MATHEMATICAL AND COMPUTER MODELING

A SOFTWARE TOOL FOR QUALITY MEASUREMENT OF BUSINESS PROCESS MODELS USING DISTANCES IN N-DIMENSIONAL SPACE

This paper considers the problem of quality measurement of business process models using the n-dimensional space distances. Business process models are graphical schemes similar to flowcharts or activity diagrams utilized in software engineering, usually represented using the BPMN (Business Process Model and Notation) or EPC (Event-driven Process Chain) notations. Business process modeling aims to capture current enterprise workflows for their analysis and then improvement using automation through IT (Information Technology) systems deployment. Therefore, designed business process models should be of high quality, so all “weak spots” of described organizational workflows could be properly identified and replaced with more efficient solutions. Whereas process models, which are not understandable, un-structured, and/or uncertain, may cause even more issues when used to improve organizational activities. Therefore, this study proposes quality measures and a software tool that can be used to detect errors in BPMN and EPC business process models using distances in n-dimensional space. The formal problem statement based on the graph-based description of business process models was given. Business process model quality characteristics and measures that consider features of both BPMN and EPC notations were proposed. Distances in n-dimensional space were suggested for usage to measure the quality of business process models. The algorithm for business process model quality measurement was proposed. The software tool to measure the quality of BPMN and EPC business process models was designed and developed. The quality of sample business process models was measured using the software tool. Obtained results were analyzed and discussed. This paper considers the Euclidean distance only, while other existing n-dimensional space distance measures or, on the opposite, similarity measures, can be used to evaluate business process model quality in further research.

Keywords: business process model, quality measure, software tool, distance measure, n-dimensional space.

A. M. KOPP, D. L. ORLOVSKYI

PROGRAMME RISHENIE DLA VIMIRU YAKOSTI MODELEJ BIZNES-PROCESOV Z VIKORISTANNYM V IDE N-VIMIRNOMU PROSTORI

У даній роботі розглядається задача вимірювання якості моделей бізнес-процесів за допомогою відстаней у n-вимірному просторі. Моделі бізнес-процесів — це графічні схеми, подібні до блок-схем або діаграм діяльності, які використовуються в розробці програмного забезпечення, зазначений за допомогою нотації BPMN (Business Process Model and Notation) або EPC (Event-driven Process Chain). Моделювання бізнес-процесів має на меті охопити поточні покоління робіт підприємства для їх аналізу та подальшого удосконалення за допомогою автоматизації через розгорнутий IT-систем (інформаційних технологій). Тому розроблені моделі бізнес-процесів мають бути якісними, щоб усі «слабкі місця» описаних організаційних бізнес-процесів могли бути правильно визначені та замінені більш ефективними рішеннями. Тоді як модельні процеси, які є нерозумілими, неструктурованими та/або невизначеними, можуть викликати ще більше проблем, якщо їх використовувати для покращення організаційної діяльності. Тому в цьому дослідженні пропонуються міри якості та програмний засіб, які можна використовувати для виявлення помилок у моделях бізнес-процесів у нотаціях BPMN та EPC з використанням відстаней у n-вимірному просторі. Дано формальну постанову задачі на основі графічного опису моделей бізнес-процесів. Запропоновано якісні характеристики моделей бізнес-процесів та міри, які враховують особливості нотацій як BPMN, так і EPC. Відстані в n-вимірному просторі було запропоновано для вимірювання якості моделей бізнес-процесів. Запропоновано алгоритм вимірювання якості моделей бізнес-процесів. Спробовано та розроблено програмний засіб для вимірювання якості моделей бізнес-процесів у нотаціях BPMN та EPC. Представлено приклади вимірювання якості моделей бізнес-процесів за допомогою розробленого програмного засобу. Отримані результати було проаналізовано та обговорено. У цій статті розглядається лише Евклідова відстань, тоді як інші існуючі відстані у n-вимірному просторі або, навпаки, міри подібності, можуть бути використані для оцінки якості моделей бізнес-процесів у подальших дослідженнях.

Ключові слова: модель бізнес-процесу, міра якості, програмний засіб, міра відстані, n-вимірний простір.

Introduction. Business process modeling is a key activity within the Business Process Management (BPM) domain [1]. Process modeling ensures IT (Information Technology) and business alignment by making easier communication between business users, such as chief officers, managers, and other stakeholders on the one side, and IT engineers who design and maintain enterprise information system(s) on the other side.

Graphical business process models are used to capture and analyze current workflows in order to find ways for their improvement by means of improvement of current IT systems or introducing new IT system modules in the case

© A. M. Kopp, D. L. Orlovskyi, 2022
considered workflows are still not automated. Hence, captured business process models should be understandable, well-structured, and may not have any uncertainties. Otherwise, it will not be possible properly analyzed current enterprise activities and suggest efficient ways for their improvement. Moreover, improper business process models may signalize improper business processes themselves, so modeling mistakes may appear because the reflected real business process has certain faults.

Considered process modeling issues may negatively affect BPM projects, since fault business process models may cause even more inefficient spots in so-called “improved” or “TO-BE” workflows. Moreover, further costs to fix arisen errors at the stage of IT system maintenance may be literally 100 times greater than at the stage of system design [2], where the business process models are created and analyzed.

This study aims to provide quality measures to improve business process modeling and analysis activities of BPM projects. A research object includes business process modeling and analysis activity, conducted by business analysts or IT management specialists. A research subject includes quality measures for BPMN (Business Process Model and Notation) [3] and EPC (Event-driven Process Chain) [4] models used nowadays the most to capture organizational workflows [5].

In order to achieve the research aim, there should be solved following tasks:

- give the formal problem statement based on the graph-based description of business process models;
- propose business process model quality characteristics and measures that consider features of both BPMN and EPC notations;
- suggest usage of distances in n-dimensional space to measure the quality of business process models;
- propose an algorithm to measure the quality of business process models;
- design and develop a software tool to measure the quality of BPMN and EPC business process models;
- measure quality of sample business process models using the software tool, analyze, and discuss obtained results.

These tasks could be solved using methods of business process modeling [6], graph theory [7], metric geometry [8], and software engineering [9].

**Problem statement.** Business process elements that deserve the most attention in business process models are tasks or also referred to as activities.

Since business process models could be described as directed graphs [7], the following formalism could be used to describe the structure of a business process diagram:

\[
BPG = \{T, SF\}. \tag{1}
\]

The \(BPG\) stands for the Business Process Graph, the formal description of a business process model, which is then should be used to analyze business process diagrams.

The \(T\) stands for the Tasks, the set of activities of a business process model. Each task could be described using the following tuple of incoming sequence flows \(i^n\) and outgoing sequence flows \(o^m\) as well:

\[
T = \{ t = \{(i^n, o^m), t \in T\} \}. \tag{2}
\]

The \(SF\) stands for the Sequence Flow, the set of sequence flows that connect tasks of a business process model and other elements, such as events and gateways. Each sequence flow could be described using the following tuple of the source task \(sf_{\text{source}}\) and target task \(sf_{\text{target}}\):

\[
SF = \{ sf = \{sf_{\text{source}}, sf_{\text{target}}, sf \in SF\} \}. \tag{3}
\]

Hence, each of \(t^n\) or \(t^m\) tuples could be formally described as following: \(t^s = \{sf, sf \in SF\}, s \in \{\text{in, out}\}\), while \(t \in T\).

The linkage between considered formalisms (1)–(3) and business process modeling elements in BPMN and EPC notations is demonstrated in Fig. 1.

Using introduced formalisms (1)–(3) and proposed quality characteristics of business process models, there could be formulated following measures.

**Business process model quality characteristics and measures.** The following quality characteristics [10] could be used to describe a business process model:

- **Complexity** – how large the business process model is, how dense, connected, and coupled are sequence flows that associate elements to each other in the business process model;
- **Structuredness** – how the business process model is structured: elements should be properly connected and the connection of business process model elements also should be consistent and coherent;
- **Uncertainty** – business process workflows, demonstrated by models, should be free of the uncertainty of depicted scenarios – business process branching should be explicitly determined and depicted without any uncertainty.

The complexity of a business process model could be described using a tuple instead of a single measure:

\[
BPG_{\text{c}} = \{i_{\text{avg}}, i_{\text{min}}, i_{\text{max}}\}. \tag{4}
\]

The \(i_{\text{avg}}\) is the average degree of a business process model activity, where degree stands for the number of incoming and outgoing sequence flows associated with a business process model element:

\[
i_{\text{avg}} = \frac{1}{|T|} \sum_{t \in T} |i^n| + |o^m|. \tag{5}
\]
The $i_{\text{min}}$ is the minimum degree of a business process model activity:

$$i_{\text{min}} = \min_{t \in T} \left\{ |i^{|in|}_t| + |i^{|out|}_t| \right\}. \quad (6)$$

The $i_{\text{max}}$ is the maximum degree of a business process model activity:

$$i_{\text{max}} = \max_{t \in T} \left\{ |i^{|in|}_t| + |i^{|out|}_t| \right\}. \quad (7)$$

The proposed measure of business process model complexity gives only relative values and, therefore, could be used to compare business process models of different sizes. The introduced measure of business process model complexity considers degrees of activities and derived sub-measures, such as average degree, minimum degree, and maximum degree of the task in a business process model. Usage of such relative measures also allows comparing business process models of different sizes to each other.

Structuredness of a business process model is a much more complex quality characteristic, which also could be described using the following tuple-based measure:

$$\text{BPG}_{S} = \left\{ \hat{i}_{\text{min}}^{(0)}, \hat{i}_{\text{max}}^{(0)} \right\}. \quad (8)$$

The $\hat{i}_{\text{in}}^{(0)}$ is the relative number of tasks (the percent of all tasks of a business process model), each of which does not have incoming sequence flows:

$$\hat{i}_{\text{in}}^{(0)} = \frac{1}{|T|} \sum_{t \in T} 1_{|i^{|in|}_t| = 0}. \quad (9)$$

The $\hat{i}_{\text{out}}^{(0)}$ is the relative number of tasks, each of which does not have outgoing sequence flows:

$$\hat{i}_{\text{out}}^{(0)} = \frac{1}{|T|} \sum_{t \in T} 1_{|i^{|out|}_t| = 0}. \quad (10)$$

Here $1_{|i^{|in|}_t| = 0}$ and $1_{|i^{|out|}_t| = 0}$ are the respective indicator functions.

Introduced relative measures of business process model structuredness demonstrate if there are present tasks, which do not have incoming or outgoing sequence flows that signalize interruptions of a business process.

Business processes should start or end with respective events, while activities that trigger business process workflow execution or serve as the ends of business process workflow may signalize inconsistent and incoherent business processes [6].

Uncertainty of a business process model is even more complex quality characteristic in comparison to structuredness, which also could be described using the following tuple-based measure:

$$\text{BPG}_{U} = \left\{ \hat{i}_{\text{in}}^{(3)}, \hat{i}_{\text{out}}^{(3)} \right\}. \quad (11)$$

The $\hat{i}_{\text{in}}^{(3)}$ is the relative number of tasks, each of which has multiple incoming sequence flows:

$$\hat{i}_{\text{in}}^{(3)} = \frac{1}{|T|} \sum_{t \in T} 1_{|i^{|in|}_t| > 1}. \quad (12)$$

The $\hat{i}_{\text{out}}^{(3)}$ is the relative number of tasks, each of which has multiple outgoing sequence flows:

$$\hat{i}_{\text{out}}^{(3)} = \frac{1}{|T|} \sum_{t \in T} 1_{|i^{|out|}_t| > 1}. \quad (13)$$

Here $1_{|i^{|in|}_t| > 1}$ and $1_{|i^{|out|}_t| > 1}$ are the respective indicator functions.

Considered relative measures also demonstrate the presence of tasks, which implement implicit branching of a business process workflow. There are exclusive and parallel branching scenarios, which could be declared without special workflow patterns using XOR-gateways and AND-gateways [6]. Such implicit decisions bring uncertainty to described business processes and should be re-designed to achieve better understandability and maintainability of business process models.

Using distances in $n$-dimensional space to measure the quality of business process models. Therefore, the business process model graph could be described using the following ordered set of measures (4), (8), and (11):

$$\text{BPG}_{M} = \left\{ \text{BPG}_{C}, \text{BPG}_{S}, \text{BPG}_{U} \right\}. \quad (14)$$

By expanding this tuple (14), we could obtain the following vector, which describes a business process model:

$$X_{\text{BPG}} = \left( i_{\text{avg}}, i_{\text{min}}, i_{\text{max}}, i_{\text{in}}^{(0)}, i_{\text{out}}^{(0)}, i_{\text{in}}^{(3)}, i_{\text{out}}^{(3)} \right). \quad (15)$$

The vector of a business process model, which does not have any flaws, which violate complexity, structuredness, and uncertainty, could be described as follows:

$$X_{0}^{\text{BPG}} = (2, 2, 2, 0, 0, 0, 0). \quad (16)$$

Hence, the quality of the business process model could be measured as the distance between the vector of a “perfect” business process model $X_{0}^{\text{BPG}}$ and the current business process model under evaluation $X^{\text{BPG}}$ [8]:

$$d \left( X^{\text{BPG}}, X_{0}^{\text{BPG}} \right) = \sqrt{\sum_{k=1}^{7} (X_{k}^{\text{BPG}} - X_{0k}^{\text{BPG}})^2}. \quad (17)$$

As it is shown, the Euclidean distance in the $n$-dimensional space ($n = 7$) could be used to measure the similarity between the “perfect” business process model and the measured business process model [8].

Therefore, business process models of high quality should demonstrate zero distance $d \left( X^{\text{BPG}}, X_{0}^{\text{BPG}} \right) = 0$ to the vector of a “perfect” business process model $X_{0}^{\text{BPG}}$. Otherwise, if $d \left( X^{\text{BPG}}, X_{0}^{\text{BPG}} \right) > 0$, there could be present violations of complexity, structuredness, or uncertainty quality features.
The algorithm for business process model quality measurement. The algorithm for business process model quality measurement may include the following steps.

Step 1. Enter the name $N$ of the measured business process model.

Step 2. Enter the name $N_i$ of the activity and its input/output sequence flow numbers $[i^\text{in}]$ and $[i^\text{out}]$.

Step 3. Return to Step 2 if one more activity should be added. Otherwise – go to Step 4.

Step 4. Build the set of activities $T$ , calculate measures defined in (4), (8), and (11), and build the vector $X^{BPG}$ for the model $N$.

Step 5. Calculate the distance value $d\left(X^{BPG}, X^{BPG}_0\right)$. If $d\left(X^{BPG}, X^{BPG}_0\right) = 0$, model $N$ does not have complexity, structuredness, or uncertainty mistakes. Otherwise – go to Step 6.

Step 6. Check the vector $X^{BPG}$ and make the following assumptions, so fault activities $N_i, t \in T$ could be detected and fixed:

- if $i^{(N)}_{\text{in}} \neq 0 \lor i^{(N)}_{\text{out}} \neq 0$, there are activity issues in model $N$;
- if $i^{(N)}_{\text{out}} \neq 0 \lor i^{(N)}_{\text{in}} \neq 0$, there are structuredness issues in model $N$;
- if $i^{(N)}_{\text{in}} \neq 0 \lor i^{(N)}_{\text{out}} \neq 0$, there are uncertainty issues in model $N$.

In order to provide the efficient quality measurement of business process models, the proposed algorithm should be implemented as a software tool. Details related to the design and development of such a software tool are outlined in the subsequent section.

Design and development of the software tool to measure business process model quality. The developed software tool is a client-server web application based on the Java platform and programming language [9]. We used Java because of its cross-platform capabilities, relative simplicity for object-oriented design, and scalability. It uses the collection of BPMN models and the Camunda BPMN library [11]. The collection of business process models locates on the application server or on the dedicated file server. Fig. 2 below demonstrates the main nodes and components of the developed software.

Let us briefly describe the main software components (see Fig. 2 above). The client-side contains the single web page and the stylesheet. The web page communicates with the Java servlet (a server-side application that processes requests and provides responses). Utility classes (see the package “util”) allow reading and measuring BPMN models given in a collection of files. Java beans (simple object representations) describe BPMN elements, models, and measured indicators.

Fig. 3 below explains the interaction of the main software components. The described workflow includes the following steps:

- the homepage accesses the backend Java code (i.e. the “MeasureServlet” class);
- the respective servlet calls the special method of the “CollectionUtil” class to read the collection of BPMN models from the server;
- the servlet then calls the special method of the “CollectionUtil” class to measure the quality of retrieved BPMN models;
- the “CollectionUtil” class returns the results to the servlet that then displays these results on the “index” page to the user.

Quality measurement of sample business process models. Fig. 4 below shows the usage of the developed software tool “BPMNmeter”. We used the Camunda repository [12] of goods dispatch process models. Models with measures different from “perfect” are colored in red, while high-quality models are highlighted in green. Models are sorted in reverse order by distance from the “perfect” model.

Fig. 5 below demonstrates one of the goods dispatch business process models of poor quality, named as “Shipping1_c87ef14a31294d689947d679015e8afb”.

Fig. 5 below demonstrates one of the goods dispatch business process models of poor quality, named as “Shipping1_c87ef14a31294d689947d679015e8afb”.

The described workflow includes the following steps:

- the user uploads a BPMN model to the server;
- the user visits the software tool’s homepages (i.e. the “index” page);
Its distance from the “perfect” BPMN model is
$$d\left( X^{BPMN}, X^{BPMN}_{perfect} \right) = 1.43.$$ Fig. 5 demonstrates the respective diagram, which contains two types of modeling mistakes:

- missing incoming flow for the “Determine shipping method” task leading to the never initiated process path;
- multiple incoming flows for the “Create package label” leading to missing synchronization.

Obtained measures prove the observations outlined above. The considered model demonstrates deviations of measures from “perfect” indicators:

- the maximum degree of a business process model activity, $$\hat{\max}^\alpha = 3 \neq 2$$;
- the relative number of tasks, each of which does not have incoming sequence flows, $$\hat{\alpha}^\alpha = 0.14 \neq 0$$;

- the relative number of tasks, each of which has multiple outgoing sequence flows, $$\hat{\alpha}^\alpha = 0.14 \neq 0$$.

These measures reflect complexity, structuredness, and uncertainty issues in the analyzed BPMN model.

Complexity issues signalize the complicated modifiability and maintainability of such a model. Structuredness issues signalize modeling errors that violate the workflow of a business process (e.g. the path beginning with the “Determine shipping method” activity may never trigger). Uncertainty issues signalize the complicated understandability of such a model and, therefore, a business process (e.g. how possible parallel flows synchronize before the “Create package label” activity).

**Conclusion.** In this study, we considered the problem of quality measurement of business process models using the n-dimensional space distances. Since process models,
which are not understandable, unstructured, and/or uncertain, may cause even more issues when used to improve organizational activities, this paper proposed quality measures and the software tool that can be used to detect errors in BPMN and EPC business process models using distances in n-dimensional space.

The formal problem statement based on the graph-based description of business process models was given. Business process model quality characteristics and measures that consider features of both BPMN and EPC notations were proposed. Distances in n-dimensional space were suggested for usage to measure the quality of business process models. The algorithm for business process model quality measurement was proposed. The software tool to measure the quality of BPMN and EPC business process models was designed and developed. The quality of sample business process models was measured using the software tool. Obtained results were analyzed and discussed.

Limitations of this paper are related to the consideration of only tasks as the most important business process modeling elements, whereas gateways (i.e. logical connectors) are also valuable and error-prone objects. Also, this paper considers the Euclidean distance only, while other existing n-dimensional space distance measures or, on the opposite, similarity measures, can be used to evaluate business process model quality. These limitations should be taken into account in future research.

References


References (transliterated)


References


References (transliterated)


References


References (transliterated)