18th PCSF 2018
18th Professional Culture of the Specialist of the Future

BUSINESS PROCESS ANALYTICAL MODELING — BUSINESS ANALYSIS BODY OF KNOWLEDGE REQUIRED COMPONENT

Sidnev Aleksandr (a)*
*Corresponding author

(a) Peter the Great St. Petersburg Polytechnic University (SPbPU), Polytechnicheskaya 29, Saint Petersburg, 195251 Russia, alex2922@gmail.com

Abstract

The analysis of the situation with the business process effectiveness quantitative evaluation revealed a clear disregard for analytical methods as opposed to simulation methods. This conclusion follows from the analysis of such a document as the Business Analysis Body of Knowledge regulating the skills required for a successful business analyst. The study of published training programs for business analysts (master programs of economic universities and short-term local courses as well) leads us to the same outcome. At the same time, the well-known truth that there is nothing more practical than a good theory, now also continues to be relevant. The evaluation of the business process performance indicators can be successfully carried out by using formal mathematical models borrowed from applied mathematics. The stochastic graph that allows cycles and loops is proposed as a formal model that provides the analytical methods for the business process analysis. The technique of obtaining an analytical expression of the moment generating function (MGF) of the business process duration is described in series of steps. After the MGF is revealed then the business process distribution law in the format of the probability density function is restored. The application sphere of the proposed methodology, its limitations and ways to overcome them are analyzed. The verification of the results obtained for a number of examples is carried out by means of the GPSS simulation model. The conclusion is made about the expediency of including the business process mathematical modeling in the sphere of business analysis application.

1. Introduction

The task of business process (BP) analysis is topical and undoubtedly takes a worthy place in the field of business analytics. In relation to this task it is worth considering the Techniques section of the BABOK guide (2015). It is already generally accepted that the formulation of the requirements for competencies and the quality of human capital in the era of industrial 4.0 society serves as the basis for the training programs formation in various fields (Alexankov, Trostinskaya, & Pokrovskaia, 2018). The above mentioned techniques section picks out the required business analyst skills, namely Decision Analysis, Decision Modeling, Process Analysis, and Process Modeling. According to the BABOK guide the distinguished skills require the analyst ability to formulate a task (to state a problem) merely in a descriptive way — in the format of tables or a standard description of the IDEF, BPMN or UML type. The problem of the formal BP model creation is not considered actual for a business analyst — he is faced with the need to rely on the mathematical background of the proposed software.

The lack of analytical BP modeling in business analysis sphere is confirmed by the following fact. (Jahangirian, Eldabi, Naseer, Stergioulas, & Young, 2010) carried out the study based on more than thousand papers review on a topic of formalisms used in the BP modeling. The conclusion is that no papers found in the Analytical Modeling section. Past analysis of the sources which are often cited show that the situation does not improve to the present time, which is confirmed by the data in Table 01. Brief description of the most cited sources on the topic of the BP modeling.

Table 01. Brief description of the most cited sources on the topic of the analytical BP modeling

Source of information	The core idea of the paper on the topic of the BP modeling	Does the topic of BP analytical Modeling present in the paper
YAkimov, Kirpichnikov, Mokshin, Kostyuhina, & SHigaeva (2014); YAkimov et al. (2014).	Association BP Editor with the BP simulation packet	No
Guseva, & Varfolomeeva (2014); YAkimov, Kirpichnikov, Mokshin, Alyautdinova, & Pajgina (2015).	BP simulation with the Arena packet	No
Barjis (2008).	Introducing the DEMO methodology based on the Petri net formal semantics with appropriate simulation engine	No
Melao & Pidd (2006).	Using component technology to develop a simulation library for	No

	the BP modeling with SIM ++ packet	
YAgudina & Novikova (2015); Kulagina & Semikin (2015); Novikova, Kurzaeva, Petelyak, Maslennikova, & Belousova (2015).	Descriptive BP models in the format of diagrams in IDEF0 and ARIS standards	No
Liu, Zhang, Li, & Jiao (2012); Khodyrev, & Popova (2014).	BP models creation based on the Process Mining idea in the Petri Nets format. Petri nets are hardly suitable to the formal analysis	No

The search for the business analysis training programs on the keywords "Business Modeling" and "Business Analytics" revealed that none of them presented methods for the BP analytical modeling. It can be stated that the business analyst's wealth of knowledge of the BP modeling is actually limited to the subject-oriented simulation software being the part of the unified methodologies like ARIS, for example, studied in the Business Processes Modeling and Optimization program by the National Research University Higher School of Economics (https://www.hse.ru/ma/bi/about). At the same time analytical BP models are the ones that allow researcher to obtain the most valuable information about the BP under investigation. First, analytical models are able to create explicit dependencies of the BP efficiency indicators on their parameters, secondly, these models generate constructive BP optimization formulations and, thirdly, they give a more accurate forecast of rare events in comparison with simulation. The importance of BP analytical modeling is shown below with reference to the evaluation of the BP temporal properties on the basis of the appropriate mathematical formalism.

2. Problem Statement

The problem is stated as follows.

1. To demonstrate the effectiveness of analytical BP models, in particular, the flowgraph algebra applied by Ren (2011) for estimating the parameters of the BP duration distribution law.

2. To suggest the BP initial description format.

3. To determine the transition from the original BP description to its mathematical model.

4. To propose the methodology for the BP duration quantitative evaluation.

5. To formulate the optimization problem variants for the minimizing the BP duration and the BP costs as well.

We will consider BP as an ordered set of activities that is subordinate to the goal achievement. Such an objective may be the fact of the BP completion with a certain result. In accordance with this definition, the initial formalization of the business process can be carried out using any of the known standards (IDEF,

BPMN], UML, etc.). As a formal mathematical model that provides the application of analytical methods for the BP analyzing, we propose a stochastic graph that allows cycles and loops (see Figure 01).
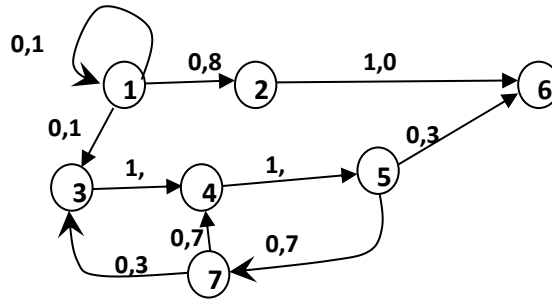

Figure 01. BP stochastic graph

A similar stochastic graph can be constructed for an arbitrary BP. Each BP activity is associated with the arc, and each node is associated with the event. If the event is accomplished, it means that all the activities-arcs entering the event are finished, and the activities-arcs going out from the event-node can be started. Each arc ij is marked by the likelihood p_{ij} that the activity ij will be started (subject to the event i accomplishment). The duration of each activity is a random variable with a distribution law given.

There are initial and final vertices in this graph, symbolizing the beginning and completion of the BP in question, respectively. In this case, they are vertices 1 and 6. The problem is to find the BP duration distribution law not by simulation.

We suggest to apply here the flowgraph algebra (Ren, 2011), which provides the BP duration MGF in a closed form. This method involves a transition from the activity duration distribution law to its MGF, which is possible for a significant number of popular distribution laws (Normal, Exponential, Triangular, Uniform, Erlangian).

The BP duration distribution law search is carried out in two steps.

1. The BP duration MGF finding out.

2. The BP duration distribution law restoration based on the basis of the BP duration MGF found earlier.

The mathematical content of this problem is the compliance with the flowgraph algebra in the first step and the search for the inverse Laplace transform, which restores the BP duration probability density function according its MGF in the final second step. It is necessary to show, on an example, the expediency of applying the proposed approach to the BP duration distribution law finding out.

3. Research Questions

- To demonstrate the steps leading to the BP duration distribution law finding out for the BP defined with the stochastic graph.

- To compare the results of analytical modeling and simulation with regard to the BP duration distribution law finding out. To make an informed conclusion about the applicability of the proposed approach.

1340

4. Purpose of the Study

- To make an assessment of the analytical methods for the BP analysis.
- To formulate the BP optimization problems for the BP synthesis.
- To offer the prospect for the BP analytical modeling.

5. Research Methods

5.1. The BP duration MGF creating in a closed form

The BP is specified by the graph presented in the Figure 01. The function $w_{ij} = p_{ij} \cdot MGF_{ij}(s)$ is associated with each activity ij. In accordance with the flowgraph algebra (Ren, 2011), the next steps are to be accomplished to obtain the BP duration MGF in a general closed form.

5.1.1. The branch transmittance matrix Q creation according to the BP stochastic graph

$$Q = \begin{bmatrix} w11 & w12 & w13 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & w26 & 0 \\ 0 & 0 & 0 & w34 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w45 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & w56 & w57 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & w73 & w74 & 0 & 0 & 0 \end{bmatrix}$$

5.1.2 Sequential matrix Q processing for the obtaining the BP duration MGF generating function creation.

$$A = I_7 - Q^T$$

$$MGF_{16} = A^{-1}(6,1)$$

The corresponding series of Matlab commands is presented further.

```
>> A = eye(size(Q,1)) - transpose(Q)

>> A_inv = inv(A)

>> MGF16 = A_inv(6,1)
```

MGF_{16} = (w12*w26 + w13*w34*w45*w56 - w12*w26*w45*w57*w74 - w12*w26*w34*w45*w57*w73)/((w11 - 1)*(w45*w57*w74 + w34*w45*w57*w73 - 1))

Substituting the actual expressions of the activity duration MGF (Table 02) in the general expression for MGF16, we provide the exact formula of the BP duration MGF.

Table 02. MGF for the set of distributions

Distribution	PDF	MGF(s)
Deterministic	$y = m = Const$	e^{sm}
Exponential	$f(y) = \dfrac{1}{m} e^{-\frac{y}{m}},\ y \geq 0$	$\left(1 - \dfrac{s}{m}\right)^{-1}$
Uniform	$f(y) = \dfrac{1}{(b-a)},\ y \in [a,b]$	$\dfrac{e^{sa} - e^{sb}}{(a-b)s}$
Erlangian	$f(y) = \dfrac{\dfrac{r}{m}\left(\dfrac{ry}{m}\right)^{r-1}}{(r-1)!} e^{-\frac{ry}{m}},\ y \geq 0$	$\left(1 - \left(\dfrac{r}{m}\right)s\right)^{-r}$
Triangular	$f(y) = \begin{cases} 0, & y < a \\ \dfrac{4(y-a)}{(b-a)^2}, & y \in \left[a, \dfrac{a+b}{2}\right] \\ \dfrac{4(b-y)}{(b-a)^2}, & y \in \left[\dfrac{a+b}{2}, b\right] \\ 0, & y > b \end{cases}$	$\left[\dfrac{2(e^{sb/2} - e^{sa/2})}{(b-a)s}\right]^2$
Normal (Gauss)	$f(y) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-m)^2}{2\sigma^2}}$	$e^{sm + (1/2)s^2\sigma^2}$

In particular, the next result is obtained for the exponential distribution law of the activity duration (in the Matlab journal format).

$MGF_{16}\_EXP= =$
(1219778560000*s^7 - 438625062400*s^6 + 66653827200*s^5 - 5427205280*s^4 + 248359264*s^3 - 6187560*s^2 + 73504*s - 270)/(3*(396428032000000*s^9 - 169693218240000*s^8 + 31804363558400*s^7 - 3389757567200*s^6 + 223841814800*s^5 - 9397605944*s^4 + 247823314*s^3 - 3882989*s^2 + 31509*s - 90)).

Similar formulas for the other activity duration distributions: Triangular, Normal, Erlangian, Uniform are also obtained the same way and used to restore the BP duration MGF in question.

5.2. The BP duration probability density function restoration according its MGF.

5.2.1. Polynomial Pade approximation of the MGF(s):

$$MGF(s) \approx \frac{U(s)}{R(s)}$$

5.2.2. Replacing s by –s in the Pade MGF approximation:

$$\frac{U(s)}{R(s)} \to \frac{U(-s)}{R(-s)}$$

5.2.3. Finding out the roots of equation $R(-s) = 0$

5.2.4. Following the Heaviside formula for the the BP duration probability density function:

$$f(y) = \sum_{k=1}^{q} \frac{U(\alpha_k)}{R'(\alpha_k)} e^{\alpha_k y}$$

This approximation of the BP duration density function was obtained following this technique for the example in question (see Figure 01) with the exponential activity duration distribution by means of the Matlab Symbolic Toolbox:

f(y_{16}) = 0.13586*exp(-0.030297*t) + 0.00081976*exp(-0.0055839*t) - 0.13672*exp(-0.037652*t)

The obtained density distribution function plot and the histogram of the GPSS simulation data sample are shown in Fig. 02. The same comparison is carried out for the triangular activity duration distribution (Fig. 03). The similarity of the theory and experimental data from the GPSS is verified for the triangular case by three goodness-of-fit tests: Kolmogorov-Smirnov, hi-square and Cramér–von Mises. All criterions confirm the null hypothesis (for significance level 0.05) about the coincidence of distributions (analytical and empirical).
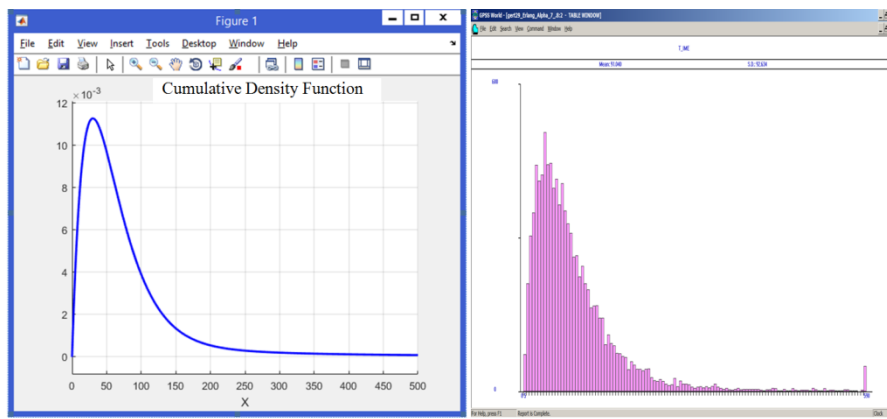

Figure 02. Analytical and empirical distributions for the exponential activity duration distribution
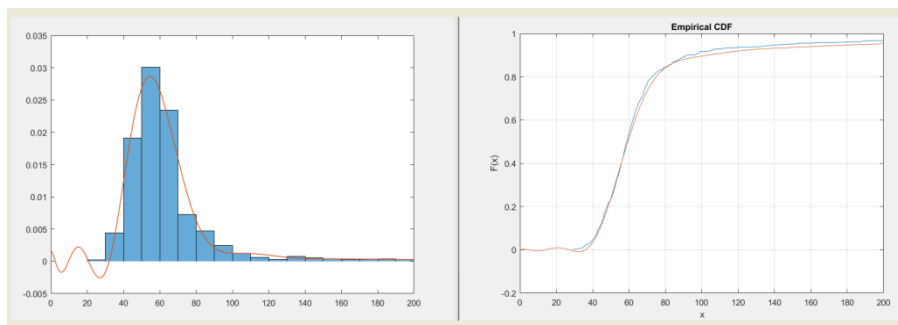

Figure 03. Analytical and empirical distributions for the triangular activity duration distribution

6. Findings

The methodology, suggested, of finding out the BP duration distribution law was tested on various business process graphs (acyclic, with loops and contours) for various activity duration distributions mentioned in Tab. 03. So we came to the conclusion that the suggested technique allows one to abandon

the simulation in favor of an analytical approach to the finding out the BP duration distribution law. All transforms and calculations were carried out in the Matlab Symbolic Toolbox of the Matlab package of versions R2015b and R2016b. The effect, which can be called the model convergence, has also been confirmed. If we consider the BP executor as the network of queues, then the time spent by the client in network is a linear function of the time spent in nodes. The same result is relevant when using the proposed technique. The next formula for the expected value of the BP duration is obtained for the example in question.

$m_{16} = M_{11}/9 + (8*M_{12})/9 + M_{13}/9 + (8*M_{26})/9 + (17*M_{34})/90 + (10*M_{45})/27 + M_{56}/9 + (7*M_{57})/27 + (7*M_{73})/90 + (49*M_{74})/270$ where M_{IJ} is the expected value of the activity ij duration. Taking into account the obtained linear subjection, we can formulate the following problems of optimal allocation of means for the BP accelerating.

$$\min T = \sum_{(ij)} C_{ij} M_{ij}(x_{ij}) \qquad \min X = \sum_{(ij)} x_{ij}$$
$$\textit{subject to} \qquad \text{or} \qquad \textit{subject to}$$
$$\sum_{(ij)} x_{ij} = X \qquad \sum_{(ij)} C_{ij} M_{ij}(x_{ij}) \leq T$$

where x_{ij} is the cost of achieving the expected value M_{ij} of the activity ij duration, X is the available amount of funds and T is the BP duration. Analytical BP modeling has revealed curious features of the influence of variation in the activity duration and the properties of the BP stochastic graph on the coefficient of variation in the BP duration. A study of the BP duration variability was made for each of the four business process variants specified by the same stochastic graph (see Figure 01), with the same set of expected activity duration, but with different coefficients of variation from zero for deterministic times to single for the exponential distribution (see Table 03). The data of Table 02 indicate that a significant reduction in the variance of the activity duration may not lead to the same decrease in the variance of the BP duration.

Table 03. BP duration variation depending on the activity variability

Duration Activity Distribution	Activity Duration Variation Coefficient	m_{16}	σ_{16}	$\dfrac{\sigma_{16}}{m_{16}}$
Exponential	1	77.863	89.455	1.149
Erlangian₃	0.577	77.863	81.603	1.048
Erlangian₁₀	0.316	77.863	78.700	1.011
Deterministic	0	77.863	77.379	0.994

It should be noted that the flowgraph algebra and the GERT methodology are limited as well. No parallel branches in the BP stochastic graph are allowed which means that concurrent activities are excluded from the BP structure which, nevertheless, is quite possible in real practice: Figure 04 shows the graph with two parallel activities between nodes 2 and 6 (the sum of the probabilities of arcs issuing from the node 2 is equal to 2).
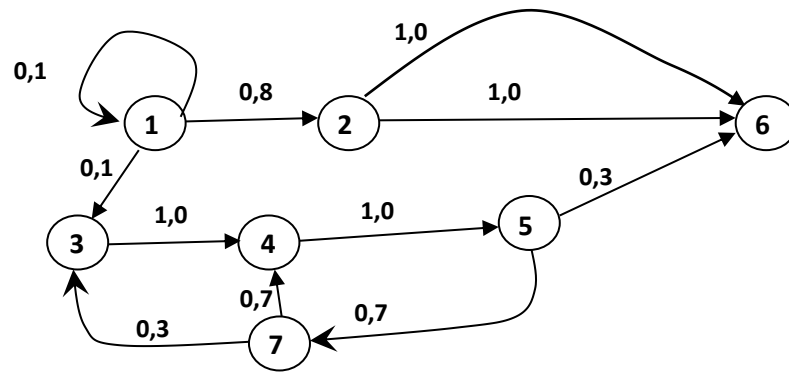
Figure 04. BP Stochastic graph with parallel arcs

This restriction can be removed using the order statistics tool - all the parallel activities are replaced by the single activity with the duration equal to the maximum of the concurrent activities durations. For this "unifying" activity it remains to find the MGF expression. This is most simply done for exponential law, for others it is somewhat more complicated. In any case, it was possible to solve this problem with Matlab Symbolic Toolbox for uniform and triangular distributions. Usually the common situation with the activity duration data is as follows: the average duration and acceptable shift are given. That is, the uniform and triangular distributions are consistent with the actual initial data. Symbolic computations in MATLAB R2016b make it possible to find the MGF expression of a maximum of several (at least 6) random variables, both with the same and with different distribution parameters. It is worth continuing research in this direction.

7. Conclusion

The experience of using the method of analytical BP modeling to assess its temporal properties allows us to make two significant conclusions. Conclusion on the advisability of using the apparatus of the flowgraph algebra by means of the Matlab system for evaluating the BP duration and the indicators directly associated with it as well. Another item relates to the idea of the BP model convergence enhancement. The problem of the BP duration estimating, as already noted, can be interpreted as the investigation of a network of queues, closed and open, depending on the context of the original problem. So it becomes possible to involve an appropriate mathematical apparatus for evaluating the BP effectiveness. Thus, we can draw a general conclusion about the need to include knowledge and skills in BP analytical modeling in the practice of business analysis. The vision of the course on the BP mathematical modeling with emphasis to analytical BP modeling was presented in Sidnev (2016).

References

Alexankov, A.M., Trostinskaya, I.R., & Pokrovskaia, N.N. (2018). Industry 4.0 Requirements For Quality Of Human Capital And Competencies Formed Within Educational Institutions. *The European Proceedings of Social & Behavioral Sciences EpSBS*, *4 (XXXIV-RPTSS),* 26-34. doi: 10.15405 / epsbs.2018.02.4

BABOK guide (2015). *A Guide to the Business Analysis Body of Knowledge*. Toronto Ontario, Canada: International Institute of Business Analysis. Retrieved from http://www.iiba.org/babok-guide.aspx

Barjis, J. (2008). The importance of business process modeling in software systems design.

Science of Computer Programming 71, 73–87 doi:10.1016/j.scico.2008.01.002

Guseva, E.N., & Varfolomeeva, T.N. (2014). Primenenie imitacionnyh modelej dlya resheniya ehkonomicheskih zadach optimizacii [The use of simulation models for solving economic problems of optimization]. *Sovremennye problemy nauki i obrazovaniya, 6,* 200. [in Rus.]

Jahangirian, M., Eldabi, T., Naseer, A., Stergioulas, L.K., & Young, T. (2010). Simulation in manufacturing and business: A review. *European Journal of Operational Research, 203,* 1–13

Khodyrev, I., & Popova, S. (2014). Discrete modeling and simulation of business processes using event logs. *Procedia Computer Science, Vol. 29*, 322–331. doi: 10.1016/j.procs.2014.05.029

Kulagina, I.I., & Semikin, D.V. (2015). Voprosy modelirovaniya biznes-processov v sootvetstvii so standartami ISO serii 9000 [The issues of modeling of business processes in accordance with the standards of ISO 9000]. *Biznes. Obrazovanie. Pravo, 1 (30).* 219-223. [in Rus.]

Liu, Y., Zhang, H., Li, C., & Jiao, R.J. (2012). Workflow simulation for operational decision support using event graph through process mining. *Decision Support Systems 52,* 685–697

Melao, N., & Pidd, M. (2006). Using component technology to develop a simulation library for business process modeling. *European Journal of Operational Research, 172,* 163–178 doi:10.1016/j.ejor.2004.09.033

Novikova, T.B., Kurzaeva, L.V., Petelyak, V.E., Maslennikova, O.E., & Belousova I.D. (2015). Opisanie upravleniya biznes-processami predpriyatiya na osnove metodologii IDEF0: trudnosti razrabotki, rekomendacii po sovershenstvovaniyu postroeniya diagramm [The description of enterprise business processes management based on IDEF0 methodology: difficulties of development, modelling improvement recommendations]. *Fundamental'nye issledovaniy,. 8-2,* 318-322. [in Rus.]

Ren, Y. (2011). *The Methodology of Flowgraph Models.* London School of Economics and Political Science. (Doctoral dissertation). Retrieved from http://etheses.lse.ac.uk/318/1/Ren_The%20Methodology%20of%20Flowgraph.pdf

Sidnev, A. (2016). The essence of the mathematical business process models discipline for computer science students. In L.G. Chova, Al. Martinez & IC. Torres (Eds), *ICERI2016 Proceedings.* Seville, Spain:IATED. doi: 10.21125/iceri.2016.0883

YAgudina, R.R., & Novikova, T.B. (2015). Biznes-modelirovanie processa "opisanie deyatel'nosti kafeteriya" s ispol'zovaniem metodologii ARIS [Modeling the business process activity "description cafeteria" using the ARIS methodology]. *Sovremennye nauchnye issledovaniya i innovacii, 11* (55)*,* 154-161. [in Rus.]

YAkimov, I.M., Kirpichnikov, A.P., Mokshin, V.V., Alyautdinova, G.R., & Pajgina L.R. (2015). Imitacionnoe modelirovanie biznes-processov v sisteme bizagi modeler [Business process simulation in Bizagi modeler system]. *Vestnik Tekhnologicheskogo universiteta, 18(9),* 236-239. [in Rus.]

YAkimov, I.M., Kirpichnikov, A.P., Mokshin, V.V., Kostyuhina, G.V., & SHigaeva T.A. (2014). Kompleksnyj podhod k modelirovaniyu slozhnyh sistem v sisteme BPWIN-ARENA [Complex approach to complicated system modeling with BPWIN-ARENA set]. *Vestnik Kazanskogo tekhnologicheskogo universiteta, 6,* 287-292. [in Rus.]

YAkimov, I.M., Kirpichnikov, A.P., Mokshin, V.V., Mahmutov, M.T., Pejsahova, M.L., Valieva, A.H., & Nizamiev, B.A. (2014). Strukturnoe modelirovanie biznes-processov v sistemah BPMN EDITOR, ELMA, RUNAWFE [Business process structural modeling in BPMN EDITOR, ELMA, RUNAWFE systems]. *Vestnik Kazanskogo tekhnologicheskogo universiteta, 17(10)*, 249-256. [in Rus.]