Using rules based Process Mining to discover PLM system processes

Abstract

Process Mining (PM) is defined by a set of techniques used in Business Process Management that combines computational intelligence and data mining with process analysis and process modeling. The growing interest in PM is based on the ability to discover, monitoring, and improve the processes model. To reach this objective knowledge is extracted from the Event Logs generated by Information Systems, and quality metrics are used to evince the quality of the matching obtained when replaying a process model against the event log. The application of PM to logs extracted from PLM systems is an almost unexplored topic in this research area. Our study enhances the application in the field of PLM with the use of business rules to filter the log, verifying the BRs impact on PM metrics in order to minimize the divergences between modeled processes and executed one and to increase the resulting quality metrics. This helps the business user to identify a line of investigation for explaining occurring misbehavior and propose alleviation/improvement measures. Our approach is finally validated on data provided by an industrial company, by confirming the impact that controlling the business process characterizations via BR can decrease the gap between the expected modeled process and the executed one.
Keywords: Process Mining, Business Rules, Business Process Assessment, Product Lifecycle Management

1. Introduction

The value of Product Lifecycle Management (PLM) is increasing, especially for manufacturing, high technology, and service industries [1]. It aims to trace and manage all the activities and flows of data and information during the product development process and after during the actions of maintenance and support [2]. PLM enables organizations to collaborate within and across the extended enterprise, integrating people, processes, and technologies and assuring information consistency, traceability, and long-term archiving [3]. For reaching an effective PLM, a company needs consistent and proper organized processes through the lifecycle of a product [1] that are followed and implemented by people and systems. The set of information needed to manage the product in relation to the procedures put into action in an organization is integrated into data, processes, and business systems.

Business Process Management is the discipline providing tools and methods for support the analysis, design, execution, monitoring and optimization of processes. It allows to better understand the different actions of the involved people and address improvements. Much BPM research is concerned with the conformance or the compliance of business cases to a reference business model or a set of rules translating directives, standards, or regulations. In BPM, deviations of processes from a reference model is traditionally studied with anomaly detection [4]. The reference to a business process model is significant as organisations are driven by normative conformity and adherence to plans. A model is an abstraction and can specify multiple occurrences for the same activity, with loops and parallel flows. This way, cases that significantly differ in the number of activities can still comply to the same model [5] In other words, model-aware analytics, such as Process Mining (PM) [6], generalise better than traditional statistics or data mining techniques in the BPM domain. Process Mining (PM) is a set of techniques used in BPM that combines computational intelligence and data mining with process analysis and process modeling. The growing interest in PM is essentially due to the enormous availability of process data and the need to improve and support business processes in rapidly changing environments [7]. Industrial contexts, where complex products are produced and the analysis of a large amount of information is required, are therefore particularly suitable for the application of PM techniques [8, 9]. The discovery, monitoring, and improvement of processes model are the main scope of PM [10, 11]. To reach this objective knowledge is extracted from the *Event Logs* generated by Information Systems.

An Event Log is a collection of *events* generated in a temporal sequence and stored according to some descriptive *attributes* such as the timestamp, the originator, and the associated resources. To completely benefit from PM techniques, also further business data obtainable from the IS of a company are connected to the observed events [12]. Events are aggregated by *case*, i.e. the end to end execution of a business processes, this representation is used to infer process-oriented models, statistical data, and performance indicators [13, 14, 15]. Among the different PM techniques Process Discovery (PD) and Conformance Checking (CC) are the most treated in literature [16]. PD enables to discover a process model from an event log of observed cases. CC is used to provide a measure of the divergences between an event log and a process model and to obtain this way a measure of the quality of models or business case executions. Common quality metrics applied in process mining are simplicity, generalization, fitness, and appropriateness. However, the literature has largely focused on *Fitness* and *Appropriateness* that offer the best generalization capacity [17].

In this study, the metrics of fitness and appropriateness are applied [18, 19], to measure the degree of covering each instance of the event logs in the process model (i.e. fitness) and the degree of accuracy in the process model description of each event logs traces (i.e. appropriateness). Both these metrics evince the quality of the matching obtained when replaying a process model against the event log. Quality assessment and results by using CC metrics have some severe limitations: 1) matching-based metrics are compensatory and it is difficult to define the variability of two different event logs respect to similar matching processes; 2) there aren't certain connections between the model-log matching and the business features of the executed process (e.g. the effectiveness or the cost). Quality metrics, therefore, suggest insufficient interpretation on how to use PM results as practicable and comprehensible business information. To overcome these drawbacks and provide practical evidence for the PLM field, our study integrates PM results with Business Rules (BR) followed by an organization [20, 21] focus on PLM and related system (PLMS).

A case study is carried out on the workflow processes of an aerospace company regarding the approval steps required to advance an item from one release status to the next during the design process. The activities are managed on a PLMS with different interactions among several organizational roles involved in product design. For the data analysis, a log file extracted by the PLMS and related to the preliminary design phase of a new product is used.

Results confirm the impact that controlling the business process characterizations via BR can decrease the gap between the expected modeled process and the executed one [22]. Besides, considering the process executions with a high degree of divergence respect to the model, it is possible to improve the awareness of the reasons generating the gap [23, 24].

The paper is organized as follows. The next Section 2 explains the rationale of the study in the PLM field. Section 3 treats an overview comprising definitions and evidences already presented in the Business Rules literature. The design is instead presented into Section 4, together with a description of the objectives, context, log file description, and a presentation of different methodologies of data analysis. A deeper and detailed data analysis is then reported into Section 5, with definition of the categories and rules, their applications, and a performance analysis discussion. Concluding remarks are finally reported into Section 6.

2. Research Rationale

Product Lifecycle Management System (PLMS) as an information technology (IT) processing system, or a set of IT-systems, enables PLM. It is a connecting technology, a collaborative backbone, that integrates products, processes, tools, and technologies and that allows people, also of different companies, to work together more effectively [1]. PLMS creates an organizational substructure that identifies and connects all the functional areas to product-centric data that the organization needs for managing engineering activities [25].

The application of PM to logs extracted from PLM systems is an almost unexplored topic. The authors in [26] have proposed PM as a method to discover inefficiencies and improvements of processes in the use of PLM systems by an automobile module manufacturer highlighting the need to extend their analysis to all the different PLM processes. A further study of Rigger and colleagues [27] presents a method to validate the alignment of IT systems and related PLM processes using enterprise architecture and process mining. Our study enhances the application in the field of PLM with the use of business rules to filter the log, verifying the BRs impact on PM metrics in order to minimize the divergences between modeled processes and executed one.

When organizations execute business processes, such as those managed in PLMS, several constraints (i.e. organizational policies, internal or external regulations, and standards) have to comply. Formal procedures have been adopted to document them in terms of conditional statements, typically referred to as Business Rules (BR). They have increasingly become a subject of interest for organizations seeking solutions to leverage business process specifications. Collecting and documenting BR or enforcing their control have been proved to support the business ecology in several ways [28]. In this view, our paper is based on the use of PD to extract a process model, the filtering of instances of execution based on BR, and then the application of CC. We can plot the change in CC results on the process executions (the log segments) selected by a given BR, i.e. focus on the specific characterization of the process lead to specific BR.

Our technique avoids applying CC on the entire event log, filtering out noise elements (processes too short, aborted, looped, never finished) and other distortions, such as executions that have delays. In this way, we support a knowledge acquisition process for the non-computer-savvy user, giving him an active role in selecting the log segments to be considered during the generation of the process model and the computation of the CC metrics. This is achieved by filtering the log based on well understood BRs. A preliminary study is described in Ceravolo et al [29]; it is a first analysis focus on a smaller dataset. This paper wants, through a case study method, to enlarge the collected evidence and better support and finalize how BRs can be used for a more valuable understanding of the process model and their feedback and impacts on the practice.

3. Theoretical Background on Business Rules

3.0.1. Definition

The main purpose of BR is to define the semantics of all the business concepts involved into a process model, such as all the conditions over tasks carried out, as well as the rights and constraints that are applied on it [30].

The literature shows that BR can be considered a key component of business process management due to the support they provide to business processes (BP) execution and monitoring [31]. Even if rules are applied to across processes and procedures, they are a clear constraint on behavior and/or they provide behavior support. As a result, a rule defines the perimeter between adequate and un-adequate business activities and related business goals and objectives. Consequently, BR may model any situation determined by states or state transitions that are mandatory, permitted, expected or forbidden in a business domain [32, 33]. Nevertheless, for the usage on the BP modeling level, the BR must be specified in a well-structured and formal language. For our analysis, a BR is defined as a logical statement composed by predicates, variables, and constants. It can be represented as the Equation 1:

$$BR\# : \textsf{\textbf{each} \underline{authorized task} \textbf{has an} \underline{Owner}} \tag{1}$$

The resulting logic statement is as follow (Equation 2):

$$Task(x) \wedge Authorized(x) \wedge Owner(y) \wedge hasOwner(x, y) = 1 \tag{2}$$

3.1. Evidences

It is well known that BRs need to be expressed in a structured formal language, able to prevent rules ambiguity while keeping good readability. Recently, the OMG [34] adopted Semantics of Business Vocabulary and Business Rules (SBVR) as the standard language for representing BR [35]. In fact, the SBVR meta-model allows business professionals to describe the organizational policies and rules clearly, unambiguously and convertible into further representations. The SBVR model has been presented as a result of the request for proposal on Business Semantics of Business Rules (BSBR) made by OMG [32]. SBVR is intended to model and capture the semantics of business facts and business rules that are expressed either explicitly or implicitly [36]. SBVR is also responsible for defining the domain concepts exploited by BR. For example, for regulating the use of data sources in the organization it is required to model concepts such as data collection, license, copyright, and patent [37].

BR are also classified according to two main modalities, called respectively *alethic* or *deontic* [38]. The first type of rules, Alethic, is used to model necessities (e.g. implied by physical laws) which cannot be violated. Deontic rules, instead, are used to model obligations (e.g., resulting from company policy) which ought to be obeyed, but may be violated in real-world scenarios.

By considering OMG, Popp and his colleague [39] presented a novel approach showing changes and main capabilities of business process models based on model transformation. In such a contribution the model transformation employs rules that transform a given source model to a related target model, according to precisely specified meta models. Therefore, the focus of their approach consists of an automated refinement of a high-level reference process. Focusing on others aspects related to the BR, Kherbouche et al [40] expressed compliance rules into a graphical, hence more readable, model in order to improve the automation process in Compliance Checking, while Caron et al [41] aimed at providing an adequate guidance to the business users that need to determine the level of compliance with directives of the business environment into BR. Generic rule patterns are classified according to their process mining perspective and their rule restriction. In that case, a user should be able to perform control effectiveness assessments on a broader spectrum of common control types. The integration between business processes and BR is also discussed in the literature. For example, in Zhao et al [42], based on the notion that different rule sets may coexist in an application with large-scale rules, the authors proposed multiple bypass processes, invoked by the first one with the corresponding variable objects, each of them responsible for the integration with a defined rule set. Such an integration was aimed for all the applications that not only hold numerous business knowledge or policies but also need the intercommunication among some distributed and heterogeneous components. Ceravolo & Zavatarelli [12] described how to relate business data with events and cases based on conjunctive statements that link these elements. A BR is represented as a query filtering an event log following defined parameters. They can constraint the events execution order, the participation of a role in an event, and the value of event outputs [43]. In a recent paper [44], BRs are identified as a key element to implement the interface between human decision-makers and AI components. Our work goes in the same direction with a focus on PM.

4. Research Design

4.1. Objectives

Companies working as OEM (Original Equipment Manufacturers), such as the first tier supplier of the aerospace supply chain, generally adopts a PLMS to track and manage all data, information, models and workflows related to items of a product design phase. Items can be the whole product system, a component, an assembly or an installation. Each component item passes different status before to be approved and sent to manufacturing. Several organizational roles are involved and contributes to evaluate, change or improve an item. The process of evaluating the different status of an item and decide to continue in the next steps of design, until is ready to be sent for manufacturing, is named Release Process. Process mining becomes relevant for understanding, monitoring and formalizing workflows, iterations between and among roles and related weight on the overall process. Business process modeling and BRs assume an important role in PLM [45, 46] as they constrain the acceptance and release procedures of the documentation supporting product development.

Several studies exist in literature describing the relevance of PM to collect and identify process behaviours [47]. These studies use PM to analyze multiple perspectives including control-flow, organizational, temporal constraints and performance results. Indicators are generated to assess cases and apply redesign procedures when the levels achieved are not satisfactory[8]. To the best of our knowledge none of the analysed studies is focalized on a log file extracted from a PLMS and useful to better understand the release process of the product design. With these premises, the overall objective of the study is to define and verify if BRs impacting on PM metrics, minimizing the divergence between modelled/known processes and practically executed ones . Furthermore, the study wants to evaluate the use of log analysis for gathering feedback about:

- lead time-frequency, average, and standard deviation for different workflows implemented in the PLMS,
- percentage of rejected items and explanation of the reject decision,
- lead time for each approval activity.

Since the proposed approach can be applied to any information system, the study wants to show practitioners how to apply process mining in an industrial context with the support of business rules.

For addressing these research objectives, a case study research is carried out. This type of research method supports the analysis of a problem, an issue or a given situation in its real context and it is particular suitable for understand how specific activities are addressed to be transferred to case with common characteristics [48, 49].

For better and sound results, the research team is composed by managerial and computer science engineers of two different Universities. Knowledge about the industrial context and the PLM are provided by the managerial engineers and are integrated with the specific knowledge on Process Mining and related algorithms of the computer science ones.

4.2. Research Context

For the case study, a log extracted by the PLMS of an Italian manufacturing company is used to assess the performed business processes. The log refers to the release process of an engineering item that can be in one of the statuses reported in Table 1.

Table 1: Process Release Status.

Status	Definition
Ready to Work	The item is created and available for working.
Initial Released	Advanced information for manufacturing is produced and available.
Frozen	Data are freezed and unchangeable to be transferred into a different system
Final Released	The item is ready to be introduced in the following process, such as manufacturing.

When a designer submits an item for review and approval, the release process has its beginning. Several roles are involved and each one checks different issues (e.g. accuracy, consistency) based on their competencies and authority. Indeed, before to release an item to subsequent steps, the roles in charge of its design assess it. The log is collected during the preliminary phase of a new product design in collaboration with a prime contractor. After the "Ready to Work" status, an item can move to the "Initial Release" status or the "Frozen" status. This choice depends on if it is needed to export the data in order to be approved using an external system by a partner company. Indeed, the engineering items are released in the intermediate Frozen status to allow the export and approval of data on the prime contractor's systems. Additionally, crucial issues could emerge by evaluating the detailed models: they can be approved or revisions can be suggested. When the "Initial Release" status or "Frozen" status are overcame, the "Final Released" status can be assigned to the item.

As illustrated in the following Table 2, the transition from one status to the other one implies a sequence of activities executed by different roles, described as follows:

- The *Designer* is the person in charge of the creation of a specific engineering item (e.g. to create the CAD model of a component).
- The *Design Leader* is the head of a team of designers working on a specific product component.
- The *Configuration Manager* (CM) is responsible for tracking the different configuration of specific product components in the different phases of development.
- The *Release Manager* is focused on the formal release of the designed parts that has to follow a set of predefined approval steps.
- The *Supervisor* collects, if it is required by quality policies in the design of the components, the results of several checks (e.g. weight and stress analysis, material planning).
- The *Process Manager* manages the process of evaluation of a release.

Depending on the specific engineering item (e.g. detail part, assembly, installation) and whether or not any suppliers are involved in the design process, different workflows are implemented in the PLMS to manage the transition from one status to another. All release workflows require two main signatures by Design Leader and Configuration Manager. If during the process an item is rejected by anyone of the involved roles, it returns in the Ready to work status. In this case, the designer needs to address the corrective actions and then re-launches the workflow.

4.3. Log File Description

The data set was about of 100 MB; it is obtained collecting, cleaning and consolidating the data in a standard event log format [1]. Table 2 provides an

[1]The eXtensible Event Stream (XES), proposed by IEEE Task Force on Process Mining, is the standard for describing event logs and event streams.

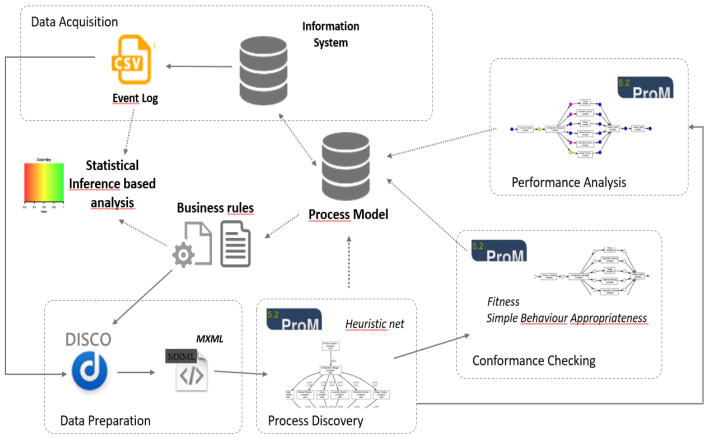


Figure 1: An overview on the techniques integrated in our research study.

extract of the event logs. Each event is described with five fields: *Case ID*, *Trace*, *TimeStamp*, *Resources*, and *Workflow type*.

4.4. Methods of Data Analysis

Data Analysis is used in case study to understand what happened in the research context and derive conclusion from the data leading also future applications [48]. In this specific case study, data analysis is based on the application of five steps in a chain of, sequential and linked, evidences that support the verification of our research objectives. Two software are used for supporting data analysis ProM and Disco; their dual use allows researchers to exploit and connect their potentialities. Therefore, data are analysed following 5 steps (Figure 1): *Data Acquisition*, *Data Preparation*, *Data Discovery*, *Conformance Checking* and then *Performance Analysis*.

The Data Acquisition step is related to gathering the event log in a CSV format by the company information system. It is followed by Data Preparation step performed with the support of DISCO [50] to extract a MXML for enabling further analysis. To support this step and filter the log also BRs are defined looking to Process model of the release approval workflows. After that, Process Discovery step is run on ProM using the *Heuristic Miner algorithm* [51]. It is the recommended algorithm while dealing with real-life data. The resulting *Heuristic Net* allows researchers to get a first view of processes really executed in the system. It can also be converted to other types of process models, such as a Petri net useful for further analysis in ProM. Petri Net is the required process model form of the *ProM Conformance Checker algorithm* used in the next step of Conformance Checking for the measurement of Fitness and Simple Behaviour Appropriateness metrics. The Petri net converted from the Heuristic net is also used as input in the performance analysis step. We use the *ProM Performance Analysis with Petri Net* plugin to assess the time performance of processes. The Process Discovery, Conformance Checking and Performance Analysis steps contributes to enlarge the knowledge on the executed process.

5. Data Analysis

5.1. Rules definition

To generate the process model, the event logs file extracted by the PLMS is used for the Process Discovery. The PM tool DISCO [50] has been used for the data preparation step. A set of BR described in SBVR was selected in collaboration with the company as a starting step. The complete set is available in Table 4. Respect to the previous preliminary study [29], having a larger dataset, we add the SBVR definition and we enlarge the set of rules.

The first two rules (BR0 and BR1) filtered out invalid or incomplete records. In particular, the BR0 rule has the effect of cleaning up the log by excluding from it instances for which the approval process was not actually carried out but which were brought to the final state directly from the PLMS. These processes are not significant and could be generated by human errors or for testing the system functionalities. BR1 instead allows to select only complete process instances, in which all the participants approved the engineering item. Four rules (BR2, BR3, BR4, and BR5) illustrate characteristics that, for the company management, have a positive impact on the quality of the items in output; they refer to the duration of a task or the whole process instance, and to the actor performing a given task. The last rules (BR6, BR7, BR8, and BR9) instead consider the type of product component on which the release workflow is applied, the involvement of suppliers, or finally they allow to select a specific release procedure implemented in the PLMS.

Figures 2 and 3 represent the model extracted from the filtered log (respectively applying BR7 and BR9) by using the Heuristic Miner algorithm of the ProM Process Mining Framework [52, 37, 53], an application tool supporting different PM techniques implemented as plug-ins.

The two selected rules allow the analysis of two specific release workflows, one of which also involves suppliers in the approval process; both refer to the same engineering item type (detail) and allow the transition from *Ready to work* to *Frozen* status. Activities have associated values that correspond to the frequencies of execution; the integer values on the arcs suggests the flow frequency, while, decimal values represent the dependency (i.e. flow's likelihood to occur in the analyzed process). In the representation, all the participating roles and their order of involvement are also identified. As an example, looking at the process, it every time begins by the designer which is the owner of the item to be released.

As illustrated in Figure 2, extracted from the log, after the first initialization represented by Start task (A) activity, the Process Creation (B) is executed. Thereafter in the case of approval of the Configuration Manager (C), a set of parallel activities can be performed and involve: Stress Analysis (D), Material Planning (E), Fabrication Check (G), Weight Analysis (F), Assembly Check (H), Design Data Check (I), then the approval of Design Leader (L) and finally, after the Release confirmation (M), an artificial End Task (N) as the end. The operator addressing one of the activities receives the approval request from the PLMS and then appraises the available information and decides to approve the release of the engineering item or to reject it providing an explanation. Different types of workflows may be characterized by the presence of additional tasks not included in the process described above. As in the process illustrated in Figure 3, the approval of the configuration manager is preceded by three distinct tasks,

Table 2: Log Example.

Case ID	Trace	TimeStamp	Resource	Workflow type
1	(A) Start	16/12/2009 10:43	Process Manager	Detail_RW-Frozen
	(B) Process Creation	17/12/2009 17:47	Designer	
	(C) Configuration Manager	18/01/2010 12:49	CM	
	(D) Stress	08/02/2010 09:54	Supervisor	
2	(A) Start	29/04/2010 09:28	Process Manager	Detail_RW-InitialRel
	(B) Process Creation	30/04/2010 15:40	Designer	
	(C) Configuration Manager	03/05/2010 09:49	CM	
	(F) Weight	04/05/2010 16:29	Supervisor	
	(L) Design Leader	11/05/2010 15:42	Supervisor	
3	(A) Start	11/05/2010 15:42	Process Manager	Detail_Frozen-FinalRel
	(B) Process Creation	13/05/2010 16:49	Designer	
	(N) ArtificialEnd Task	14/05/2010 16:40	Process Manager	
4	(A) Start	14/05/2010 16:40	Process Manager	Detail_RW-Frozen
	(B) Process Creation	17/05/2010 17:40	Designer	
	(C) Configuration Manager	20/05/2010 11:16	CM	
	(E) Material Planning	20/05/2010 17:11	Supervisor	
	(G) Fabrication Checker	21/05/2010 09:59	Supervisor	
	(I) Design Checker	22/05/2010 11.00	Supervisor	
	(M) Release state	23/05/2010 16.30	Release Manager	
	(N) ArtificialEnd Task	25/05/2010 12.55	Process Manager	
5	(A) Start	27/05/2010 09:43	Process Manager	Detail_Supplier_RW-Frozen
	(B) Process Creation	28/05/2010 12.34	Designer	
	(C) Configuration Manager	29/05/2010 11.26	CM	
	(D) Stress	30/05/2010 09:51	Supervisor	
	(G) Fabrication Checker	31/05/2010 13.11	Supervisor	
	(I) Design Checker	09/06/2010 15.23	Supervisor	
	(L) Design Leader	10/06/2010 12.50	Supervisor	

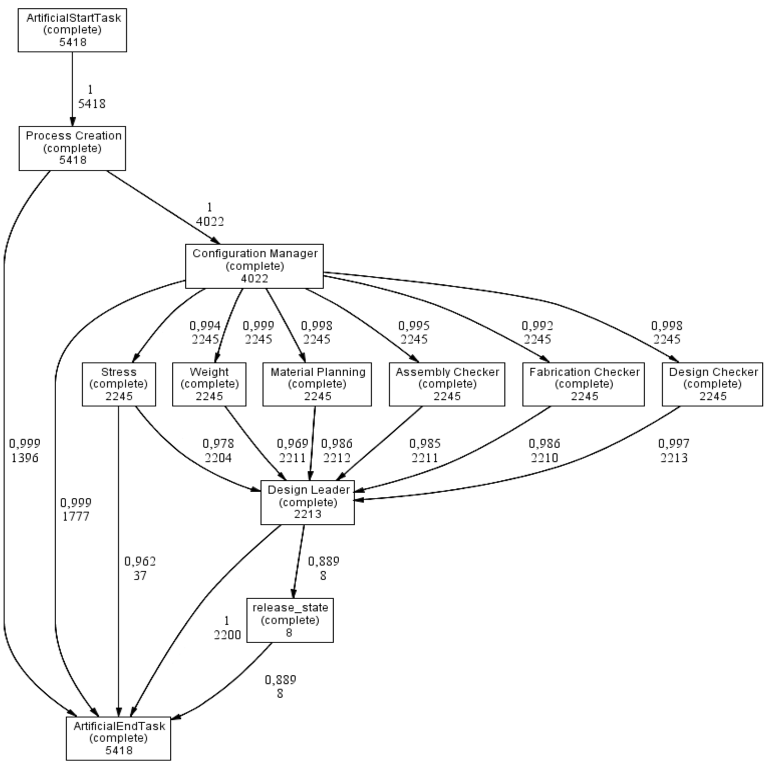


Figure 2: Model deriving from the 'ready to work to the frozen status' log (BR7).

namely the approval of the work-package leader and two roles of the supplier (designer and stress analyst).

SBVR provides the vocabulary to describe the business rules. It is represented by different concepts, as terms, names, verbs and specific keywords. A Term is a noun concept and it is represented by a word or a group of words used to represent a business entity. Consequently, a Name is an individual concept and it is represented by a word or a group of words which can be used to represent an instance of a particular term. A Verb establishes a relationship between the terms and/or the names. Finally, Keywords correspond to linguistic symbols to construct statements. An extraction of the vocabolary used for the implemented BR is reported in Table 3.

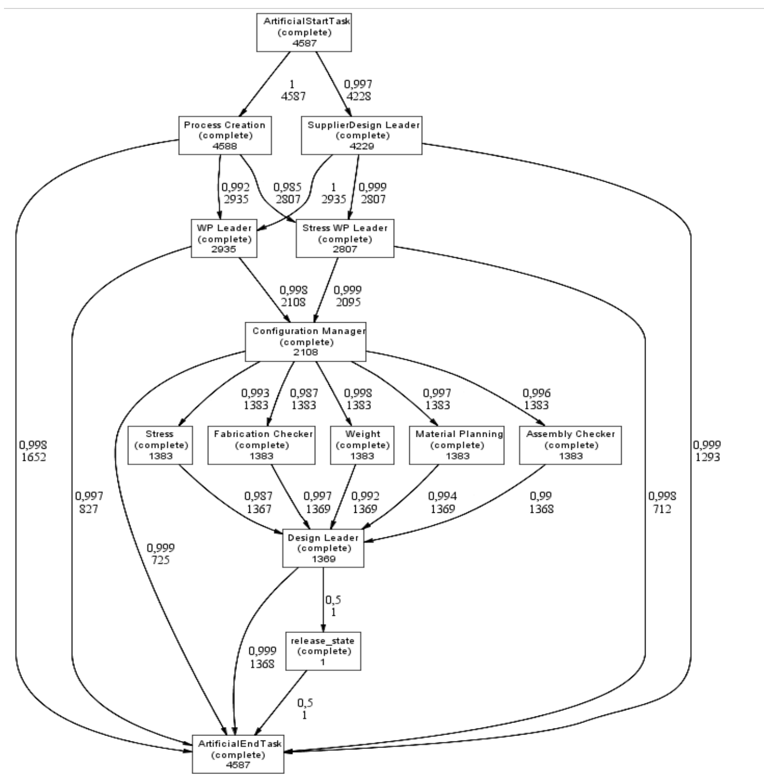


Figure 3: Model deriving from the ‘ready to work to the frozen status’ workflow involving suppliers (BR9).

Table 3: SBVR Vocabulary Definition.

Terms (Names)	Verbs	Keywords
process instance	is followed by	in, each
activity (Process_Creation, Release_State, Configuration_Manager)	is the	who, the, last
duration (1 minute, 1 day, 10 days)	is defined	between .. and..
person (Resource X, Resource Y)	must be	first,
role (CM, Process_Manager)	is shorter than	more often
workflow_name (Deatil_RW-Frozen, Detail_Supplier_RW-Frozen)	is performed by	
	has	
	contains	
	is equal to	

Table 4: Business Rule Definition.

Business Rule (Filter)	Selection Process Description	SBVR
BR0	Processes closed without approval request	In each process instance the activity 'Process creation' is followed by the activity 'Release_state'.
BR1	Completed processes: all actors approved the item.	BR1: the last activity is 'Release_state'.
BR2	Process with rational time elapsed: in the range of 1 minute to 10 days.	BR2: each process instance duration is defined between 1 minute and 10 days
BR3	First process activity ended by one day (time elapsed from start activity and the first signature).	BR3: in each process instance the first activity must be shorter than 1 day
BR4	The CM activity is executed by the user that most frequently performs this step.	BR4: the CM activity is performed by *'ResourceX'* *'ResourceX'* is the person who more often has the role of Configuration Manager
BR5	The process is started by the resource that most frequently performs this step.	BR5: the Process Creation activity is performed by *'ResourceY'* *'ResourceY'* is the person who more often has the role of Process Manager
BR6	Select Processes only applied to Detail Product Types.	BR6: in each process instance the workflow_name contains 'Detail'.
BR7	Select only Processes with Procedure Name = 'Detail_IW-Frozen'	BR7: in each process instance the workflow_name is equal to 'Detail_RW-Frozen'
BR8	Select the Processes that involve 'Supplier'.	BR8: in each process instance the workflow_name contains 'Supplier'.
BR9	Select only Processes with type = 'Detail_Supplier_IW-Frozen'	BR9: in eachprocess instance the workflow_name is equal to 'Detail_Supplier_RW-Frozen'

5.2. Rule Categories

Essentially, the PD is first proceeded by a data preparation step. In the analysis, the cases recorded in the event log are filtered by the BR available in Table 4 by applying specific filters in Disco and categorizing BRs based on two dimensions. BR can be categorized, in fact, according to their process mining perspective and their rule restriction focus. Process Mining Perspective Dimension [41] refers to the following four different perspectives on business process modeling:

- Functional process perspective, that deals with the process elements (such as activities, events, etc.) that are being performed in a process instance, as well as the relevant process artifacts linked to these process elements (e.g. an invoice artifact for a paid activity).
- Control-flow process perspective that refers to the ordering of activity in a process instance (i.e. this includes conditions on complex decisions and entry & exit criteria).
- Organizational process perspective related to the organization leading the business process (e.g. the performers that are involved).
- Data process perspective (also known as informational perspective) that represents the informational elements (e.g. event data and case date) that are used, produced or manipulated during the process, as well as relationships among them.

Secondly, business rules can be classified along their main rule restriction focus [41] as described below:

- Cardinality-based rules are business rules that restrict the number of allowed instances of a specific process element type in a specific process instance.
- Coexistence rules can be defined as business rules that restrict the coexistence of process elements of different types over the execution of a specific process instance.

- Dynamic data-driven rules specify the influence of certain data elements (i.e. case or event data) and their value on the occurrence of process elements in a specific process instance.
- Relative time rules focus on specifying a time restriction on process elements relative to certain points in a process execution (e.g. start of a process or completion of a specific activity).
- Static property rules deal with specifying a specific property for a particular type of process element at a predefined process state.

The case study rules categorization and association with filter type is shown in Table 5. In this table BR0 is not considered as it simply filter out irrelevant cases. The first three columns identify the business rules by a description and an association with the rule category defined in the case study. The classification of the BRs based on the Process Mining Perspective Dimension and on rule restriction dimension proposed in [41] is also provided in the fourth and fifth column. The last column of the table shows the filter used in DISCO [54] in order to apply BR on the event log. The Endpoints filter allows determining what should be the first and the last event in the process and it removes all incomplete cases. Applying BR1 requires setting the 'release confirmation task' as the last activity. The Performance filter is a case filter that allows focusing on cases in data set according to certain performance criteria (e.g. in BR2 it is case duration that is the time between the first and the last event in each case). The Follower filter specifies a simple process pattern based on the so-called follower relation. This requires that a certain activity (or other event value) must follow the reference event value directly afterward in the same case. Another requirement can be added to the Follower filter based on another dimension, e.g. the time between the matching event must be shorter (or longer) than a specific value [54]. In the process of the case study, the start activity is followed by the first signature; the time between these was taken into account in the BR3. The Attribute filter allows filtering out events or cases based on arbitrary attributes (but also activity name and resources) in the data set. It

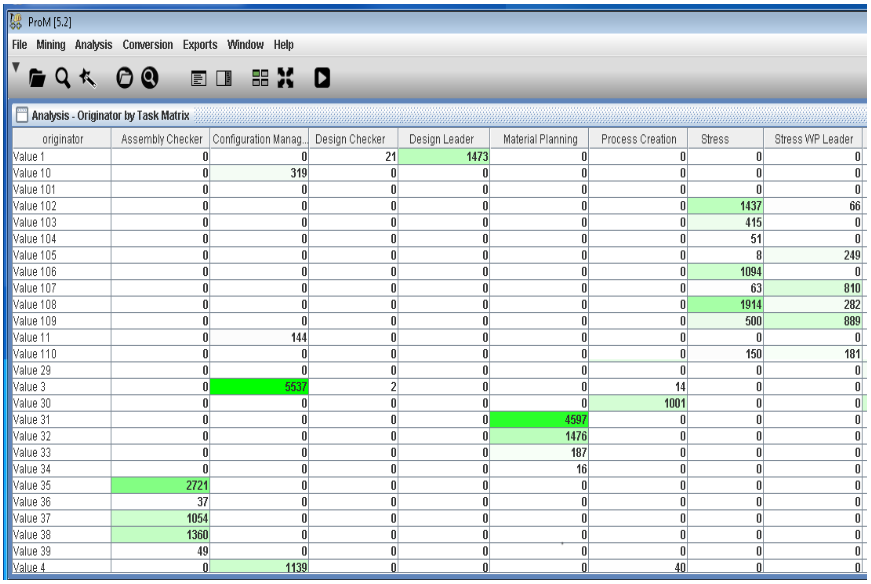
Table 5: Rules Categorization.

Category	Process Description	Rule Category	PM Perspective Dimension	Rule Restriction Focus Dimension	DISCO Filter
BR1	Completed processes: all actors approved the item.	Process Complete	Control-flow Process Perspective	Coexistence Rules	Endpoint
BR2	Process with rational time elapsed: in the range of 1 minute to 10 days.	Process Duration	Control-flow process perspective	Relative time rules	Performance
BR3	First process activity ended by one day (time elapsed from start activity and first signature).	Process Duration	Data process perspective	Relative time rules	Follower
BR4	The CM activity is executed by the user that most frequently performs this step.	Resource	Organizational process perspective	Static property rules	Attribute (*)
BR5	The process is started by the resource that most frequently performs this step.	Resource	Organizational process perspective	Static property rules	Attribute (*)
BR6	Select Processes only applied to Detail Product Types.	Product & Process Type	Data process perspective	Coexistence rules	Attribute
BR7	Select only Processes with Procedure Name = 'Detail_RW-Frozen'	Product & Process Type	Data process perspective	Coexistence rules	Attribute
BR8	Select the Processes that involve 'Supplier'.	Product & Process Type	Data process perspective	Coexistence rules	Attribute
BR9	Select only Processes with type = 'Detail_Supplier_RW-Frozen'	Product & Process Type	Data process perspective	Coexistence rules	Attribute

removes events by attribute, eliminating all events that do not have the selected value of the specific attribute. The workflow names (BR6, BR7, BR8, BR9) and the resources performing the activity (BR4, BR5) have been used as selection attributes. Before applying the filters corresponding to the BR4 and BR5 rules it has been necessary to identify the resource that most frequently executes the *Configuration Manager* and *Process Creation* activities. This has been identified by applying the Originator by Task Matrix *ProM* 5.2 plug-in [55]. The result (see Figure 4) enables to identify which originators perform the same tasks in the log dataset.

5.3. Rules Application

After PD step, ProM Conformance Checker plug-in is applied to the different portions of the event log obtained by filtering and to the corresponding process



originator	Assembly Checker	Configuration Manag...	Design Checker	Design Leader	Material Planning	Process Creation	Stress	Stress WP Leader
Value 1	0	0	21	1473	0	0	0	0
Value 10	0	319	0	0	0	0	0	0
Value 101	0	0	0	0	0	0	0	0
Value 102	0	0	0	0	0	0	1437	66
Value 103	0	0	0	0	0	0	415	0
Value 104	0	0	0	0	0	0	51	0
Value 105	0	0	0	0	0	0	8	249
Value 106	0	0	0	0	0	0	1094	0
Value 107	0	0	0	0	0	0	63	810
Value 108	0	0	0	0	0	0	1914	282
Value 109	0	0	0	0	0	0	500	889
Value 11	0	144	0	0	0	0	0	0
Value 110	0	0	0	0	0	0	150	181
Value 29	0	0	0	0	0	0	0	0
Value 3	0	5537	2	0	0	14	0	0
Value 30	0	0	0	0	0	1001	0	0
Value 31	0	0	0	0	4597	0	0	0
Value 32	0	0	0	0	1476	0	0	0
Value 33	0	0	0	0	187	0	0	0
Value 34	0	0	0	0	16	0	0	0
Value 35	2721	0	0	0	0	0	0	0
Value 36	37	0	0	0	0	0	0	0
Value 37	1054	0	0	0	0	0	0	0
Value 38	1360	0	0	0	0	0	0	0
Value 39	49	0	0	0	0	0	0	0
Value 4	0	1139	0	0	0	40	0	0

Figure 4: OriginatorByTaskMatrix.

models generated in the previous step. The goal is to identify the BR that isolate segments of the event log giving better CC results.

Table 6 proposes the results emerged from the application of the BR to the event log. It reports the values of two CC metrics, namely Fitness and Simple Behavioral Appropriateness, for the diverse segments isolated by BR. In detail, for some segments, the filter considers the application of a single BR and for others, multiple BRs are analyzed in the same run. Disco provides also information about log dimension and case duration [54]. For the duration, statistics are suggested in terms of hours (h), days (d) or minutes (m). Considering the results, a positive impact on CC results is provided by BR1, BR2, BR3, BR4, and BR5; while Product & Process Type rules (from BR6 to BR9) don't influence positively the results. However, the best results are achieved by the combination of business rules and in particular by combining the process type rules with the process duration ones.

To study the possible combination of the segments generated by BR, we

Table 6: Experiments.

Filter Applied over Frozen	Case #	Case Duration Mean	Median	Min.	Max.	Mean Events case	Fitness Value	Simple Behav. Appropriat.
No filter	24858	61 h	7 s	0 s	1 y, 300 d	3	0.552	1.067
BR1	13194	4.2 d	8 h	0 s	1 y, 300 d	5	0.9442	1.077
BR2	10539	59.2 h	42 h	59 m, 1 s	9 d, 23 h	6	0.862	1.077
BR3	16290	48.8 h	119.7 m	0 s	1 y, 300 d	4	0.900	1.077
BR4	7539	5.5 d	51 h	1 s	1 y, 300 d	6	0.878	1.077
BR5	3168	35.8 m	1 s	0 s	9 d, 22 h	1	0.910	0.958
BR6	10091	55.9 h	21 h	0 s	77 d, 19 h	4	0.712	1.077
BR7	5418	60.2 h	22 h	0 s	78 d, 19 h	4	0.792	0.923
BR7 + BR1	2210	5.2 d	3.6 d	1 h, 2 s	56 d, 7 h	9	0.933	0.856
BR7 + BR2	3777	56.5 h	31 h	59 m, 9 s	9 d, 3h	5	0.804	0.959
BR7 + BR3	2936	63.3 h	24 h	1 s	48 d, 6 h	5	0.815	0.894
BR7 + BR1 + BR2	1997	3.5 d	3.1 d	1h, 2 s	9 d, 3h	9	0.933	0.921
BR7 + BR1 + BR2 + BR3	1532	3.1 d	55h	1h, 2 s	9 d, 3 h	9	0.933	0.815
BR7 + BR1 + BR2 + BR3 + BR4	1250	3.2 d	65 h	1 h, 2 s	9 d, 23 h	9	0.933	0.815
BR7 + BR1 + BR2 + BR3 + BR5	154	3.1 d	57.5 h	2 h, 13 s	9 d, 1 h	9	0.733	0.861
BR8	5215	55.2	21 h	0 s	47 d, 16 h	5	0.821	0.996
BR9	4587	51.5 h	20 h	0 s	47 d, 16 h	6	0.995	1.083
BR9 + BR1	1369	3.8 d	4.6 d	2 h, 17 s	47 d, 16 h	12	1	0.917
BR9 + BR2	2990	57.7 h	41 h	59 m, 1 s	9 d, 23 h	7	0.995	1.091
BR9 + BR3	4110	53.8 h	22 h	0 s	47 d, 16 h	5	0.829	0.962
BR9 + BR1 + BR2	1246	3.7 d	3.1 d	2 h, 17 s	9 d, 21 h	11	1	0.900
BR9 + BR1 + BR2 + BR3	1246	3.7 d	3.1 d	2 h, 17 s	9 d, 21 h	11	1	0.900
BR9 + BR1 + BR2 + BR3 + BR4	1064	3.6 d	3 d	2 h, 17 s	9 d, 15 h	11	1	0.900
BR9 + BR1 + BR2 + BR3 + BR5	304	3.8 d	3.8 d	5 h, 59 m	9 d, 21 h	11	1	0.900

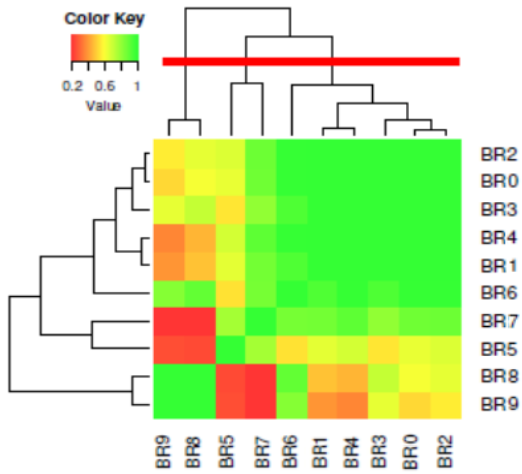


Figure 5: Comparison Table.

applied a Statistical Inference-based Analysis. First, a hierarchical clustering procedure is applied to segments, to create a significant combination of BR, then an analysis based on statistical inference is applied for characterizing the distribution of activities as manifested in segments, offering a justification of their similarities or dissimilarities. The statistical inference adopted is based on the Jensen_Shannon Distance as is illustrated with details in [56].

Figure 5 shows the results of the clustering method applied to compare BR. The Kendall's test is used as a metric for clustering segments, as reported in [56]. The value of similarity is defined in the range $[0, 1]$: 0 corresponds to the "lowest similarity", 1 to the "highest similarity". This last one is represented in Figure 5 by the green areas. The thick red line in the figure helps to cut the dendrogram and gives the group of segments that construct a cluster. By adjusting the height of cut level we can have more or less detailed group of segments in each cluster. Figure 6 shows the heuristic models discovered for each cluster. Two of the discovered models show more informative and simple Process Models.

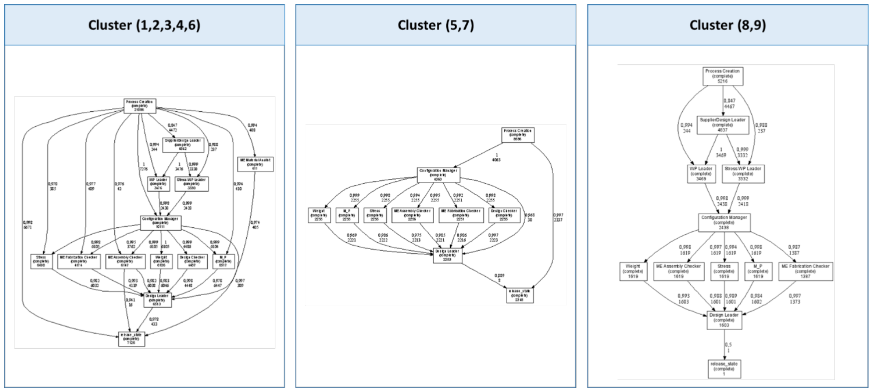


Figure 6: Discovered Process Models of clusters.

5.4. Performance Analysis

Business rules definition and analysis allow turning process mining result in useful business information. BR clusters identification has demonstrated that rules concerning product & process type (e.g. BR7, BR8, BR9) return a more simplified and informative model. On the other hand, the results reported in Table 6 reveal that not all types of rules affect positively CC results, but the best results are achieved by the combination of business rules and in particular by combining a process rules type with the process duration ones.

Once identified the BR needed to extract the portion of the entire event log on which concentrate the analysis, process performances have been evaluated. Table 7 highlights the performance report of the activities elaborated in the *Ready to work* - *Frozen* processes, that have been selected by filtering the log, using respectively BR7 and BR9. Observing the results, it is possible to carry out two analysis: 1) a quantitative one about the percentage of all the activities, rejected for each activity type (e.g. the *Fabrication Checker* and the *Assembly Checker* signature), and 2) a qualitative one, exploring the reasons for a reject. In addition, the average duration of the approval of the tasks is calculating with Petri net through the ProM plugin Performance Analysis. Almost all the rejected activities are determined by the first two roles, respectively, the *Process Manager* in the *Process Creation* activity, and the *Configuration Manager*, in the corresponding activity. The main reasons are about wrong configuration data yet at the beginning (i.e. it could generate inconsistencies in the subsequent design process) as also critical situations, coming from analysis and design activities (i.e. weight and stress analysis, fabrication and assembly procedures definition, etc.) that are specified and blocked before to be executed. Anew, in hardly any cases the *Design Leader* interrupts the process.

In brief, it is possible to claim that BR application allows characterizing a business process with higher quality. The application of BR allows obtaining a benefit consisting of clear and concrete information about the average process duration. In fact, thanks to the application of filters, the calculated values are more aligned to standards. Our claim was validated by process owners that included our approach in their protocol of analysis.

6. Conclusions

This paper introduces an approach for assessing the performance of PLM processes using PM algorithms. PM is executed on specific characterizations of a business process generated by filtering event logs using BRs. The proposed approach presents both practical and theoretical implications.

From a practical viewpoint, the approach represents a demonstration of the applicability of PM in the PLM scenario, but it can be applied to different industrial sectors and IT systems. The activities are developed in order to simplify the management of processes mining and involve a user through the definition of useful rules in the process analysis. The approach also focuses on finding the BR that maximize the value of CC metrics. From this point of view, the paper contributes to three critical issues: 1) minimize the gap between PM analysis and Business Processes characterization; 2) enhance the knowledge about the origin of dysfunctional behavior recorded in the event logs; 3) making PM results intelligible for business users.

From a theoretical viewpoint, the approach is based on rules to eliminate anomalies (identify undesired behaviors that emerge), e.g. by checking from

Table 7: Example of a Ready to Work - Frozen Performance Report.

Process Type : Detail_RW-Frozen			
Activity	% Rejected Cases	Description	Mean Approval Time (hrs)
Process Creation			
Configuration Manager	25,766	Item is rejected by the CM due to inconsistent configuration data	22.7
Weight Stress Design Checker Fabrication Checker Assembly Checker Material Planning	32,798	Item is rejected by one of the parallel signature roles due to possible inconsistencies in the design data (geometric data, weight or stress analysis results, materials, property of manufacturability or assembly of the item)	17.91 54.15 55.28 38.37 17.51 32.89
Design Leader	0.59	Item rejected by Design leader due to data inconsistency detected in an overall check of the work package	7.85
State	40.8	Completed Release Process (no rejection)	

Process Type : Detail_Supplier_RW-Frozen			
Activity	% Rejected Cases	Description	Mean Approval Time (hrs)
Process Creation			
Supplier Design Leader	7.826	Item is rejected due to missing supplier approval	3.53
WP Leader	28.188	Item is rejected by the WP leader due to inconsistent work package data	19.56
Sress WP Leader	28.188	Item rejected due to missing approval on work package stress data	23.45
Configuration Manager	18.029	Item is rejected by the CM due to inconsistent configuration data	17.35
Weight Stress Fabrication Checker Assembly Checker Material Planning	15.806	Item is rejected by one of the parallel signature roles due to possible inconsistencies in the design data (geometric data, weight or stress analysis results, materials, property of manufacturability or assembly of the item)	17.58 27.03 36.16 28.34 11.82
Design Leader	0.3	Item rejected by Design leader due to data inconsistency detected in an overall check of the work package	8.2
State		Completed Release Process (no rejection)	

whom a process activity is most frequently performed, eliminating processes that are too short or too long (exceptional processes), and performing a series of consequent actions. The approach thus eliminates the possible causes of deviation for a process.

The multi-disciplinary academia team involved in the paper has allowed to mix competencies and expertise of managerial and computer science with the company's practice. Supporting the emergence of final results that merge the context and systems knowledge with the managerial implications and process mining techniques. The achieved results are also discussed in the company and validated with the process owners reinforcing their validity: Thereby, The approach can be used for real-time analysis of process execution to intervene directly, or as a retrospect, analysis to identify recurring patterns of undesired process behaviors. The approach can be also applied by others industrial practitioners to explore and better understand their PLMS workflows and leading improving actions for product design. Academia can also replicate the case study in others context in order to predict similar results or to contrast them enlarging the state of art of industrial applications of Process Mining.

Possible future developments could consider the application of the approach to other more complex case studies, for example by increasing the number of rules to be considered, and their complexity, and considering new metrics for assessing the goodness of the approach. The definition of new rules could also regard different data and process perspetives, such as costs or revenue of the cases, or based on resource skills. Moreover, also the design and development of the automation phase of the transition from rules to their implementation on data could be particularly useful. In practice, in this approach, the work would include the automation of the generation, starting from SBVR, of a filter on PROM. A significant upgrade of the current approach could also include the definition of a rule writing tool able to support the automatic translation of rules into filters to be applied to a considered process. Finally, an interesting future development could consider the optimization of the Interaction currently in use with different tools.

References

[1] A. Saaksvuori, A. Immonen, Product lifecycle management, Springer Science & Business Media, 2008.

[2] S. Terzi, A. Bouras, D. Dutta, M. Garetti, D. Kiritsis, Product lifecycle management“ from its history to its new role, International Journal of Product Lifecycle Management 4 (4) (2010) 360–389.

[3] A. Corallo, M. E. Latino, M. Lazoi, S. Lettera, M. Marra, S. Verardi, Defining product lifecycle management: A journey across features, definitions, and concepts, ISRN Industrial Engineering.

[4] W. M. P. van der Aalst, Process Mining: Data Science in Action, 2nd Edition, Springer, Heidelberg, 2016.

[5] M. Boltenhagen, T. Chatain, J. Carmona, Generalized alignment-based trace clustering of process behavior, in: International Conference on Applications and Theory of Petri Nets and Concurrency, Springer, 2019, pp. 237–257.

[6] F. Bezerra, J. Wainer, W. M. van der Aalst, Anomaly detection using process mining, in: Enterprise, business-process and information systems modeling, Springer, 2009, pp. 149–161.

[7] W. Van Der Aalst, A. Adriansyah, A. K. A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. van den Brand, R. Brandtjen, J. Buijs, et al., Process mining manifesto, in: Business process management workshops, Springer, 2011, pp. 169–194.

[8] M. Cho, M. Song, M. Comuzzi, S. Yoo, Evaluating the effect of best practices for business process redesign: An evidence-based approach based on process mining techniques, Decision Support Systems 104 (2017) 92–103.

[9] A. Corallo, M. Lazoi, F. Striani, Process mining and industrial applications: A systematic literature review, Knowledge and Process Management 27 (2020) 225–233.

[10] W. V. der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, Springer Heidelberg, Berlin, 2011.

[11] J. Li, H. Wang, Z. Zhang, J. Zhao, A policy-based process mining framework: mining business policy texts for discovering process models, Information Systems and e-Business Management 8 (2) (2010) 169–188.

[12] P. Ceravolo, F. Zavatarelli, Knowledge acquisition in process intelligence, in: Information and Communication Technology Research (ICTRC), 2015 International Conference on, IEEE, 2015, pp. 218–221.

[13] M. De Leoni, W. M. van der Aalst, M. Dees, A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs, Information Systems 56 (2016) 235–257.

[14] van der Wmp Wil Aalst, M. M. Pesic, M. H. Schonenberg, Declarative workflows: Balancing between flexibility and support, Computer Science - Research and Development 23 (2) (2009) 99–113.

[15] R. P. J. C. Bose, W. M. P. van der Aalst, Trace clustering based on conserved patterns : towards achieving better process models, in: International Conference on Business Process Management, 2009, pp. 170–181.

[16] W. M. van der Aalst, Business process management: A comprehensive survey, ISRN Software Engineering 2013.

[17] J. C. Buijs, B. F. van Dongen, W. M. van der Aalst, Quality dimensions in process discovery: The importance of fitness, precision, generalization and simplicity, International Journal of Cooperative Information Systems 23 (01) (2014) 1440001.

[18] A. Rozinat, W. M. van der Aalst, Conformance testing: measuring the fit and appropriateness of event logs and process models, in: Business Process Management Workshops, Springer, 2005, pp. 163–176.

[19] W. Van der Aalst, A. Adriansyah, B. van Dongen, Replaying history on process models for conformance checking and performance analysis, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2 (2) (2012) 182–192.

[20] A. Azzini, P. Ceravolo, E. Damiani, F. Zavatarelli, Knowledge driven behavioural analysis in process intelligence, in: Proceedings of the International Workshop on Algorithms & Theories for the Analysis of Event Data, ATAED 2015, Satellite event of the conferences: 36th International Conference on Application and Theory of Petri Nets and Concurrency Petri Nets 2015 and 15th International Conference on Application of Concurrency to System Design ACSD 2015, Brussels, Belgium, June 22-23, 2015., 2015, pp. 97–111.

[21] A. Awad, G. Decker, M. Weske, Efficient compliance checking using bpmn-q and temporal logic, in: BPM '08 Proceedings of the 6th International Conference on Business Process Management, 2008, pp. 326–341.

[22] F. Calabrese, G. Di Dio, A. R. Fasolino, P. Tramontana, Business processes characterisation through definition of structural and non-structural criteria, in: Proceedings of the Advanced Int'L Conference on Telecommunications and Int'L Conference on Internet and Web Applications and Services, AICT-ICIW '06, IEEE Computer Society, Washington, DC, USA, 2006.

[23] W. Samek, T. Wiegand, K. Müller, Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models, Computer Science, Artificial Intelligence abs/1708.08296 (2017) 8.

[24] M. G. Core, H. Lane, M. Lent, D. Gomboc, S. Solomon, M. Rosenberg, Building explainable artificial intelligence systems., in: Proceedings of the twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference, American Association for Artificial Intelligence, www.aaai.org, 2006.

[25] G. Githens, Product lifecycle management: Driving the next generation of lean thinking by michael grieves, Journal of Product Innovation Management 24 (3) (2007) 278–280.

[26] S.-I. Lee, K.-Y. Ryu, M. Song, Process improvement for pdm/plm systems by using process mining, Transactions of the Society of CAD/CAM Engineers 17 (2012) 294–302.

[27] E. Rigger, T. Vosgien, S. Bitrus, P. Szabo, B. Eynard, Enterprise architecture method for continuous improvement of plm based on process mining, in: IFIP '20, Proceedings of the 17th International Conference on Product Lifecycle Management, 2020.

[28] B. Von Halle, Business rules applied: building better systems using the business rules approach, Wiley Publishing, 2001.

[29] P. Ceravolo, A. Azzini, E. Damiani, M. Lazoi, M. Marra, A. Corallo, Translating process mining results into intelligible business information, in: Proceedings of the The 11th International Knowledge Management in Organizations Conference on The changing face of Knowledge Management Impacting Society, 2016, p. 14.

[30] R. G. Ross, The business rules manifesto, Business Rules Group. Version 2.

[31] T. Skersys, L. Tutkute, R. Butleris, The enrichment of bpmn business process model with sbvr business vocabulary and rules, CIT. Journal of Computing and Information Technology 20 (3) (2012) 143–150.

[32] P. Ceravolo, C. Fugazza, M. Leida, Modeling semantics of business rules, in: Digital EcoSystems and Technologies Conference, 2007. DEST'07. Inaugural IEEE-IES, IEEE, 2007, pp. 171–176.

[33] F. Arigliano, P. Ceravolo, C. Fugazza, D. Storelli, Business metrics discovery by business rules, in: Emerging Technologies and Information Systems for the Knowledge Society, Springer, 2008, pp. 395–402.

[34] VvAa, Object management group, http://www.omg.org/hot-topics/iot-standards.htm (2019).

[35] VvAa, Semantics of business vocabulary and rules, http://www.omg.org/spec/SBVR (2019).

[36] S. Team, et al., Semantics of business vocabulary and rules (sbvr), Tech. rep., Technical Report dtc/06–03–02, Object Management Group, Needham, Massachusetts (2006).

[37] R. Crerie, F. A. Baião, F. M. Santoro, Discovering business rules through process mining, in: Enterprise, Business-Process and Information Systems Modeling, Springer, 2009, pp. 136–148.

[38] E. Damiani, P. Ceravolo, C. Fugazza, K. Reed, Representing and validating digital business processes, in: International Conference on Web Information Systems and Tecnologies, Springer Verlag, 2008, pp. 19–32.

[39] R. Popp, H. Kaindl, Automated refinement of business processes through model transformations specifying business rules, in: Research Challenges in Information Science (RCIS), 2015 IEEE 9th International Conference on, IEEE, 2015, pp. 327–333.

[40] O. M. Kherbouche, A. Ahmad, H. Basson, Formal approach for compliance rules checking in business process models, in: Emerging Technologies (ICET), 2013 IEEE 9th International Conference on, IEEE, 2013, pp. 1–6.

[41] F. Caron, J. Vanthienen, B. Baesens, Business rule patterns and their application to process analytics, in: 2013 17th IEEE International Enterprise Distributed Object Computing Conference Workshops, IEEE, 2013, pp. 13–20.

[42] Y. Zhao, D. Ma, Y. Zhao, Z. Li, Integrating business processes and business rules, in: Services Computing Conference (APSCC), 2011 IEEE Asia-Pacific, IEEE, 2011, pp. 493–497.

[43] W. van der Aalst, K. van Hee, J. M. van der Werf, A. Kumar, M. Verdonk, Conceptual model for online auditing, Decision Support Systems 50 (3) (2011) 636 – 647, on quantitative methods for detection of financial fraud.

[44] Y. Duan, J. S. Edwards, Y. K. Dwivedi, Artificial intelligence for decision making in the era of big data–evolution, challenges and research agenda, International Journal of Information Management 48 (2019) 63–71.

[45] G. Schuh, H. Rozenfeld, D. Assmus, E. Zancul, Process oriented framework to support plm implementation, Computers in industry 59 (2) (2008) 210–218.

[46] O. Budde, G. Schuh, J. Uam, Holistic plm model–deduction of a holistic plm-model from the general dimensions of an integrated management, in: International Conference on Product Lifecycle Management, Bremen, Germany, 2010.

[47] M. Arias, R. Saavedra, M. R. Marques, J. Munoz-Gama, M. Sepúlveda, Human resource allocation in business process management and process mining, Management Decision.

[48] P. Runeson, M. Host, A. Rainer, B. Regnell, Case Study Research in Software Engineering: Guidelines and Examples, Wiley Publisher, 2012.

[49] R. K. Yin, Case Study Research and Applications Design and Methods, SAGE Publisher, 2017.

[50] C. W. Günther, A. Rozinat, Disco: Discover your processes., BPM (Demos) 940 (2012) 40–44.

[51] A. Weijters, W. Aalst, A. Medeiros, Proces mining with the heuristics miner-algorithm, Cirp Annals-manufacturing Technology 166.

[52] VvAa, Process mining research tools application, http://www.processmining.org/ (2019).

[53] A. Rozinat, W. M. van der Aalst, Decision mining in prom, in: International Conference on Business Process Management, Springer, 2006, pp. 420–425.

[54] VvAa, Process Mining in Practice, Fluxicon.com, 2019.

[55] W. M. Van der Aalst, B. F. van Dongen, C. W. Günther, A. Rozinat, E. Verbeek, T. Weijters, Prom: The process mining toolkit., BPM (Demos) 489 (31) (2009) 2.

[56] P. Ceravolo, E. Damiani, M. Torabi, S. Barbon, Toward a new generation of log pre-processing methods for process mining, in: J. Carmona, G. Engels, A. Kumar (Eds.), Business Process Management Forum, Springer International Publishing, Cham, 2017, pp. 55–70.