DEVELOPING A TOOL THAT CONVERTS A PROCESS MODEL INTO A BPMN DATA MODEL BY MINING FROM MULTIPLE PERSPECTIVES

SÜREÇ MODELİNİ ÇOKLU PERSPEKTİFLE MADENLEYEREK BPMN VERİ MODELİNE ÇEVİREN BİR ARAÇ GELİŞTİRME

MERVE NUR TİFTİK

ASSOC. PROF. DR AYÇA KOLUKISA TARHAN

Supervisor

Submitted to

Graduate School of Science and Engineering of Hacettepe University

as a Partial Fulfillment to the Requirements

for the Award of the Degree of Master of Science

in Computer Engineering

April 2021

To my family…

ABSTRACT

DEVELOPING A TOOL THAT CONVERTS A PROCESS MODEL INTO A BPMN DATA MODEL BY MINING FROM MULTIPLE PERSPECTIVES

Merve Nur TİFTİK

Master of Science, Computer Engineering Department

Supervisor: Assoc. Prof. Dr. Ayça KOLUKISA TARHAN

April 2021, 63 pages

Process mining is mainly focused on process discovery, and also replay techniques to check conformance and analyze bottlenecks. It is further applied to mine the other process perspectives, such as time, data, and resources, by replaying the events in event logs over the initial process model. When process mining is extended far beyond discovering the control-flow models; roles, bottlenecks, amounts of time passed, guards, and routing probabilities in the process can be identified. Such extensions are considered under the topic of multi-perspective process mining, which makes the discovered process model more understandable. In this study, a framework for applying multi-perspective process mining and creating a BPMN process model as the output is introduced. The framework, which uses a recently developed API for storing the BPMN Data Model in blockchain in a secure and immutable way, has been developed as a plugin to the ProM tool. In doing so, it integrates a number of techniques for multi-perspective process mining in literature, for the perspectives of control-flow, data, and resource; and represents a holistic process model by combining the outputs of these in the BPMN Data Model. In this thesis, we explain technical details of the framework and also demonstrate its usage over a case in medical domain.

Anahtar Kelimeler: process mining, multi-perspective, BPMN, process model, block-chain, framework

ÖZET

SÜREÇ MODELİNİ ÇOKLU PERSPEKTİFLE MADENLEYEREK BPMN VERİ MODELİNE ÇEVİREN BİR ARAÇ GELİŞTİRME

Merve Nur TİFTİK

Yüksek Lisans, Bilgisayar Mühendisliği

Danışman: Doç. Dr. Ayça KOLUKISA TARHAN

Nisan 2021, 63 sayfa

Süreç madenciliği esas olarak süreç keşfine ve ayrıca, uygunluğu kontrol etmek ve darboğazları analiz etmek için tekrar tekniklerine odaklanır. Olay günlüklerindeki kayıtları ilk süreç modeli üzerinden tekrarlayarak zaman, veri ve kaynak gibi diğer süreç perspektiflerini keşfetmek için uygulanır. Süreç madenciliği ile kontrol akış perspektifini keşfetmenin yanı sıra; süreçteki roller, darboğazlar, geçen süre, korumalar ve yönlendirme olasılıkları belirlenebilir. Bu yetenekler, keşfedilen süreç modelini daha anlaşılır kılan, çok perspektifli süreç madenciliği kapsamında ele alınmaktadır. Bu çalışmada, çok perspektifli süreç madenciliği uygulamak ve çıktı olarak bir BPMN süreç modeli oluşturmak için bir yazılım çatısı tanıtılmaktadır. Bu çatıda, BPMN Veri Modelini güvenli ve değişmez bir şekilde blok zincirinde depolamak için yakın zamanda geliştirilen bir API kullanılmış ve çatı, ProM aracına bir eklenti olarak geliştirilmiştir. Bunu yaparken kontrol akış, veri ve kaynak perspektifleri için literatürde yer alan, çok perspektifli süreç madenciliği için önerilmiş bir dizi teknik bütünleştirilmiş ve bunların çıktıları BPMN Veri Modeli'nde

birleştirilerek bütünsel bir süreç modeli oluşturulmuştur. Bu tezde, çatının teknik ayrıntıları açıklanmakta ve ayrıca tıbbi alanda bir vaka üzerinden kullanımı gösterilmektedir.

Anahtar Kelimeler: süreç madenciliği, çoklu perspektif, BPMN, süreç modeli, blok zinciri, çatı, araç, eklenti, ProM

ACKNOWLEDGEMENTS

First and foremost, I would like to thank to my supervisors Assoc. Prof. Dr. Ayça Kolukısa Tarhan, Dr. Tuğba Gürgen Erdoğan, and Assist. Prof. Dr. Adnan Özsoy for giving me the opportunity to work together and always leading me to the right path.

Furthermore, I would like to thank my thesis committee members Assoc. Prof. Dr. Oumout Chouseinoglou, Assoc. Prof. Dr. Ayça Kolukısa Tarhan, Assoc. Prof. Dr. Murat Aydos, Assist. Prof. Dr. Ebru Gökalp, Assist. Prof. Dr. Özden Özcan Top for reviewing my thesis and their valuable comments.

I deeply thank, my mother and my father; Sıttıka and Osman, for always carry me forward with their support throughout my educational life.

I would like to thank my older brother and his wife, Hasan Burak and Ayşe Pınar, and my younger brother and his wife, Abdullah Furkan and Merve, who always helped me increase my motivation and provided me with a working environment. I would also like to thank my dear and beautiful nephews Elif Başak, Zeynep Cansu and Cemre who helped me increase my energy in my works.

In addition, I would like to thank my friend Büşra Deniz DUMAN, who always supported me in my thesis work in both library and remote studies, and Sena Hilal ARSLAN, whom I know she is always with me.

Finally, I would like to thank my old workplaces Innova and HAVELSAN and also now ASELSAN for giving me school leave during my thesis work.

CONTENTS

FIGURES

TABLES

ABBREVIATIONS

BIS	Business Information Systems
BPM	Business Process Management
MPCM	Multi-perspective Chain Miner
MPPM	Multi-perspective Process Mining
PM	Process Mining
MPM	Multi-perspective Process Mining
EPC	Engineering, Procurement and Construction
TU/e	Eindhoven University of Technology
ProM	Process Mining Tool
UML	The Unified Modeling Language
XML	Extensible Markup Language

1. INTRODUCTION

Business Process Management (BPM) is the discipline of designing, analyzing, executing and monitoring business processes in order to improve their performance in quality, time and cost dimensions [1]. Process mining is the techniques used primarily in the discovery, conformance checking and enhancement of business processes based on event logs in business information systems (BIS) to support BPM activities. BIS automations keep detailed information about the activities performed during the execution of the business processes they support and the order of these activities in their own databases [1,2]. This data is converted into event logs to start process mining.[3].

Each event log contains the activity that is part of the process. Event logs store information in the database, such as the person or organization that initiated and performed the activity, the start and end time of the activity, and the data element such as age, gender, size, and comments. With this information, it is possible to explore processes from different perspectives such as control-flow, resource, data, function, conformance, and time [4]. Also, there are various discovery algorithms that provide mining results separately for these perspectives. The perspectives, however, are not isolated and are all related to each other. Since the perspectives require different kinds of data to mine as specific to them, the formats of the produced outputs vary, e.g., Petri-net, Heuristic net, Fuzzy model, BPMN, Petri-net with data, and XML. In order to obtain a holistic process model, firstly, the control-flow perspective is discovered by using algorithms such as Inductive Miner [5]. A skeleton model is created by using these algorithms; and other perspectives such as data, resources, and performance, are added to this skeleton model. The results of the mined perspectives are represented in Petri-net [6] or BPMN [7]. Both notations have token-based semantics, and allow one to integrate the outputs of the multiple perspectives mined into a single process model. However, as different from Petri-net, BPMN is able to hold the output of the resource perspective easily due to its elements to represent different roles; and this makes it an attractive notation for process mining experts.

Based on the motivation stated above, in this thesis, we introduce a framework for applying multi-perspective process mining and creating a BPMN process model as an output that is kept in blockchain. To the best of our knowledge, this is the first study that provides an extended representation of BPMN for multi-perspective process mining, and keeps it in a secure and immutable way. More specifically, the main contributions of this research are:

(1) Proposing a plugin [8] to ProM [9] for multi-perspective process mining;

(2) Combining distinct outputs of control-flow, data, and resource perspectives in a single, multi-perspective process model in BPMN;

(3) Enriching the representation of Data Petri-net produced by the data perspective with percentage of transition (relative frequency), average time and instance count in-formation;

(4) Using a BPMN Data Model, which has been developed recently [10], in order to keep the multi-perspective process model in the blockchain in a secure and immutable way;

(5) Validating the correctness of the multi-perspective process mining proposed by the framework over a pilot study in medical domain.

In this thesis, process mining algorithms are examined in detail. Then these algorithms are grouped according to the perspectives discussed. At this stage, the inputs and outputs of different perspective algorithms were examined and the usability of these algorithms with each other was studied. When these different perspective algorithms are used together, it is aimed to examine the behaviors and show the results on a single process model. According to the results obtained from matching algorithms and perspectives, it is aimed to develop a tool to apply different perspective algorithms one after the other, and to store and display the resulting output (multi-perspective process data) in BPMN format.

As a research approach, a path from general to specific has been followed as described below. After examining the existing methods, it is aimed to produce an integrated new solution. Improvements have been made to the existing algorithms.

1. Investigation of process mining algorithms.
2. Investigation of multi perspective process mining algorithms.
3. Grouping process mining algorithms according to perspectives.

4. Application of different perspectives with each other.
5. Tool development on displaying different perspective algorithms applied on the event log in a single model.
6. Making improvements on the already existing algorithms.

The remaining of this thesis is organized as follows: Section 2 gives general information about process mining domain and Section 3 summarizes the related studies. Section 4 explains briefly the BPMN Data Model. Section 5 describes the proposed framework in detail. Section 6 presents the example application of the framework over a pilot study. Section 7 presents validation of the framework over the pilot study. Section 8 summarizes validation results of the framework. Finally, Section 9 concludes the thesis by providing overall results and suggestions for future work.

2. BACKGROUND INFORMATION

2.1. Workflow Management

Workflow Management (WFM) mainly focuses on business process automation. It enables users in the same or different disciplines to send information, request a job, predict and manage results. Thanks to workflow management systems, companies can follow all their business data and resources in digital environments and get efficient outputs. Ultimately, workflow management is for achieving a result [11].

2.2. Business Process Management

Business Process Management (BPM) is a discipline that combines information technologies and management science knowledge and applies it to operational business processes. We can also say that it is the extended version of WFM for BPM.

BPM has a wide scope. This area includes from process automation and process analysis to process management and business organization. From another point of view, BPM is related to software management, control, and operational support processes. BPM aims to improve operational business processes and tries to do this without using new technologies [11].

2.3. Data Science and Process Mining

Although data science is a term that we have heard a lot in recent years, it has become a discipline that is very beneficial to the business world. This discipline, which gains more importance with the increase of data, is frequently used to provide competitive power in the business world. Process mining, as a sub-field of data science, is based on the databases of the records of the transactions performed in the information systems and the event recording logs kept in those databases. As it is understood from here, the raw material of process mining is the process data recorded instantly. Extracting and analyzing these process data records from data warehouses are included in the “data science”.

Software applications used in process mining have Import / Export options. However, if you want to customize the data you want to use, that is, if you want to adjust it to cover the parameters you want, the concept of “data mining” is included here. Data mining allows you to customize the data of the process you want to examine [11].

Machine Learning has lately become a very popular field of study under “data science”. Algorithms are at the heart of machine learning. The part of machine learning related to process mining is that predictions can be made with the help of the algorithms used and the “discovery” function, one of the three main functions of process mining, can be used. These algorithms make process discovery very efficient. For example, the alpha-algorithm [12] is an algorithm used in process mining that aims to reconstruct causality from a series of events. [11]

2.4. Process Mining

Process mining acts as a bridge between model-based process representations and data-centered analysis techniques in business process management. Process mining aims to discover meaningful information about ongoing processes based on event records to explore processes, control data, institution and social structure.

Process mining is a technology that takes inputs the disciplines mentioned above and transforms these inputs into positive outputs. With process mining, companies discover the processes they use in detail, create new process models, and improve their existing processes. In this way, they have the capacity to increase their efficiency and profitability.

In order to understand process mining, it is necessary to grasp the subject “Business Process Management Life Cycle” as a preliminary step. The life cycle concept here describes the different stages of managing a particular business process [11].

The Business Process Management Life Cycle consists of 6 stages, as follows [11]:

1. *Design :* It covers the design of the process. This model then transforms into a configuration and implementation model.
2. *Configuration / Implementation :* If the model has already been run and is running on WFM or BPM systems, this step takes less time. However, if the model is ready on paper like in informal and traditional projects, the time taken for this step will increase.
3. *Enachement / Monitoring :* This stage is after the processes have been designed and supported by systems. In this section, observation is carried out while the processes continue to run. The needs for change are aimed to be determined.
4. *Adjustment :* Some of the planned changes are made at this stage. However, redesign of processes is not applied or new software is not integrated. Only pre-designed and defined settings are performed.
5. *Diagnosis / Requirements :* In this part, processes and urgent requirements are aimed to be determined. Changing new conditions, their analysis and evaluation are carried out (e.g., changing policies, rules, laws, etc.).

As the days go on, hours or even minutes pass, hundreds of data are recorded in information system databases. These data are usually generated as a result of an activity. These activities contain information such as the person performing the activity, and the timestamp when the activity has occurred. These data are generated by certain activities, by certain people, in a certain order. After this amount of data is created, it is possible to verify and improve whether the running process is working correctly or not with process mining. In fact, process mining allows us to view and improve the working process model by revealing the mystery behind the processing of data in the database. In this way, institutions get the opportunity to evaluate the process that is currently running. Thanks to process mining, inconsistencies that occur in the process are learned and precautions are taken beforehand by making improvements in the process. Using process mining benefits organizations both in terms of time and improvement of the running process. It will be extremely beneficial for the company to discover the defects or bottlenecks in the process beforehand. It will also save the company time to take action. For this reason, while there is so much data and processing, the use of process mining is important for companies.

Process Mining (PM) is a process management technique that has been used extensively for BPM in the last decade by combining data-driven techniques with traditional process modeling techniques [1,13]. Process mining is applied for three main aims [14]: i) process discovery, ii) conformance checking, and iii) process enhancement.

Table 2.1 contains the basic questions and answers that can be asked on process mining. Thanks to this table, a quick overview of process mining can be made and the information to be obtained in this area is collected in a single table [11].

The most important activity of process mining is process discovery as it provides the basis for conformance and enhancement. During process discovery, event log is used as an input to process discovery algorithms and thus the process model is created automatically [14]. In this way, organizations can improve their processes due to the discovery of the differences between the real process model and the expected process model. To summarize, the discovery technique takes an event log and is the section where a model is tried to be created without using any prior knowledge. For example, it could be the alpha algorithm [12]. Alpha algorithm takes an event log as a input and creates Petri-nets from this log. This helps us understand the behavioral state of logging of the network in event log.

In conformance checking, the process model and its flow discovered from the event logs are analyzed and it is checked whether the process has been carried out as identified in the model [14]. Conformance checking measures the differences between the actual process and the process model specifications. The main aim of this technique is to identify the areas that need improvement using the information gained from the actual process [15]. To summarize, conformance is the second major scenario of process mining. Here, an existing transaction model is compared to an event log of the same process. The conformity check is carried out to check whether the logged reality is suitable for the model. It is also the opposite. Conformity, bias control, detection, and explanation of deviations can also be used to measure the severity of these deviations.

Table 2.1. Questions and Answers for Process Mining

QUESTION	ANSWER
What are the three maximum qualities required for process mining?	Event Log Case-Event ID Time Stamp
How many stages does Process Mining consist of?	Process Discovery: What is really happening? Conformity Check: We apply the accepted ones or are there any things that need to be added? Performance Analysis: Where are the bottlenecks? Transaction Forecast: Will the event be delayed? Process Development: How can we redesign these processes?
What do we discover?	Discovery of process models (Petri nets, EPCs, BPMN), Organizational Models, Social networks, Sequence diagrams, Business Rules, Bottlenecks, Simulation Models
What are Process Discovery Techniques?	Algorithmic Techniques Genetic Process Mining Zone Based Process Mining
How many basic principles is process mining based on?	Workflow Management Business Process Management Data Science Machine Learning

Process enhancement is the improvement of the process model by considering the event logs. This can be accomplished by adding more perspectives to the process model using event data, an activity also known as ‘extension’. Another enhancement is improving the quality of the process model using event data and defining a new repaired model [16].

The main idea is to extend or improve the current process model using information about the actual process recorded in the event log. While the conformity check measures model and reality fit, the third function of process mining, “enhancement and improvement”, aims to expand and change the possible model.

Enhancement and improvement consists of two parts:

- ***Repair:*** It tries to make the current model better and more reflective. For example, if two activities are modeled sequentially, but in reality, these operations can be performed in any order without any problems, the model can be repaired to correct this complexity.
- ***Extension:*** It is the process of creating new perspectives by applying cross-correlation [1] with event log records to process models. For extension, “performance data” can be given as an example. By using the time stamps in the event log, the expansion process can be applied to show the bottlenecks and service levels, times, and frequencies. Additionally, Control-flow Perspective, Organizational Perspective, Event Perspective, and Time Perspective can be added in this activity.

2.5. Multi-perspective Process Mining

In the multi perspective-process mining (MPM) technique, there are many perspectives; e.g., control-flow, resource, data, time, and function [14]. The control-flow perspective focuses on the sequence of activities and discovery of the process model, and aims to find the best process model for all possible paths with its different algorithms. The data perspective enables the analysis of the data that has an impact on the formation of the activities in the process model. The case perspective focuses on the definition of cases and the factors affecting the data. The resource / organizational perspective is concerned with actors such as people, systems, and roles and the relationships between them to classify the activities in the process model according to roles and organization. The time perspective deals with the time and frequency of activities occurring and helps to identify bottlenecks, monitor the use of resources, and estimate the time remaining for ongoing events.

The performance perspective allows to observe the performance problems in the process model to be exposed, and can be based on time information as in the time perspective.

Further details are provided in the following paragraphs with regard to the perspectives used in this thesis.

- ***Control-flow perspective:*** Its name is also known as process perspective. In this perspective, the hidden workflow in the event log is expressed through notations such as Petri-net [6], BPMN [7], Heuristic net by using process mining discovery algorithms. Alpha Miners [12], Inductive Miner [17], and Heuristics Miner [2] can be given as examples to the algorithms used in this perspective.
- ***Data perspective:*** It plays a role in showing the conditions (data guards) in transitions on the discovered control-flow output through notations such as Data Petri-net or BPMN. In discovering this perspective, Data Aware Discovery algorithm [18] can be used.
- ***Resource/organizational perspective***: If there are roles performing activities in the event log, one can use this perspective in the discovery phase. Role-activity matching can be accomplished through the Organizational Miner algorithm [9], which can be exported as XML output. Social Network Analysis plugin [19] can also be used for this perspective, but the resulting output cannot be exported except for pdf and image display.
- ***Time perspective:*** If there is timestamp data in the event log, this perspective becomes available. In this way, bottlenecks can be discovered and resource usage can be measured. Thus, possible future problems in the use of resources are determined in advance and precautions are taken. While discovering this perspective, Dotted Chart [20] algorithm can be used.
- ***Performance perspective:*** It is the perspective that is effective in discovering the performance problems in the business process. With this perspective, it can be determined how well the performance of the business process is and measures can be taken in advance for future problems with performance measurement. While discovering this perspective, Multi-perspective Process Explorer [21] can be used.

To bring all process mining perspectives together, the flow is started with the control-flow perspective which finds all possible paths of activities from the event log in the discovery step [13]. Then the generated process model is checked for whether the reality conforms to the model in the conformance checking step. Finally, additional perspectives are applied to the process model in the enhancement step [22].

Various software products have process mining capabilities [23], e.g., ProM (TU/e) [9], Disco (Fluxicon) [24], PALIA-ER [25], CELONIS [26], pMineR [27] and bupaR [28]. ProM is the most comprehensive PM tool, which provides a standard environment incorporating a generic open-source framework for implementing process mining techniques. Although its user interface is difficult to understand for beginners in comparison to other tools, it is the most preferred tool especially by the researchers due to its large number of plugins [29]. Also, it supports several multi-perspective PM plugins [4,30]. Disco is a popular process mining toolkit, which is powerful, easy-to-use, and fast. The revolutionary commercial process mining technology in Disco helps researchers to create visual maps and can run case and time perspectives together. On the other hand, bupaR is an R-library tool, and it also allows to combine different profiles of the process with information about frequencies and performance.

2.6. Process Mining Event Log

In order for process mining algorithms to be implemented, the data must contain some information in certain formats. This data type is called “event log”. Before applying algorithms in process mining, some preliminary processes should be performed on the online or offline data already created in the database and event logs should be created. Event log consists of events, and it contains information such as case id, activity, time, and resource in the events. For each case, the activities performed by the case existing in the database can be followed with the process model. With the time information, the events in the process are sorted. If there are different perspectives in the process model, this information should also be included in the event log. For example, model branches can be made more meaningful with diagnostic data in a hospital for data perspective. With the time information, the average duration of each transition in the process model can be shown. Alternatively, with resource information, people who perform the activities can be included in the process model. For this reason, by storing such information in the event log, the process can be examined from different perspectives.

2.7. Petri-net vs BPMN

Petri-net [6] offers a graphical notation for step-by-step processes that require selection, iteration, and concurrent operation, similar to industry standards such as UML [31], BPMN [7], and EPC [32]. But beyond these standards, it can model the operation of processes related to the mathematical theory developed for process analysis with mathematical precision. BPMN is typically designed to provide a standard notation that can be easily understood by all business stakeholders. Therefore, BPMN helps prevent communication gaps that may arise between the stakeholders in a project that adopts a common language to explain business processes. A well-modelled business process in BPMN can be understood by people who do not know BPMN at all, because BPMN 2.0 [33] is a high-level notation. In other words, a process model in BPMN can be thought of as "high level" representation of the process, which can be used to display a low level, mathematically formalized process description in Petri-net [6]. A comparison of representations in Petri-net and BPMN is given in Table 2.2.

Table 2.2. Petri-net vs BPMN

Petri-net	**BPMN**
Low-level	High-level
Domain: Computer Science	Domain: Business Process Management
Mathematically formalized	Formalized in XML

Petri-net output may be sufficient for control-flow representation of a process. It is also possible to convert it to BPMN if a higher-level representation is desired. In this way, a more understandable demonstration is provided. For the data perspective, the Data Petri-net output is obtained. Since this output is placed on the skeleton model that has already been discovered with control-flow perspective; the more comprehensible the Control-flow representation is, the more comprehensible the data perspective becomes. For the resource perspective, pools are created by establishing an activity and originators relationship through BPMN lanes. It is typically not preferred to represent the resource perspective in Petri-net. As a result, BPMN becomes an attractive notation for process mining experts. BPMN can be used for MPM process discovery solutions since it enables to include the outputs of control-flow, data, resource, performance, and time perspectives in a single process model. An integrated multi-perspective solution can be obtained by converting these outputs to a single BPMN model by running different perspective algorithms/methods.

2.8. Quality Dimensions for Process Discovery Algorithms and Outputs

A process model created by process discovery step of PM should describe the observed behavior and not explain more of that observed behavior. In addition, this model should contain unobserved but possible behavior without unnecessary complexity. There are four dimensions that show the quality of the process model created by process discovery: fitness, simplicity, precision, and generalization [34]. There are also trade-offs among these dimensions, and this is enough to make process discovery a challenging task.

When process discovery algorithms are grouped according to the quality dimensions [34] as shown in Table 2.3, it is seen that the models produced by the algorithms used in process discovery can be represented in different notations such as Petri-net, EPC, BPMN, and Heuristic Net. The process model that is created by process discovery algorithms can be addressed to one of four situations [14]: (1) the model can only contain frequent behaviors, which is the most ideal process model, (2) it may include some of the frequent and infrequent behaviors, which is the non-fitting model, (3) it includes only frequent and infrequent behaviors, which is the overfitting model and cannot cover new situations, or (4) it contains all behaviors, including the uncommon behavior, which would be an underfitting model.

The ideal representation of the discovered process model with correct representational bias is as important as the comprehensibility of the process model. Representational bias defines the search space by having the ability to represent concurrency, loops, silent actions, duplicate actions, OR-splits/joins, non-free-choice behavior and hierarchy [19]. It helps capture the quality and ideal representation for discovered model of the process while exploring the process in the event log. Each process discovery method should have its own most effective representational bias. Because each notation has its own capabilities, and thanks to the correct algorithm and notation matching, the most ideal process models can be obtained in a way that ideally covers the event logs. In other words, representation of the process with different notations and representational bias are not the same thing. When the relevant process model is not represented with its ideal notation, this may lead to some losses.

Table 2.3. Quality dimensions of process discovery algorithms and notations used for their outputs [34]

Quality Dimension	Algorithm	Output
Replay Fitness	Region Based Algorithms Multi-phase Miner	Petri-net Event-driven process chain
Simplicity	α-algorithms	Petri-net
Precision	Region Based Algorithms	Petri-net
Generalization	Fuzzy Miner Heuristics Miner	Fuzzy Model Heuristic Net

3. RELATED WORK

In this section, the literature is examined by focusing on the perspectives of process discovery algorithms, and then the existing multi-perspective process mining studies are examined in terms of the perspectives they address. Since the process model created as a result of process discovery is based on the control-flow perspective, there are a lot of studies on mining control-flow perspective in the literature. Dakic et al. [35] report that 94% of the PM studies are based on the control-flow perspective. The examination of the most widely used control-flow algorithms can be started with the alpha algorithm proposed by Aalst et al. [12]. Mohamed [19] mentions that this algorithm fills the gap between the event log and the process model but it has a flaw in producing a reliable model since it ignores frequencies, and therefore, does not guarantee completeness and soundness of the model. In the same article, the Heuristic Miner algorithm proposed by Weijters et al. [2], which can deal with noise, is mentioned as an improvement of the alpha algorithm. It is also mentioned that Heuristic Miner algorithm can filter noisy or infrequent behaviors because it takes frequencies into account, but it remains short too in guaranteeing sound process models. Another control-flow perspective study by Aalst et al. [36] presents a new approach based on genetic algorithms. With this approach, noise and deficiency are dealt with the genetic process mining approach using causal-matrix. Weber et al. [37] states that existing algorithms currently used in the control-flow perspective in general does not discover the probabilistic nature of the process. In this study, a framework is proposed and applied on the alpha algorithm, and it is shown how much data is required to achieve a certain accuracy by using the structures in the model. In another study, Ghawi [38] mentions that the models formed as a result of the Inductive Miner algorithm developed by Leemans et al. [17] correspond to sound, block-structured workflow (WF) net systems, and they always fit.

As the control-flow perspective studies have been emerging, Rozinat et al. [39] have conducted a study on the data perspective by using machine learning techniques, data dependencies, and decision-point analysis. Following that, DecisionMiner algorithm has been implemented in ProM. The article by Leoni et al. [18] is based on the fact that the algorithms used for the control-flow perspective have discovered the decision points, but not because they have discovered the rules required for the data-flow. With the latest developments in conformance-checking, it is mentioned in the study that data-flow could also be discovered and added to the process model with an adaptation process.

While examining other perspectives in process discovery, it is seen that the perspectives such as organizational, time and performance are also included in the studies. In Table-3.1 under this title, studies in the literature on process discovery are marked according to their mined perspectives. Mans et al. [40] mention that PM techniques can be used to determine patient care-flow, and the health process has been examined from different perspectives such as control, organizational and performance. Bozkaya et al. [41] propose a method on process discovery, where the process has been examined from various perspectives such as control-flow, performance (Dotted Chart Analysis [20]), and organizational (Organizational Miner and Social Network Analysis [19]. Aalst [42] explains that it is possible to apply the control-flow perspective by converting the logs obtained when web services are communicating with each other into event logs, and that different perspectives can be added to the model by replaying event logs. This study gives information about the process mining manifesto and provides information about the situations that may be encountered while doing service mining. Gupta [43] has studied about the comparison and operation of PM algorithms in order to guide the companies in choosing the most suitable algorithm for their needs. Mohamed [19] mentions in his study that the process model can be expanded by using Social Network analysis to show the organizational perspective.

When the related studies in Table-3.1 are examined, it can be seen which perspective is used in which study. With this information, a road map has been drawn about which algorithm could be used for which perspective. Therefore, the review summarized in the table has helped in selecting algorithms that explore different perspectives that can be used for MPM. It is also seen from the related literature that the methods used for process discovery are based on the control-flow perspective. In other words, the process model is first discovered using the control-flow perspective, then other perspectives are added on this process model.

Table 3.1. Process Mining perspectives in scientific literature

Related Study	**Year**	**Control-Flow**	**Data**	**Performance/ Time**	**Resource/ Organiza-tional**
[36]	2005	+			
[39]	2006		+		
[40]	2008	+		+	+
[44]	2009	+		+	+
[37]	2011	+			
[42]	2013	+	+	+	+
[18]	2013	+	+		
[43]	2015	+			
[19]	2016	+	+	+	+

In our study, the algorithms used in developing the MPM framework have been selected from the studies in Table-3.1 by considering their advantages and drawbacks. As an example, since the Alpha algorithm does not address frequencies, it would be more advantageous to use the Heuristic Miner algorithm, which is an improved version of the Alpha algorithm and addresses the frequencies. As a result, different perspective algorithms have been integrated to provide a single, multi-perspective process mining solution, with some additional features such as data guards with average time, percentage of transition, and instance count information. More specifically; percentage, average time, and instance count properties have been added to the branches in the process model formed by the Data Aware Explorer algorithm in the data perspective. If the user wishes, s/he will be able to use this feature and display these properties on Petri-net or BPMN diagrams; that is, this part is left optional.

After examining the studies that mine different perspectives on the process model, multi-perspective process mining studies have also been examined. The studies in the literature on MPM are marked according to their included perspectives in Table-3.2.

Table 3.2. Multi-perspective Process Mining in scientific literature

Related Study	Year	Control-Flow	Data	Performance / Time	Resource / Organizational	Functional
[21]	2015	+	+	+		
[50]	2015	+		+	+	
[51]	2016	+	+	+	+	
[4]	2018	+	+	+	+	+
[30]	2019	+	+		+	
[48]	2019	+	+		+	
[49]	2019	+			+	

Folino et al. [45] propose a new approach that can be used while exploring the process model. In this approach, they mention the control-flow perspective as structural, and other data such as activity executors, parameter values, and time-stamps, as non-structural. It is also mentioned in the study that the process would be explored by characterizing it from structural and non-structural viewpoints. With this approach, the process could be discovered in a multi-perspective way. Mannhardt et al. [21] define the Multi-perspective Process Explorer (MPE) tool in which they have integrated already existing PM techniques such as data-aware discovery, conformance checking, and performance analysis. Schönig et al. [46] propose a framework to discover Multi-Perspective Declare Models, where a process is analysed from different perspectives such as control-flow, data, time, and organizational. Sturm et al. [47] offer a distributed mining framework called Multi-perspective DECLARE, which includes an efficient big data technology for MP-DECLARE models. Mannhardt [4] proposes a multi-perspective process mining method considering problems in multiple interacting process perspectives of control-flow, data, resources, time, and function. He makes two contributions regarding multi-perspective process discovery: Data-aware heuristic process discovery and the Guided Process Discovery. Jablonski et al. [48] develop multi-perspective process trace clustering approach with tool support, which enables reducing the complexity of mined process models in terms of improving the homogeneity of trace subsets. Kalenkova et al. [30] integrate various process mining algorithms for the discovery of multi-perspective hierarchical BPMN models, and bridge the gap between the PM techniques and BPMN tools. Sikal et al. [49] propose a variability discovery approach for resource perspective in configurable processes, and use the dependencies between the activity and the resource variability. Pini et al. [50] present a multi-perspective visualization framework to compare processes. In this article, the authors propose a design approach that allows to address the multiple perspectives like control-flow, time, performance, and resource.

From the studies in Table-3.2, multi-perspective PM solutions in the literature can be seen. Considering the integrated solutions, in our study, the alternatives for mining the control-flow perspective are presented in the MPM framework and this way, the opportunity to experiment and work with the algorithm most suitable for the dataset is provided. In addition, perspective selection is enabled prior to representing the process model in BPMN. As also different from the studies in this table, additional features have been added to represent the data perspective and the performance perspective (including information of percentage, average time, instance count) together and separately in Petri-net or BPMN.

When the existing studies on multi-perspective process mining are examined, it is seen that there is no solution that includes the perspectives of control-flow, data with performance information, and resource, which is applied directly on the event logs to create a process model in BPMN. Therefore, the proposed framework contributes to the body of knowledge and practice by addressing these features. In addition, the BPMN model is kept in blockchain in our solution via BPMN Data Model [52], which can be considered a timely and motivating contribution for the following similar proposals and also for future MPM applications.

In the light of the studies examined within the scope of this thesis, the perspectives studied in the literature were analyzed and it was found appropriate to work with control-flow, data (with average time and performance (percentage)), and resource perspectives.

4. BPMN DATA MODEL IN BLOCKCHAIN FOR MULTI-PERSPECTIVE PROCESS MINING

The Data Model has been proposed to keep the outputs from the low-level process models, which are produced via multi-perspective process mining techniques, together in a new model that is extended from BPMN [10,52]. Since each technique used in an additional perspective produces a different low-level model, an operation to modify the results of analysis methods to proper forms and to transfer them to the Data Model is required. Accordingly, three basic perspectives (i.e., control-flow, data, and resource) are applied one by one, and the Data Model is enriched during this process by creating the necessary elements that correspond to what the perspectives describe. What those elements are and how they are determined are explained below in detail as specific to each perspective.

The construction process of the Data Model includes the activities of log creation, log preparation [53], control-flow analysis, decision analysis, and role analysis, as shown in Figure 4.1. In control-flow analysis, a process model is discovered based on pre-processed event logs and is produced in a Petri-net file as an outcome. This is the first type of process mining analysis where control-flow is used as backbone, and the other elements will be attached to the control-flow backbone. This Petri-net file is used as an input in addition to the event logs, in order to discover process model from data or resource perspectives. While creating the process skeleton, the Inductive Miner algorithm [17] is used as control-flow algorithm. In data perspective, Decision Miner algorithm [39], which extracts the knowledge about decision rules, is applied for decision analysis. In resource perspective, Organizational Miner Algorithm, which derives the organizational concepts of the process, is used for role analysis. After these steps, the Data Model gets the last and most meaningful form which is capable to describe what the actual process looks like, i.e., representing which characteristics or attributes belong to case influence decisions in the process, who executes or performs activities, and what the organizational roles are.

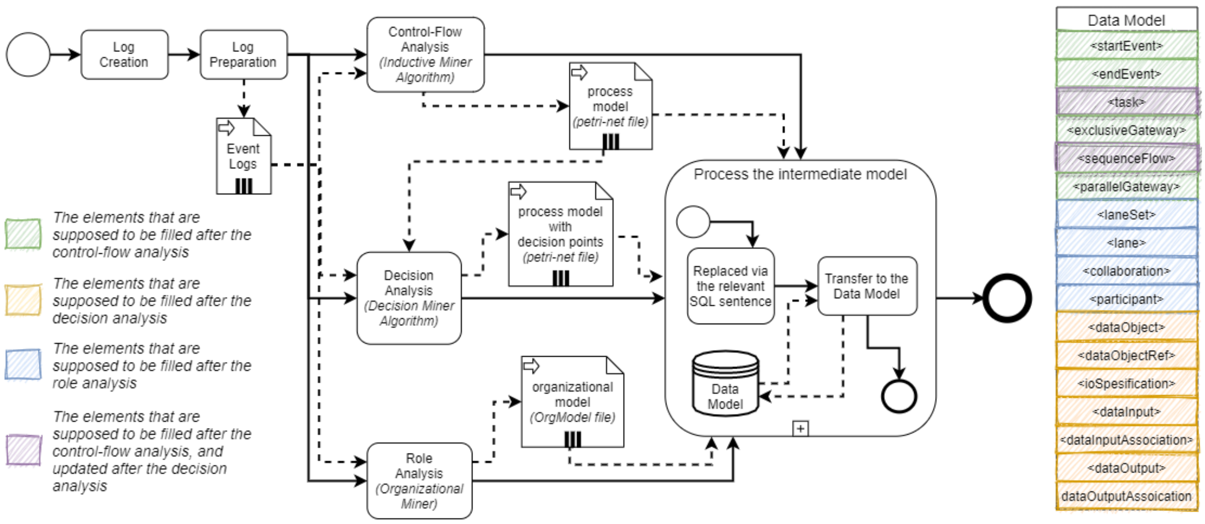


Figure 4.1. The construction process of BPMN Data Model [10]

In Figure 4.1, how these perspectives work together and which elements are taken from the relevant analyses into the Data Model are also illustrated. After control-flow analysis, startEvent, endEvent, task, exclusiveGateway, parallelGateway, and sequenceFlow elements are filled in the Data Model. After data analysis, the dataObject, dataObjectRef, ioSpesification, dataInput, dataInputAssociation, dataOutput, dataOutputAssociation elements are transferred to the Data Model. Aside from these elements, the task and sequenceFlow elements are updated in this analysis. After resource analysis, the laneSet, lane, collaboration, and participant elements are filled in the Data Model. Further information on the Data Model and data transformations can be obtained from [52].

5. FRAMEWORK FOR MULTI-PERSPECTIVE PROCESS MINING

This section introduces the framework presented in this thesis. It provides an overview about supported PM perspectives and algorithms, then shows the details of plugin functionalities to explain its usage.

This integrated discovery approach is implemented as a plugin to ProM, the most widely used open-source process mining framework. As a pre-requisite, BPMN Data Model API must be deployed to Docker before running this plugin. Our plugin is available in a new package called Multi-Perspective Chain Miner (MPCM) which is adaptable to ProM. It is aimed to contribute to the body of solutions for multi-perspective process mining. With this solution, the process model mined from the event logs for multiple perspectives can be integrated and stored in a single BPMN-extended data file in blockchain, and be visualized including the information of control-flow, resource, data, and performance. While developing MPCM, perspective algorithms used in process mining have been examined, as reported in Section 3. In the light of these algorithms, it has been investigated how these perspectives could be integrated with each other. In this context, the algorithms supported by the framework are given in Figure 5.1.

As shown in Figure 5.1, within the scope of the MPCM, the process is mainly analyzed from three perspectives. These are the control-flow, data, and resource perspectives. From the control-flow perspective, the skeleton of a process model can be generated using different algorithms/plugins, namely the Alpha algorithms [12], IMPetriNet [5], HybridILPMinerPlugin [54], FlexibleHeuristicsMinerPlugin [55], CNetMiner [56], IMProcessTree [5], and FlowerMiner [57]. User can experiment how the outputs of these algorithms are integrated with the outputs of other perspectives by selecting and trying different control-flow algorithms. The process skeleton discovered by the control-flow perspective can be extended with the data or resource perspective. Data perspective helps to understand the data inputs to the branches in the process skeleton. For the data perspective, Data Aware Explorer plugin [18] is run and the process skeleton is enriched with information such as average time, percentage of cases that followed the transition, and instance count. In other words, this information can be added to the process model while mining from the data perspective, if selected by the user. The resource perspective that allows to see the roles who perform the activities can also be added to the process skeleton (or to the skeleton previously enriched by the data perspective) by executing Organizational Miner Algorithm [19].

In the control-flow analysis, the output of the mining is converted to a PNML file, while a DataPNML file is created in the data perspective analysis. In the resource perspective analysis, the OrgMiner file, in which activate-resource mapping is made, is created as an XML file. These files are used to obtain a BPMN diagram (e.g., the one shown in Figure 6.7) which puts together the outputs of the control-flow, data, and resource perspectives.

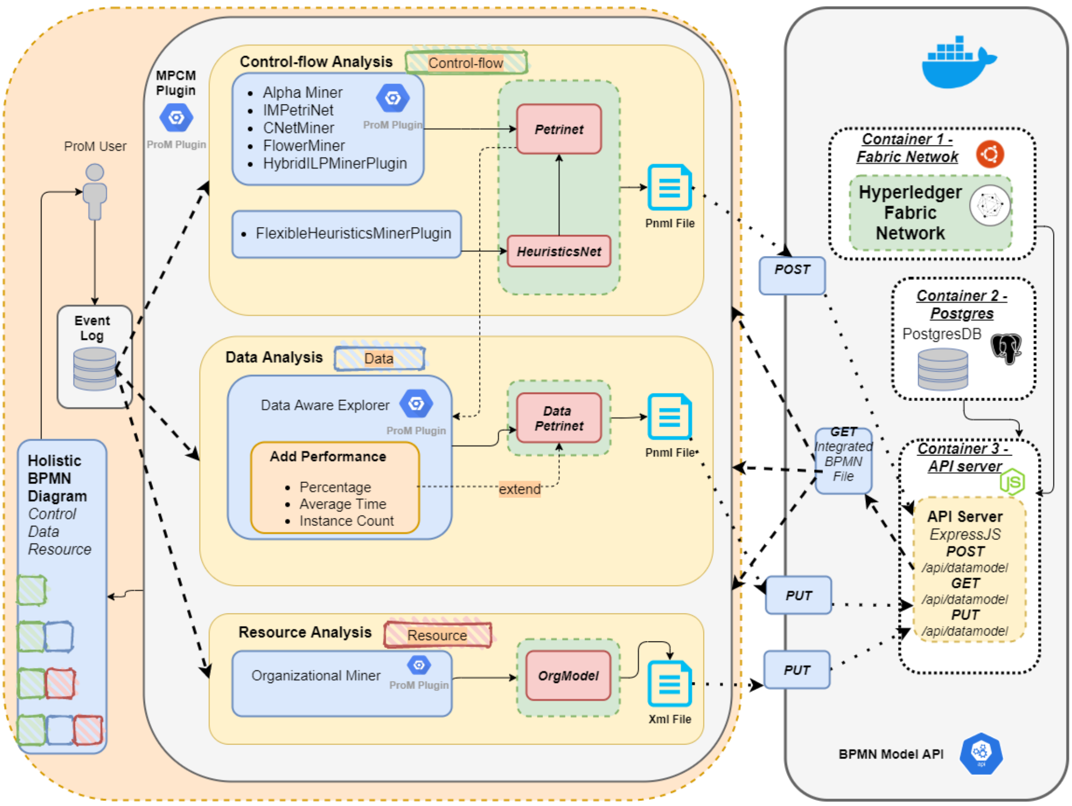


Figure 5.1. MPCM (Multi-Perspective Chain Miner) Plugin Overview

The use-case diagram given in Figure 5.2. illustrates the main interactions between the user and the MPCM framework. Once event logs include with data, resource attributes in the event-level, we can combine control-flow, data and resource perspectives into a single BPMN Diagram. In the first part, after importing an event log (typically including the fields of activity, timestamp, data, and resources), the user selects the control-flow algorithm. Then, data and resources perspectives to be explored are selected. If the data perspective is selected, the information about the percentage of transition, average time, and instance count can also be added in front of the data guards and branches displayed in the Data Petri-net. In the second part, the user can proceed with the creation of a BPMN Diagram with regard to the process perspectives mined. The framework allows the user to visualize the intermediate models in the control-flow and data perspectives, or in the control-flow and resource perspectives, depending on the user's selection. Finally, the user can export the integrated process model in the format of BPMN or Data Petri-net.

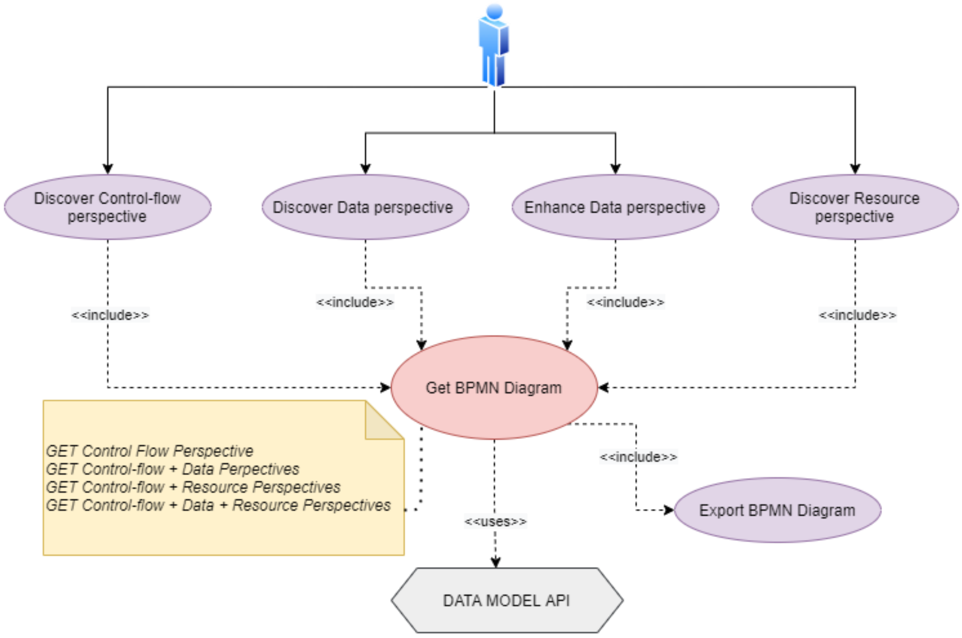


Figure 5.2. Use-case Diagram of MPCM Framework

5.1. Development Steps of the Plugin

The process model in the core of the XLog file imported by the user of the ProM framework can be discovered from different perspectives by selecting the MPCM plugin. At this stage, firstly the perspectives to be used in process discovery are selected, and then the perspectives to be transformed into BPMN Diagram are selected. Lastly, exported BPMN Diagrams can be viewed and exported from the ProM's Import section. In Figure 5.3, the activities that are done and should be done after XLog file is imported can be observed in ProM.

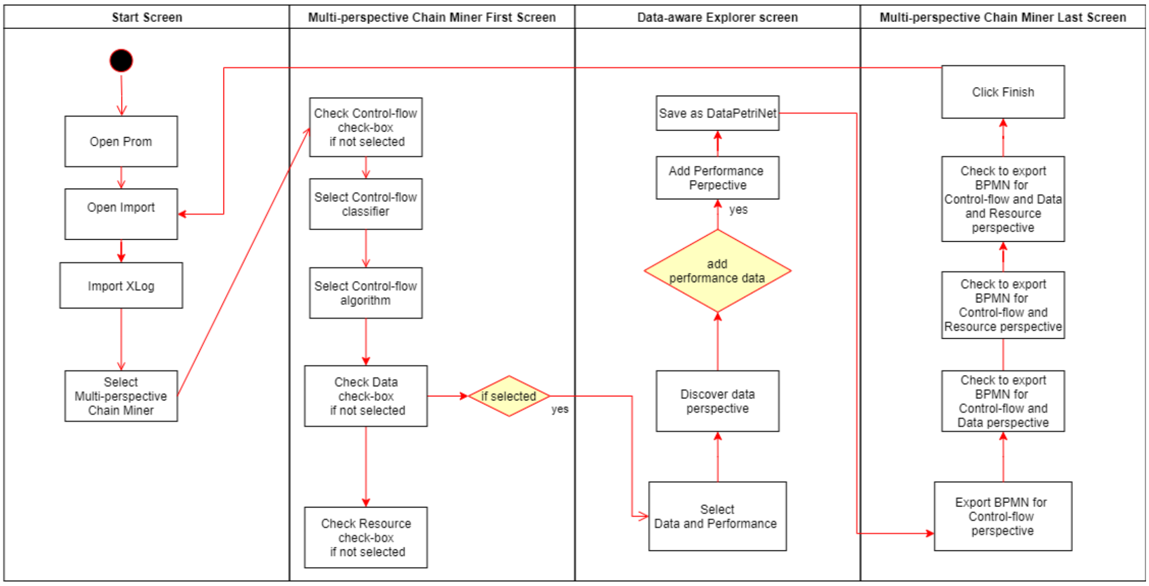


Figure 5.3. MPCM Activity Diagram for ProM usage

BPMN Data Model API is used to combine the outputs of the different perspective algorithms applied with the MPCM plugin, which holds the process model developed within the scope of this work from different perspectives. The integration of MPCM with the BPMN Data Model API is shown in Figure-5.4.

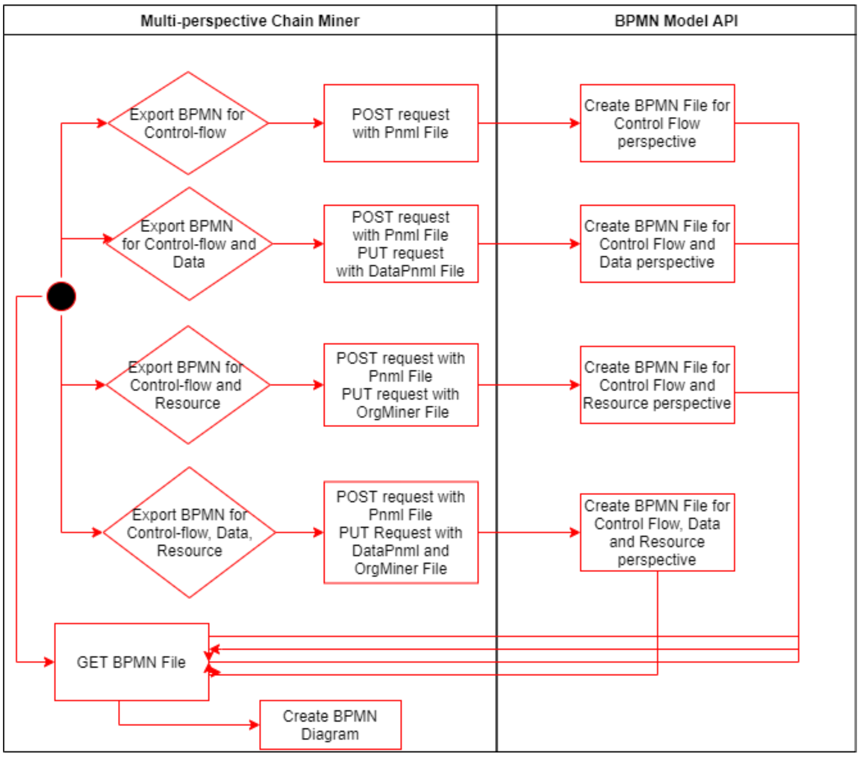


Figure 5.4. Activity Diagram of MPCM and BPMN Data Model API Integration

5.1.1. Used Algorithms in the Plugin

While developing MPCM, perspective algorithms used in process mining were examined. In the light of these algorithms, it has been investigated how these perspectives can be integrated with each other. In this context, the algorithms used can be seen in the table below.

Table 5.1. Used Perspectives and Algorithms in MPCM

Perspectives	**Algorithms**	
Control Flow Perspective	Alpha algorithms	AlphaClassicMiner
		AlphaPlusMiner
		AlphaPlusPlusMiner
		AlphaRobustMiner
	IMPetriNet	
	HybridILPMinerPlugin	
	FlexibleHeuristicsMinerPlugin	
	CNetMiner	
	IMProcessTree	
	FlowerMiner	
Data Perspective	Data Aware Explorer	
Resource Perspective	Organizational Miner	
Enrichment of Data with percentage, average time, instance count	Data Aware Explorer	

5.2. Screens of the Plugin

Within the scope of this thesis, the MPCM Plugin integrated into the ProM Framework is selected as shown in Figure 5.5. For this, the XLog file must be imported from the ProM's Import section beforehand. Control-flow checkbox appears as selected on the screen shown in Figure 5.6. Because with this perspective, the skeleton model of other perspectives is created. Then the classifier and control-flow algorithm should be selected according to the properties of the relevant dataset. Other perspectives to be applied can also be selected through the checkboxes available on the screen.

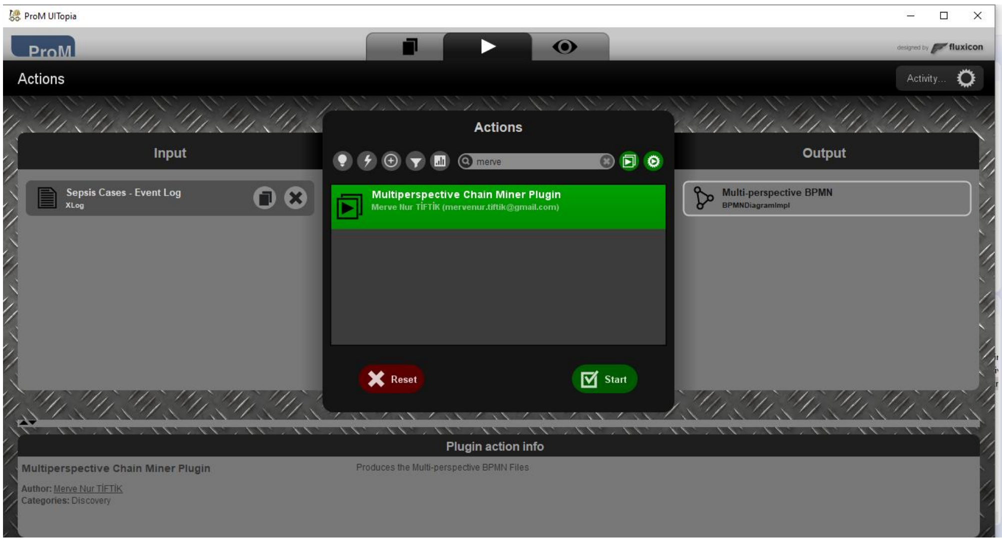


Figure 5.5. Plugin selection in ProM

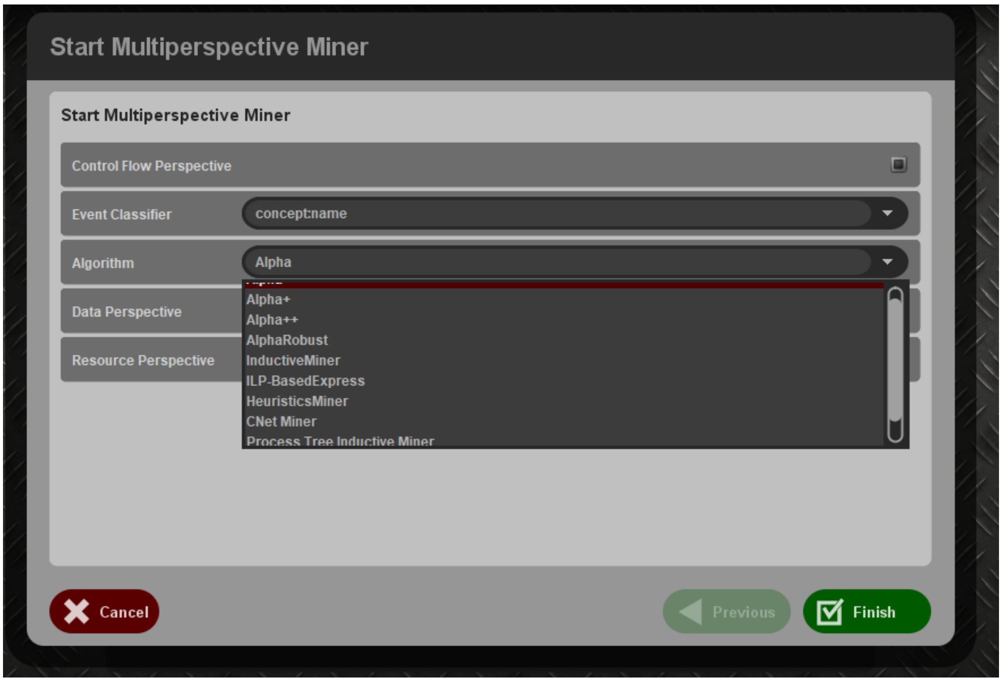


Figure 5.6. Perspective and Control-flow algorithm selection in MPCM in ProM

Finally, the outputs resulting from the algorithms applied with the screen shown in Figure 5.7 are integrated using the BPMN Data Model API according to the selections. BPMN files are obtained according to the respective selection. Then, the obtained BPMN Data Model can be accessed, viewed, and exported from the ProM's Import tab.

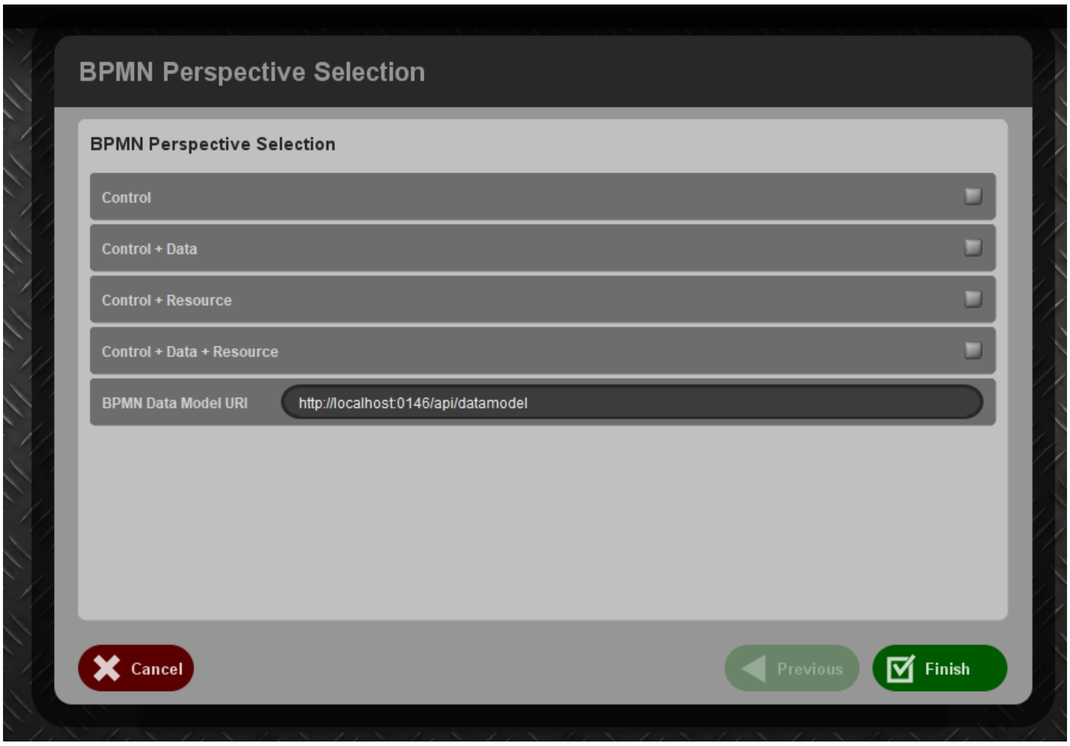


Figure 5.7. BPMN perspective selection in MPCM in ProM

5.3. Steps for Discovering the Perspectives

In this thesis, mainly three perspectives have been studied. These perspectives are control-flow, data, and resource perspectives. In addition, data perspective guards are enhanced with percentage, average time, and instance count information.

5.3.1. Control Flow Perspective

In this thesis, the algorithms used in Control-flow perspective in process model discovery were examined. After working of these algorithms used in the control-flow perspective, the outputs generated are converted into a Petri-net file for use in the BPMN Data Model API. The steps required to add this perspective to the BPMN Data Model are shown in Figure 5.8.

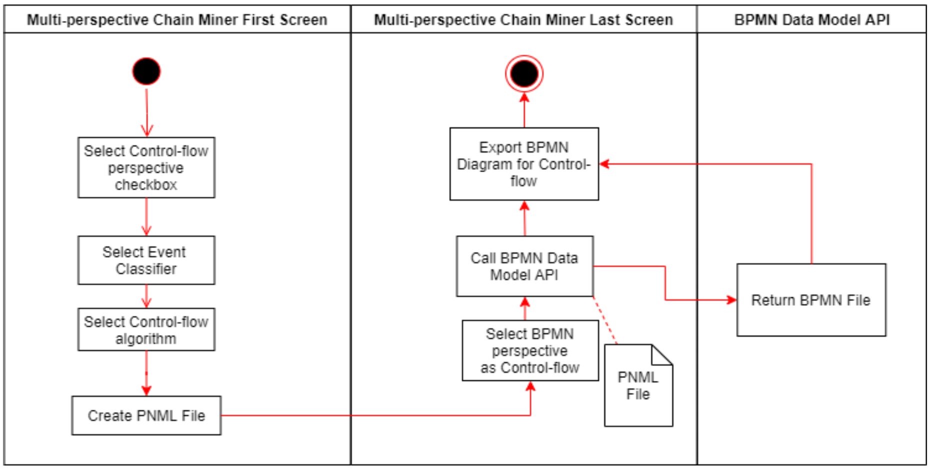


Figure 5.8. Activity Diagram for Control-flow Perspective in MPCM

5.3.2. Control Flow and Data Perspectives

In this thesis, algorithms used in Control-flow and data perspectives in process model discovery were examined. What has been done for the control-flow perspective is mentioned in Section 5.3.1. Data Aware Explorer Plugin is used in ProM framework for data perspective exploration. Here, data guards are added to the process model created as a result of control-flow perspective algorithms. In addition, percentage, average time, and instance count information can be added to the process model if desired by the user. Finally, DataPNML file is created to be used in BPMN Data Model API. The steps required to add these perspectives to the BPMN Data Model are shown in Figure 5.9. Also, Petri-net output of the data perspective can be seen in Figure 5.10.

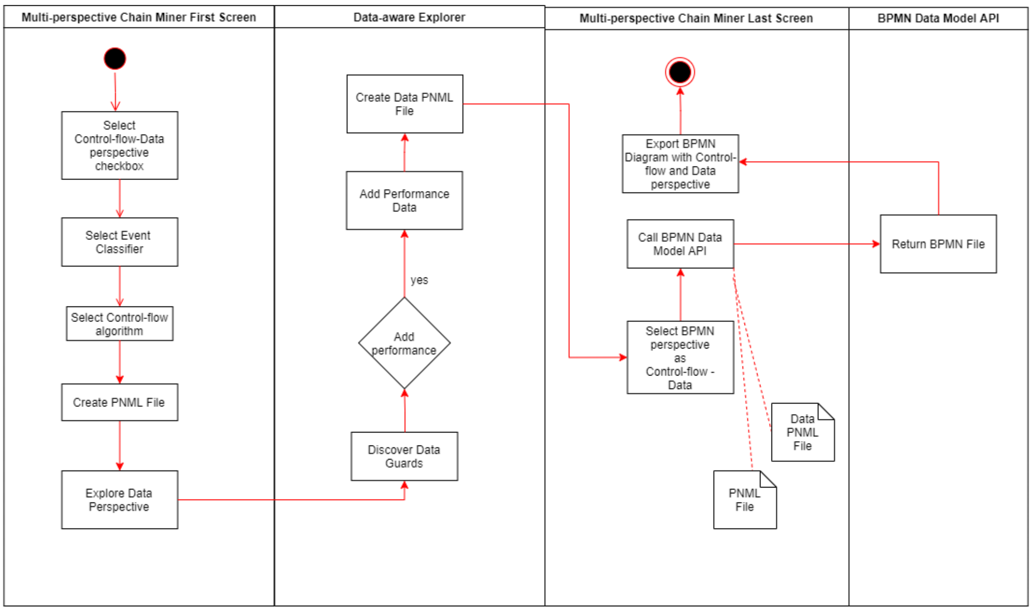


Figure 5.9. Activity Diagram for Control-flow and Data Perspectives in MPCM

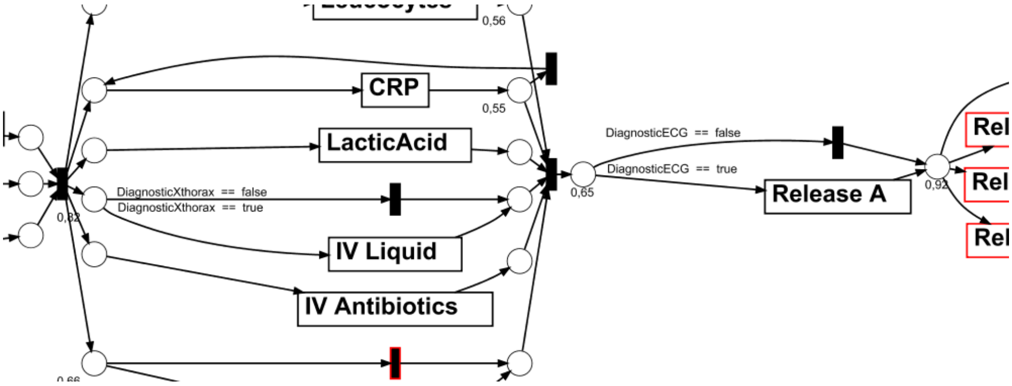


Figure 5.10.   Petri-net output of Sepsis Dataset [58] In DataAwareExplorer in MPCM

5.3.3. Control Flow and Resource Perspectives

In this thesis, the algorithms used in Control-flow and resource perspectives in process model discovery were examined. What has been done for the control-flow perspective is mentioned in Section 5.3.1 Organizational Miner Plugin is used in ProM 5 Framework for resource perspective exploration. The OrgMiner file, whose activitity-resource mapping is made, is created as an XML File to be used in the BPMN Data Model API. The steps required to add these perspectives to the BPMN Data Model are shown in Figure 5.11.

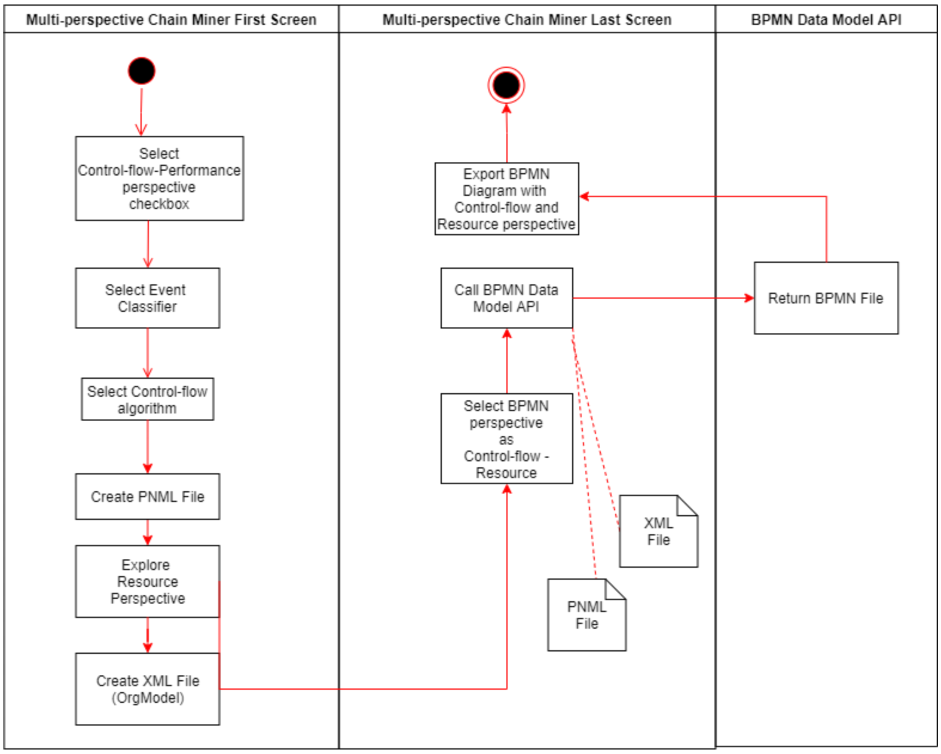


Figure 5.11.   Activity Diagram for Control-flow and Resource Perspectives in MPCM

5.3.4. Control Flow, Data and Resource Perspectives

Within the scope of this thesis, PNML, DataPNML, and XML OrgMiner files are created to be used in BPMN Data Model API with the algorithms mentioned in Sections 5.3.1, 5.3.2, and 5.3.3 for Control-flow, data, and resource perspectives in process model discovery, respectively. The steps required to add these perspectives to the BPMN Data Model are shown in Figure 5.12.

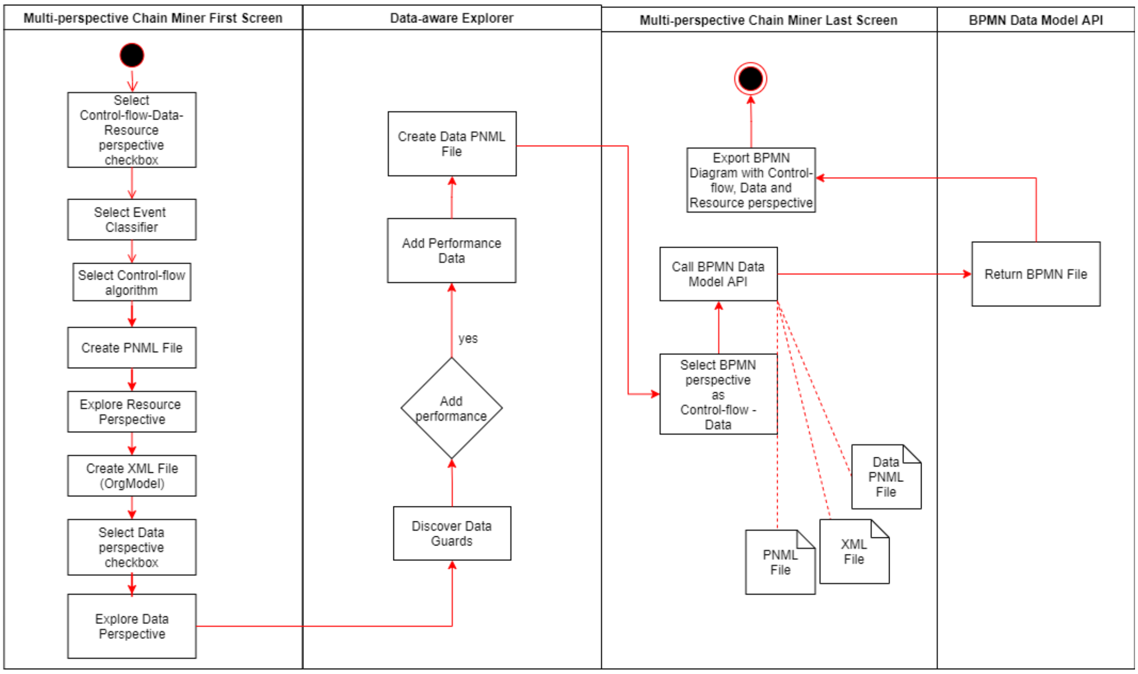


Figure 5.12. Activity Diagram for Control-flow, Data, and Resource Perspectives in MPCM

6. EXAMPLE APPLICATION OF FRAMEWORK IN MEDICAL DOMAIN

This section explains how the MPCM framework can be used to generate a holistic process model for a case in medical domain.

Sepsis data [58] is analyzed from different perspectives and represented in an integrated and automatically generated BPMN Diagram by examining the results of different control-flow algorithms. The applicability of the plugin developed within the scope of this thesis can be seen with this dataset, because it contains resource information that performs activities and some data fields (e.g., DiagnosticXthorax, DiagnosticECG etc.) that relate to these activities, in addition to the mandatory fields of case id, activity and timestamp information. This anonymized event log covers traces of 1050 patients' trajectories that are admitted to the emergency ward by displaying symptoms of Sepsis, over the course of 1.5 years in the hospital information systems [58]. It contains 15,214 events for 16 activities which are categorized into medical activities and logistic activities. There are 846 process variants, which indicates the inherited complexity and flexibility problems of healthcare processes [59], although the dataset focuses on a specific sub-group of patients.

The steps required to put together the results from the control-flow, data and resources perspectives to the BPMN Data Model are shown in Figure 6.1. The user needs to deploy the BPMN Model API [52] before using MPCM framework in order to keep the multi-perspective process model in blockchain. With respect to the steps shown in the figure, multi-perspective process mining starts with the control-flow analysis, and continues with the data perspective analysis and resource analysis.

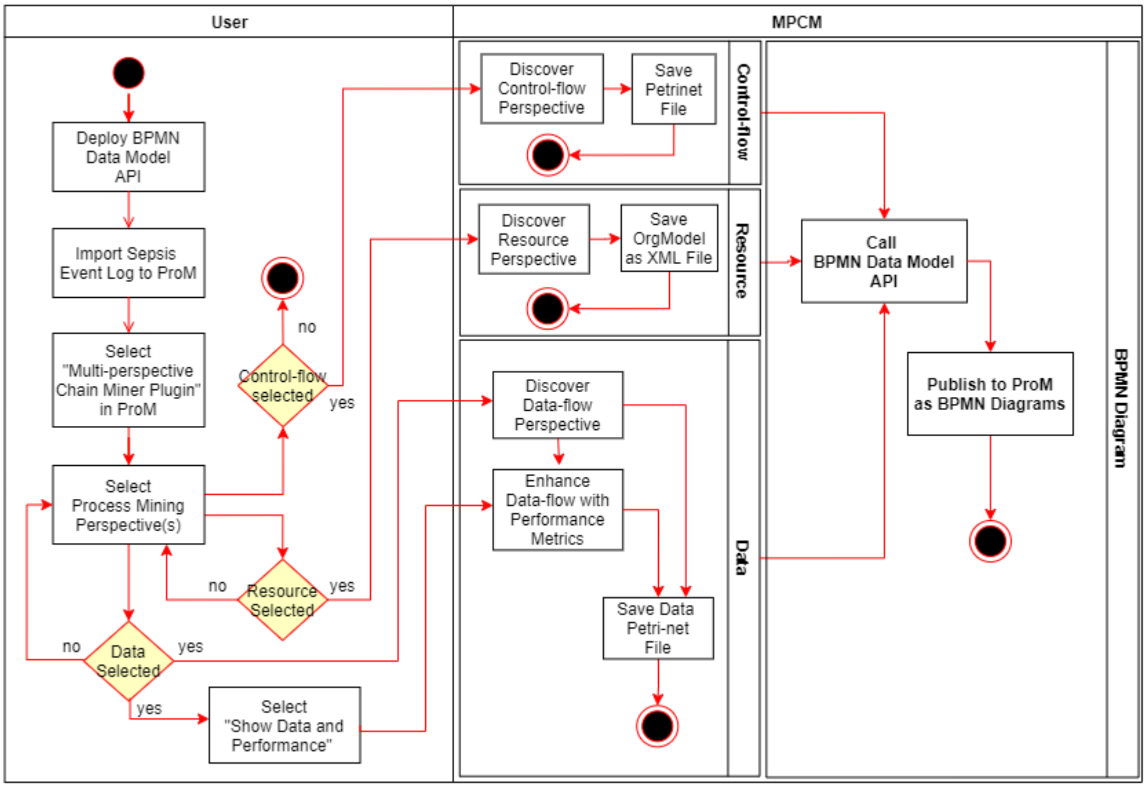


Figure 6.1. Activity Diagram for Operational Usage of MPCM with BPMN Model API

Table 6.1. Precision, Generalization, and Fitting Metrics of Plugins in ProM 6.10 [60]

Plugin Name	**Precision**	**General-ization**	**Perfectly-fit-ting cases / Cases**
Alpha Miner	0	0	1035/1050
Mine Petri net with Inductive Miner	0.50288	0.99858	318/1050
Mine for a Heuristics Net using Heuris-tics Miner	0	0	1035/1050
ILP-Based Process Discovery (Express)	0.86297	1	0/1050
Mine Petri net using Flower Miner	0.16858	1	0/1050

6.1. Discover Control-flow Perspective

This analysis starts with selecting control-flow algorithms. According to the comparison of the algorithm metrics on ProM 6.10 as given in Table 6.1, it was decided that the best alternative to be chosen for Sepsis data can be Inductive Miner algorithm. So, for the Sepsis case, we selected Inductive Miner (IMPetriNet) from the screen as shown in Figure 6.2, and discovered the process model from the control-flow perspective. IMPetriNet generates a Petri-net which can be difficult to understand by the medical professionals. The discovered Petri-net was automatically converted to a BPMN diagram as shown in Figure 6.4. According to the figure, the Sepsis Process starts with registration or triaging activities; continues with measuring of leukocytes, CRP, and lactic acid; and ends with some types of discharge (Release C-D) or Return ER activities.

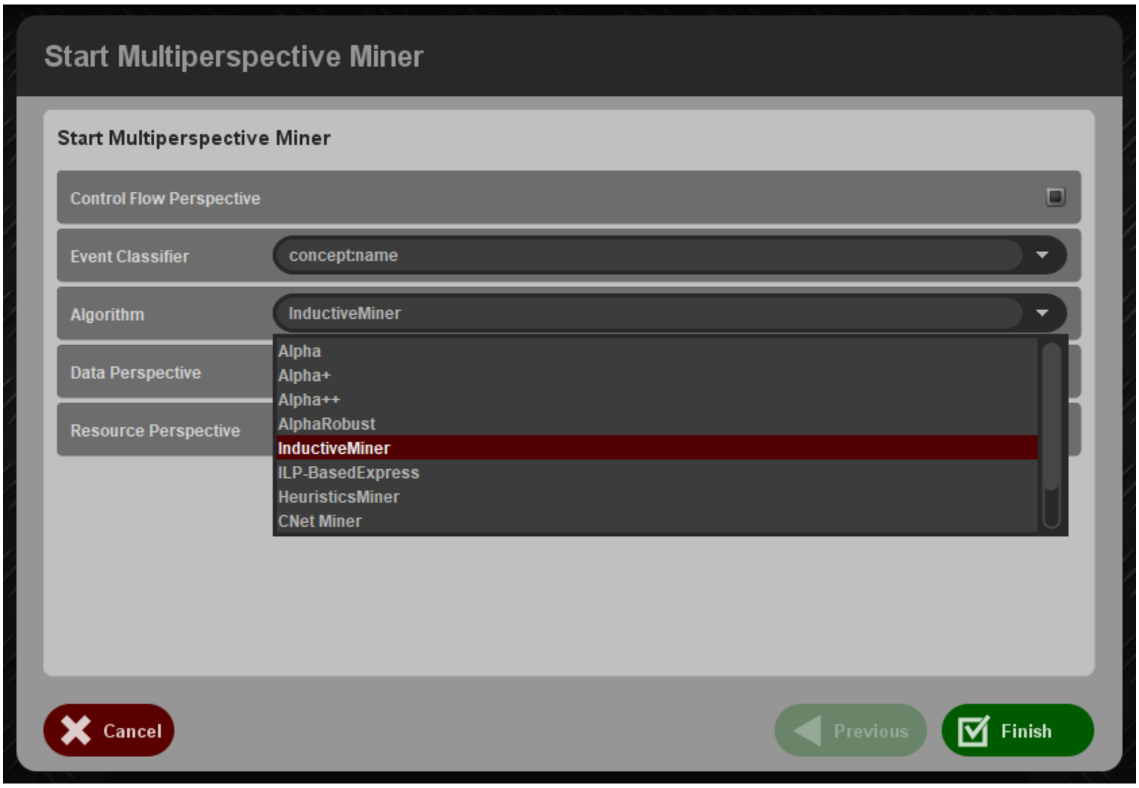


Figure 6.2. Selection of Perspectives and Control-flow Algorithms for Sepsis Process

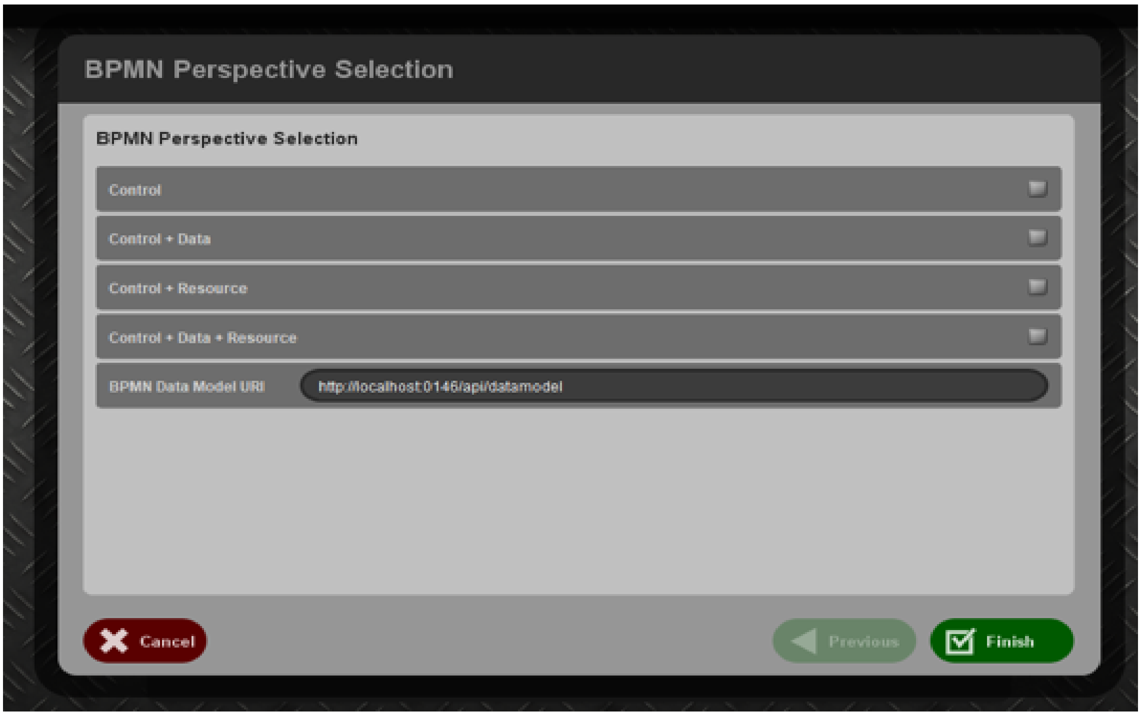


Figure 6.3. Selection of Perspectives for an Output in BPMN for Sepsis Process

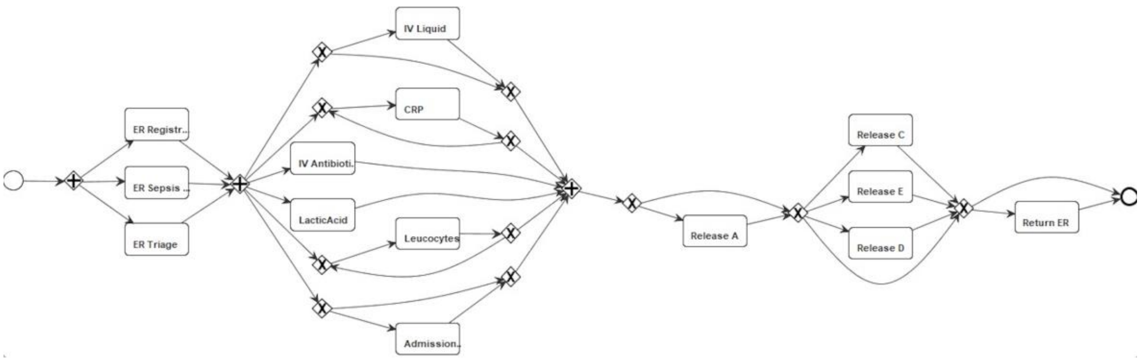


Figure 6.4. BPMN Diagram of Sepsis Process Model from Control-flow Perspective

6.2. Discover Data Perspective

In multi-perspective process mining, discovered process model in BPMN or Petri-net can be extended with the data perspective and represented in a Data Petri-net. When the user checks the data perspective in the perspective selection screen (in Figure 6.2), Data Aware Explorer Plugin is executed by MPCM plugin. Data guards are added to the previously discovered process model that is a skeleton for other perspectives. The discovered model of Sepsis Process in Data Petri-net is denoted in Figure 6.5.

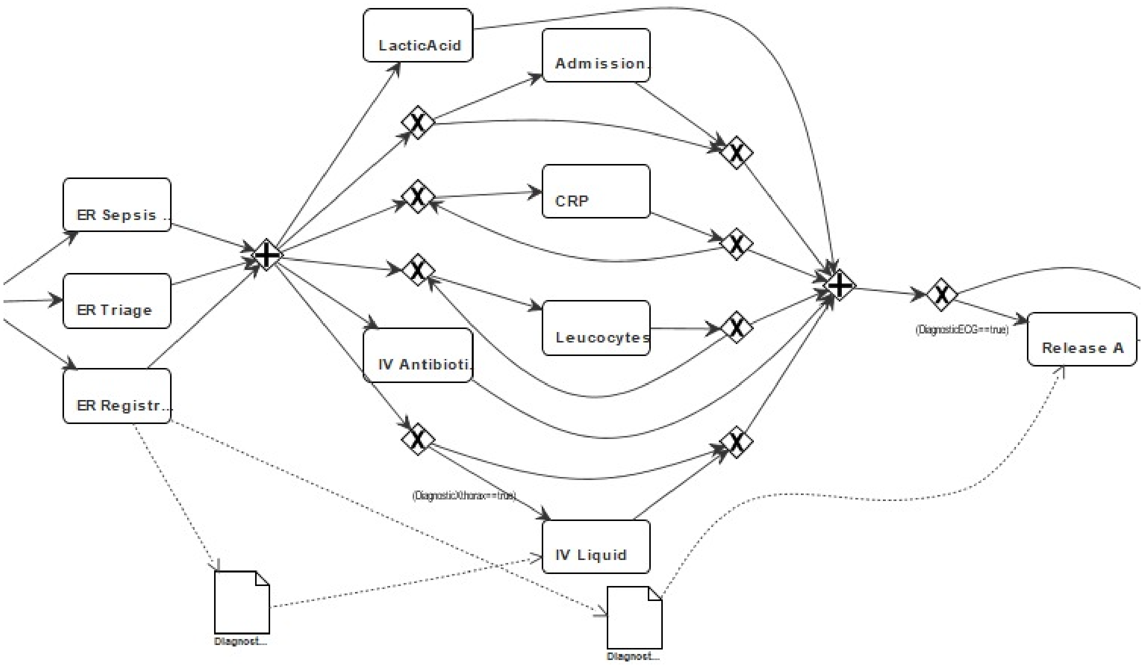


Figure 6.5. BPMN Diagram for Data Petri-net of Sepsis Process Model (Control-flow Perspective and Data Perspectives)

6.3. Enhance Data Perspective

The Data Petri-net created by the data perspective can be enhanced if the user wants to add information from the performance perspective, such as percentage of transition, average time and instance count, to the discovered model. The enhanced process model in Data Petri-net is shown in Figure 6.6 and in Figure 6.7 (with data guards). We discovered the data guards in the Sepsis Process and visualize the multi-perspective process in the BPMN diagram in Figure 6.8.

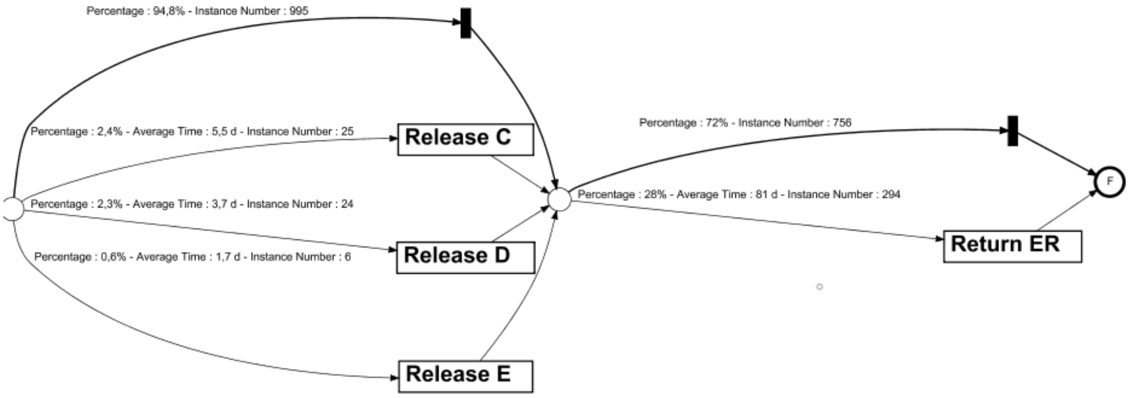


Figure 6.6. Percentage, average time and instance count information in transitions for Sepsis Process Model

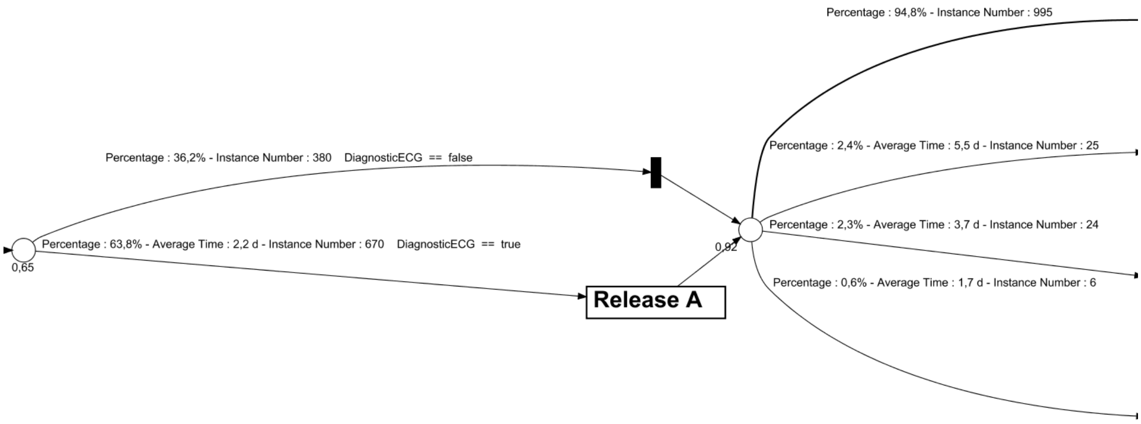


Figure 6.7. Percentage, average time, instance count and data guards information in transitions for Sepsis Process Model

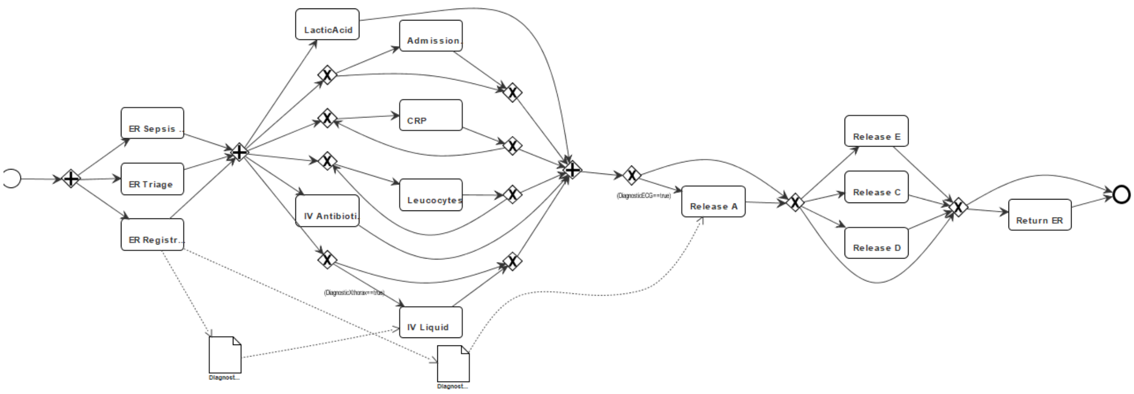


Figure 6.8. Full view of BPMN Diagram of Sepsis Process Model from Control-flow and Data Perspectives

6.4. Discover Resource Perspective

As an alternative, the framework integrates resources information to the BPMN Data Model, and visualizes the roles that perform the activities in addition to the information from the control-flow perspective. Organizational Miner Plugin (in ProM 5) is used for the exploration of the resource perspective. We discovered the roles in the Sepsis Process and visualize the multi-perspective process in the BPMN diagram in Figure 6.9.

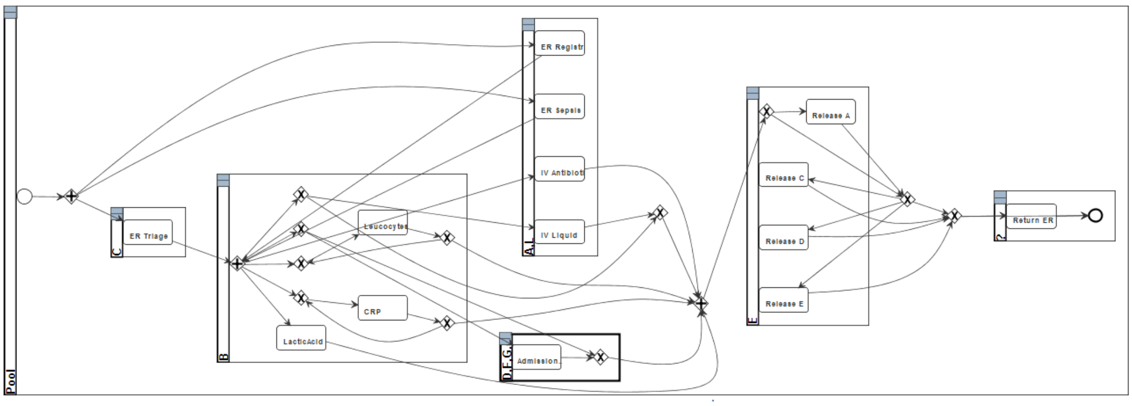


Figure 6.9. BPMN Diagram of Sepsis Process Model from Control-flow and Resource Perspectives

6.5. Combine Control Flow, Data and Resource Perspectives

As an another alternative, the framework integrates data guards and resources information to the BPMN Data Model, and visualizes the roles that perform the activities in addition to the information from the control-flow and data perspectives. Inductive Miner algorithm is used for control flow perspective, and Data aware explorer algorithm (in ProM 5) is used for data perspective and Organizational Miner Plugin (in ProM 5) is used for the exploration of the resource perspective. We discovered the roles in the Sepsis Process and visualize the multi-perspective process in the BPMN diagram in Figure 6.10.

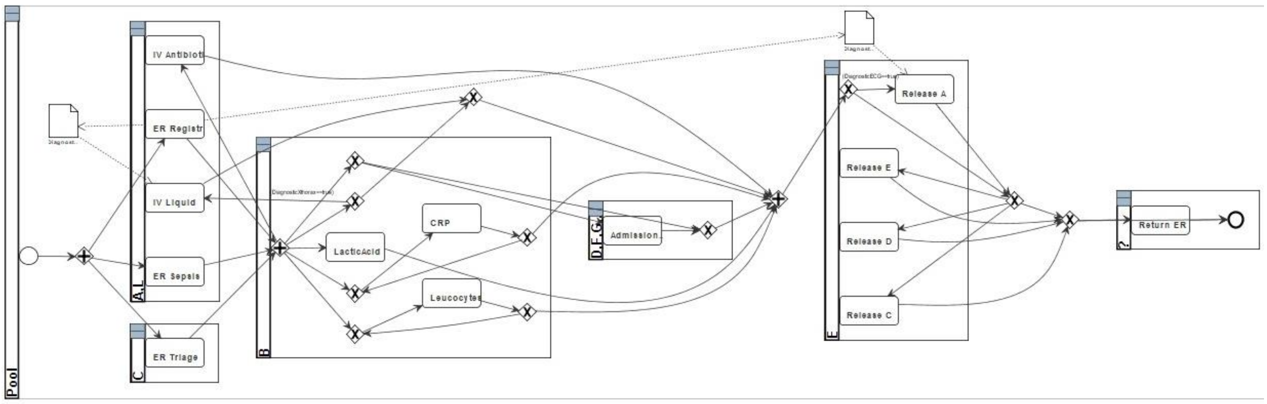


Figure 6.10. BPMN Diagram of Sepsis Process Model from Control-flow, Data, and Resource Perspectives

6.6. Get BPMN Diagram

The outputs of the control-flow, data and resource perspectives are integrated using the BPMN Data Model API according to the user selections displayed in the screen in Figure 6.3. The multi-perspective process model is stored in the BPMN Data Model in block-chain in a secure and immutable way. Then, the process model in BPMN can be accessed, viewed, and exported from the Import tab in ProM.

6.7. Observations and Lessons Learned by the Example Application

ProM tool is a framework used in process mining. It differs from other tools because it is open source. In this way, it is open to continuous development and improvement. Within the scope of this thesis, we studied ProM process discovery algorithms and examined algorithms that elicit the process model from different perspectives. In the light of the algorithms we obtained, we aimed to implement a Multi-perspective Process Mining Plugin compatible with the ProM framework that works in integration with the BPMN Data Model API. At this stage, however, the BPMN Data Model API requested outputs of the algorithms in a certain format. Therefore, the outputs of some algorithms (e.g., FlexibleHeuristicsMinerPlugin output had to be converted to Petri-net representation). Another case was that the Organizational Miner algorithm was not included in the last version of ProM (6), so this algorithm was integrated from the ProM's older version (ProM 5). Nevertheless, this led to a slowdown in implementation speed. Thanks to this thesis, it was learned that there may be new plugins that are developed for ProM. An examples is the adaptation of the Organizational Miner to the latest ProM version, i.e., the plugin that chooses the most accurate Control-flow algorithm according to the quality metrics for the event log.

In addition, within the scope of this thesis, the compatibility of the medical data used for the example application and the multi-perspective process discovery solutions to the health sector has been observed. Within the scope of this application, the process model could be viewed in multi-perspectives, thanks to the data fields included in the dataset.

In this thesis, by researching algorithms from different perspectives, a tool that enables the algorithms of control-flow, data and resource perspectives to be run together and displayed in a single BPMN model has been developed. While exploring the data perspective, time and performance (percentage) information is also recorded in the output file. It has been observed that the recorded data can be viewed in the BPMN model.

7. VALIDATION OF THE FRAMEWORK

In this section, the outputs for the Sepsis dataset [58] are compared with the versions run in MPCM and run in ProM. The MPCM outputs for this dataset are detailed in Section 6. Printouts are taken for the Sepsis dataset, the BPMN Data Model API is triggered by using Postman [61] to convert the obtained algorithm outputs to BPMN, and the outputs are displayed through the ProM's "Select BPMN Diagram" plugin.

7.1. Validation of Control-flow Perspective

First of all, Sepsis data was mined according to the control flow perspective using the "Mine Petri net with Inductive Miner" Plugin in ProM 6.10. This plugin was run with configurations in Figure 7.1. In this way, our Petri-net model in Figure 7.2 that would be used in the BPMN Data Model was created. Postman was used to convert the generated Petri net-file (exported from ProM 6.10) to the control-flow BPMN output. With this tool, a POST request was made as in Figure 7.3.a and then the BPMN model was downloaded with the GET request in Figure 7.3.b. The BPMN Model formed after all these steps was imported into ProM 6.10 and the results were observed with the "Select BPMN Diagram" Plugin in Figure 7.4.

Figure 7.1. ProM 6.10 “Mine Petri net with Inductive Miner” Plugin configuration

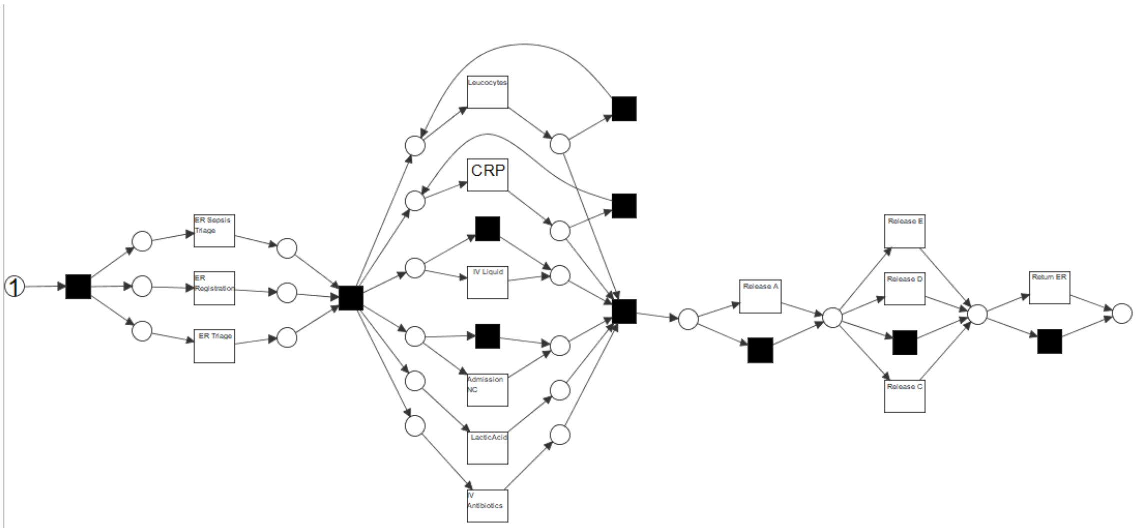
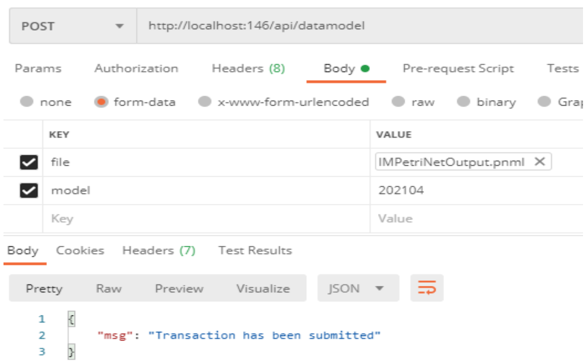


Figure 7.2. ProM 6.10 "Mine Petri-net with Inductive Miner" Plugin Output in Petri-net



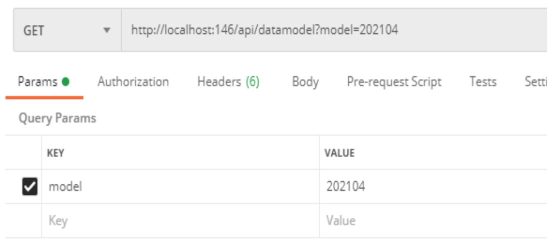


Figure 7.3.a. Postman POST Request

Figure 7.3.b. Postman GET Request

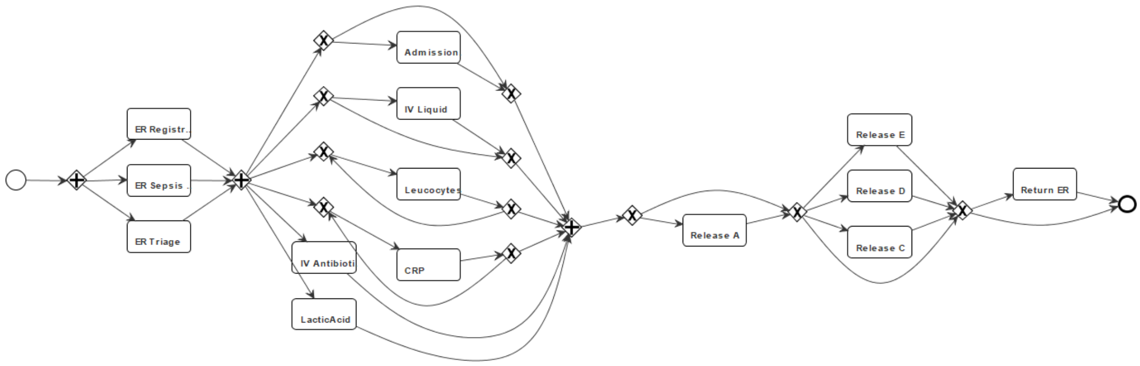


Figure 7.4. Created BPMN Diagram with using BPMN Data Model API

7.2. Validation of Control-flow and Data Perspective

Then, using the “Multi-perspective Process Explorer” Plugin in ProM 6.10, the Sepsis data was mined according to the data perspective. In this way, our Data Petri-net file, which would be used to add the data perspective to the BPMN Data Model, was created as in Figure 7.5. Postman was used to convert the resulting Data Petri-net file (exported from ProM 6.10) to add data perspective BPMN output. Through this tool, PUT request was made as shown in Figure 7.6.a and then BPMN model was downloaded with GET request in Figure 7.6.b. The BPMN Model formed after all these steps was imported into ProM 6.10 and the results were observed with the "Select BPMN Diagram" Plugin in Figure 7.7.

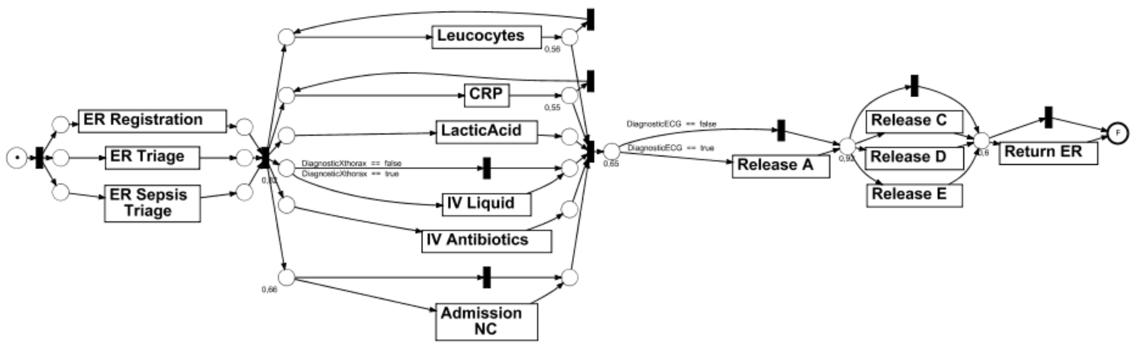
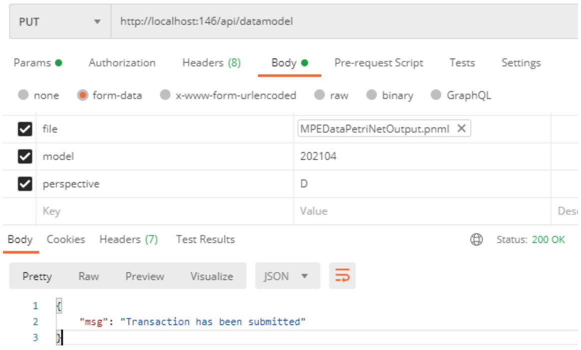


Figure 7.5. Created Data Petri net Model with using “Multi-perspective Process Explorer”



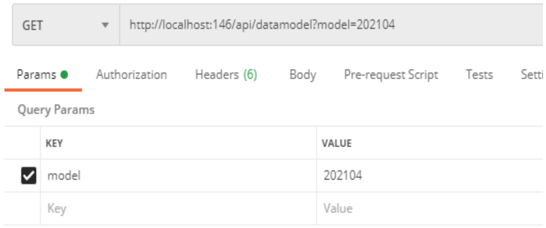


Figure 7.6.a. Postman PUT Request Figure 7.6.b. Postman GET Request

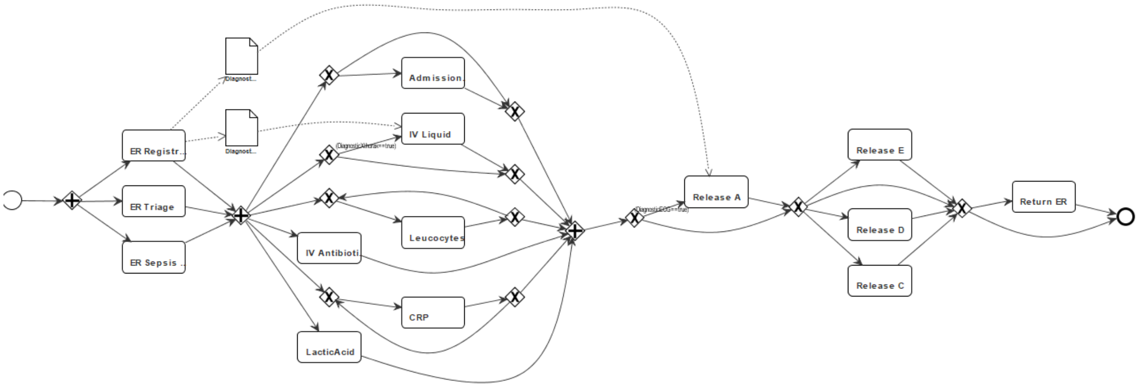


Figure 7.7. Created BPMN Diagram with using BPMN Data Model API

7.3. Validation of Control-flow and Resource Perspective

Then, using the “Organizational Miner” Plugin in ProM 5, the Sepsis data was mined according to the resource perspective. In this way, our Org Model file, which would be used to add the resource perspective to the BPMN Data Model, was created. Postman was used to convert the resulting Org Model file (exported from ProM 5) to add resource perspective to BPMN output. Through this tool, POST request was made as in Figure 7.8.a for control flow perspective, PUT request was made as in Figure 7.8.b for resource perspective, and then BPMN model was downloaded with GET request in Figure 7.8.c. The BPMN Model formed after all these steps was imported into ProM 6.10 and the results were observed with the "Select BPMN Diagram" Plugin in Figure 7.9.

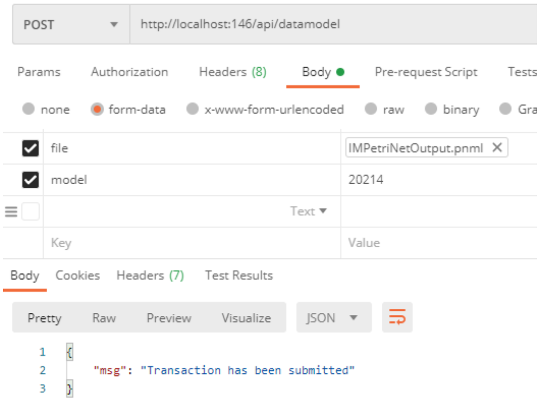


Figure 7.8.a. Postman POST Request

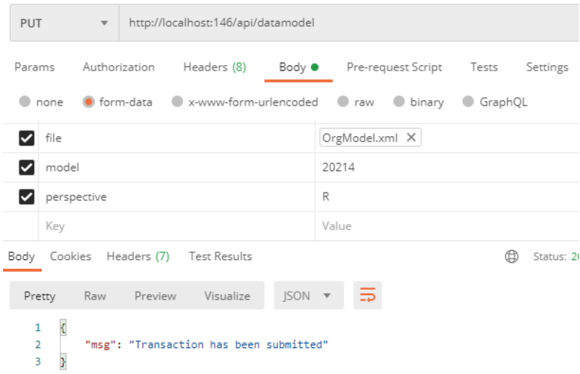


Figure 7.8.b. Postman PUT Request

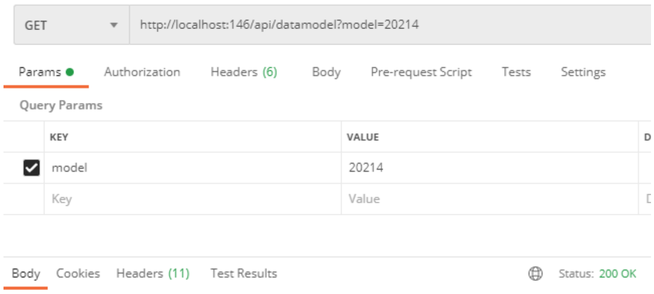


Figure 7.8.c. Postman GET Request

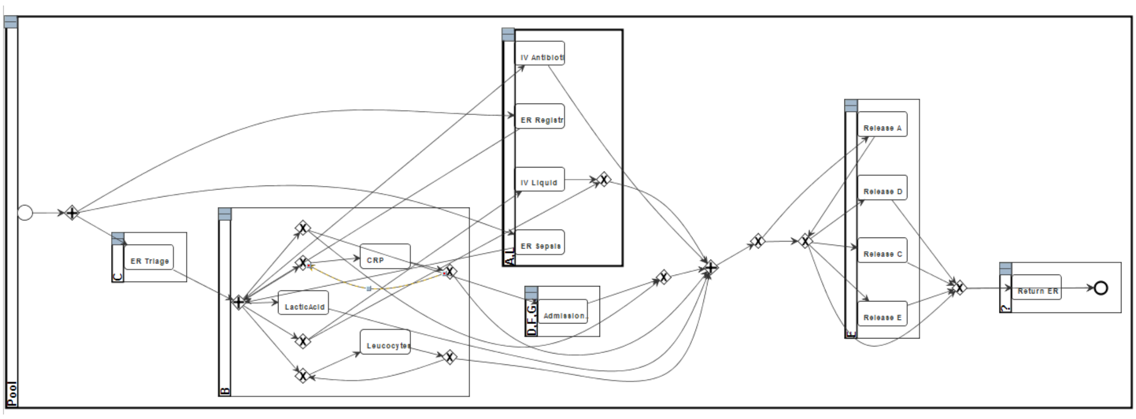


Figure 7.9. Created BPMN Diagram with using BPMN Data Model API

7.4. Validation of Control-flow, Data, and Resource Perspective

Finally, using the "Organizational Miner" Plugin in ProM 5, Sepsis data was mined according to the resource perspective. This plugin was run with configurations in Figure 7.10. In this way, our Org Model file (exported from ProM 5), which would be used to add resource perspective to the BPMN Data Model, was created. Postman was used to convert the resulting Org Model file to add resource perspective BPMN output. Through this tool, PUT request was made as in Figure 7.11.a, and then BPMN model was downloaded with GET request in Figure 7.11.b. The BPMN Model formed after all these steps was imported into ProM 6.10 and the results were observed with the "Select BPMN Diagram" Plugin in Figure 7.12.

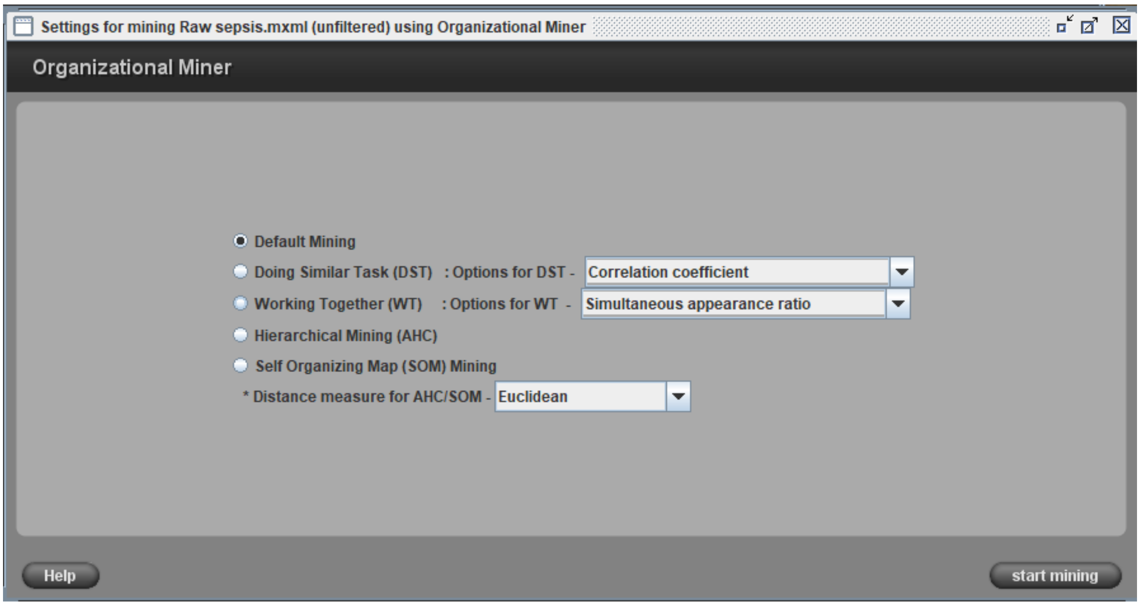
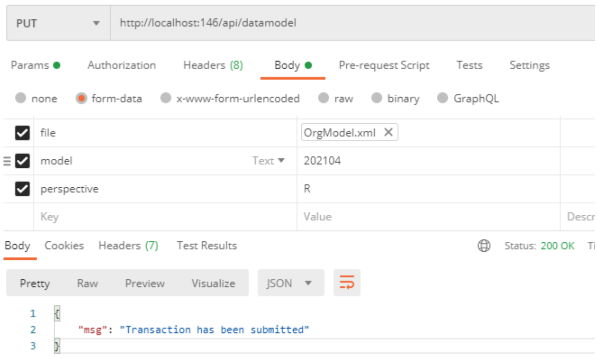


Figure 7.10. ProM 5 “Organizational Miner” Plugin configuration



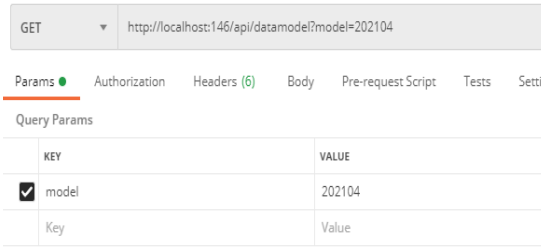


Figure 7.11.a. Postman PUT Request

Figure 7.11.b. Postman GET Request

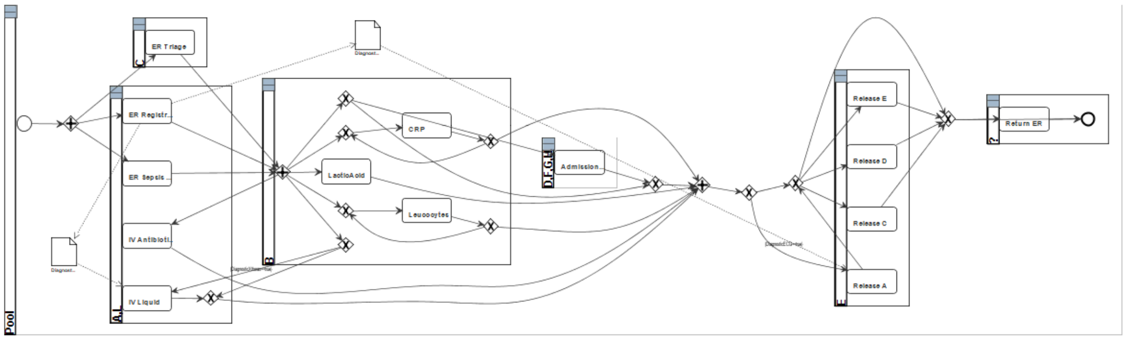


Figure 7.12 Created BPMN Diagram with using BPMN Data Model API

8. OVERVIEW OF FRAMEWORK VALIDATION RESULTS

In this section, comparison results of the outputs obtained in Sections 6 and 7 are given, together with a discussion at the end.

8.1. Comparison for Control-flow Perspective

The visualized versions of the outputs are given below.

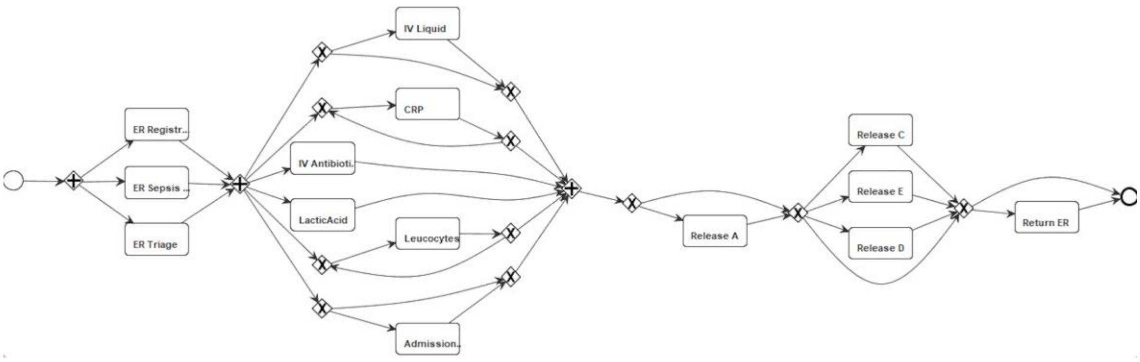


Figure 8.1. MPCM Output of Control-flow perspective (copied from Figure 6.4)

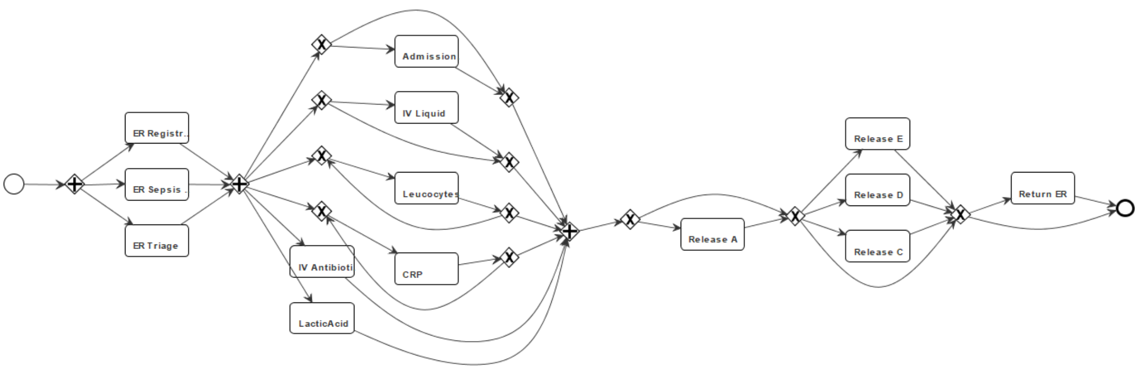


Figure 8.2. ProM Output of Control-flow perspective (copied from Figure 7.4)

Figure 8.1 above includes the control-flow perspective output from MPCM, and Figure 8.2 contains the control-flow perspective output obtained by using ProM and Postman. Although the locations of the activities in the visual are not the same, it can be seen that the outputs are compatible with each other.

At this stage, if the ProM tool is considered as the basic model, it can be considered that the MPCM Plugin is valid for the control-flow perspective, since the output of the MPCM Plugin is compatible with the ProM tool output.

8.2. Comparison for Control-flow and Data Perspective

The visualized versions of the outputs are given below.

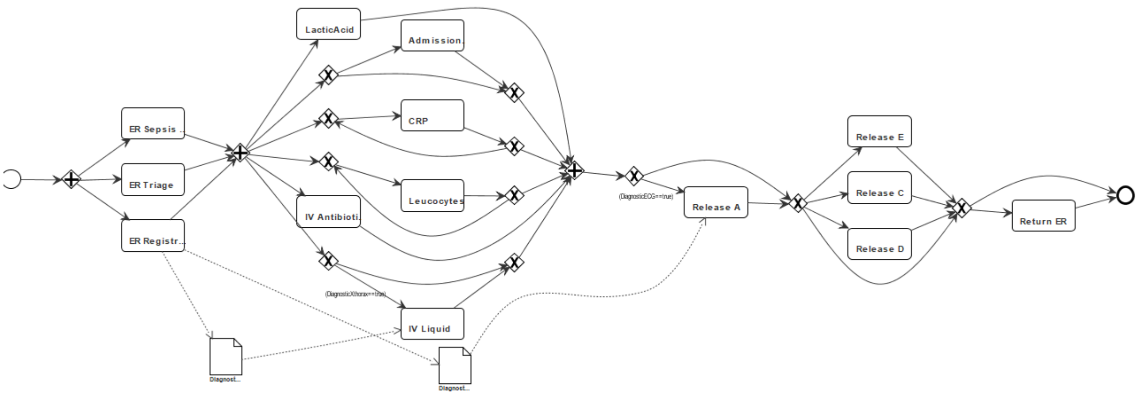


Figure 8.3 MPCM Output of Control-flow and Data perspective (copied from Figure 6.8)

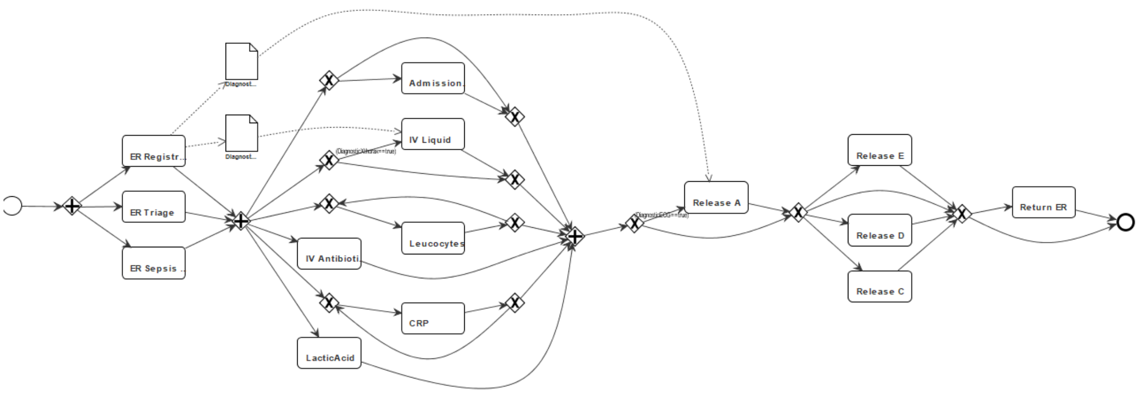


Figure 8.4 ProM Output of Control-flow and Data perspective (copied from Figure 7.7)

Figure 8.3 above includes the control-flow and data perspective output from MPCM, and Figure 8.4 contains the control-flow and data perspective output obtained by using ProM and Postman. Although the locations of the activities in the visual are not the same, it can be seen that the outputs are compatible with each other.

At this stage, if the ProM tool is considered as the basic model, it can be considered that the MPCM Plugin is valid for the control-flow and data perspective, since the output of the MPCM Plugin is compatible with the ProM tool's output

8.3. Comparison for Control-flow and Resource Perspective

The visualized versions of the outputs are given below.

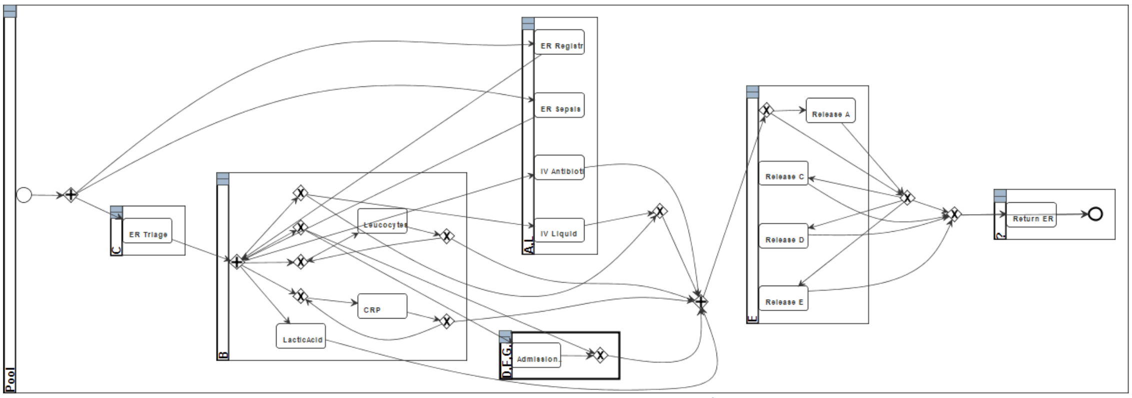


Figure 8.5. MPCM Output of Control-flow and Resource perspective (copied from Figure 6.9)

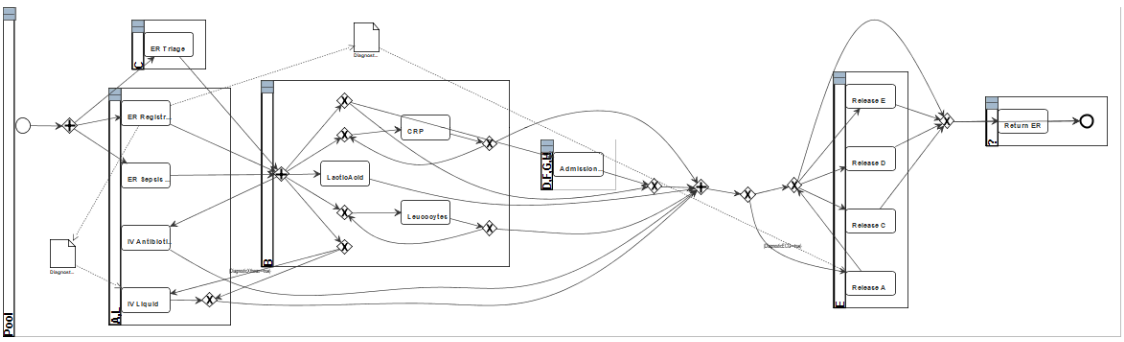


Figure 8.6. ProM Output of Control-flow and Resource perspective (copied from Figure 7.9)

Figure 8.5 above includes the control-flow and resource perspective output from MPCM, and Figure 8.6 contains the control-flow and resource perspective output obtained by using ProM and Postman. Although the locations of the pools and activities in the visual are not the same, it can be seen that the outputs are compatible with each other.

At this stage, if the ProM tool is considered as the basic model, it can be considered that the MPCM Plugin is valid for the control-flow and resource perspective, since the output of the MPCM Plugin is compatible with the ProM tool's output.

8.4. Comparison for Control-flow, Data, and Resource Perspective

The visualized versions of the outputs are given below.

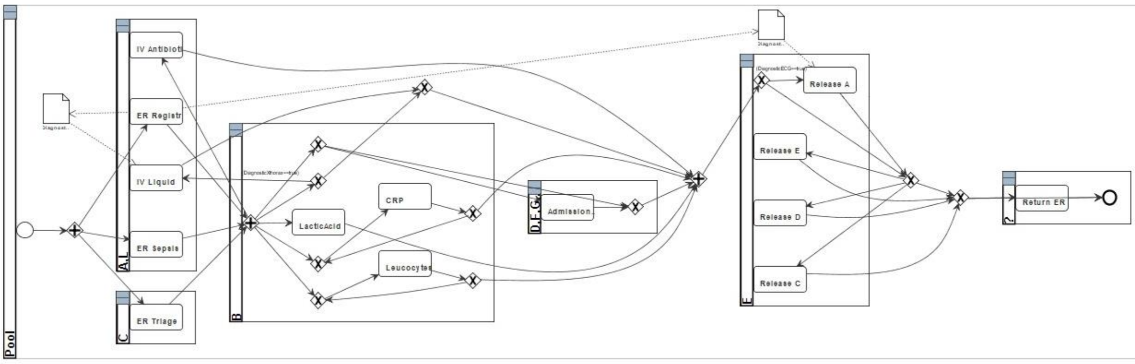


Figure 8.7. MPCM Output of Control-flow, Data, and Resource perspective (copied from Figure 6.10)

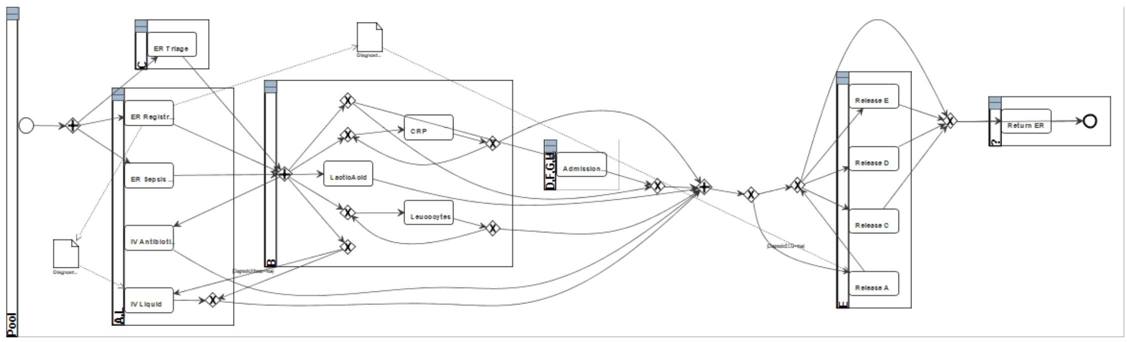


Figure 8.8. ProM Output of Control-flow, Data, and Resource perspective (copied from Figure 7.12)

Figure 8.7 above includes the control-flow, data, and resource perspective output from MPCM, and Figure 8.8 contains the control flow, data, and resource perspective outputs obtained by using ProM and Postman. Since the pool feature of the BPMN is used in the resource perspective stage and the pools can be positioned in different places during the output, the output image looks different. However, when the activities and pools of the formed models are taken to similar places, it can be seen that the outputs are compatible with each other. In other words, although the locations of the activities and pools in the visual are not the same, it can be seen that the outputs are compatible with each other.

At this stage, if the ProM tool is considered as the basic model, it can be considered that the MPCM Plugin is valid for the control-flow, data, and resource perspective since the output of the MPCM Plugin is compatible with the ProM tool's output.

8.5. Assumption and Limitations on the Results Obtained

Some assumptions and limitations on the results obtained in this thesis are included under this sub-section.

8.5.1. Examination of BPMN Outputs

When the BPMN Model output for the Sepsis dataset is examined, it is seen that the activity is re-operated after the activity operation in some places. When the event log is examined in detail to understand whether the situation is related to the algorithm or the event log, it is seen that some activities can be repeated several times for a case. In this way, it was understood that this situation observed in the BPMN model was due to data. Some tests in the medical field can be requested by doctors every other day. Since such repetitive activities may occur in the medical field, it is possible to have such situations in the dataset.

When the BPMN Model output for the Sepsis dataset is examined, it is also seen that some activities (medical measurements) are omitted in some places. Medical measurements are not required for some examinations. So this is a possible situation for cases in the medical domain.

It is also possible that a case in the medical domain has more than one parallel process. Therefore, it is rational that there are parallel gateways in the diagram.

Also, it is observed that some data guards are not in transitions. However, this situation is related to the capability of the algorithm. In addition, there must be a distinctive data input value that allows the activity to be branched in the event log. The branching activities in the event log are discovered with the control flow algorithm, but if the data used to distinguish in the event log is not clearly included, this situation is beyond the capability of the algorithm. It is, again, a situation related to the dataset. Even the most suitable algorithm was chosen for the Sepsis dataset, such a situation was encountered. In real datasets, sometimes factors other than the capabilities of the algorithms (e.g., the dataset may not have enough data to contain each transition's guard) can affect the output model in this way. For example, as seen in the event log [58] , when the "Release A" activity occurs, the value of the "DiagnosticECG" data is true, so this distinction is included on the model. In order for other transitions to include data guards, there must be such a separation in the dataset. During the conformance checking phase and during the compliance with the process model, such situations can be recognized and inconsistent situations can be corrected by making process improvements in the organization.

With this plugin, all operations specific to BPMN are performed with the BPMN Model API. And for all changes (adding, removing, improving) to be made on the BPMN side, this API needs to be updated. In addition, since the BPMN is validated via ProM, these new or improved features can be brought to the ProM algorithms and contributed to the ProM tool.

8.5.2. Examination of event log content

The names of the activities are clear to people working in the medical domain. However, it can be considered by people who are not working on that domain that it is not like the activity name. For this reason, the activity names for this Sepsis dataset can be considered as follows [58], and a more accurate result will be achieved by preparing the event logs with the action and activity verbs.

Table 8.1. Number of activities and Explanation for Sepsis Dataset

#Activities	**Explanation**
3	In the emergency ward • Registration • Triaging
3	Measurements of • Leukocytes • CRP • lactic acid
2	Admission or transfer to • Normal care • Intensive care
5	Hospital discharge types
1	Returning patients later

8.5.3. About blockchain

Blockchain technology is used indirectly in this plugin thanks to the BPMN API. In this context, it is aimed to keep an immutable copy of the converted BPMN Diagrams on the blockchain. Even if the ProM tool is closed, access can be easily provided to the last model that was converted, in a way that is guaranteed not to change.

9. CONCLUSION AND FUTURE WORK

In this thesis, process mining algorithms have been examined from different perspectives, and a framework called Multi-Perspective Chain Miner (MPCM) has been proposed as a plugin to ProM tool. While the plugin has been developed, open-source algorithms in ProM have been used, and with additional improvements in the Data Aware Explorer plugin, performance and time information have been added in the Petri-net file with the data expression generated in the branches. After the application of the algorithms, it has been ensured that the outputs generated by the control-flow, data, and resource perspectives have been combined in a single process model in BPMN using an API that has generated the BPMN Data Model in blockchain. Then, the BPMN process model has been exported with the plugin developed within the scope of this study.

In this work, it has been observed that the difference of control-flow algorithms applied according to the event logs is important and affects the results. The selected algorithm for the control-flow perspective directly affects the results in terms of quality metrics. The quality of the skeleton obtained by the control-flow algorithm also affects the quality of the model enriched with data and resource perspectives. For this reason, before using the MPCM plugin on an example Sepsis data in medical domain, the algorithm and metric comparisons were made for the control-flow perspective and thus, the most suitable algorithm was chosen for the creation of the skeleton model. Since other perspectives were based on this skeleton, their quality was also positively affected in this way.

The advantages of the plugin, which is compatible with the ProM tool and developed within the scope of this thesis, compared to the multi-perspective plugins currently available in ProM can be mentioned as follows. Control-flow algorithms can be selected with the Multi-perspective Chain Miner plugin. Here, the event log is aimed to reach the most accurate skeleton process model for the user with the option of running different control-flow algorithms. Thus, when other perspectives are applied on this correct skeleton, a more correct solution is achieved. In addition, the feature of enhancing the data perspective with percentage, average time and instance count information makes the model more understandable and prevents possible problems that may occur in business process performance in the future. Moreover, perspective options are also offered within the scope of this plugin. Perspective outputs can be taken separately as BPMN models. With this feature, it is possible to take perspective printouts in the desired perspective combination. Also, this thesis is the first plug-in that allows the use of the ProM tool to keep process models as secure and immutable by integrating blockchain technology, thanks to the BPMN Data Model API. As a constraint in this study, there was a problem due to the lack of Organizational Miner in the latest versions of ProM, and for this reason, the Organizational Miner algorithm in the older ProM version was used. Since there are code design differences in the old and new versions of Prom, problems were encountered that would cause the plugin to run slowly. For this reason, if the resource perspective is chosen, the application has to work slowly. When the validation of the outputs is examined on the basis of ProM tool, it is seen that the outputs are compatible with MPCM. The output looks different because the pool feature of the BPMN is used and the pools can be positioned in different places during the output. However, when the activities and pools of the formed models are taken to similar places, it can be seen that the outputs are compatible with each other.

During the implementation of this plugin, it is ensured that the outputs obtained are kept in a way that cannot be changed with a certain id, by using blockchain technology indirectly, if not directly. The results of using the BPMN model API used in this way were automatically displayed. Manual steps have been automated using the BPMN Model API. In addition, a plugin has been developed that stores the BPMN model versions of the different perspective outputs resulting from the application of different algorithms on the blockchain. Accordingly, the unchanged state of the last output created is stored in the blockchain. This way, the security that the data does not change is guaranteed with the help of blockhain technology. If this data was stored in standard databases, certain controls would have to be added to understand that it would not change, but thanks to the blockhain's distributed structure, it can guarantee that the data will not change.

In future studies, it is planned to work on determining bottlenecks in process models with the help of Smart Contracts to be implemented in blockchain. As another future work, the

Organizational Miner algorithm can be added to the new version of ProM. In this way, the latest version of ProM can be used also for the resource perspective, and the performance problem can be solved by using the plugin. In addition to these, and finally, a plugin for selecting the most appropriate control-flow algorithm according to quality metrics can be developed in ProM. Then, this plugin can be added at the MPCM entry and enable the selection of the most appropriate control-flow algorithm according to the best quality metrics. Also, it is contemplated that inclusive gateways can be added to BPMN displays. However, this needs to be handled as a future work in the BPMN Model API. In addition, since the BPMN Model API is validated with the ProM tool, this improvement should also be made in the plugin that is actively used in the ProM tool. In other words, these improvements can be considered as future works in terms of contributing to both the BPMN Model API and the ProM tool. As the last observation, improvements can be made in the BPMN Model in order to use the BPMN features more (e.g. use of mentioned BPMN capabilities, Exclusive Gateway, Parallel Gateway, and Inclusive Gateway), which can also be considered for future works.

REFERENCES

[1] M. Dumas, M. La Rosa, J. Mendling, H.A. Reijers, Fundamentals of Business Process Management, 2013.

[2] A.J.M.M. Weijters, W.M.P. van der Aalst, A.K.A. de Medeiros;, Process Mining with the HeuristicsMiner Algorithm, Beta Work. Pap. 166 (2006) 1–34.

[3] M. De Leoni, W.M.P. Van Der Aalst, M. Dees, A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs, Inf. Syst. 56 (2016) 235–257. https://doi.org/10.1016/j.is.2015.07.003.

[4] F. Mannhardt, Multi-perspective Process Mining, in: BPM 2018 Diss. Award. Demonstr. Ind. Track, Sydney, 2018.

[5] J.M. Colom, J. Desel, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 7927 LNCS (2013). https://doi.org/10.1007/978-3-642-38697-8.

[6] J. Peterson, Petri Net Theory and the Modeling of Systems, Prentice Hall PTR, Upper Saddle River, NJ, USA, 1981.

[7] R.M. Dijkman, M. Dumas, C. Ouyang, Semantics and analysis of business process models in BPMN, 50 (2008) 1281–1294. https://doi.org/10.1016/j.infsof.2008.02.006.

[8] M.N. Tiftik, MultiperspectiveChainMiner, (2021). https://github.com/mervemisin/MultiperspectiveChainMiner (accessed April 13, 2021).

[9] W.M.P. van der Aalst, B.F. van Dongen, C.G.¨unther, A. Rozinat, H.M.W. Verbeek, A.J.M.M. Weijters, ProM: The Process Mining Toolkit, in: Proc. Bus. Process Manag. Demonstr. Track (BPMDemos 2009), Germany, 2009.

[10] B. Ekici, T. Erdogan, A. Tarhan, BPMN Data Model for Multi-Perspective Process Mining on Blockchain (under review), Data Knowl. Eng. (2021).

[11] Veri Bilimi Okulu, Veri Bilimi Okulu, (n.d.). https://www.veribilimiokulu.com/surec-madenciligine-giris/ (accessed April 21, 2021).

[12] W. Van Der Aalst, T. Weijters, L. Maruster, Workflow mining: Discovering process models from event logs, IEEE Trans. Knowl. Data Eng. 16 (2004) 1128–1142. https://doi.org/10.1109/TKDE.2004.47.

[13] T.G. Erdogan, A. Tarhan, A Goal-Driven Evaluation Method Based On Process

Mining for Healthcare Processes, Appl. Sci. 8 (2018) 894. https://doi.org/10.3390/app8060894.

[14] W. Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, Springer, 2011.

[15] T. Erdogan, A. Tarhan, Process mining for healthcare process analytics, in: Proc. - 26th Int. Work. Softw. Meas. IWSM 2016 11th Int. Conf. Softw. Process Prod. Meas. Mensura 2016, 2017: pp. 125–130. https://doi.org/10.1109/IWSM-Mensura.2016.027.

[16] M. Rovani, F. Maggi, M. De Leoni, W. Van der Aalst, Declarative process mining in healthcare, Expert Syst. Appl. 42 (2015) 9236–9251. https://doi.org/10.1016/j.eswa.2015.07.040.

[17] J.M. Colom, J. Desel, Discovering Block-Structured Process Models From Event Logs - A Constructive Approach, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 7927 LNCS (2013). https://doi.org/10.1007/978-3-642-38697-8.

[18] M. De Leoni, W.M.P. Van Der Aalst, Data-aware process mining: Discovering decisions in processes using alignments, Proc. ACM Symp. Appl. Comput. (2013) 1454–1461. https://doi.org/10.1145/2480362.2480633.

[19] Ahmed Adel Fares Gadelrab Mohamed, Process mining application considering the organizational perspective using social network analysis, 2016.

[20] M. Song, W.M.P. Van Der Aalst, Supporting process mining by showing events at a glance, WITS 2007 - Proceedings, 17th Annu. Work. Inf. Technol. Syst. (2007) 140–145.

[21] F. Mannhardt, M. De Leoni, H.A. Reijers, The multi-perspective process explorer, CEUR Workshop Proc. 1418 (2015) 130–134.

[22] W.M.P. van der Aalst, Process Mining: Data Science in Action, 2016.

[23] IEEE Task Force on Process Mining, Process Mining Manifesto, Bus. Process Manag. Work. (2011) 169–194. https://doi.org/10.1007/978-3-642-28108-2_19.

[24] R.N. A Rozinat, CW Günther, Process Mining and Automated Process Discovery Software for Professionals - Fluxicon Disco., (2017). https://fluxicon.com/disco/ (accessed April 13, 2021).

[25] E. Rojas, C. Fernández-Llatas, V. Traver, J. Munoz-Gama, M. Sepúlveda, V. Herskovic, D. Capurro, PALIA-ER: Bringing question-driven process mining closer to the emergency room, CEUR Workshop Proc. 1920 (2017) 1–5.

[26] About Celonis|The Leader & Innovator in Process Mining, (2018). https://www.celonis.com/company (accessed April 13, 2021).

[27] R. Gatta, J. Lenkowicz, M. Vallati, E. Rojas, A. Damiani, L. Sacchi, B. De Bari, A. Dagliati, C. Fernandez-Llatas, M. Montesi, A. Marchetti, M. Castellano, V.

Valentini, pMineR: An innovative R library for performing process mining in medicine, in: AIME 2017, 2017: pp. 351–355. https://doi.org/10.1007/978-3-319-59758-4_42.

[28] G. Janssenswillen, B. Depaire, M. Swennen, M. Jans, K. Vanhoof, bupaR: Enabling reproducible business process analysis, Knowledge-Based Syst. 163 (2019) 927–930. https://doi.org/10.1016/j.knosys.2018.10.018.

[29] U. Çelik, E. Akçetin, S. Madenciliği Araçları Karşılaştırması ÖZ, A. Kelimeler, V. Madenciliği, İ. Zekası, S. Madenciliği, Process Mining Tools Comparison, Online Acad. J. Inf. Technol. 2018-Cilt. 9 (2018) 34. https://doi.org/10.5824/1309.

[30] A. Kalenkova, A. Burattin, M. de Leoni, W. van der Aalst, A. Sperduti, Discovering high-level BPMN process models from event data, Bus. Process Manag. J. 25 (2019) 995–1019. https://doi.org/10.1108/BPMJ-02-2018-0051.

[31] D. Bell, UML basics : An introduction to the Unified Modeling Language A little background, Ration. Softw. (2003) 1–11. http://www.therationaledge.com/content/jun_03/f_umlintro_db.jsp.

[32] J. Mendling, M. Nüttgens, EPC markup language (EPML): An XML-based interchange format for event-driven process chains (EPC), Inf. Syst. E-Bus. Manag. 4 (2006) 245–263. https://doi.org/10.1007/s10257-005-0026-1.

[33] M. Geiger, S. Harrer, J. Lenhard, G. Wirtz, BPMN 2.0: The state of support and implementation, Futur. Gener. Comput. Syst. 80 (2018) 250–262. https://doi.org/10.1016/j.future.2017.01.006.

[34] J.C.A.M. Buijs, B.F. Van Dongen, W.M.P. Van Der Aalst, On the role of fitness, precision, generalization and simplicity in process discovery, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 7565 LNCS (2012) 305–322. https://doi.org/10.1007/978-3-642-33606-5_19.

[35] D. Dakic, D. Stefanovic, I. Cosic, T. Lolic, M. Medojevic, Business process mining application: A literature review, Ann. DAAAM Proc. Int. DAAAM Symp. 29 (2018) 866–875. https://doi.org/10.2507/29th.daaam.proceedings.125.

[36] W.M.P. Van Der Aalst, A.K.A. De Medeiros, A.J.M.M. Weijters, Genetic process mining, 2005. https://doi.org/10.1007/11494744_5.

[37] P. Weber, B. Bordbar, P. Tiňo, A principled approach to the analysis of process mining algorithms, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 6936 LNCS (2011) 474–481. https://doi.org/10.1007/978-3-642-23878-9_56.

[38] R. Ghawi, Process Discovery using Inductive Miner and Decomposition, 2016. http://arxiv.org/abs/1610.07989.

[39] A. Rozinat, W.M.P. Van Der Aalst, Decision mining in ProM, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 4102 LNCS (2006) 420–425.

https://doi.org/10.1007/11841760_33.

[40] R.S. Mans, M.H. Schonenberg, M. Song, W.M.P. Van Der Aalst, P.J.M. Bakker, Application of process mining in healthcare - A case study in a Dutch Hospital, Commun. Comput. Inf. Sci. 25 CCIS (2008) 425–438. https://doi.org/10.1007/978-3-540-92219-3_32.

[41] M. Bozkaya, J. Gabriels, J.M. Van Der Werf, Process diagnostics: A method based on process mining, Proc. - Int. Conf. Information, Process. Knowl. Manag. EKNOW 2009. (2009) 22–27. https://doi.org/10.1109/eKNOW.2009.29.

[42] W. Van Der Aalst, Service mining: Using process mining to discover, check, and improve service behavior, IEEE Trans. Serv. Comput. 6 (2013) 525–535. https://doi.org/10.1109/TSC.2012.25.

[43] E. Gupta, Process mining algorithms, Int. J. Adv. Res. Sci. Eng. (2015).

[44] M. Bozkaya, J. Gabriels, J.M. van der Werf, Process Diagnostics: A Method Based on Process Mining, in: 2009 Int. Conf. Information, Process. Knowl. Manag., IEEE, 2009: pp. 22–27. https://doi.org/10.1109/eKNOW.2009.29.

[45] F. Folino, G. Greco, A. Guzzo, L. Pontieri, DISCOVERING MULTI-PERSPECTIVE PROCESS MODELS, (2008) 70–77. https://doi.org/10.5220/0001705700700077.

[46] S. Schönig, C. Di Ciccio, F.M. Maggi, J. Mendling, Discovery of multi-perspective declarative process models, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 9936 LNCS (2016) 87–103. https://doi.org/10.1007/978-3-319-46295-0_6.

[47] C. Sturm, S. Schönig, C. Di Ciccio, Distributed multi-perspective declare discovery, CEUR Workshop Proc. 1920 (2017) 1–6.

[48] S. Jablonski, M. Röglinger, S. Schönig, K. Wyrtki, Multi-Perspective Clustering of Process Execution Traces, Enterp. Model. Inf. Syst. Archit. 14 (2019) 1–22. https://doi.org/10.18417/emisa.14.2.

[49] R. Sikal, H. Sbai, L. Kjiri, Promoting resource discovery in business process variability, ACM Int. Conf. Proceeding Ser. Part F1481 (2019). https://doi.org/10.1145/3320326.3320380.

[50] A. Pini, R. Brown, M.T. Wynn, Process visualization techniques for multi-perspective process comparisons, Lect. Notes Bus. Inf. Process. 219 (2015) 183–197. https://doi.org/10.1007/978-3-319-19509-4_14.

[51] S. Schönig, C. Di Ciccio, F.M. Maggi, J. Mendling, Discovery of Multi-Perspective Declarative Process Models Stefan, Int. Conf. Serv. Comput. 9936 (2016) 157–170. https://doi.org/10.1007/978-3-319-46295-0.

[52] B. Ekici, A BPMN Data Model to Keep a Multi-Perspective Process Model on the Blockchain, Hacettepe University, 2021.

[53] B. Ekici, A. Tarhan, A. Ozsoy, Data Cleaning for Process Mining with Smart Contract, UBMK 2019 - Proceedings, 4th Int. Conf. Comput. Sci. Eng. (2019) 324–329. https://doi.org/10.1109/UBMK.2019.8907140.

[54] T. Slaats, Flexible Process Notations for Cross-organizational Case Management Systems, 2015. https://pure.itu.dk/portal/files/81016376/PhD_thesis_final_version_Tijs_Slaats.pdf#page=75.

[55] A.J.M.M. Weijters, J.T.S. Ribeiro, Flexible heuristics miner (FHM), 2011. https://doi.org/10.1109/CIDM.2011.5949453.

[56] J.P. Katoen, B. König, Causal Nets: A Modeling Language Tailored towards Process Discovery, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 6901 LNCS (2011). https://doi.org/10.1007/978-3-642-23217-6.

[57] W.M.P. Van Der Aalst, V. Rubin, H.M.W. Verbeek, B.F. Van Dongen, E. Kindler, C.W. Günther, Process mining: A two-step approach to balance between underfitting and overfitting, Softw. Syst. Model. 9 (2010) 87–111. https://doi.org/10.1007/s10270-008-0106-z.

[58] F. Mannhardt, D. Blinde, Analyzing the trajectories of patients with sepsis using process mining, CEUR Workshop Proc. 1859 (2017) 72–80.

[59] Á. Rebuge, D. Ferreira, Business process analysis in healthcare environments: A methodology based on process mining, Inf. Syst. 37. (2012).

[60] N. Gupta, K. Anand, A. Sureka, Pariket: Mining business process logs for root cause analysis of anomalous incidents, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 8999 (2015) 244–263. https://doi.org/10.1007/978-3-319-16313-0_19.

[61] POSTMAN, POSTMAN, (n.d.). https://www.postman.com/ (accessed April 26, 2021).