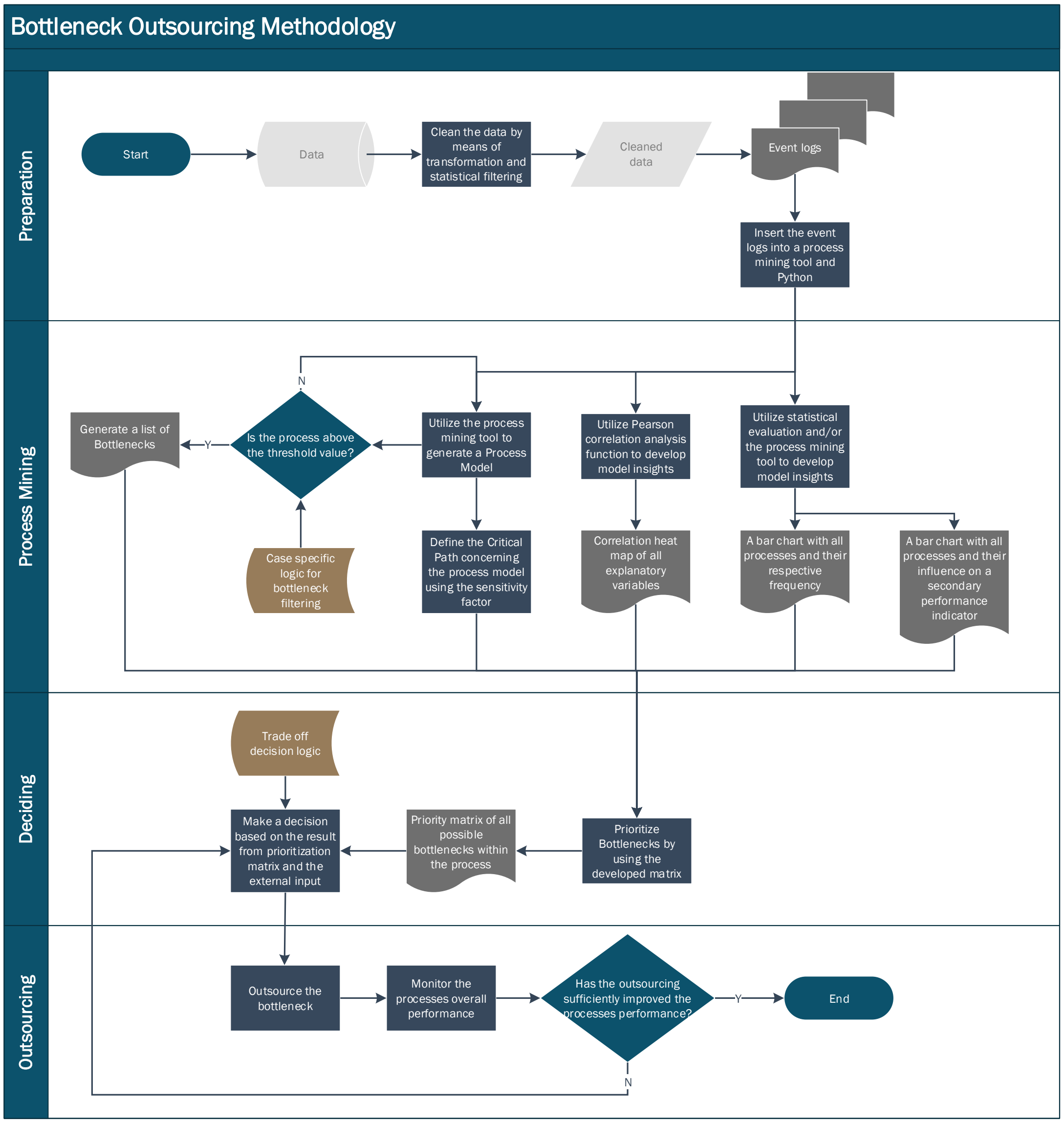

Fig. 1. Schematic overview of the developed model for outsourcing prioritization for bottleneck processes using process mining techniques. A bottleneck prioritization matrix has been created as an extension for this model. This schematic serves as an extension of the process explained above. Note that here, Preparing and Outsourcing have to be performed in-depth, and they do not fit into the scope of this case study. Thus, they will not be discussed in the demonstration.

– **Process Mining**
  (i) Hereafter, a discovered model is created based on the event logs using the process mining tool. The model represented the current processes and visualized how the processes were currently running.
  (ii) The next step is to generate a list of bottlenecks considering the logic filtering and the boundary conditions.
  (iii) After this, the critical path concerning the process model was defined. A critical path is the sequence of activities that determines how long the process must take to be completed in the shortest amount of time [7].
  (iv) Regarding the event logs and the developed process model, an in-depth analysis is performed. The analysis can be conducted using Python and a process mining tool. As a result, it is determined how long activities take, how frequent the process steps are, how activities correlate with each other and whether a process influences a secondary performance indicator.

– **Deciding**
  (i) Using the results from the analysis mentioned above, the next step is to prioritize the list of bottlenecks. For this, a bottleneck prioritization matrix has been created (see

BOTTLENECK PRIORITIZATION MATRIX

	Criteria	Is the bottleneck correlated with another bottleneck?		Is the bottleneck correlated with more than one other bottleneck? How many (in total)?		Does the bottleneck process affect a secondary performance indicator?		Does the bottleneck impact the main process?		Does the bottleneck have high frequency?		Is the bottleneck on the critical path?		TOTAL
№	List of bottlenecks	Value	Points	Value	Points	Value	Points	Value	Points	Value	Points	Value	Points	Points
1	pickedUpRegion1-AGV:1	0,9	8	1	7	1,017	5	0	0	1422	5	1	1	26
2	droppedOffRegion2-AGV:1	0,9	8	1	7	1,585	7	0	0	1419	4	1	1	27
3	assignedToVehicleRegion3-AGV:2	0,5	7	1	7	0,491	2	0	0	1685	7	1	1	24
4	pickedUpRegion3-AGV:2	0,9	8	2	8	1,148	6	0	0	1683	6	1	1	29
5	droppedOffRegion3-AGV:2	0,9	8	1	7	1,998	8	0	0	1687	8	1	1	32
6	droppedOffRegion2-HDF:1	0	6	0	6	0,756	3	0	0	1201	3	0	0	18
7	pickedUpRegion3-HDF:2	0,9	8	1	7	0,476	1	0	0	1183	2	0	0	18
8	droppedOffRegion3-HDF:2	0,9	8	1	7	0,901	4	0	0	1182	1	0	0	20

Fig. 2. Illustration of the developed decision matrix for outsourcing bottleneck prioritization. This table serves as an extension of the process explained above. Note that here, droppedOffRegion3-AGV:2 should be the first bottleneck to be considered for outsourcing as it has the most total points.

Table 1. Frequency of the events and average decay per process

Activity	Decay	Frequency	Activity	Decay	Frequency	Activity	Decay	Frequency
arrivalAtSource	0.0	7497	assignedToVehicleRegion3-UAV:6	0.0243	1548	pickedUpRegion3-AGV:2	1.1482	1683
productCallsForTransportRegion1	0.0	7497	assignedToVehicleRegion3-HDF:2	0.2796	1182	droppedOffRegion2-UAV:1	0.4495	1639
productCallsForTransportRegion3	0.0	7499	assignedToVehicleRegion3-AGV:2	0.4914	1685	droppedOffRegion2-UAV:2	0.4484	1618
startProcessingRegion2	0.0	7500	pickedUpRegion1-UAV:1	0.1153	1639	droppedOffRegion2-UAV:3	0.4407	1623
finishedProcessingRegion2	0.0	7499	pickedUpRegion1-UAV:2	0.1187	1618	droppedOffRegion2-HDF:1	0.7561	1201
assignedToVehicleRegion1-UAV:1	0.0063	1638	pickedUpRegion1-UAV:3	0.1119	1623	droppedOffRegion2-AGV:1	1.5855	1419
assignedToVehicleRegion1-UAV:2	0.0089	1617	pickedUpRegion1-HDF:1	0.4022	1201	droppedOffRegion3-UAV:4	0.5186	1538
assignedToVehicleRegion1-UAV:3	0.0052	1622	pickedUpRegion1-AGV:1	1.0179	1422	droppedOffRegion3-UAV:5	0.5348	1547
assignedToVehicleRegion1-HDF:1	0.1406	1201	pickedUpRegion3-UAV:4	0.1691	1538	droppedOffRegion3-UAV:6	0.5365	1549
assignedToVehicleRegion1-AGV:1	0.1666	1422	pickedUpRegion3-UAV:5	0.1774	1546	droppedOffRegion3-HDF:2	0.9016	1182
assignedToVehicleRegion3-UAV:4	0.0185	1538	pickedUpRegion3-UAV:6	0.1709	1548	droppedOffRegion3-AGV:2	1.9988	1687
assignedToVehicleRegion3-UAV:5	0.0226	1546	pickedUpRegion3-HDF:2	0.4762	1183			

Figure 2). The matrix consists of a list of bottlenecks and six criteria. Points are distributed to all bottlenecks, and the total score is derived.

(ii) After this, possible external factors should be considered. These include but are not limited to budget, technologies, resources, companies and deadlines.

(iii) Based on the external factor and the results from the bottleneck prioritization matrix, a trade-off decision can be made as to which bottleneck should be outsourced.

– **Outsourcing**

(i) Having identified the areas for improvement and made the decision; the next step is to improve these processes. The improvement is made by outsourcing the bottleneck process.

(ii) The final step is to monitor and evaluate whether the new process is working as expected. An analysis may be performed, or the procedure can be repeated to enhance the process.

3.1 On the developed Bottleneck Prioritization Matrix

The goal of the prioritization matrix is to use the list of bottlenecks and the insights gained from the process mining to fill in the matrix and then assign points for every bottleneck and criteria. The points can be from 0 to *n*, the total number of bottlenecks. After the values for all bottlenecks and given criteria are placed in the matrix, they are compared, and the bottleneck with the highest value gets the most points, namely *n*. The bottleneck with the second highest value gets *n-1* points. If two or more bottlenecks have equal values, they are assigned the same number of points. If the criteria is a yes/no question, then only 0 and 1 points are assigned. The matrix consists of the following criteria:

– **Is the bottleneck correlated with another bottleneck?**
For this, a Pearson correlation analysis should be performed. The correlation coefficient value between two bottlenecks describes the extent to which the two move in concert with each other [24]. As a result, if two bottlenecks have a positive correlation coefficient value, improving one will also improve the other. Note that here, the biggest correlation coefficient is taken into account.

– **Is the bottleneck correlated with more than one other bottleneck? How many (in total)?**
For this, a check with how many bottleneck processes a bottleneck is positively correlated with should be performed. In order to have a weakly, moderately, strongly or perfectly positive correlation, the coefficient should be bigger than 0.2 [24]. For this, the aforementioned correlation can be used.

– **Does the bottleneck process affect a secondary performance indicator?**

For this, an analysis of the data and the event logs should be performed.
- **Does the bottleneck impact the main process?**
For this, the main process should be identified. The main process is usually the activity for which the business was established. The aforementioned correlation analysis can be used to determine if the bottleneck impacts the main process.
- **Does the bottleneck have a high frequency?**
For this, the number of times the bottleneck occurs should be counted. This can be done with the help of process mining tools like ProM or Disco.
- **Is the bottleneck on the critical path?**
For this, a check if the bottleneck is one of the activities from the critical path should be made. This can be done in Python.

4 RESULTS

This section explains the performed case study that demonstrates the proposed method. The case study is based on the previous research work by Bemthuis et al. (2020) [4], in which a conceptual agent-based simulation framework for analyzing and learning from emergent behaviour in a logistics context is presented. The data is from a factory that processes four types of products that are transported around by three types of vehicles: human-driven forklifts (HDF), automated guided vehicles (AGV), and unmanned aerial vehicles (UAV). The factory is divided into three regions: region 1, region 2, and region 3. Normally, products were transferred from region 1 to region 2 and from region 2 to region 3. The data set contains data from 27 different scenarios in which specific dispatching rules, known as Vehicle-initiated Product-initiated rules, had been applied [21]. The case study used the data from the first scenario, where three UAVs, one HDF and one AGV have been used.

4.1 Definitions

Note that for this case study, the results have been based on certain assumed definitions, elaborated upon as follows:

- **Bottleneck** - in this case study, a bottleneck is a process that takes more than expected time, which is 15 seconds. Bottleneck processes from the process model (see Figure 4): pickedUpRegion1-AGV:1 (18,5 sec), droppedOffRegion2-AGV:1 (30 sec), assignedToVehicleRegion3-AGV:2 (17,8 sec) pickedUpRegion3-AGV:2 (29,1 sec), droppedOffRegion3-AGV:2 (37,5 sec), droppedOffRegion2-HDF:1 (15 sec), pickedUpRegion3-HDF:2 (15 sec), droppedOff-Region3-HDF:2 (20 sec). The bottlenecks have been identified with the help of the process model created by Disco.
- **Main activity** - the main activity in this case study is the processing of the products, which from the process model is: finishedProcessingRegion2 (see Figure 4).
- **Critical path** - critical path in the process model (see Figure 4): arrivalAtSource, productCallsForTransport- Region1, assignedToVehicleRegion1-AGV:1, pickedUp Region1-AGV:1, droppedOffRegion2-AGV:1, startProcessingRegion2, finishedProcessingRegion2, productCallsForTransportRegion3, assignedToVehicleRegion3-AGV:2, pickedUpRegion3-AGV:2, droppedOff-Region3-AGV:2. The critical path have been discovered with the help of the process model and Python.
- **Secondary performance indicator** - in this case study, this is the decay of the products. The execution of every process increases the decay of the products. Note that here, the quality of the products when they arrive at the source is 100, and it decreases while they are being transported in regions 1 and 3.

4.2 Case Study

The case study discusses only the Process Mining and the Deciding part of the method. The results from the case study can be seen in Figure 2. The tools used for the case study are Python and Disco. Disco has been used to create the process model, detect the bottlenecks, and calculate the processes' frequency. Python has been used for the correlation analysis and the analysis of the secondary performance indicator and also to discover the critical path.

(0) **Assumptions and boundary conditions**:
   - Although, finishedProcessingRegion2 takes 3 minutes mean time to be completed, the activity cannot be in the generated bottleneck list because it has been defined as the main activity.

(1) **Developed approach**:

   - **Process Mining**
     (i) First, Disco was used to generate a process model based on the event logs. The model depicts how the processes are operating. An illustration of the discovered model can be found in Figure 4.
     (ii) Second, a list of bottlenecks is created using the process model and taking boundary conditions and logic filtering into account. For every process was checked whether it is the main process (boundary condition) and whether it takes more than 15 seconds for completion (filtering logic). In this study eight bottlenecks were identified. The list of bottlenecks can be seen in Figure 2.
     (iii) Third, the critical path regarding the process model was discovered. For this, Python and the sensitivity factor of the processes, which in this case is the mean time for completion of every process, were used. The critical path can be found in Figure 4.
     (iv) Then, an in-depth analysis was conducted regarding the event logs and the developed process model. The analysis was performed using Python and Disco. The outcome from the correlation analysis can be seen in Figure 3. The frequency of the processes and the results from the analysis of the secondary performance indicator can be found in Table 1. From this, it can be seen how frequent the process steps are, how activities correlate with each other and whether and how much a process influences a secondary performance indicator.
   - **Deciding**
     (i) Hereafter, the created bottleneck prioritization matrix was used to compare the identified bottlenecks. The results can be found in Figure 2. It has been important to note that the bottleneck with the highest total points is droppedOff-Region3-AGV:2, and it should be the first to be considered for outsourcing.

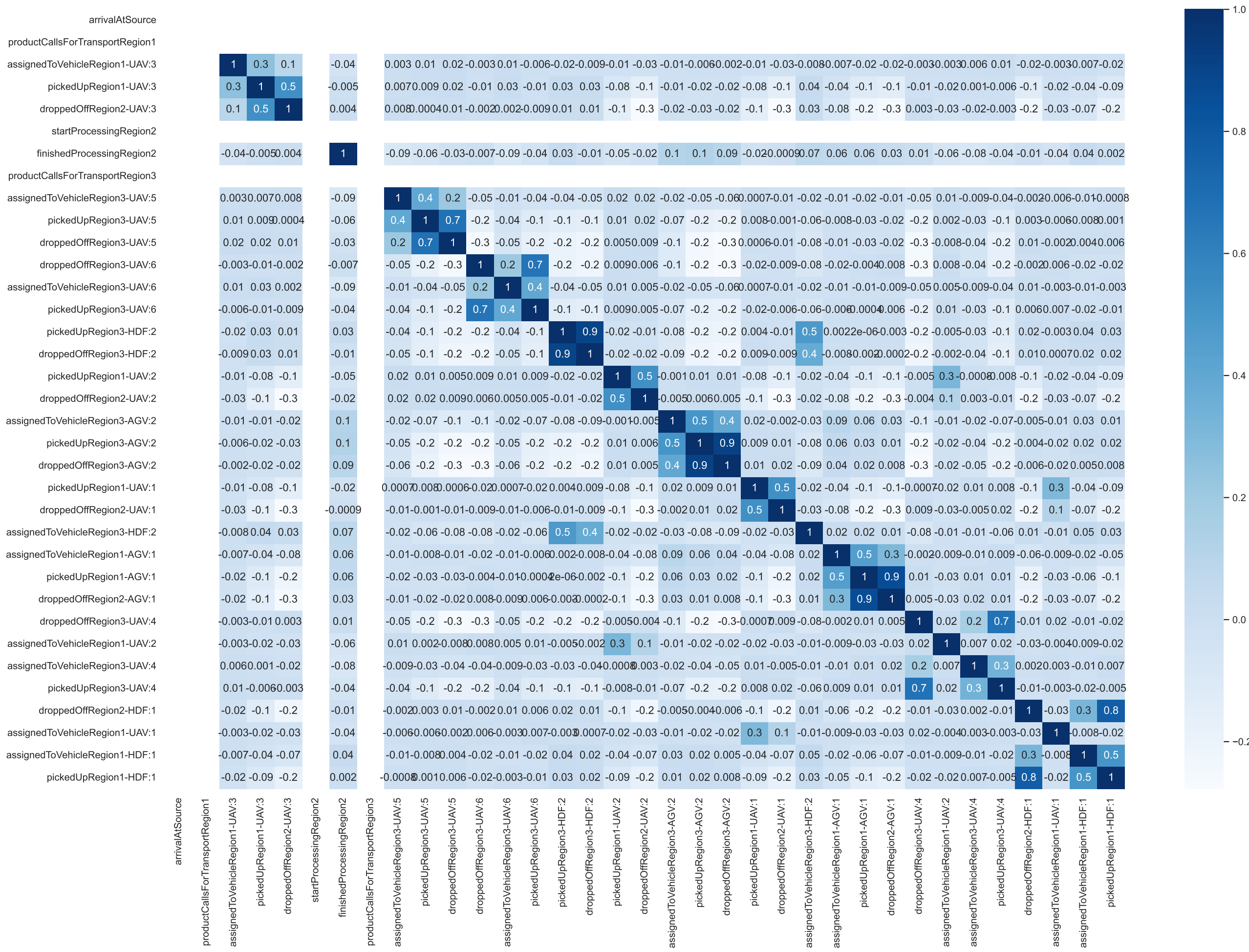

Fig. 3. Illustration of the Pearson correlation heat map, concerning all relevant processes and vehicles. The correlation analysis used the delta time. The sign of the coefficient defines the direction of the relationship. Interpretation of Correlation Coefficient: -1.0 - Perfectly negative, -0.8 - Strongly negative, -0.5 - Moderately negative, -0.2 - Weakly negative, 0.0 - No association, +0.2 - Weakly positive, +0.5 Moderately positive, +0.8 - Strongly positive, +1.0 - Perfectly positive [24]. Note that if a row or a column is empty, then the process is not correlated with other processes.

(ii) The next phase in the decision-making process is to consider potential external influences such as budget, technology, and resources. For this, in-depth research should be done; however, this does not fit into the scope of this case study.

(iii) Finally, a trade-off decision is made on which bottleneck should be outsourced based on the external factors and the outcomes from the bottleneck prioritization matrix.

4.3 Bottleneck Prioritization Matrix

- **Is the bottleneck correlated with another bottleneck?** pickedUpRegion1-AGV:1, droppedOffRegion2-AGV:1, pickedUpRegion3-AGV:2, droppedOffRegion3-AGV:2, pickedUpRegion3-HDF:2, droppedOffRegion3-HDF:2 have a positive correlation of 0.9 and they are assigned with 8 points. assignedToVehicleRegion3-AGV:2 has a correlation of 0.5 and it is assigned with 7 points. droppedOffRegion2-HDF:1 is not correlated with another bottleneck so it is assigned with 6 points.
- **Is the bottleneck correlated with more than one other bottleneck? How many (in total)?** pickedUpRegion3-AGV:2 is correlated with two other bottlenecks, namely assignedToVehicleRegion3-AGV:2 and droppedOffRegion3-AGV:2, while droppedOffRegion2-HDF:1 is not correlated with any other bottlenecks. The rest of the bottlenecks are correlated with only one other bottleneck. The respective assigned points can be seen in Figure 2
- **Does the bottleneck process affect a secondary performance indicator?** The average decay that the bottleneck processes cause can be seen in Table 1 and the respective assigned points can be seen in Figure 2. droppedOffRegion3-AGV:2 affect the product quality

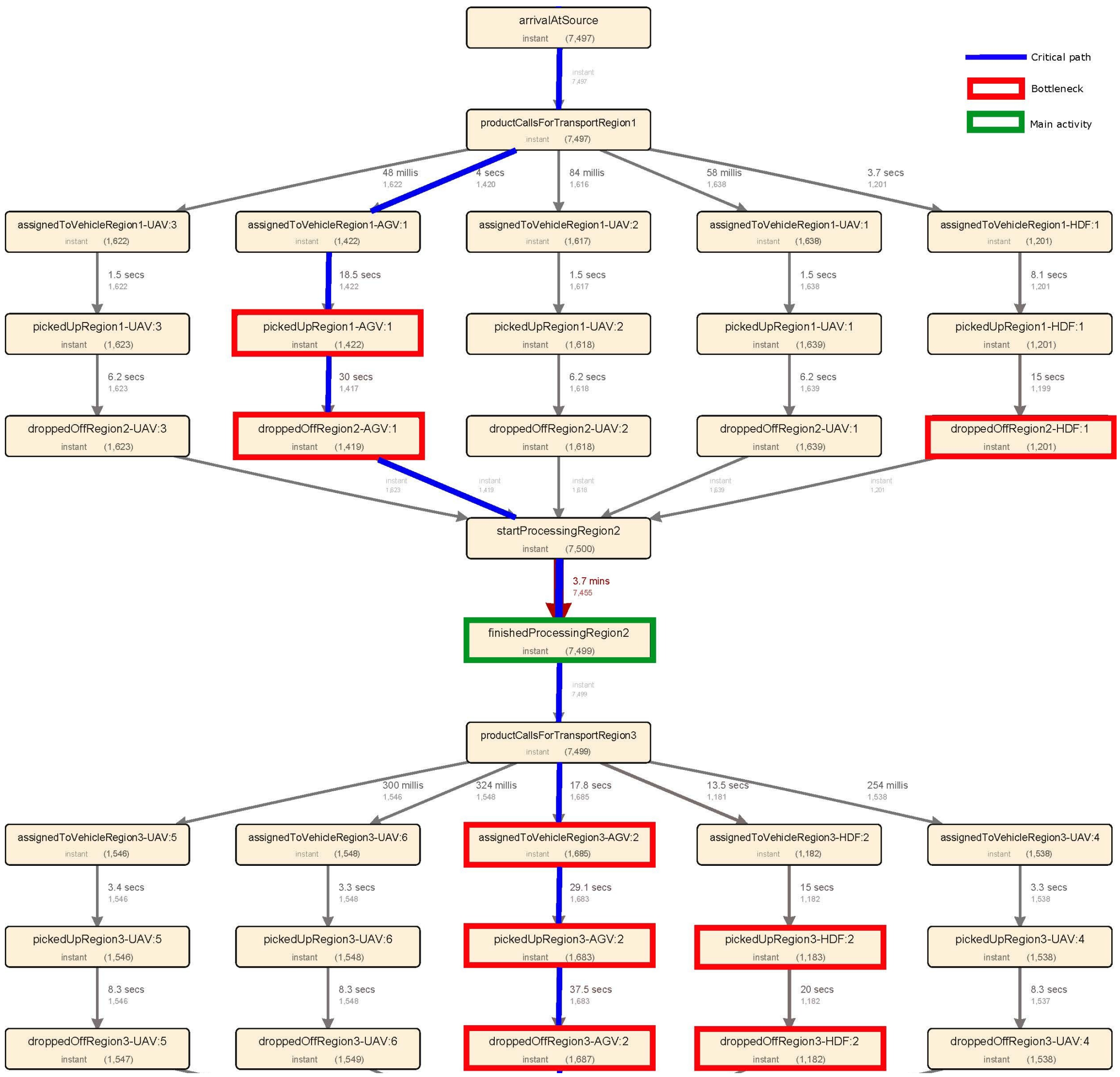

Fig. 4. Illustration of the Process Model, concerning all relevant processes and vehicles. The values indicate the mean duration. The smaller numbers under the mean duration (also the number in the brackets under the processes) indicate the absolute frequency of the processes. The blue line shows the critical path. The red rectangles indicate the bottleneck processes.

the most, so it is assigned with 8 points, while pickedUpRegion3-HDF:2 affects the product quality the less, so it is assigned with 1 point.

- **Does the bottleneck impact the main process?**
  Moreover, it can be seen that neither of the bottlenecks affects the main process, so all bottlenecks get equal points for this criteria.
- **Does the bottleneck have a high frequency?**
  The frequency of the events can be found in Table 1 and the respective assigned points can be seen in Figure 2.
- **Is the bottleneck on the critical path?**
  From Figure 4 it can be seen that only five bottlenecks are part from the critical path namely pickedUpRegion3-AGV:2, droppedOffRegion3-AGV:2, pickedUpRegion1-AGV:1, droppedOff-Region2-AGV:1 and assignedToVehicleRegion3-AGV:2. As a result, they are assigned with 1 point, while droppedOffRegion2-HDF:1, pickedUpRegion3-HDF:2 and droppedOffRegion3-HDF:2 are assigned with 0 points.

In conclusion, droppedOffRegion3-AGV:2 has the highest total point, so it is the first bottleneck to be considered for outsourcing. A possible solution in this case study would be to change all automated guided vehicles 2 (AGV:2) with a faster type of vehicle.

4.4 Reliability of the developed method

With the aim of establishing the reliability of the developed method, a similar experiment was developed concerning the Bemthuis et al. (2020) [4] data set. Under the same constraints and conditions, the data from the tenth scenario was utilized, where three UAVs, two HDFs and two AGVs have been used. The results of identified bottlenecks have been presented in table 2.

Table 2. The obtained results subsequent to testing on the Bemthuis et al. (2020) [4] data set.

№	List of bottlenecks	Points
1	droppedOffRegion3-AGV:4	40
2	droppedOffRegion3-AGV:3	37
3	pickedUpRegion3-AGV:4	35
4	droppedOffRegion2-AGV:1	34
5	droppedOffRegion2-AGV:2	31
6	pickedUpRegion3-AGV:3	31
7	droppedOffRegion3-HDF:3	27
8	droppedOffRegion2-HDF:1	25
9	droppedOffRegion3-HDF:4	25
10	droppedOffRegion2-HDF:2	23

Noticeable from table 2, has been the similarity in identified vehicle type concerning the bottleneck processes, where just as the case with the initial data set, AGVs in region 3 were identified as high priority bottleneck vehicles. The two data sets produced similar results under the same conditions, confirming the reliability of the approach.

5 DISCUSSION

This section addresses the developed approach and its validity, generalizability, applicability, reliability, and scalability. Although the results of the previous section suggest adequate adoption of the proposed method to logistic applications, several implications should be discussed.

First, the **validity** of the approach depends heavily on the quality of the data-set used; this has been due to the fact that data in the real world is not always of favourable quality and quantity. Moreover, the **quality** of the results depends not only on the results of the prioritization matrix but also on the (sensitivity) study parameters defined by the user. For this, it is the responsibility of the user to evaluate the external factors -, identify the (process-specific) constraints -, and set clear boundary conditions for bottlenecks to be successfully outsourced.

The developed method was **generalized** as much as possible, with the aim of handling large amounts of data from different domains. Although the approach did not contain any case-specific parameters, the possibility remains that it is still specific for the used data set. Further examples and more complex applications of the proposed approach should be explored to determine the **scalability** and **applicability** of the method in different domains. Although the case study performed on a simple data set concerning a logistical process provided meaningful findings and insights, the **reliability** of the method was tested by applying it to a similar data set, where the only difference was the total number of vehicles. The reliability of the approach was confirmed by the similarity in the results from the two experiments. Still, more extensive reliability testing could be performed.

6 CONCLUSIONS, LIMITATIONS AND FUTURE WORK

The **objective** of this study has been to propose an approach for minimizing bottleneck processes utilizing process mining and outsourcing techniques. An approach for outsourcing bottlenecks and a bottleneck prioritization matrix have been created. A brief proof-of-concept demonstration, using a data set from logistic processes, has been performed to validate the proposed approach and show how the bottleneck prioritization matrix could be used in practice. A literature review was performed to analyze the state-of-the-art of bottleneck mitigation techniques and define bottleneck prioritization criteria. The review's findings provided information on state-of-the-art of bottleneck mitigation strategies. Most publications have focused on identifying bottlenecks; little study has been done on predicting bottlenecks and recommending actions.

There are several **limitations** regarding this study. First, further validation of the designed approach is required. In this research, a brief case study has been performed; however, a complete study is necessary. Furthermore, the potential incompleteness of the approach might be another limitation. Additional criteria might be added to the bottleneck prioritization matrix. For instance, the utilization of a process can be considered, which might give another perspective on the frequency or the cost of the activities may be included. Moreover, the points distribution in the prioritization matrix and the final score determination could use more testing. A more suitable way for assigning points and calculating the final score could be developed. For instance, it is possible to add weight to the criteria. The goal of this research was to provide an approach to prioritize bottlenecks rather than to offer a definitive approach.

This research leads to several potential **future studies**. To begin, more research may be conducted to verify the suggested approach, for instance, by performing more in-depth case studies to verify the bottleneck prioritization matrix. A data set from logistic processes was utilized in this study. Thus, a possible future study would be to experiment with databases from different industries to get more precise information. Secondly, as the publications regarding predicting bottlenecks are much less than those about detecting bottlenecks, developing an approach that focuses on predicting bottlenecks might be a potential future study.

ACKNOWLEDGMENTS

I would like to express my appreciation to my supervisors dr. Faiza Bukhsh and Rob Bemthuis for guiding me during this research. Additionally, I would like to thank Rogier Anijs for all the help and support throughout this research.

REFERENCES

[1] Wil van der Aalst, Arya Adriansyah, Ana Karla Alves de Medeiros, Franco Arcieri, Thomas Baier, Tobias Blickle, Jagadeesh Chandra Bose, Peter van den Brand, Ronald Brandtjen, Joos Buijs, et al. 2011. Process mining manifesto. In *International conference on business process management*. Springer, 169–194.

[2] Aicha Aguezzoul. 2014. Third-party logistics selection problem: A literature review on criteria and methods. *Omega* 49 (2014), 69–78.

[3] Mohammadreza Akbari. 2018. Logistics outsourcing: a structured literature review. *Benchmarking: An International Journal* (2018).

[4] Rob Bemthuis, Martijn Mes, Maria-Eugenia Iacob, and Paul Havinga. 2020. Using agent-based simulation for emergent behavior detection in cyber-physical systems. In *2020 Winter Simulation Conference (WSC)*. IEEE, 230–241.

[5] Rob Bemthuis., Niels van Slooten., Jeewanie Arachchige., Jean Piest., and Faiza Bukhsh. 2021. A Classification of Process Mining Bottleneck Analysis Techniques for Operational Support. In *Proceedings of the 18th International Conference on e-Business (ICE-B)*. INSTICC, SciTePress, 127–135. https://doi.org/10.5220/0010578601270135

[6] B Blanchard and W Fabrycky. 2008. Reliability-centered maintenance guide: for facilities and collateral equipment.

[7] Gilberto Aleces dos Santos, Luiz Fernando Puttow Southier, and Edson E Scalabrin. 2020. Method to Reduce Lead-time of Business Process discovered. In *2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 840–845.

[8] Cleiton dos Santos Garcia, Alex Meincheim, Elio Ribeiro Faria Junior, Marcelo Rosano Dallagassa, Denise Maria Vecino Sato, Deborah Ribeiro Carvalho, Eduardo Alves Portela Santos, and Edson Emilio Scalabrin. 2019. Process mining techniques and applications–A systematic mapping study. *Expert Systems with Applications* 133 (2019), 260–295.

[9] Shereazad Jimmy Gandhi, Alex Gorod, and Brian Sauser. 2012. Prioritization of outsourcing risks from a systemic perspective. *Strategic Outsourcing: An International Journal* (2012).

[10] Tugba Gurgen Erdogan and Ayca Tarhan. 2018. A goal-driven evaluation method based on process mining for healthcare processes. *Applied Sciences* 8, 6 (2018), 894.

[11] Gwangjin Heo, Jinsung Lee, and Jae-Yoon Jung. 2018. Analyzing bottleneck resource pools of operational process using process mining. *ICIC express letters. Part B, Applications: an international journal of research and surveys* 9, 5 (2018), 437–441.

[12] Olumuyiwa Ibidunmoye, Francisco Hernández-Rodriguez, and Erik Elmroth. 2015. Performance anomaly detection and bottleneck identification. *ACM Computing Surveys (CSUR)* 48, 1 (2015), 1–35.

[13] Zafar Iqbal and Aasim Munir Dad. 2013. Outsourcing: A review of trends, winners & losers and future directions. *International Journal of Business and Social Science* 4, 8 (2013).

[14] Alexander Kinast, Karl F Doerner, and Stefanie Rinderle-Ma. 2022. Combing metaheuristics and process mining: Improving cobot placement in a combined cobot assignment and job shop scheduling problem. *Procedia Computer Science* 200 (2022), 1836–1845.

[15] Tibor Kremic and Oya Tukel. 2006. *Assisting Public Organizations in Their Outsourcing Endeavors: A Decision Support Mode.* Technical Report.

[16] Bojan Krstić and Ernad Kahrović. 2015. Business process outsourcing as a tool for improving enterprise efficiency. *Ekonomika* 61, 3 (2015), 31–41.

[17] Soila Pertet and Priya Narasimhan. 2005. Causes of failure in web applications (cmu-pdl-05-109). *Parallel Data Laboratory* (2005), 48.

[18] Jean Paul Sebastian Piest, Jennifer Alice Cutinha, Rob Henk Bemthuis, and Faiza Allah Bukhsh. 2021. 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*. 290–296.

[19] Christoph Roser, Kai Lorentzen, and Jochen Deuse. 2015. Reliable shop floor bottleneck detection for flow lines through process and inventory observations: the bottleneck walk. *Logistics Research* 8, 1 (2015), 1–9.

[20] Silvia Jaqueline Urrea-Contreras, Brenda L Flores-Rios, María Angélica Astorga-Vargas, and Jorge E Ibarra-Esquer. 2021. Process Mining Perspectives in Software Engineering: A Systematic Literature Review. In *2021 Mexican International Conference on Computer Science (ENC)*. IEEE, 1–8.

[21] J.H. van Midden. 2021. Using process mining and event log analysis for better business strategy decision-making. http://essay.utwente.nl/87378/

[22] Donald Waters. 2021. *Logistics An Introduction to supply chain management*. Palgrave macmillan.

[23] Wenwen Zhu, Stephen CH Ng, Zhiqiang Wang, and Xiande Zhao. 2017. The role of outsourcing management process in improving the effectiveness of logistics outsourcing. *International Journal of Production Economics* 188 (2017), 29–40.

[24] Kelly H Zou, Kemal Tuncali, and Stuart G Silverman. 2003. Correlation and simple linear regression. *Radiology* 227, 3 (2003), 617–628.