Structural Similarity Assessment of Business Process Graph Using GED-Greedy

by Universitas Internasional Semen Indonesia

Structural Similarity Assessment of Business Process Graph Using GED-Greedy

Kristina
Department of Informatics
Institut Teknologi Sepuluh November
Universitas Widya Dharma Pontianak
Surabaya, Indonesia
7025221016@mhs.its.ac.id

Ary Mazharudin Shiddiqi*
Department of Informatics
Institut Teknologi Sepuluh November
Surabaya, Indonesia
ary.shiddiqi@if.its.ac.id

Daniel Siahaan
Department of Informatics
Institut Teknologi Sepuluh November
Surabaya, Indonesia
daniel@if.its.ac.id

Hidayatul Munawaroh
Department of Information System
Universitas Internasional Semen Indonesia
Gresik, Indonesia
hidayatul.munawaroh@uisi.ac.id

***Abstract*—Process automation in business process redesign is a critical issue. One of the problems is related to retrieving similar business process models (BPMs) stored in a repository based on implicit business process model queries extracted from user feedback. Previous studies in graph similarity focus on one-to-one business process model (BPM) similarity; however, this approach is considered time-consuming. This study aims to identify the possibility of using the graph approach and Graph Edit Distance Greedy (GED-Greedy) as structural similarity methods for measuring similarity between BPM queries and BPMs in the repository. The structure of a BPM can be represented as the inter-structure and intra-structure of a graph. The structural similarity value of the two graphs' BPM final result is 0.91. It means that the proposed methods from the case study could be used for measuring structural similarity between BPMs.**

***Index Terms*—Business process model, structural similarity, inter-structural similarity, intra-structural similarity, GED-Greedy**

I. Introduction

A business process is a collection of interrelated, complex, and complete activities within a unit or organization to produce functional value for stakeholders [1]. A complete business process helps the organization in achieving its primary goals of the organization. This goal can be achieved by visually modeling business processes using Business Process Modeling Notation (BPMN), which can be done via formal documents or informal documents [2].

Research business models are currently developing and exciting issues to be carried out by many parties [3]. The reason that supports research on business process models is to build a clear framework for an organization to achieve organizational goals. The business process model is an organizational work model with strategic value and systematic dimensions. Business processes can be modeled using BPMN [4].

The search for similarities between business processes is a process to measure similarities between business processes with the aim of modeling business processes automatically and practically. The similarity of business processes can be measured in terms of semantics, structure, and behavior [5]. Fuertes et al. [6] measures the similarity of business processes from a semantic point of view with the aim of reusing existing business processes within the company. Sarno et al. [7] measures the similarity of business processes can also be carried out using a graph approach. The graph approach has been widely used in solving pattern recognition in various fields of computing [8].

Several previous studies have used various methods to measure semantic, behavioral, and structural similarities between business processes, such as the weighted graph-based method [9], Probabilistic Latent Semantic Analysis (PLSA) and Weighted Directed Acyclic Graph (WDAG) methods [10], Extended Transition Relation (ETR) method [11], System Application and Product (SAP) Reference Model [12], structural method of weighted-tree declarative pattern models [13] and Complete Firing Sequences (CFS) [14]. These studies use the Natural Language Processing (NLP) approach to find behavioral and semantic similarities between business processes. Meanwhile, Tangkawarow et al. [5] looked for structural similarities using the NLP approach, namely the Jaccard method.

Structural similarity measurement using Unified Modelling Language (UML) Class Graph (UCG) approach has been carried out using a UML diagram case study namely class diagrams [15] and graph approach namely states chart [16], [17], and use case diagram [18]. In this research, we propose an approach for modeling business process into two types of graphs, namely inter-structural and intra-structural and calculating the structural similarity between BPMs graphs using the GED-Greedy method.

Structural similarity between the two business process models can be measured based on inter-structural and intra-structural similarities. This research applied the graph-based GED-Greedy method to measure the structural similarity between two business processes. Measurement of similarity between business processes using NLP approach and graph approach has been carried out in Tangkawarow et al. [5], Ehrig et al. [12] and Sarno et al. [7]. While the measurement of similarity between graphs using the GED-Greedy approach with the case study is the UML diagram has also been carried out in Munawaroh et al. [16], Fauzan et al. [17], and Zulfa et al. [18]. Tangkawarow et al. [5] using the NLP approach to measure the similarity of business process models emphasized the semantic similarity between business process models. Setiawan et al. [9] using the graph approach only measured the similarity of the business process model as a whole but had not divided the graph into inter-structural and intra-structural. Where as Munawaroh et al. [16] using inter-structural and intra-structural graphs have only been applied to UML diagrams. This study uses the case of business process modeling using BPMN notation. Each BPMN would be modeled on inter-structural and intra-structural graphs to measure the structural similarity values of the two BPMs originating from two different organizations.

II. Methodology

This research consists of several stages, namely (1) Literature study, (2) BPM data collection, (3) Data preprocessing, (4) Measuring structural similarities between graphs, and (5) Compiling research results reports. Fig. 1. shows a flowchart of the research process to measure the structural similarity between the two business process models using the graph approach and GED Greedy.

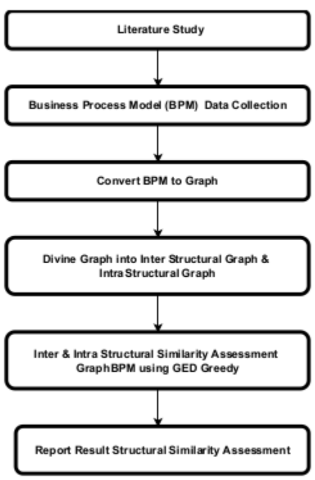

Fig. 1. The flowchart of this study.

A. BPM Data Collection

The BPM dataset used in this paper is obtained from interviews at two different libraries [5]. Each BPM of the two library units would be compared for similarity. The BPMs used is the first BPM university for borrowing regular books as shown in Fig. 2. BPM_1 and the second BPM university for borrowing regular books as shown in Fig. 3. BPM_2 would be compared for similarities.

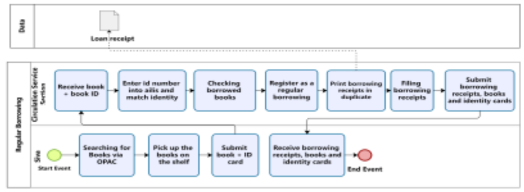

Fig. 2. The BPM_1 regular borrowing of $University^{1st}$.

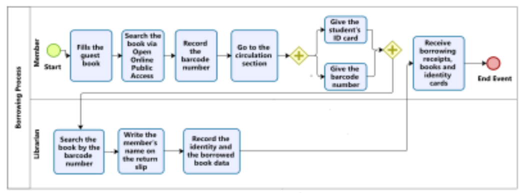

Fig. 3. The BPM_2 regular borrowing of $University^{2nd}$.

B. Data Processing

Each BPM would be converted into Extensible Markup Language Process Definition Language (XPDL) form and modeled into a graph form. Fig. 4 shows the XPDL form of BPM. Each BPM must be modeled in graph form before measuring the structural similarity between the two BPMs. BPM modeling into graphs would be carried out after labelling each state and transition from the BPM. Business process modeling uses BPMN notation [19]. The types of labels that have been prepared can be seen in Table I.

TABLE I
THE BPM GRAPH ELEMENT LABELS.

Type	Name	Label
Node	Start	Vst
Node	End	Ved
Node	State	Vs
Node	Sequence Transition	Vts
Node	Message Transition	Vtm
Node	Association Transition	Vta
Node	Exclusive OR (XOR)	Vxor
Node	AND/Paralel	Vand
Node	OR/Inclusive	Vor
Node	Data Store	Vds
Node	Data Object	Vtrd
Node	Message Event	Vtrm
Node	Timer Event	Vtrt
Edge	Sequence Flow	es
Edge	Message Flow	em
Edge	Association	ea

C. Two BPM Structural Similarity Assessment Method

BPM structural similarity assessment would be divided into two types: inter-structural similarity and intra-structural similarity. The structural similarity measured based on the state

```xml
    <Pool Id="103a9b38-21d4-44a0-b6e3-93cd42e34630" Name="regular borrowing"
Process="c6dbdce8-4ac4-43dc-b7ce-7adf1563e30f" BoundaryVisible="true">
      <Lanes>
        <Lane Id="28f94255-0699-4c35-ad79-83fa5bc7f4c4" Name="Circulation
Service Section" ParentPool="103a9b38-21d4-44a0-b6e3-93cd42e34630">
          <NodeGraphicsInfos>
            <NodeGraphicsInfo ToolId="BizAgi_Process_Modeler" Height="151"
Width="1284" BorderColor="-11513776" FillColor="-1">
              <Coordinates XCoordinate="50" YCoordinate="0" />
              <TextDirection xsi:nil="true" />
            </NodeGraphicsInfo>
          </NodeGraphicsInfos>
          <ExtendedAttributes />
        </Lane>
        <Lane Id="0851b6a9-296d-4213-9176-08d96733463b" Name="Siva"
ParentPool="103a9b38-21d4-44a0-b6e3-93cd42e34630">
          <NodeGraphicsInfos>
            <NodeGraphicsInfo ToolId="BizAgi_Process_Modeler" Height="136"
Width="1284" BorderColor="-11513776" FillColor="-1">
              <Coordinates XCoordinate="50" YCoordinate="151" />
              <TextDirection xsi:nil="true" />
            </NodeGraphicsInfo>
          </NodeGraphicsInfos>
```


Fig. 4. A snapshot of XMI file from BPM.

relationship in each BPM is called inter-structural similarity. Meanwhile, the structural similarity is measured based on the attributes of each state and transition in each BPM, called intra-structural similarity. The calculation of the structural similarity between the two BPMs (strucSim (bp_1, bp_2)) can be seen in Eq. (1).

$$strucSim(bp_1, bp_2) = w_{inter} \times interSim(bp_1, bp_2) + w_{intra} \times intraSim\ (bp_1, bp_2) \quad (1)$$

Where
bp_1 = Business Process Model First Graph,
bp_2 = Business Process Model Second Graph,
w_{inter} = Inter-structural similarity's weight,
w_{intra} = Intra-structural similarity's weight,
interSim = inter-structure BPM similarity,
intrasim = intra-structure BPM similarity.

D. Assessment Inter-Structural Similarity of BPM

Business process modeling in the form of a graph for the measurement of similarity is only concerned with the relationship between states and ignores the attributes that exist in the state and transition. First, do the BPM into inter graph forms. Then, do the measurement of inter-structural similarity using the GED-Greedy method.

E. Cost Matrix of Inter-Structural Similarity Assessment

C matrix is the label of the cost matrix that contains the needed cost to make a change from the first graph as shown in Fig. 5. become the second graph as shown in Fig. 6. Cost is the needed value to change from node to node. The size of the cost matrix is $(m+n) \times (m+n)$, where m is the amount vertices from the first graph and n is the amount vertices from the second graph. The case study shows that the second graph has 28 nodes and the second graph has 27 nodes. Therefore, the resulting cost matrix would be 28+27= 55 nodes so the size of the resulting cost matrix is 55 x 55. After getting the size of the cost matrix, the cost matrix would be divided into four parts as shown in Eq. (2).

$$C = \left[\begin{array}{c|c} Q_1 & Q_3 \\ Q_2 & Q_4 \end{array}\right] \quad (2)$$

Where
Q_1 = Substitution cost from V_t to V'_j
Q_2 = Delete cost V_t (from V_t to ε)
Q_3 = Add cost V'_f (from ε to V'_f)
Q_4 = Substitution cost from ε to ε
V_t = node in the first graph
V'_j = node in the second graph

Each quadrant of the cost matrix contains the costs formed when changing node per node in the graphs being compared. The GED algorithm calculates the costs incurred when making changes to the graph, while the Greedy algorithm minimizes the number of steps that must be carried out when making changes to the graph. The choice of the cost quadrant in the cost matrix is based on the number of nodes owned by the graph to be compared.

If the number of nodes of the first graph is more than the number of nodes of the second graph (m > n), then the first cost selection is the second quadrant (Q_2). If the number of nodes of the first graph is less than the number of nodes of the second graph (m < n), then choose the third quadrant (Q_3) for the first cost selection. If the number of nodes of the first graph equals the number of nodes of the second graph (m=n), so choose the first quadrant (Q_1) for the first cost selection.

After that, choose the fourth quadrant Q_4 be the second cost selection because Q_4 is the minimum value that contains the value "0". When all values in Q_4 have been selected, the selection step of costs in Q_4 is already complete. The final step is to calculate the total cost of the selected costs.

F. Assessment Intra-Structural Similarity of Business Process Model

Assessment of intra-structural similarity is based on the attributes possessed by states and transitions. The intra-structural is divided into two groups, namely state and transition. The formula used to calculate intra-structural similarity values, *intraSim*(bp_1, bp_2), is shown in Eq. (3).

$$\begin{aligned} \mathrm{intraSim}\,(bp_1, bp_2) &= w_{st} \times \mathrm{stSim}\,(bp_1, bp_2) \\ &+ w_{tr} \times \mathrm{trSim}\,(bp_1, bp_2) \end{aligned} \quad (3)$$

Where
bp_1 = Business Process Model First Graph,
bp_2 = Business Process Model Second Graph,
w_{st} = State similarity's weight,
w_{tr} = Transition similarity's weight,
stSim = BPM state similarity,
trSim = BPM transition similarity

The formula for calculating intra-structural similarity values in the BPM state, *stSim*(bp_1, bp_2), is shown in Eq. (4).

$$stSim(bp_1, bp_2) = \frac{(n \times maxCostst) - costS}{n \times maxCostst} \quad (4)$$

Where
bp_1 = Business Process Model First Graph,
bp_2 = Business Process Model Second Graph,
n = the number of cost values from the cost matrix S,
maxCostSt = maximum cost from the cost matrix S

costS = GED cost taken from the cost matrix S.

The value of intra-state similarity (Cost S) is obtained from the application of the greedy algorithm. The Greedy algorithm is used to determine the permutation with the minimum cost value from the set of costs calculated by the GED. Predetermined values would be stored in the cost matrix S in Eq. (5).

$$S = \begin{bmatrix} Vs_1Vs_1 & \cdots & Vs_tVs_j \\ \vdots & \ddots & \vdots \\ Vs_tVs_j & \cdots & Vs_xVs_y \end{bmatrix} \quad (5)$$

Where
x = BPM first diagram's number of state nodes
y = BPM second diagram's number of state nodes

Next, the intra-structural similarity transition's value between two BPMs, $trSim(bp_1, bp_2)$, can be calculated using Eq. (6).

$$trSim(bp1, bp2) = \frac{(n \times maxCostTr) - costT}{n \times maxCostTr} \quad (6)$$

Where
bp_1 = Business Process Model First Graph
bp_2 = Business Process Model Second Graph
n = the number of cost values taken from the cost matrix T
maxCostTr = maximum cost that might occur in the cost matrix T.

TThe transitional intra-structural similarity value (CostT) is obtained by applying the greedy algorithm. The Greedy algorithm is used to determine the permutation with the minimum cost value from the set of costs calculated by the GED. Predetermined values would be stored in the cost matrix T in Eq. (7).

$$T^* = \begin{bmatrix} Vt_1Vt_1 & \cdots & Vt_tVt_j \\ \vdots & \ddots & \vdots \\ Vt_tVt_j & \cdots & Vt_uVt_v \end{bmatrix} \quad (7)$$

Where
u = BPM Diagram first Graph's number of transition nodes.
v = BPM Diagram Second Graph's number of transition nodes.

III. Result and Analysis

This analysis uses two business process model diagrams (BPM_1 and BPM_2) as examples to asses structural similarity using the inter and intra structural graph proposes method. The diagrams compared are from different universities.

A. Assessment Result of Two BPM Diagram's Inter Structural Similarity

To calculate the inter-structural similarity between two BPMs, the BPM must be modeled in the form of a graph. The inter graph model of borrowing regular books from BPM_1 and BPM_2, which would be compared, can be seen in Fig. 5. and Fig. 6. After being modeled into a graph, the two BPMs would be compared, and the similarity values of the structures would be calculated. Each cost would be entered into the cost C matrix as shown in Eq. (8).

$$C = \begin{bmatrix} 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & 2 & 2 \end{bmatrix} \quad (8)$$

From the calculation results, the minimum permutation

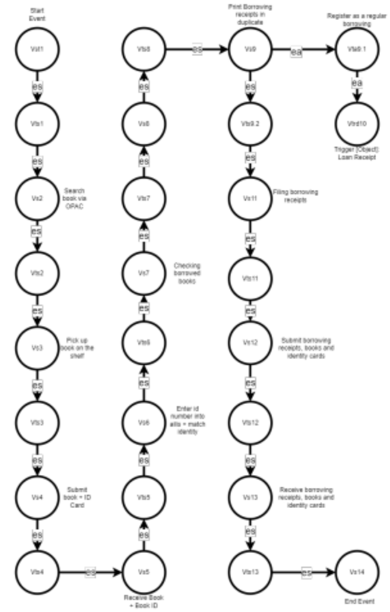

Fig. 5. The inter-structural graph of BPM_1.

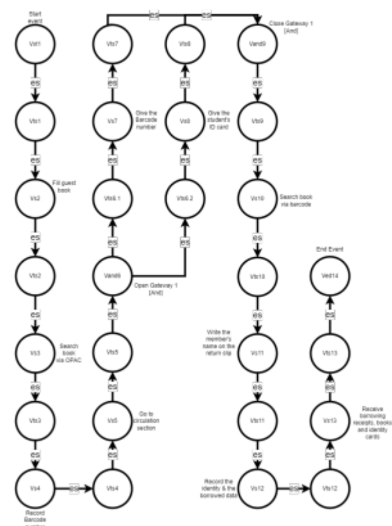

Fig. 6. The inter-structural graph of BPM_2.

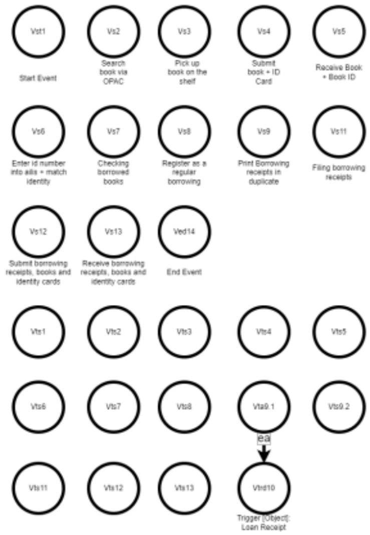

Fig. 7. The intra-structural graph of BPM_1.

based on the Greedy algorithm are 1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 13, 9, 15, 14, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 23, 54, 55, 28, 29, 30, 31, 32, 33, 34, 35, 36, 10, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53., so the total cost is 11. The GED cost needed to change the inter graph of BPM_1 to the inter graph of BPM_2 is 11. Thus it can be concluded that the inter-structural similarity value between the inter-graph of BPM_1 to the inter graph of BPM_2 is 0,89.

B. Assesment Result of Two BPM Diagram's Intra Structural Similarity

To measure the two BPMs Diagram intra structural similarity values, the BPM must be modeled in the form of an intra graph. The intra graph of the two BPMs are shown in Fig. 7 and Fig. 8.

The intra graph of BPM_1 as shown in Fig. 7 has 13 states and 13 transitions, while the intra graph of BPM_2 as shown in Fig. 8 has 14 states and 14 transitions. There are 182 state pairs and 182 transition pairs that need to be calculated using GED-Greedy when a change is made from the intra graph of BPM_1 to the intra graph of BPM_2. The calculations are the same as in the inter-structure calculations, using the cost S matrix. The cost S matrix is the cost obtained when calculating the intra graph for state nodes.

Based on the cost S matrix as shown in Eq. (9), the result of Greedy permutation is 1,2,3,4,5,6,7,8,9,10,11,12,13 with a total cost of 3. From that permutation, the value of the GED cost is 3. It means the value of GED cost is 3 when doing a calculation of the values of intra-structural similarity for state nodes based on ($stSim(Bp_1, Bp_2)$) from the intra-structural graph of BPM_1 to the intra-structural graph of graph BPM_2. The value of intra-structural similarity for states between the two graphs by following Eq. (4) is 0.88.

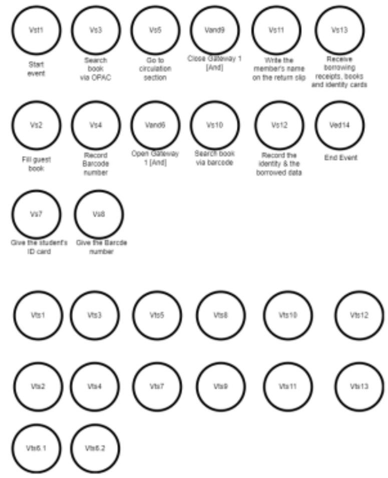

Fig. 8. The intra-structural graph of BPM_2.

$$S = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \quad (9)$$

While the cost T matrix is the cost obtained when calculating the intra-graph for the transition nodes. Based on the cost T matrix as shown in Eq. (10), the result of Greedy permutations is 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, and 9. From that permutation, the result total cost is 3. It means the value of GED cost is 3 when doing a calculation of the value of intra-structural similarity for transition nodes based on $(\operatorname{tr}\operatorname{Sim}(bp_1, bp_2))$ from graph BPM_1 to graph BPM_2 . The final value of intra-structural similarity for the transition between the two graphs by following Eq. (6) is 0.98.

$$T = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (10)$$

Assuming that the intra-structural similarity value for node state (W_{st}) and node transition (W_{tr}) is 0.5, respectively. The result of calculating the intra-structural similarity value of the state nodes is 0.88, and the transition nodes are 0.98. The intra-structural similarity value of BPM_1 and BPM_2 graphs using Eq. (3) is 0.93.

Based on the inter-structural similarity value is 0,89 and the intra-structural similarity value is 0.93 and assuming the weight of inter-structural similarity (w_{inter}) and the weight of intra-structural similarity (w_{intra}) is 0.5 for each weight, then the final result of structural similarity value of graph BPM_1 to graph BPM_2 using Eq. (1) is 0.91.

IV. Conclusion

This study evaluates the possibility of using the GED-Greedy and graph approach to measure the similarity between two BPMs. The BPMs are represented as directed graphs with labeled edges. Each business process graph can be viewed in two forms, i.e. inter-structural and intra-structural graphs. This study uses the two views to measure the BPM structural similarity.

Our case study indicates that the approach could be used to measure the similarity between two BPMs. The approach can represent the necessary structural element of BPM and uses it for measuring the structural similarity between two BPMs. Our future research would be directed to measuring the scalability of this approach in our targeted application, i.e., BPM redesign.

Acknowledgment

This research is a collaboration among Institut Teknologi Sepuluh November, Universitas Widya Dharma Pontianak, and Universitas Internasional Semen Gresik Indonesia.

References

[1] F. Amalina and Y. Handayati, "Business process analysis and improvement in selling process using business process modelling notation (bpmn) at locarvest," *KnE Social Sciences*, pp. 459–477, 2020.

[2] F. A. Sheikha and D. Inkpen, "Learning to classify documents according to formal and informal style," *Linguistic Issues in Language Technology*, vol. 8, 2012.

[3] D. Andreini, C. Bettinelli, N. J. Foss, and M. Mismetti, "Business model innovation: A review of the process-based literature," *Journal of Management and Governance*, vol. 26, no. 4, pp. 1089–1121, 2022.

[4] P. Wiśniewski, K. Kluza, and A. Ligeza, "An approach to participatory business process modeling: Bpmn model generation using constraint programming and graph composition," *Applied Sciences*, vol. 8, no. 9, p. 1428, 2018.

[5] I. Tangkawarow, R. Sarno, and D. Siahaan, "Id2sbvr: A method for extracting business vocabulary and rules from an informal document," *Big Data and Cognitive Computing*, vol. 6, no. 4, p. 119, 2022.

[6] G. Fuertes, M. Alfaro, M. Vargas, A. Espinoza, D. Galvez, and J. P. Sepúlveda-Rojas, "Measure of semantic likeness among business process activities in a telecommunication company," *IEEE Access*, vol. 8, pp. 32 332–32 340, 2019.

[7] R. Sarno, K. R. Sungkono, and R. Septiarakhman, "Graph-based approach for modeling and matching parallel business processes," *International Information Institute (Tokyo). Information*, vol. 21, no. 5, pp. 1603–1614, 2018.

[8] N. Deo, *Graph theory with applications to engineering and computer science*. Courier Dover Publications, 2017.

[9] Y. Setiawan, K. R. Sungkono, and R. Sarno, "A new similarity method based on weighted graph models for matching parallel business process models," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 5, pp. 267–276, 2020.

[10] I. G. Anugrah and R. Sarno, "Business process model similarity analysis using hybrid plsa and wdag methods," in *2016 International Conference on Information & Communication Technology and Systems (ICTS)*. IEEE, 2016, pp. 231–236.

[11] C. Liu, Q. Zeng, H. Duan, S. Gao, and C. Zhou, "Towards comprehensive support for business process behavior similarity measure," *IEICE TRANSACTIONS on Information and Systems*, vol. 102, no. 3, pp. 588–597, 2019.

[12] M. Ehrig, A. Koschmider, and A. Oberweis, "Measuring similarity between semantic business process models," in *APCCM*, vol. 7, 2007, pp. 71–80.

[13] C. S. Wahyuni, K. R. Sungkono, and R. Sarno, "Novel parallel business process similarity methods based on weighted-tree declarative pattern models," *International Journal of Intelligent Engineering and Systems*, vol. 12, no. 6, pp. 236–248, 2019.

[14] Z. Dong, L. Wen, H. Huang, and J. Wang, "Cfs: A behavioral similarity algorithm for process models based on complete firing sequences," in *On the Move to Meaningful Internet Systems: OTM 2014 Conferences: Confederated International Conferences: CoopIS, and ODBASE 2014, Amantea, Italy, October 27-31, 2014, Proceedings*. Springer, 2014, pp. 202–219.

[15] Z. Yuan, L. Yan, and Z. Ma, "Structural similarity measure between uml class diagrams based on ucg," *Requirements Engineering*, vol. 25, pp. 213–229, 2020.

[16] H. Munawaroh, D. O. Siahaan, R. Fauzan, and E. Triandini, "Structural similarity measurement using graph edit distance-greedy on state chart diagrams," in *2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS)*. IEEE, 2020, pp. 1–6.

[17] R. Fauzan, D. Siahaan, S. Rochimah, and E. Triandini, "An automated statechart diagram assessment using semantic and structural similarities," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 11, no. 6, 2021.

[18] F. Zulfa, D. O. Siahaan, R. Fauzan, and E. Triandini, "Inter-structure and intra-structure similarity of use case diagram using greedy graph edit distance," in *2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS)*. IEEE, 2020, pp. 1–6.

[19] U. Campos, A. Lopes, E. Oliveira, B. Gadelha, and T. Conte, "e-vol bpmn: a technique to support the evolution and learning of bpmn diagrams," *IET Software*, vol. 14, no. 1, pp. 50–58, 2020.

Structural Similarity Assessment of Business Process Graph Using GED-Greedy

ORIGINALITY REPORT

10%	0%	10%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Hidayatul Munawaroh, Daniel Oranova Siahaan, Reza Fauzan, Evi Triandini. "Structural Similarity Measurement using Graph Edit Distance-Greedy on State chart Diagrams", 2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS), 2020 Publication	7%
2	Fatimatus Zulfa, Daniel Oranova Siahaan, Reza Fauzan, Evi Triandini. "Inter-Structure and Intra-Structure Similarity of Use Case Diagram using Greedy Graph Edit Distance", 2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS), 2020 Publication	3%

Exclude quotes	On	Exclude matches	< 3%
Exclude bibliography	On		