# Scalability Measurement of Business Process Model Using Business Processes Similarity and Complexity

Muhammad Ainul Yaqin[1], Riyanarto Sarno[2], Abd. Charis Fauzan[3]
[1,2,3]*Informatics Department, Institut Teknologi Sepuluh Nopember*
[1]*Informatics Department, State Islamic University of Maulana Malik Ibrahim*
[1]*yaqin16@mhs.if.its.ac.id,* [2]*riyanarto@if.its.ac.id,* [3]*fauzan16@mhs.if.its.ac.id*

**Abstract – The scalability of business process model shows the ability of a system to handle the amount of provided business process activity. In business process management, scalability is used to compare the size of a business process model to other business process models. This research focused on measuring the scalability of business process models. Every business processes were modeled using Petri Net. Petri Net was used as a business process model because of its simple model notation so it can be analyzed mathematically. The scalability of business processes had been done by comparing the similarity of some business processes and their scales. The similarity measurement method proposed were based on structural and behavioral using Jaccard coefficient, whereas the number of elements of the business process model and control flow complexity were used to measure its scale. Then, the scale value (0..1) and the similarity of the business process model was calculated by the specified formula. The result is a scalability value where the greater the value of the scalability, then the growth potential of the business process is also greater.**

**Keywords – scalability; scale; model business processes; structural similarity; behavior similarity; petri net; complexity; control flow.**

## I. INTRODUCTION

System scalability is the ability of a system to handle its growth on the number of jobs or its potential to handle that growth [1]. Similarly, business process scalability can be defined as the ability of a business process to handle the growth on the number of processes or the potential of the business processes to be enlarged.

System scalability is determined by scalability metric value. Scalability metrics can be determined based on throughput, resource utilization, and costs [2,3,4,5,6]. Research on scalability measurements that had been performed have not yet measured the scalability metrics based on the similarity and complexity of business processes.

Therefore, this paper defines the attributes which make the business process is scalable. The business process scalability is measured by comparing it with other business processes. The scalability measurements used were structural similarity, behavioral similarity, and complexity parameters.

## II. LITERATURE REVIEW

### *A. Business Process Model*

The business process is a set of activities that work under coordination in an organizational and technical environment [7]. Business process model is a set of models activities and execution limits between them [7]. Business process models can be represented in various ways, namely Petri net, BPMN, and graph [7,8]. In this research, the business processes were modeled with petri net.

This paper used four business process models that represents several conditions. The business process model A is the simplest business process model as shown in Fig. 2. The business process model B is a more complex business process model. The business process model B has more transition numbers than the business process model A and sequenced sequentially as shown in Fig. 3. The business process model C shows the process model with the addition of two "and" branches as shown in Fig. 4. The business process model D shows the process model by replacing the logic of one of the "and" branch with "xor" branch as shown in Fig. 5.

### *B. Business Process Scalability*

The scalability of business processes is the ability of the process to handle the growing number of business processes or to accommodate its growth. Fig. 1 illustrates the scalability of the business process model, where the business process A is part of the business process B, and the business process A can grow into the business process B. In other words, the business process A is scalable to the business process B. According to Yong Chen et al. [4], Scalability is the efficiency (E) ratio of the processor work in two cases that can be formulated as (1)

$$\psi(A,B) = \frac{E(A)}{E(B)} \quad (1)$$

978-1-5386-0549-3/17/$31.00 ©2017 IEEE

While Wei-Tek Tsai et al. [4] proposed that scalability metrics in SaaS applications are obtained by measuring the Performance Resource Ratio (PRR) combined with data mining techniques. For distributed systems, Prasad Jogakelar et al. [3] proposes the use of throughput, the mean value of each response, and cost. The results are compared with comparability scalability. Previously, Xian-He Sun et al. [6] has stated that when measurements are made in a homogeneous environment, the scale factor is the number of processors, and it will get speedup, efficiency, and scalability

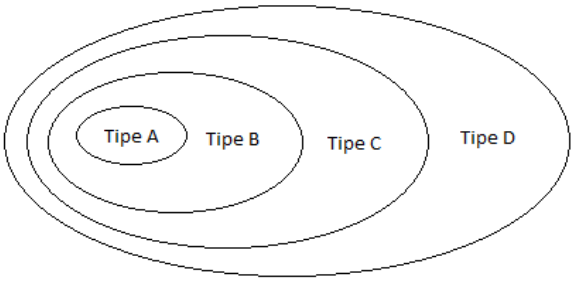


**Fig. 1Illustration of model business process scalability**

### C. *Workflow Similarity*

Process similarity is the process of discovering the similarity of several business processes. The discovery of this similarity is based on research that has been done by Marc Ehrig, Agnes Koschmider, and Andreas Oberweis [9]. The similarities of the workflow to another workflow were calculated after the workflow has been fragmented. The similarity measured or calculated are :

- Elements of a process model consists of:
  - Semantic similarity
  - Syntactic similarity
  - Contextual similarity
- Label matching
- Structural
- Behavioral

The above measurement aspects were raised by several researchers, namely Michael Becker and Ralf Laue [10], Remco Dijkman et.al. [11], and Riyanarto Sarno et.al. [12].

In this research, workflow similarities were calculated based on structural and behavioral aspects only. Workflow similarities were calculated by Jaccard coefficient method.

### D. *Structural Similarity*

The structural similarity is the similarity of two graph structures [13]. The graph structure is all the elements in the graph. Methods for calculating structural similarities are numerous. In this research Jaccard coefficient method was used.

*Jaccard coefficient* [14] was used to compare a set of the limited sample size similarity and define the combined set of samples intersection. Jaccard coefficient is formulated as (2)

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

### E. *Behavioral Similarity*

A similarity on behavior is a measure of similarity of business processes based on the relationship between the behavior of business processes [15]. The value of behavioral similarity is derived from the relationships between the existing business process activities. Those activity relations can also be interpreted as a possible sequence of executions. The method used to calculate the similarity of business process behavior is similar to those that have been used to calculate the structural similarities.

### F. *Control Flow Complexity*

Process complexity is the difficulty level of analysis, understanding, or explanation of a process. The metrics used to analyze business process complexity are Control Flow Complexity (CFC) [16,17]. The CFC measurements were performed by summing all split builders and calculated as (3)

$$CFC(X) = \sum CFC_{XOR-split}(X) + \sum CFC_{OR-split}(X) + \sum CFC_{AND-split}(X) \tag{3}$$

Where

$CFC_{XOR-split}(X) = fanout(X)$

$CFC_{OR-split}(X) = 2^{fanout(X)} - 1$

$CFC_{AND-split}(X) = 1$

The greater the CFC value means the overall structural complexity is also higher [17].

## III. PROPOSED METHOD

This research procedure consists of four stages, namely business process modeling with net entry, calculates workflow similarity, calculate workflow scalability, and scalability analysis in extreme cases. Details of explanations are as follows.

### *A. Business Process modeling with Petri-Net*

At this stage, the business process modeling was done with petri net. Four business processes were modeled as in figure 2 to 5. The business process model A has four transitions, and the process model B, C, and D consist of eight transitions. The process model A and B are sequence process models, while the process model C and D are with branch.

The process model A and B were compared to determine the effect of adding transitions and the scalability of sequence process models. The comparison of process model B and C were done to know the effect of the insertion of "and" logic with the same number of transitions on scalability. The process model C and D were compared to find out the effect of "xor" logic insertion on scalability.

### *B. Calculation of similarity of workflow*

Firstly, the business process model is obtained, then calculate the similarity between the workflows formed based on the business process model. The calculation of workflow similarity was done in two aspects, namely structural aspects and behavioral aspects.

The structural similarity is calculated by the Jaccard coefficient method as (2). Jaccard coefficient method was chosen because it is more fair and simple in calculating the similarity index when it is compared to Cosine, Dice and Overlap calculation formula. The calculation processes are as follows:

Given :
Model A = p1, p2, p3, p4, p5, A, B, C, D, p1-A, A-p2, p2-B, B-p3, p3-C, C-p4, p4-D, D-p5.
The business process model A has 17 elements
Model B = p1, p2, p3, p4, p5, p6, p7, p8, p9, A, B, C, D, E, F, G, H, p1-A, A-p2, p2-B, B-p3, p3-C, C-p4, p4-D, D-p5, p5-E, E-p6, p6-F, F-p7, p7-G, G-p8, p8-H, H-p9
The business process model B has 33 elements. The numbers were inserted into the calculation of similarity formula. Based on the workflow in Figures 2 to 5, then the workflows were processed the same way with the above. The results of calculations were obtained in Table 1.

**Table 1 Results of the structural similarity calculation with Jaccard Coefficient**

| | A | B | C | D |
|---|---|---|---|---|
| **A** | 1 | 0.51 | 0.31 | 0.12 |
| **B** | 0.51 | 1 | 0.41 | 0.25 |
| **C** | 0.31 | 0.41 | 1 | 0.6 |
| **D** | 0.12 | 0.25 | 0.6 | 1 |

$$J(A,B) = \frac{17}{(17+33)-17} = 0.51$$

The next calculation is the calculation of behavior similarity. Based on the workflow in Figures 2 to 5, the behavioral similarity was calculated by Transition Adjacency Relations (TAR) Jaccard Coefficient method. The calculations are as follows:

Given :
Model A = AB, BC, BD.
The business process model A has 3 elements.
Model B = AB, BC, BD, DE, EF, FG, GH.
The business process model B has 7 elements. The numbers will be used for the calculation of similarity formula.

$$J(A,B) = \frac{3}{(3+7)-3} = 0{,}43$$

Based on the workflow in Figure 2 – 5, then the workflows were processed the same way with the above. The result of calculations are obtained in Table 2.

**Table 2 Results of the behavioral similarity calculation with Jaccard Coefficient**

| | A | B | C | D |
|---|---|---|---|---|
| **A** | 1 | 0.43 | 0.13 | 0.13 |
| **B** | 0.43 | 1 | 0.23 | 0.23 |
| **C** | 0.13 | 0.23 | 1 | 1 |
| **D** | 0.13 | 0.23 | 1 | 1 |

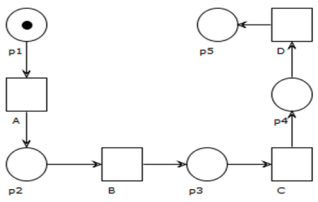


**Fig. 2 Business process model A**

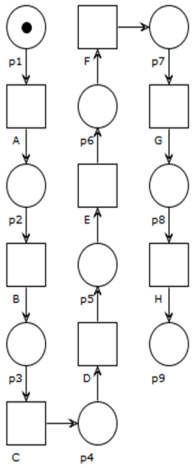


**Fig. 3 Business process model B**

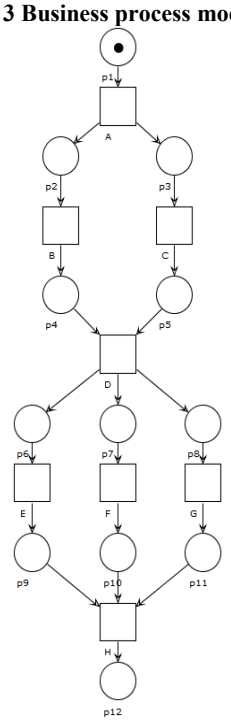


**Fig. 4 Business process model C**

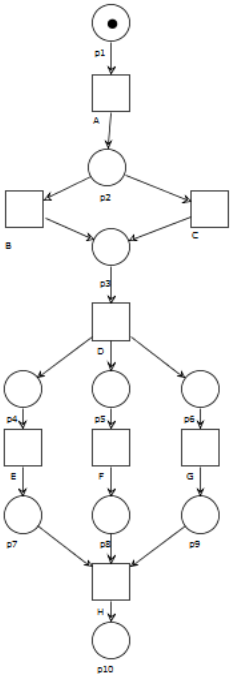


**Fig. 5 Business process model D**

### *C. Calculation of scalability*

The scalability index as found in (7) is obtained through a mathematical calculation involving several parameters, i.e., simulated workflow *sim(A,B)* and *scale(A,B)*. *sim(A,B)* obtained from the average value of structural and behavioral similarities which were formulated in (4). The results of the calculation of the average value of similarities are shown in Table 3. Structural and behavioral similarities were used as scalability parameters. The value of the similarity determines whether a business process is a subset of other business processes or not.

$$sim(A,B) = average\big(simS(A,B), simB(A,B)\big) \qquad (4)$$

**Table 3 The average value of structural and behavioral similarity**

| | A | B | C | D |
|---|---|---|---|---|
| A | 1 | 0.47 | 0.22 | 0.125 |
| B | 0.47 | 1 | 0.32 | 0.24 |
| C | 0.22 | 0.32 | 1 | 0.8 |
| D | 0.125 | 0.24 | 0.8 | 1 |

Scale value (*scale(A)*) was obtained by multiplying the number of business process model elements with a Control Flow Complexity (CFC) value as shown in (5).

$$scale(A) = E(A) * CFC(A) \quad (5)$$

The result of the scale calculations obtained are shown in Table 4

**Table 4 Results of calculations of scale value**

| Business Process Model | Control Flow Complexity | Element Number | Scale Value |
|---|---|---|---|
| A | 1 | 17 | 17 |
| B | 1 | 33 | 33 |
| C | 2 | 42 | 84 |
| D | 3 | 38 | 114 |

Then, the value of the scale is compared with the scale value in the other business process model, so that (6) is obtained

$$scale(A,B) = \frac{scale(A)}{scale(B)} \quad (6)$$

Scalability value (A, B) was obtained by comparing the scale of process models A and B multiplied by the value of similarity, so that it can be formulated in (7). The scalability value is in the range of 0 to 1. If the scalability value is equal to zero, it means that the process model being compared is exactly the same, so it cannot be grown anymore. Conversely, if the scalability value is equal to one then it means that the two process models are completely different, so that it can be grown infinitely.

$$scalability(A,B) = 1 - \big(scale(A,B) * sim(A,B)\big) \quad (7)$$

After the calculation was performed, the results are on the Table 5:

**Table 5 Scalability of business process models**

| | A | B | C | D |
|---|---|---|---|---|
| **A** | 0 | 0.758 | 0.955 | 0.981 |
| **B** | - | 0 | 0.874 | 0.931 |
| **C** | - | - | 0 | 0.411 |
| **D** | - | - | - | 0 |

*D. Scalability analysis in extreme cases*

At this stage, business process scalability is analyzed for extreme cases. The first extreme case is to compare the business process model with the same structure but the behavior is different. For this case we use the process model in Fig. 6.

The calculation of similarity in both process models yields that the structural similarity = 1 and the behavior similarity = 0. Thus, the average mean = 0.5, scalability of scale (A, B) = 1, and scalability (A, B) = 0.5.

The second case is comparing business process models with the same behavior but with completely different structure. This case is impossible to happen, because with the same behavior of business process models, the compared business process models must have structural similarity.

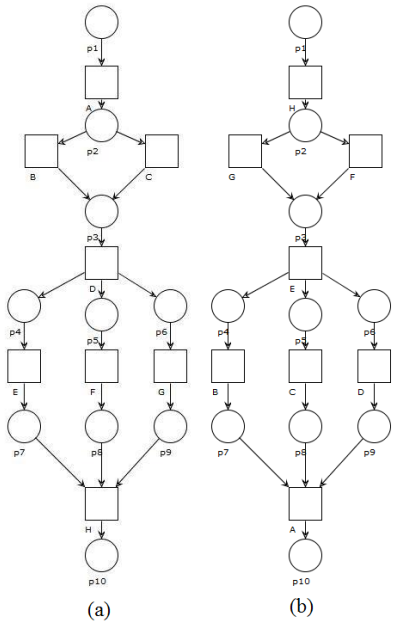


**Fig. 6 Business process model with the same structure and different behavior**

## IV. DISCUSSION

The first stage of the experiment is modeling business processes using petri net. The result of this stage is the four business process models as shown in Figure 2 to 5. The process models were used as the material for business process scalability calculation.

In the second stage, the workflow similarity is calculated. The results of similarity calculations are shown in Table 3. These numbers show that the business process model A is a subset of the business process model B. The business process model B is a subset of the business process model C. The business process model C is a subset of the business process model D. Thus, when the business process models are not similar, then they are completely different so the business process models are not scalable. Similarly, if the process models that were compared are the same, the value of similarity and scale is equal to one, and the value of scalability is zero. This means that the process model cannot be grown anymore.

In the third stage, scalability is calculated. Based on the experimental results in Table 5 it can be

seen that the effect of adding sequence transitions on the business process model B resulted in a scalability of 0.758 to the business process model A. The effect of "and" logic insertion on the business process model C resulted in a scalability of 0.874 to the business process model B. The effect of logical insertion "xor" on the business process model D yields a scalability of 0.411 to the business process model C. The results indicate that the insertion of logic is more influential to the magnitude of the scalability than the addition of sequence transitions.

Business process scale is the comparison between a business process with another business process. In this paper, the calculations of business process scale were done to the model. The scale of a business process was defined as the multiplication of the number of elements with the complexity of its flow control. Thus, business process scalability can be calculated by multiplying the scale of business processes with similarities.

From the experimental results of calculating the scalability values in Table 5, it is found that the process model with a large similarity index will result in a large scalability value as well. The value of scalability obtained shows that the greater the value of scalability, the greater the potential value of growth.

Table 5 also shows that the results of scalability calculations are only in the upper triangle, as they fit the definition of scalability that describes the growth potential of a business process to become larger. In the lower triangle, the figures depict the potential of depreciation of the business process. This violates the definition of scalability, so that on the bottom triangle is not calculated.

In the fourth stage, an analysis of the effect of similarity of the business process model was carried out on two business process models. Both process models have absolute structural similarities ($simS(A,B) = 1$) and do not similar at all in terms of their behavior ($simB(A,B) = 0$). The structural and behavioral similarity values produce an average value of 0.5. Then, scalability metrics were calculated and the result is 0.5. In fact, both business process models are completely different. Scalability metrics for different models should be zero. Therefore, in such cases there are peculiarities. The calculation formula of scalability metrics cannot be used to calculate scalability metrics. In reality, this case will never exist.

## V. CONCLUSION AND FUTURE WORKS

From the experiments and the discussions that had been done can be inferred that:

1. The value of the scalability is influenced by the similarity index, the amount of elements in the business processes model, and the complexity of the flow control.
2. Index of similarity on measuring scalability is used to ensure that the compared business process models have intersections.
3. The scale value is very important on measuring scalability.
4. Scalability metric calculation formulation cannot be used on the case when business process models has the same structure but has a completely different behavior or vice versa.

The future works are:

1. The business process model used on this research only involve “and” and “xor” logic, the future works must include “or” logic.
2. Scalability measurement is utilized for simulating the performance of business process growth.

## VI. REFERENCES

[1] Andre B. Bondi, "Characteristics of scalability and their impact on performance," in *Proceedings of the 2nd international workshop on Software and performance*, New York, 2000, pp. 195--203.

[2] Greg Barish, "Scalable and High-Performance Web Applications," in *Building Scalable and High-Performance Java Web Applications Using J2EE Technology*.: Addison-Wesley Professional, 2001, ch. 1, p. 17.

[3] Prasad Jogakelar and Murray Woodside, "Evaluating the scalability of distributed systems," *IEEE Transactions on parallel and distributed systems*, vol. 11, no. 6, pp. 589--603, 2000.

[4] Yong Chen and Xian-He Sun, "STAS: A scalability testing and analysis system," in *2006 IEEE International Conference on Cluster Computing*, 2006, pp. 1 - 10.

[5] Wei-Tek Tsai, Yu Huang, and Qihong Shao, "Testing the scalability of SaaS applications," in *2011 IEEE International Conference on Service-Oriented Computing and Applications (SOCA)*, 2011, pp. 1 - 4.

[6] Xian-He Sun and Lionel M. Ni, "Scalable problems and memory-bounded speedup," DTIC Document, ICASE Report doi:

10.1006/jpdc.1993.1087, 1992.

[7] Mathias Weske, *Business Process Management Concepts, Languages, Architectures*. Berlin: Springer, 2007.

[8] Riyanarto Sarno, Bandung Arry Sanjoyo, Imam Mukhlash, and Hanim Maria Astuti, "Petri Net Model of ERP Business Process Variations for Small and Medium Enterprises," *Journal of Theoretical and Applied Information Technology*, vol. 54, no. 1, pp. 31-38, Oct. 2013.

[9] Marc Ehrig, Agnes Koschmider, and Andreas Oberweis, "Measuring Similarity Between Business Process Model," in *Proceeding APCCM '07 Proceedings of the fourth Asia-Pacific conference on Conceptual Modelling*, 2007, pp. 71-80.

[10] Michael Becker and Ralf Laue, "A comparative survey of business process similarity measures," *Computers in Industry*, vol. 63, no. 2, pp. 148-167, February 2012.

[11] Remco Dijkman, Marlon Dumas, Boudewijn van Dongen, Reina Käärik, and Jan Mendling, "Similarity of business process models: Metrics and evaluation," *Information Systems*, vol. 36, no. 2, pp. 498-516, April 2011.

[12] Riyanarto Sarno, Endang Wahyu Pamungkas, Dwi Sunaryono, and Sarwosri, "Business process composition based on meta models," in *2015 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2015, pp. 315--318.

[13] Bandung Arry Sanjoyo, Subiono, and Riyanarto Sarno, "Pemodelan Komposisi Web Service Dengan Menggunakan Petri Net.," in *Seminar Nasional Matematika dan Pendidikan Matematika 2011*, Yogyakarta, 2011.

[14] Michael Levandowsky and David Winter, "Distance between sets," *Nature*, vol. 234, no. 5323, pp. 34--35, 1971.

[15] Indra Gita Anugrah, Riyanarto Sarno, and Ratih Nur Esti Anggraini, "Decomposition Using Refined Process Structure Tree (RPST) and Control Flow Complexity Metrics," in *2015 International Conference on Information & Communication Technology and Systems (ICTS)*, 2015.

[16] Elvira Rolón, Jorge Cardoso, Félix García, Francisco Ruiz, and Mario Piattini, "Analysis and Validation of Control-Flow Complexity Measures with BPMN Process Models," in *Enterprise, Business Process and Information Systems Modeling*, Terry Halpin et al., Eds. Amsterdam, Netherland: Springer Berlin Heidelberg, 2009, pp. 58 - 70.