A Method for Business Process Model Analysis and Improvement

Andrii Kopp[0000-0002-3189-5623] and Dmytro Orlovskyi[0000-0002-8261-2988]

National Technical University "KhPI", Kyrpychova str. 2, 61002 Kharkiv, Ukraine
{kopp93, orlovskyi.dm}@gmail.com

Abstract. Since business process modeling is considered as the foundation of Business Process Management, it is required to design understandable and modifiable process models used to analyze and improve depicted business processes. Therefore, this article proposes a method for business process model analysis and improvement. The lifecycle of Business Process Management from business process modeling to applying the Business Intelligence and process mining techniques is considered. Existing approaches to business process model analysis are reviewed. Proposed method is based on best practices in business process modeling, process model metrics, and corresponding thresholds. The usage of business process model metrics and thresholds to formalize process modeling guidelines is outlined, as well as the procedure of business process model analysis and improvement is shown. The application of Business Intelligence techniques to support the proposed method is demonstrated.

Keywords: Business Process Management, Business Process Modeling, Process Model Analysis, Process Model Improvement.

1 Introduction

Today Business Process Management (BPM) is one of the most popular management concepts. It is based on the set of methods and tools used to design, analyze, improve, and automate organizational business processes. In its turn, business process is a structured set of activities that takes one or more kinds of input and produces a product or service valuable for a particular customer [1].

According to professor van der Aalst [2], BPM combines knowledge from information technology and knowledge from management sciences and applies this to operational business processes. It has received considerable attention over the last decade due to its potential for significantly increasing productivity, saving costs, and reducing flow-time. Business Process Intelligence (BPI) is a concept that can be described as the application of Business Intelligence (BI) techniques in BPM in order to understand organization's business processes [3]. BI tools are used to integrate transactional data generated by business processes in a Data Warehouse (DWH) and consolidate this raw data into Key Performance Indicators (KPIs) that serve as basis for business process improvement decisions [4].

The fundamental technique of BPM is business process (process) modeling. It is used to understand, document (e.g., for instructing people), analyze (e.g. to find errors and measure performance), and improve the business processes they describe [2]. Therefore, it is required to design such business process models that can be easily understood and modified both during business process execution and the transformation from “as-is” to “to-be” according to improvement decisions obtained by application of BPI.

The object of this research is the procedure of business process structure design and analysis using various modeling notations. The subject of this research is development of the method for business process model analysis and improvement. The aim of this research is to eliminate violations of business process model correctness that affects its understandability and modifiability.

2 Related Work

2.1 Business Process Management Lifecycle

According to [2], the BPM lifecycle includes steps related to business process modeling, implementation, and monitoring (Fig. 1).

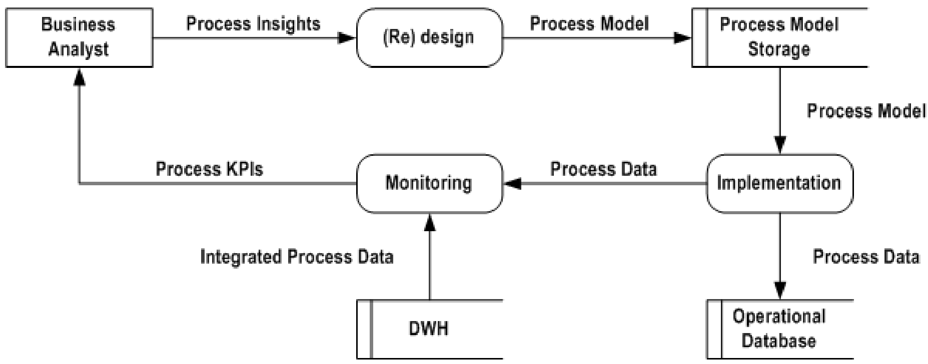

Fig. 1. BPM lifecycle

Business process modeling. Business process models are widely used in documentation of business operations. There are various business process modeling notations of different perspectives for this purpose [5]. Latest survey demonstrates that Business Process Model and Notation (BPMN) models are used by 64% of organizations that support BPM initiative. Event-driven Process Chain (EPC) models are used by 18% of respondents, while IDEF-based techniques IDEF0 and Data Flow Diagram (DFD) are used by 4% of survey participants [6].

Analysis and improvement of business process models. The main goal of business process modeling is to provide high quality diagrams that show understandable and modifiable structure of described business process. This goal might be achieved by applying Plan-Do-Check-Act (PDCA) method for the control and continuous improvement of business process models designed during BPM projects (Fig. 2) [7].

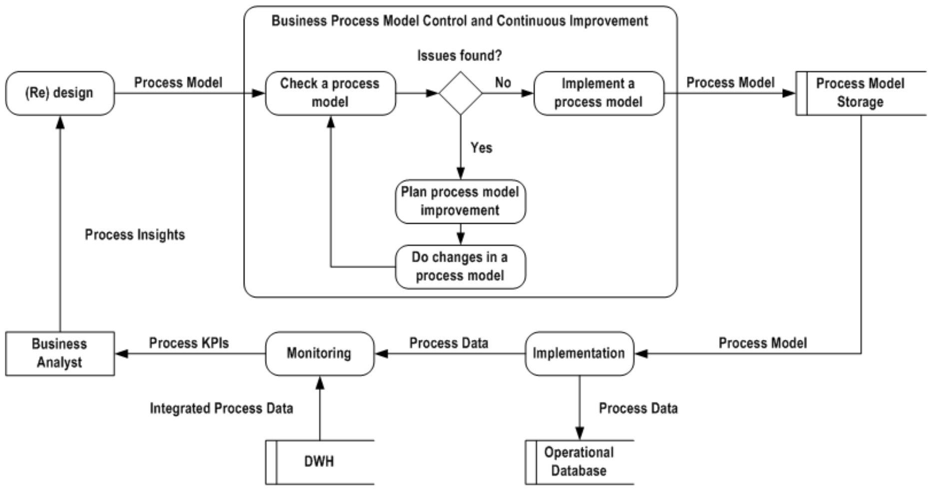

Fig. 2. Control and continuous improvement of business process models

Storage for business process models. Reuse of business process models is a way to reduce the cost of modeling business processes from scratch by using existing process models. A process model repository offers a central location for collecting and sharing process knowledge for future reuse [8]. Repository is a specialized, extensible database application that adds value to a database system by being tailored to a specific domain. A business process model repository enables stakeholders to retrieve process models for understanding business operations; updating, simulating and analyzing business process models; and reusing process models [9].

Interchange of business process models. Although there is no standardized formats to interchange process models described using IDEF0 and DFD notations, there was an attempt by Mendling and Nuttgens [10] to provide XML-based interchange format for EPC models. The ARIS Markup Language (AML) is the proprietary file format the ARIS Toolset uses when a model is exported to a file. However, the AML is non-EPC specific format [10]. The XML Process Definition Language (XPDL) is the format proposed by the Workflow Management Coalition (WfMC) to interchange process definitions between different modeling tools and management systems [11]. The XPDL is widely used as the file format for exchange of BPMN diagrams. However, BPMN 2.0 notation introduced its own XML-based interchange format.

Business process intelligence and process mining. As the application of BI techniques to BPM, BPI refers to methods that use event data to support decision making in the field of business processes, e.g., Business Activity Monitoring (BAM) and Complex Event Processing (CEP). Nevertheless, even mature data mining capabilities offered by BI tools are not process-centric, i.e., the focus is on data and local decision making, rather than end-to-end processes [12]. Thus, the process mining is proposed to bridge the gap between BI and BPM. The goal of process mining is to automatically generate a business process model using process-related event data [2, 12]. However, business process models discovered from event data still should be understandable and modifiable for its further use in BPM lifecycle, e.g., to check the conformance of a given model by comparing it with reality.

2.2 Business Process Model Analysis

Authors of research [13] have analyzed several approaches to make a business process model understandable, reliable, and reusable. They classified them based on their main research topic (Fig. 3):

- Approaches focused on improving business process design through the suggestion of modeling guidelines.
- Approaches which identify business process model metrics to evaluate model correctness.
- Approaches which establish thresholds for the identified metrics.

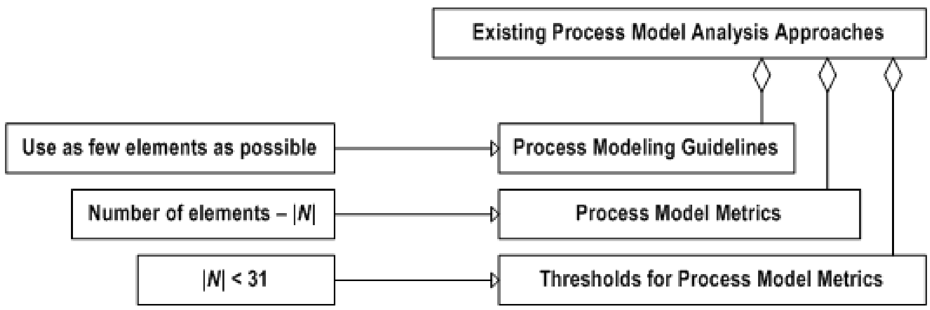

Fig. 3. Current state-of-the-art of business process models analysis

Process modeling guidelines. Process modeling guidelines (7PMG) by Mendling et. al. are supposed to guide the modeler in designing understandable models that are less prone to errors. Guidelines recommend to use as few elements as possible (G1), minimize the degree of elements (G2), use one start and one end event (G3), make sure that every split connector matches a respective join connector of the same type (G4), avoid OR split and join connectors (G5), use verb-object activity labels (G6), and decompose the model if it has too many elements (G7) [13, 14].

Process model metrics. Rolon et. al. [15] have proposed metrics (e.g., total number of sequence flows, events, gateways etc.) based on software metrics to evaluate the complexity of BPMN models. Cardoso [16] has proposed metric to measure the complexity of BPMN-based process models from a control flow perspective. Mendling et. al. [17] have proposed complexity metrics for EPC models used to predict errors by applying a logistic regression model. As for IDEF0 and DFD models, a balance coefficient is used to measure unevenness of arcs distribution in process diagrams [18].

Thresholds for process model metrics. Sanchez-Gonzalez et. al. [19] and Mendling et. al. [20] have identified thresholds for structural metrics (e.g., number of elements, gateway mismatch, density, connectivity, control flow complexity etc.) by analyzing their impact on the complexity and error probability of business process models.

As the result of the review, we can conclude that existing approaches were designed separately by different authors with various visions. Therefore, it is quite difficult to find correspondence between certain guidelines, metrics, and thresholds. Also there is lack of approaches used to provide recommendations on business process model improvement, e.g., which nodes to add, remove or replace, or how to interconnect them, etc. Thus, a method outlined in this paper is intended to fill this gap.

3 Business Process Model Analysis and Improvement Method

3.1 Formalization of Business Process Modeling Best Practices

According to Mendling et. al. [17] a business process model might be formalized as a coherent, directed graph:

$$BPModel = (N, l, A). \tag{1}$$

Where:

- $N = F \cup E \cup C \cup V$ is the set of nodes which represent various elements of the business process model described using pairwise disjoint and finite subsets: functions F, events E, connectors C, and other notation-specific elements V (e.g., data stores $D \subseteq V$ and external entities $X \subseteq V$ for DFD models, and interfaces $I \subseteq V$ for IDEF0 models);
- $C = S \cup J$ is the subset of connectors which, in its turn, consists of the subsets of split S and join J connectors;
- $l: C \to \{and, or, xor\}$ is the mapping that defines types of connectors;
- $A \subseteq N \times N$ is the binary relation A that represents arcs of the process model.

Use as few elements as possible or decompose the model if it has too many elements. It is not recommended to use more than 31 elements in EPC and BPMN, 7 elements in DFD, and 6 elements in IDEF0 diagrams. At the same time, IDEF0 diagrams must consist of at least 3 functions, while other process models (EPC, BPMN, and DFD) – of at least 1 function [14, 18]:

$$|N| \leq \begin{cases} 31, & EPC \vee BPMN, \\ 7, & DFD, \\ 6 + |V|, & IDEF0, \end{cases} \tag{2}$$

$$|F| \geq \begin{cases} 1, & EPC \vee BPMN, \\ 1, & DFD, \\ 3, & IDEF0. \end{cases} \tag{3}$$

Minimize the degree of an element in the business process model. The higher the degree of elements in the process model the harder it becomes to understand the model [14]. The following equations based on the balance coefficient [18] might be used to measure compliance with this guideline:

$$K_b^C = \frac{1}{|C|} \cdot \sum_{k=1}^{|C|} |d(c_k) - \delta_C| = 0, \tag{4}$$

$$K_b^F = \frac{1}{|F|} \cdot \sum_{q=1}^{|F|} \sum_{t \in T_A} \left|d^t(f_q) - \delta_F^t(f_q)\right| = 0. \tag{5}$$

Where:

- K_b^C is the balance coefficient of connectors (for EPC and BPMN models);
- K_b^F is the balance coefficient of functions;
- c_k is the k-th connector of the business process model;
- f_q is the q-th function of the business process model;
- $d(c_k)$ is the number of arcs connected to the k-th connector;
- $d^t(f_q)$ is the number of arcs of t-th type connected to the q-th function, $t \in T_A$:

$$T_A = \begin{cases} \{in, out\}, & EPC \vee BPMN \vee DFD, \\ \{in, con, out, mech\}, & IDEF0; \end{cases} \tag{6}$$

- δ_C is the recommended number of arcs per connector [13], $\delta_C = 3$;
- $\delta_F^t(f_q)$ is the recommended number of arcs of t-th type connected to the q-th function (one arc of each type for EPC and BPMN models [13], not more than 3 arcs of each type for IDEF0 and DFD models [18]):

$$\delta_F^t(f_q) = \begin{cases} 1, & EPC \vee BPMN, \\ \max\left\{\delta_{min}^t, \min\left\{d^t(f_q), 3\right\}\right\}, & IDEF0 \vee DFD. \end{cases} \tag{7}$$

- δ_{min}^t is the required number of arcs of t-th type, $t \in T_A$ (this equation defines that it is possible that functions on IDEF0 diagrams may not have any input arcs [18]):

$$\delta_{min}^t = \begin{cases} 0, & t = in \wedge IDEF0, \\ 1, & else. \end{cases} \tag{8}$$

Use one start and one end event. According to [13, 14], the number of start and end events (in EPC and BPMN models) is positively connected with the increase in error probability, and models that satisfy this requirement are easier to understand:

$$|E_s| = 1, |E_e| = 1. \quad (9)$$

Where:

— $E_S = \{e \in E \middle| d^{in}(e) = 0 \wedge d^{out}(e) > 0\}$ is the subset of start events;
— $E_e = \{e \in E \middle| d^{in}(e) > 0 \wedge d^{out}(e) = 0\}$ is the subset of end events.

Make sure that every split connector matches a respective join connector of the same type. It is required to model business processes as structured as possible. Unstructured models are more likely to have errors, as well as less understandable [20]:

$$MM = \sum_{i \in \{xor, or, and\}} |SC_i - JC_i| = 0. \quad (10)$$

Where:

— MM is the coefficient of connectors mismatch (for EPC and BPMN models);
— $SC_i = |\{c \in S | l(c) = i\}|$ is the number of split connectors of the i-th type;
— $JC_i = |\{c \in J | l(c) = i\}|$ is the number of join connectors of the i-th type;
— $S = \{c \in C \middle| d^{in}(c) = 1 \wedge d^{out}(c) > 1\}$ is the subset of split connectors;
— $J = \{c \in C \middle| d^{in}(c) > 1 \wedge d^{out}(c) = 1\}$ is the subset of join connectors.

It is recommended to avoid OR routing elements. Models that do not have OR split or join connectors are less error-prone [14, 20]:

$$|C_{or}| = 0. \quad (11)$$

Where $C_{or} = \{c \in C | l(c) = or\}$ is the subset of OR routing elements (for EPC and BPMN models), both splits and joins.

3.2 Business Process Model Analysis and Improvement Procedure

The main steps of the proposed method for business process model analysis and improvement are shown in Fig. 4. The method is based on process modeling best practices, metrics, and corresponding thresholds described in the previous section. Mathematical models M_1, M_2, M_3, M_4 are intended to represent optimization problems used to elaborate recommendations in order to eliminate found violations of business process modeling guidelines. Therefore, the following models should be formulated:

- Optimization problem M_1 used to define structural changes of a process model in order to obtain desired values for the following metrics $\langle |N|, |F|, |E_s|, |E_e|, |C_{or}| \rangle$.
- Optimization problem M_2 used to define structural changes of a process model in order to obtain a desired value for the K_b^C metric.
- Optimization problem M_3 used to define structural changes of a process model in order to obtain a desired value for the K_b^F metric.
- Optimization problem M_4 used to define structural changes of a process model in order to obtain a desired value for the MM metric.

The underlying idea of the described models refers to parametric optimization. Thus, they will be used to find best values of $\{|N|, |F|, |E_s|, |E_e|, |C_{or}|\}$, $\{d(c_k) \big| k = \overline{1, |C|}\}$, $\{d^t(f_q) \big| q = \overline{1, |F|}, t \in T_A\}$, and $\{SC_i, JC_i \big| i \in \{xor, or, and\}\}$ respectively.

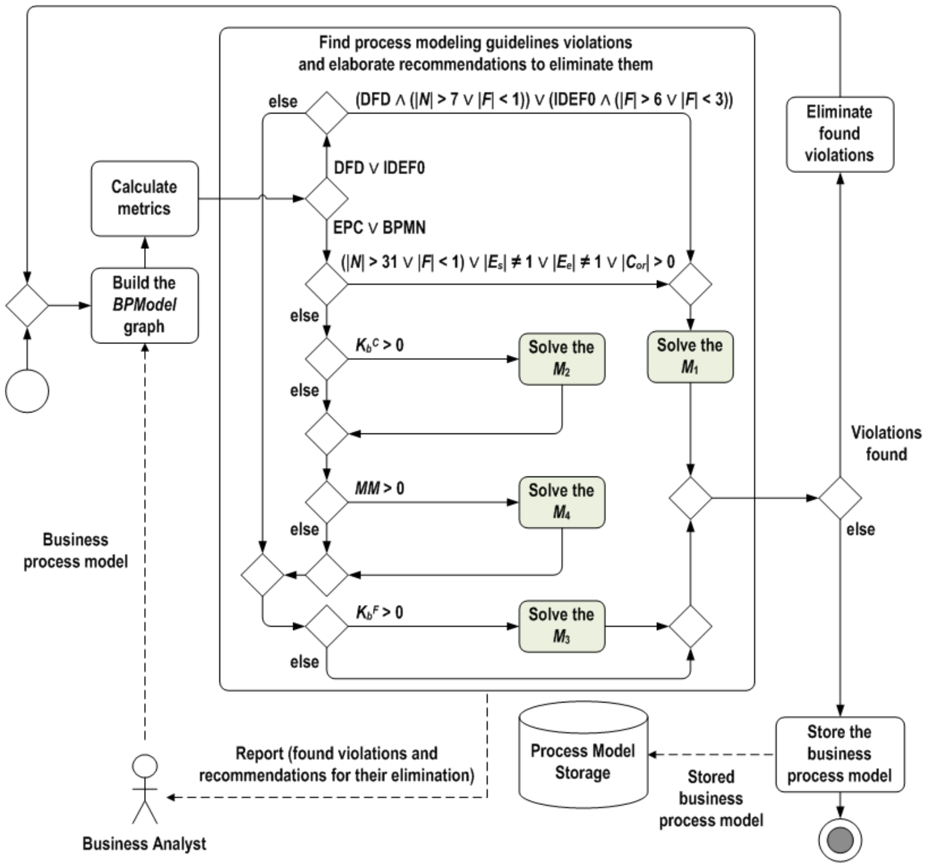

Fig. 4. Procedure of business process model analysis and improvement

3.3 Applying Business Intelligence Techniques

Implementation of the proposed method requires processing the considerable amounts of business process models. Hence, we propose the usage of the well-known BI techniques such as integration and consolidation of raw data into metrics which serve as the basis for management decisions [4]. The proposed data flow (Fig. 5) includes the following elements:

- **Data Sources.** Some BPM tools use databases to store process models (e.g., Bizagi Studio uses Microsoft SQL Server Express to store business process data). Besides, business process models might be stored using XPDL or BPMN 2.0 formats.

- **Business Process Model Analysis and Improvement.** It is required to develop the software used to extract data from various data sources, calculate metrics, and plan changes of a business process model structure if it is necessary.
- **Data Storage**. Calculated metrics and planned changes should be stored in a DWH for analytic purposes. Relational databases (e.g., SQL Server or MySQL) might be used to build such DWH.
- **Data Visualization**. Visualization (e.g., using Microsoft Power BI) of the DWH content is necessary to support decisions on business process model correctness.

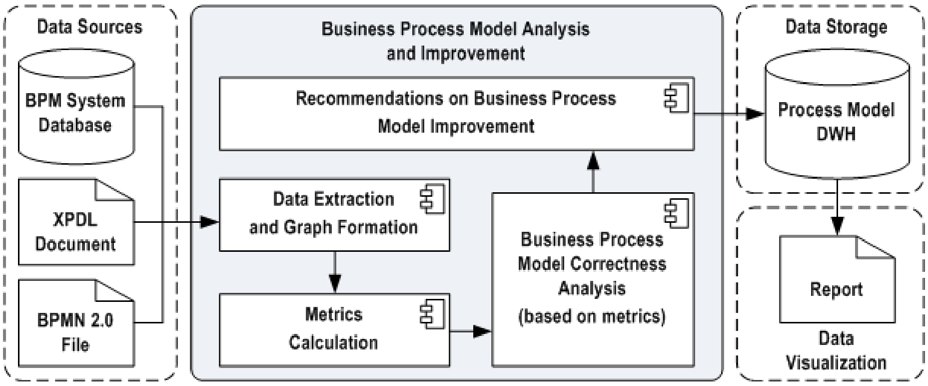

Fig. 5. Applying BI techniques to support the proposed method

4 Conclusion and Future Work

In this paper we have proposed a method for business process model analysis and improvement. This method is based on formalization of process modeling best practices using metrics and corresponding thresholds in order to find and eliminate violations of business process model correctness that affects its understandability and modifiability. Existing guidelines and metrics designed mostly for EPC and BPMN models were extended also for IDEF0 and DFD process diagrams. Future work includes formulation of the optimization problems used to elaborate recommendations for business process model improvement, and implementation of the proposed method (Fig. 4) by using BI techniques and tools as it is shown in Fig. 5. It is also planned to elaborate an evaluation criteria for the proposed method and apply it to a set of business process models described using various modeling notations.

References

1. Hammer, M., Champy, J.: Reengineering the Corporation: A Manifesto for Business Revolution. Zondervan (2009).
2. Van der Aalst, W. M. P.: Business process management: a comprehensive survey. Hindawi Publishing Corporation: ISRN Software Engineering (2013).

3. Linden, M., Felden, C., Chamoni, P.: Dimensions of business process intelligence. In: International Conference on Business Process Management, pp. 208-213. Springer, Berlin, Heidelberg (2010).
4. Bucher, T., Gericke, A., Sigg, S.: Process-centric business intelligence. Business Process Management Journal 3(15), 408-429 (2009).
5. Krogstie, J.: Perspectives to process modeling. In: Business Process Management, pp. 1-39. Springer, Berlin, Heidelberg (2013).
6. Harmon, P.: The State of Business Process Management 2016. BPTrends (2016).
7. What is the Plan-Do-Check-Act (PDCA) cycle?, https://asq.org/quality-resources/pdca-cycle, last accessed 2019/01/24.
8. Shahzad, K., Elias, M., Johannesson, P.: Requirements for a business process model repository: A stakeholders' perspective. In: International Conference on Business Information Systems, pp. 158-170. Springer, Berlin, Heidelberg (2010).
9. Elias, M.: Design of business process model repositories: requirements, semantic annotation model and relationship meta-model: Doctoral thesis. Department of Computer and Systems Sciences, Stockholm University (2015).
10. Riehle, D. M. et al.: Towards an EPC Standardization–A Literature Review on Exchange Formats for EPC Models. In: Proceedings of the Multikonferenz Wirtschaftsinformatik (MKWI 2016). Ilmenau, Germany (2016).
11. Van der Aalst, W. M. P.: Patterns and XPDL: A critical evaluation of the xml process definition language. BPMcenter (2003).
12. Van der Aalst, W. M. P.: Using Process Mining to Bridge the Gap between BI and BPM. IEEE Computer 12(44), 77-80 (2011).
13. Corradini, F., Ferrari, A., Fornari, F., Gnesi, S., Polini, A., Re, B., Spagnolo, O.: Quality Assessment Strategy: Applying Business Process Modelling Understandability Guidelines. University of Camerino, Italy (2015).
14. Krogstie, J.: Quality of business process models. In: Quality in Business Process Modeling, pp. 53-102. Springer, Cham (2016).
15. Rolon, E., Ruiz, F., Garcia, F., Piattini, M.: Applying software metrics to evaluate business process models. CLEI-Electronic Journal 1(9) (2006).
16. Cardoso, J.: Process control-flow complexity metric: An empirical validation. In: IEEE International Conference on Services Computing (SCC'06), pp. 167-173. IEEE (2006).
17. Mendling, J., Verbeek, H. M. W., van Dongen, B. F., van der Aalst, W. M. P., Neumann, G.: Detection and Prediction of Errors in EPCs of the SAP Reference Model. Data & Knowledge Engineering 64(1), 312-329 (2008).
18. Godlevskyi, M. D., Orlovskyi, D. L., Kopp, A. M.: Structural analysis and optimization of IDEF0 functional business process models. Radio Electronics, Computer Science, Control 3, 48-56 (2018).
19. Sánchez-González, L., Garcia, F., Mendling, J., Ruiz, F.: Quality assessment of business process models based on thresholds. In: OTM Confederated International Conferences "On the Move to Meaningful Internet Systems", pp. 78-95. Springer, Berlin, Heidelberg (2010).
20. Mendling, J., Sánchez-González, L., Garcia, F., La Rosa, M.: Thresholds for error probability measures of business process models. Journal of Systems and Software 5(85), 1188-1197 (2012).