From Modeling to Automated Discovery of Project Processes: Tools for Improving the Project Processes Performance in Digital Era

Constanta-Nicoleta Bodea
Department of Economic Informatics and Cybernetics
The Bucharest University of Economic Studies
Bucharest, Romania
bodea@ase.ro

Radu Ioan Mogos
Department of Economic Informatics and Cybernetics
The Bucharest University of Economic Studies
Bucharest, Romania
mogos.radu@gmail.com

Mihail Cazacu
Doctoral School on Economic Informatics
The Bucharest University of Economic Studies
Bucharest, Romania
mihail.cazacu@gmail.com

Abstract — **Business process management (BPM) tools allow the analysis and improvement of the actual business processes in the organization, for making them more efficient and effective. The paper presents how the discipline of BPM can be applied to the project processes, by adopting the automation in different stage of the project process management, especially in the process analysis and discovery. The automated project process discovery is possible due to the extended IT infrastructure for project process implementation, available in the digital era. While the process simulation for project process analysis is already applied on large scale, the project process mining is still not so much adopted and exploited.**

The authors investigate how the existing BPM tools can be used for project process modeling and ~~the~~ automated discovery. In the last part of the paper, the authors present some proposals for future development of these applications.

Keywords — *project processes, process modeling, process simulation, process mining, process automated discovery*

I. INTRODUCTION

The organization competitiveness is strongly related to the efficiency of its business processes ([1, 2]). A business process represents "a set of one or more linked procedures or activities which collectively realize a business objective or policy goal, normally within the context of an organizational structure defining functional roles and relationships" ([3]). According to the same source, the workflow represents "that part of the process definition which comprises the automatable activities".

Business Process Management (BPM) represents the identification, description, analysis, improvement, execution and monitoring of the business processes inside the organization. BPM is usually described as multiple iterations, each including several phases (Fig. 1). The process discovery means process identification, usually by observing the current workflows inside the organization. Most of the times, the process discovery is performed by business analysts, in a manual manner. As a result of this BPM

phase, a map of interrelated executed processes is defined.

During the second phase, different process models are developed, by using process modeling languages in order to represent in a formally manner the existing processes. They are so-called "as-is" process models. The process models allow a better understanding and analysis of the current business processes. The process analysis can be improved by using the simulation features of different BPM tools. The simulation results are used for adopting process re-design decisions, intending to improve the performance of the processes, mainly meaning less needed resources and shorter execution time. The re-designed/improved processes are implemented and executed.
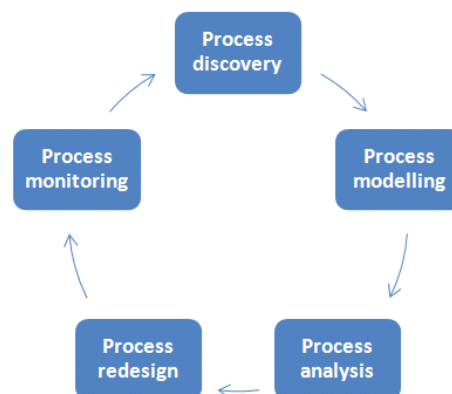

Fig. 1. The business process management cycle

Monitoring and controlling the process are allowing the identification of changes/deviations in process execution. These changes are analyzed and, if it is needed, new process models are developed/induced from the actual practices in the organization. The new process models are replacing/updating the previous ones. We can consider this as being the discovery of the new/changed processes, actually in place in the organization. Some BPM tools can perform an

automated process discovery, based on the monitoring data, such as the process event file logs.

The paper describes the main aspects of BPM and how it is applied for the processes that are executed within the projects, highlighting the phases of process simulation and process automated discovery. While the process simulation for the analysis purpose is more often applied, the process mining is still less exploited. The paper presents how the simulations can contribute to the project process analysis and the processes re-design and how the process mining is used for the automated process discovery. In the last part of the paper there are presented the main conclusions and future developments.

In their operations and projects, organizations are executing processes. The ISO 9000 standards promote the process-orientation in organizations [4]. This family of standards defines a process as "a set of interrelated or interacting activities which transforms inputs into outputs". It is argued that "a desired result is achieved more efficiently when activities and resources are managed as process".

Projects are organized and executed as flows of processes [5]. The product and the support processes must to be managed in order to achieve the desired results. The process management activity is the responsibility of the project manager. The project management can also be described as a flow of processes, from the project initiation to the project closure.

In the context of extended IT infrastructure of projects, the execution of project processes generates a high data volume, as project traces. For this reason, the project organizing and execution can extensively benefit from the latest achievements of the BPM discipline.

II. THE PROJECT PROCESS MODELING, ANALYSIS AND RE-DESIGN

A. Comparative analysis of Process Modeling Languages

Some of the most applied processes modeling languages are: Business Process Model and Notation 2.0 (BPMN 2.0), Event-driven Process Chain (EPC) and Yet Another Workflow Language (YAWL). In Tables I and II is presented a comparative analysis of these modeling languages.

TABLE I. THE CHARACTERISTICS OF THE PROCESS MODELING LANGUAGES (ADAPTED FROM [6])

Characteristics	BPMN 2.0	EPC	YAWL
Multiple perspectives (functional, behavioral, organizational, informational and operational)	√	√	√
Modeling hierarchical	—	√ (organiza	√ (parameter
structure		tional chart)	)
Modeling objects (process, pathway, cost, data, document, organizational unit, role, resource)	√	√	√
Modeling location	√ (pool)	√	—
Assignment of roles and activities	√	√	√ (variable)
Event modeling (triggers, user decisions, time, message and exception events)	√	√	√ (transition)
Priorities (flows and schedule tasks)	√	√ (annotation)	√
Transition probabilities	√	√	√
Execution of workflow models	√	—/√ (by automatic conversion in another language, such as BPMN 2.0)	√

Legend: √ means available characteristics; — means unavailable characteristics

The process modeling tools provide functions and methods for the workflows analysis, modeling and simulation. Many tools are capable of simulating the execution of business processes, in order to verify their correctness and efficiency. The simulation produces quantitative results about the processes execution, such as: statements cycle times, occupancy rates or workloads, prior to actual execution. Based on these results, different alternative process versions can be tested and decisions about the process redesign are taken.

TABLE II. THE MAIN ADVANTAGES (+) AND DISAVANTAGES (-) OF THE PROCESS MODELING LANGUAGES (ADAPTED FROM [6])

	BPMN 2.0	EPC	YAWL
+	• Established standard, with continuous development • Wide spread • Executable models • Easily comprehensible process models	• Wide spread and long-lasting market success • Multiple description view (ARIS concept) • Comprehensive range of symbols • Intuitive, easily comprehensible and learnable • Transformation in BPMN 2.0 and other workflow languages	• Dynamic resource management • Full workflow pattern support • Executable models
-	• Only a few	• No official	• Only a few

possibilities of graphical modeling actors and resources • No description of activities performed by different actors and resources at the same time • Complex constructs and symbols (many different concepts of activities and events) • Programming skills are required	standard • No executable models • Few modeling tools are available which support EPC modeling • Limited modeling functional structures	modeling tools which support YAWL are available • Only functional view on process can be modeled • Only a few symbols are available • Many parameterizations necessary • High programming effort • Less intuitive • No interoperability

B. Project Process Analysis through Simulation

The main simulation objective is to identify potential restructuring measures and to examine the effects of such measures from different views. For allowing the project process simulation, the modeling functionalities should include not only the description and visualization of the relevant components of the processes, such as: activities, tasks, their sequence, possible alternative paths the process could take, the involved resources, including people, organizational structures, temporal aspects, monetary costs, events and documents, but also the definition of probabilities and priorities for the execution of certain tasks.

There are many simulation algorithms implemented into the BPM tools ([7]):

- *Path analysis algorithm*, which calculates key performance indicators for the entire business process/an execution path and offers answers to several questions, such as: "How many possible paths exist for a particular process?", "Which is the probability for each path?", "What is the default path and can it be improved?", "Which path has the longest cycle time, the longest processing time?" and "Which is the critical path?".
- *Capacity analysis algorithm*, which simulates one/many business processes in a working environment and offers answers to several questions, such as: "What are the requirements for manpower and other resources?", "What are the fixed and variable costs of process?", "Which is the average cost of the process?", "What is the basis of internal cost accounting or the input for my external price model?" and "Which activities are the most expensive ones?".
- *Workload analysis algorithm* (for the steady state or fixed time period), which simulates one/many business processes in a working environment and offers answers to several questions, such as: "How can I optimally use my employees and other resources?", "Where do resource bottlenecks occur?", "Which waiting times are caused by the bottlenecks?" and "How do the bottlenecks affect the process?".

The process evaluation mechanisms allow further evaluation of the simulation results, by using ADL queries. The evaluation queries are similarly performed for different simulation algorithms. The project process models applied in the simulations have an extended number of components and several indicators defined for each model component, which are progressively refined. The initial version of such a model has only the basic components, without any associated indicators. The successive model versions include additional components, such as: variables for setting the transition conditions and probabilities and associated macros and indicators values for different components (for example classification type, order, task stack, time period, time indicators). Some examples of time indicators are: resting time, execution time, transport time, tolerance waiting time etc. In order to set the indicators' values for resources, time, performers etc., each component of the model has an attached table that has to be filled out. The transition conditions and the transition probabilities are used by the simulation algorithms to determine under which conditions or probabilities a path will be chosen. The transition conditions are evaluated during the simulation, to establish the process cost and to plan the human resource. They consist of an expression, a probability or a macro. The transition probabilities are applied during the analytical evaluation to determine paths. Macros are globally valid.

Although the simulation aspects (features) are present almost at all process software modeling tools, there are still differences between different tools. Table III presents a comparative analysis of four workflow software modeling tools: Adonis, Bonita BPM, jBPM and Signavio.

Bonita BPM is a workflow modeling tool offered by BonitaSoft company (https://www.bonitasoft.com/bonita-platform). There is available a free limited version of the product, which can be downloaded and tested.

jBPM (Java Business Process Model) is an open-source workflow modeling tool offered under the ASL by the JBoss company (https://www.redhat.com/en/technologies/jboss-middleware).

Signavio is part of the Business Transformation Suite offered by Signavio company (https://www.signavio.com). Table II presents the comparative analysis of these workflow software modeling tools.

TABLE III. THE COMPARISON OF SIMULATION FEATURES
(ADAPTED FROM [6])

Simulation features	ADONIS	Bonita BPM	jBPM	Signavio
Duration	√	√	√	√
Costs	√	√	√	√
Probabilities	√	√	√	√
Capacities	√	√	—	√
Workload	√ (R)	√	—	√
Indicators	√ (R)	√	—	√
Resource	—	√	√	√

|			
requirements (including human resources)			
Bottleneck detection	—	√ (R)	—	√

Legend: √ means available feature; √ (R) means restricted available feature (only at the purchased version); — means unavailable feature.

C. A Case Sudy

The case study is based on the project customers' orders process and it is presented to students enrolled in 3rd year at the bachelor degree programme in economic informatics ([8,9]). Figure 2 presents the workflow model, as the initial version (Fig. 2.a), without the elements required by the simulations and as the version, ready for running simulation (Fig. 2.b).
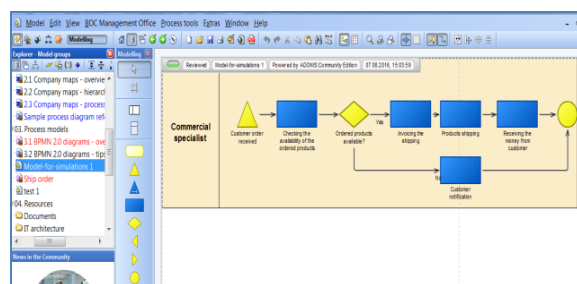
a) The initial process model (without the simulation featuers)
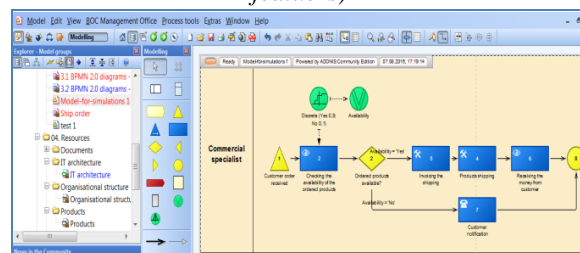
b) The Process model with the simulation features

Fig. 2. Two models of the customer's orders process

The "Macro1" replaces "X=Yes" transition condition, while the "Macro2" replaces the "X=No" condition.

Let's consider that the path analysis algorithm has been chosen for simulation. The path analysis allows the evaluation of the workflow without taking into account the working environment. In order to run the simulation, it is required to set the simulation parameters (i.e. the number of simulations, working days per years, hours per working days, simulation type, start value and the log file) and to define the simulation agents.

Three agents were defined, for the following activities: checking the availability of the order and customer notification (agent 1), product shipping (agent 2) and invoicing the shipping and receiving money from customers (agent 3). Figure 3 presents the BPM diagram with the agents included in the model.
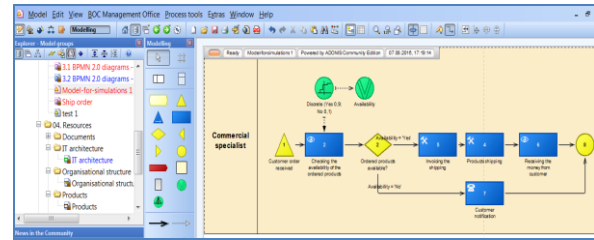

Fig. 3. Process model with the simulation agents

The panel with the simulation results, named *"Path Analysis – results"* is presented in Fig. 4.
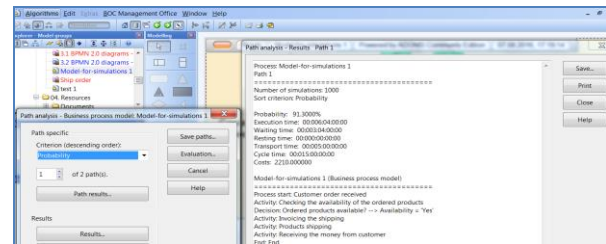

Fig. 4. The simulation results

The simulation results show that by reducing the waiting and transport times for all sub-tasks with 1% lead to a reduction of the process execution time with 10%. Also, by increasing the resources allocated to different tasks, this will lead to a significant decrease of the project execution time (1% increasement of the allocated resources leads to a decrease of the process execution time with 4%).

D. Process redesign

Based on the conclusions of simulations, several re-design decisions can be adopted. This represents an exemplary case of data-driven decision making.

III. PROCESS MONITORING AND PROCESS AUTOMATED DISCOVERY

The process mining (or process autometed discovery) represents the analysis of data generated during the process monitoring (most of the time, this data is represented by streams of process events), in order to discover significant patterns in the process behavior over time. As examples of what can be identified/discovered by using process mining techniques are the unplanned changes of certain sequences of activities/events during the process execution. The identification of these unplanned process changes can be used for the prediction and avoidance of their negatively impacts over the project performance.

The identification of unplanned process changes should be associated with the interpretation of the changes, in terms of: the impact on the process performance (time and resources), impact on other process components/activities, at a larger level, and so on [10]. The identification and the characterization of

the process changes support the understanding of the current process performance at project level.

This better understanding is enhanced when the BPM tools allow the induction of new process models, based on abstractization of the low-level relations between activities.

The process discovery tools induce process models using heuristic, fuzzy or genetic techniques for the process visualization, understanding and evaluation. Table 4 presents some examples of tools which can be used for the process mining and discovery.

TABLE IV. EXEMPLES OF PROCESS MINING TOOLS

Tool	Main functionality	Comments
OBS Studio	Video Screen Capture	-
TensorFlow + Keras	Object recognition in videos	Framework for various types of ML algorithms, especially popular for all types of artificial neural networks
NLTK + TextBlob	Natural Language Processing (for final reports written without CybOX and STIX)	TextBlob offers a friendly interface to the tools in NLTK
Blender	Video Editing (for Content Generation Subsystem)	Bleneder allows scripted editing via Python scripts
scikit learn	ML toolkits	ML toolkit written in Python
Weka	Idem	ML toolkit written in Java, with a graphical interface
PyAutoGUI	Automation (for Activity Replay and Content generating Subsystems)	GUI automation tool written in Python
Selenium WebDriver	Idem	Automation tool for browser-based applications
Click	Idem	Python package for creating command-line interfaces

Taking as an example the process of managing the customer's orders, as it was presented in the section II, we consider that the monitoring activity revealed constant and significant delays between the product shipping and the payment. This suggests that a control point should be added in the process, in order to check the time spent between the arrival of the products at the customer warehouse and the payment date. The inclusion of this new task has to be associated with a new clause in the contract, in order supplier to be entitled to ask for penalties when the delay is higher than the agreed number of days. This

new version of the process is implemented and the process mining will automatically update the process model, based on the process log, without any additional human involvement.

IV. CONCLUSIONS

The paper presents the main aspects of the BPM, highlighting the project process simulation and automated discovery aspects. Comparative analysis are made for some of the most applied process modeling languages (BPMN 2.0, EPC and YAWL) and four process software modeling tools (Adonis, Bonita BPM, jBPM and Signavio). A case study based on the process of the customer orders is presented.

The simulation process is playing a very important role for a successful business based on its advantages. The main advantages of the simulation process consist in producing quantitative results about the processes execution (statements cycle times, occupancy rates or workloads, prior to actual execution), helping the management to identify when to increase or reduce the resources usage, measuring the impact of any change that might appear. Taking into account these advantages it can be said that the simulation process is a very important part of any informational system.

The automated process discovery is based on the process mining that extract processes patterns/ knowledge from process log data.

Organizations can substantially benefit by adopting process simulations and automated process discovery tools in order to visualize, understand and evaluate the executed processes in order to take actions for reducing the execution time and costs of these processes.

The main contribution of the authors is that they propose the application of BPM in order to manage better the project management processes. The paper includes examples which validate the proposed approach.

REFERENCES

[1] S. Bulgacs, "The first phase of creating a standardized international technological implementation framework/software application", Int. J. Business and Systems Research. 7 (3): 250, doi:10.1504/IJBSR.2013.055312, 2013.

[2] D. Kroenke, MIS Essentials, (Fourth ed.), Boston, Pearson, 2015.

[3] Workflow Management Coalition, "Terminology & Glossary", Document Number WFMC-TC-1011, 1996, available online: http://www.aiai.ed.ac.uk/project/wfmc/ARCHIVE/DOCS/glossary/glossary.html

[4] International Organization for Standardisation, "ISO 9004:2018; Quality management -- Quality of an organization - Guidance to achieve sustained success", 2018, available at: https://www.iso.org/standard/70397.html

[5] R. Wagner, "Processes of ISO 21500", in: N. Grau, C. N. Bodea (eds), ISO 21500 Project Management Standard, Shaker Verlag GmbH, Germany, 2013.

[6] J. Neumann, M. Rockstroh, S. Vinz, T. Neumuth, "Technical report", v1.1, Faculty of Medicine, Leipzig University, Germany, 2015.

[7] BOC Group, "The Business Process Management Toolkit ADONIS Strengths and application scenarios", White paper, BOC Information Technologies Consulting AG, Vienna, 2015

[8] C.N. Bodea, "Sisteme informationale economice", Lecture notes, Bucharest University of Economic Studies, 2019.

[9] R. I. Mogos, "Economic Information Systems", Lecture notes, Bucharest University of Economic Studies, 2019.

[10] A. Ostovar, "Business Process Drift: Detection and Characterization", Dissertation, Science and Engineering Faculty, Queensland University of Technology (QUT), Australia, 2019.

.