ICIC Express Letters
Part B: Applications
Volume **7**, Number **1**, January **2016**

ICIC International ©2016 ISSN 2185-2766
pp. 111–116

THE ANALYSIS OF CHANGE REGION IN BUSINESS PROCESS BASED ON THE DEGREE OF BEHAVIOR SIMILARITY

Xianwen Fang, Fang Zhao, Xiangwei Liu and Huan Fang

School of Science
Anhui University of Science and Technology
No. 168, Middle Shungeng Road, Huainan 232001, P. R. China
{ 280060673; 1012377428 }@qq.com; lxw7710@tom.com; fanghuan0307@163.com

Received June 2015; accepted September 2015

Abstract. *In the business process modeling, it is an important task to find the behavior similarity of the process models. The existing methods to search for similarity are mainly from the behavior relations or semantic equivalence of the adjacent transitions and never considering it from the angle of non-adjacent transitions. In order to study the similarity degree of two business process models, the degree of behavior similarity is proposed on the basis of existing theory. In view of behavioral profiles and simulated behavioral profiles, the transitions which have direct behavior relationship and indirect behavior relationship are considered. According to a certain algorithm, the degree of behavior similarity of two business process models is calculated. Then the theories about the change region of the target model are found. Finally, the effectiveness of this method is verified through specific procurement process instance.*
Keywords: Degree of behavior similarity, Simulated behavioral profiles, Direct behavior relationship, Indirect behavior relationship

1. **Introduction.** In the field of the existing commercial business process modeling, it is necessary to consider the models' match from the perspective of similarity of proposed model. Based on different objectives, the system needs to build some similar models to meet the needs of more people. For this reason, many people engaged in the research of consistency or similarity.

M. Weidlich et al. described the behavior profile of a model which captured the basic behavior relations between the models and gave the judgment standard of the consistency [1]. [2] introduced a method called measure searching, and then used it to find the similar model and it was suitable for searching large data sets. [3] described a measure method – m^3, providing an abstract process behavior based on behavioral profiles, and this method could effectively search out similar models in the process model's warehouse. [4] proposed the method of finding a model or a fragment which was similar to the given process model. [5] gave the metrics and evaluation about the similarity of business process models.

All of them analyze the similarity of two models from the angle of direct behavior relations or semantics, they have not considered the indirect behavior relations, so the credibility is very low.

Under this background, according to definition of the direct behavior relations, indirect behavior relations, behavioral profiles and simulated behavioral profiles, we consider the relations of the transitions from the angle of the direct behavior relations and the indirect behavior relations. On the basis of the definition of behavior similarity, we calculate eight kinds of it. By giving a weight to each kind of similarity degree, we can calculate the final behavior similarity degree of two models. Finally, based on the comparison analysis of two models, we can find out the suspicious areas. Then we use the algorithm to find the region which affects the similarity of the behavior, namely the change region.

2. **Motivation for the Instance.** Figure 1 and Figure 2 give us two instances of the purchase process in the commercial sector procurement. One is a source model, and the other is a target model.

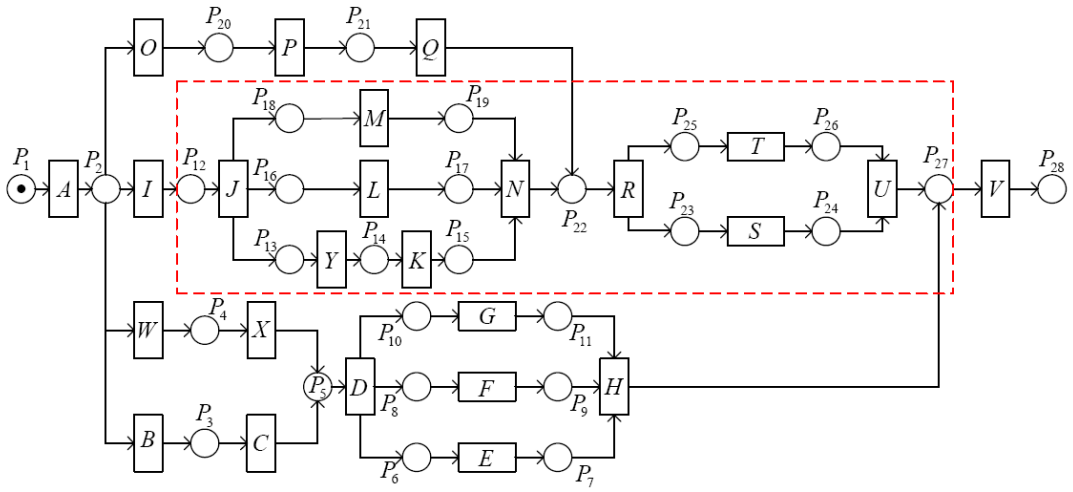

FIGURE 1. The purchase process graph of the source model

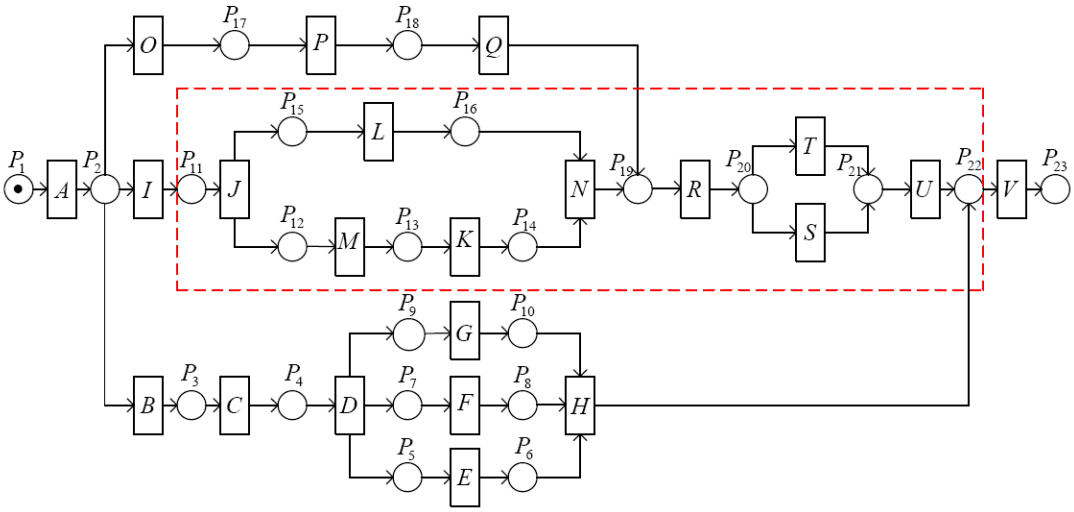

FIGURE 2. The purchase process graph of the target model

3. **The Basic Definition.** This part mainly introduces some basic concepts, as to the Petri nets about the process model, the weak order relations, dead-lock, live-lock and the behavioral profiles; please refer to reference [1].

Definition 3.1. *(A sound free-choice net): Let $N = (P, T, F)$ be a Petri net, if there do not exist circulation, dead-lock or live-lock, then N is a sound free-choice net.*

Definition 3.2. *(Direct behavior relationships, Indirect behavior relationships): Let $N = (P, T, F)$ be a sound free-choice net, and then $\forall t_i, t_j \in T$, if $t_i^\bullet = {}^\bullet t_j$, ${}^\bullet t_i = {}^\bullet t_j$, $t_i^\bullet = t_j^\bullet$, ${}^\bullet({}^\bullet t_i) = {}^\bullet({}^\bullet t_j)$ or $(t_i^\bullet)^\bullet = (t_j^\bullet)^\bullet$, then t_i and t_j have the direct behavior relationships, or else they have the indirect behavior relationships.*

On the basis of the behavioral profiles, we define the simulated behavioral profiles.

Definition 3.3. *(Simulated behavioral profiles): Let $N = (P, T, F)$ be a sound free-choice net, x and $z \in T$ are two single transitions, $y = \{t_1, t_2 \cdots t_k\} \subset T$ and $(\{x\} \cup \{z\}) \cap y = \Phi$, so x and z have the indirect behavior relationships, and the simulated behavioral profiles should satisfy one of the following relationships.*

(1) Simulated strict order relations: $x \Rightarrow z$, if $\exists y \subset T$ with $x \rightarrow y$ and $y \rightarrow z$.

(2) Simulated exclusiveness relations: $x \Updownarrow z$, if $\exists y \subset T$ with $x + y$ and $y \rightarrow z$, also x and z are neither in strict order relations nor in simulated strict order relations.

(3) Simulated interleaving order relations: $x \Leftrightarrow z$, if $\exists y \subset T$ with $x//y$ and $y \rightarrow z$, also x and z are neither in strict order relations nor in simulated strict order relations.

As for the inverse simulated strict order relations: $x \Leftarrow z$, if $\exists y \subset T$ with $x \leftarrow y$ and $y \leftarrow z$, then we use Figure 3 to express these relations.

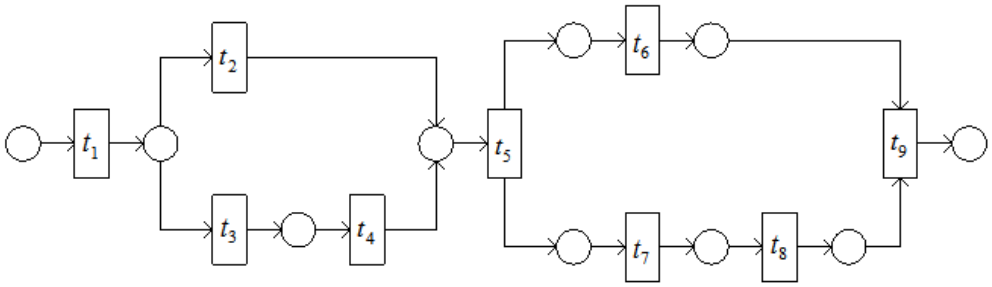

FIGURE 3. Simulated behavioral profiles

As shown in Figure 3, t_1 and t_i $(i = 2, 3)$ are in strict order relations, t_1 and t_i $(i = 4, 5, \cdots, 9)$ are in simulated strict order relations, t_2 and t_3 are in exclusiveness relations, t_2 and t_4 are in simulated exclusiveness relations, t_6 and t_7 are in interleaving order relations, and t_6 and t_8 are in simulated interleaving order relations.

The nets in the following content are all sound free-choice nets by default.

Definition 3.4. *(The degree of behavior similarity in behavioral profiles and simulated behavioral profiles): Let $N_1 = (P_1, T_1, F_1)$ be a Petri net of a source model, and $N_2 = (P_2, T_2, F_2)$ be a Petri net of a target model. Then*

Strict order similarity:

$$sim_{\rightarrow}(N_1, N_2) = \frac{|(\rightarrow_{N_1}) \cap (\rightarrow_{N_2})|}{|(\rightarrow_{N_1}) \cup (\rightarrow_{N_2})|} \tag{1}$$

Simulated strict order similarity:

$$sim_{\Rightarrow}(N_1, N_2) = \frac{|(\Rightarrow_{N_1}) \cap (\Rightarrow_{N_2})|}{|(\Rightarrow_{N_1}) \cup (\Rightarrow_{N_2})|} \tag{2}$$

Exclusiveness similarity:

$$sim_{+}(N_1, N_2) = \frac{|(+_{N_1}) \cap (+_{N_2})|}{|(+_{N_1}) \cup (+_{N_2})|} \tag{3}$$

Simulated exclusiveness similarity:

$$sim_{\Updownarrow}(N_1, N_2) = \frac{|(\Updownarrow_{N_1}) \cap (\Updownarrow_{N_2})|}{|(\Updownarrow_{N_1}) \cup (\Updownarrow_{N_2})|} \tag{4}$$

Interleaving order similarity:

$$sim_{//}(N_1, N_2) = \frac{|(//_{N_1}) \cap (//_{N_2})|}{|(//_{N_1}) \cup (//_{N_2})|} \tag{5}$$

Simulated interleaving order similarity:

$$sim_{\Leftrightarrow}(N_1, N_2) = \frac{|(\Leftrightarrow_{N_1}) \cap (\Leftrightarrow_{N_2})|}{|(\Leftrightarrow_{N_1}) \cup (\Leftrightarrow_{N_2})|} \tag{6}$$

Inverse strict order similarity:

$$sim_{\leftarrow}(N_1, N_2) = \frac{|(\leftarrow_{N_1}) \cap (\leftarrow_{N_2})|}{|(\leftarrow_{N_1}) \cup (\leftarrow_{N_2})|} \tag{7}$$

Inverse simulated strict order relations:

$$sim_{\Leftarrow}(N_1, N_2) = \frac{|(\Leftarrow_{N_1}) \cap (\Leftarrow_{N_2})|}{|(\Leftarrow_{N_1}) \cup (\Leftarrow_{N_2})|} \tag{8}$$

Definition 3.5. *(The change region): Let $N_1 = (P_1, T_1, F_1)$ be a Petri net of a source model, and $N_2 = (P_2, T_2, F_2)$ be a Petri net of a target model. If R_2 is a suspicious area of N_2, R_1 is a suspicious area of N_1, $\forall a, b \in R_2$ respond to $c, d \in R_1$, if a, b do not meet the behavioral profiles and simulated behavioral profiles relations of c, d, then $\{a, b\}$ are the change region in N_2.*

4. **The Analysis of Change Region in Business Process Based on the Degree of Behavior Similarity in Petri Nets.** By introducing the definitions, we calculate the similarity degree of the source and target models from the angle of direct behavior relations and indirect behavior relations. Also we assign a weight to each of the eight similarity degrees, and then calculate the behavior similarity degree of the source and target models. Based on the behavioral profiles and simulated behavioral profiles of the corresponding transitions, we can find out the change region of the target model. To this end, we give the following algorithms.

Algorithm 1: The degree of behavior similarity in source and target models.

Input: the Petri net of a source model $N_1 = (P_1, T_1, F_1)$, the Petri net of a target model $N_2 = (P_2, T_2, F_2)$.

Output: the degree of behavior similarity in source and target models: $sim(N_1, N_2)$.

(1) If the transition in N_1, N_2 has the same meaning, then we use the same letter to denote it.

(2) According to Definition 3.3 and Equation (1), we use $|(\rightarrow_{N_1}) \cap (\rightarrow_{N_2})|$ to denote the number of the same letters in the strict order relations in N_1, N_2 , and use $|(\rightarrow_{N_1}) \cup (\rightarrow_{N_2})|$ to denote the union set of the letters in the strict order relations in N_1, N_2, and then the strict order similarity degree is $sim_{\rightarrow}(N_1, N_2)$.

(3) Then according to Definition 3.3 and Definition 3.4, we can obtain the simulated strict order similarity degree, the exclusiveness similarity degree, the simulated exclusiveness similarity degree, the interleaving order similarity degree, the simulated interleaving order similarity degree, the inverse strict order similarity degree and the inverse simulated strict order similarity degree.

(4) According to steps (2) and (3), we assign a weight w_1 to $sim_{\rightarrow}(N_1, N_2)$, w_2 to $sim_{\Rightarrow}(N_1, N_2)$, w_3 to $sim_{+}(N_1, N_2)$, w_4 to $sim_{\Updownarrow}(N_1, N_2)$, w_5 to $sim_{//}(N_1, N_2)$, w_6 to $sim_{\Leftrightarrow}(N_1, N_2)$, w_7 to $sim_{\leftarrow}(N_1, N_2)$, w_8 to $sim_{\Leftarrow}(N_1, N_2)$, and

$$w_1 + \cdots + w_8 = 1 \tag{9}$$

(5) Then calculate the degree of behavior similarity in source and target models:

$$\begin{aligned} sim(N_1, N_2) = {} & w_1 \cdot sim_{\rightarrow}(N_1, N_2) + w_2 \cdot sim_{\Rightarrow}(N_1, N_2) + w_3 \cdot sim_{+}(N_1, N_2) \\ & + w_4 \cdot sim_{\Updownarrow}(N_1, N_2) + w_5 \cdot sim_{//}(N_1, N_2) + w_6 \cdot sim_{\Leftrightarrow}(N_1, N_2) \\ & + w_7 \cdot sim_{\leftarrow}(N_1, N_2) + w_8 \cdot sim_{\Leftarrow}(N_1, N_2) \end{aligned} \tag{10}$$

Based on the algorithms, we can obtain the degree of behavior similarity in source and target models.

The following algorithm is used to analyze the reason why the degree of behavior similarity is lower than 1.

Algorithm 2: Look for the change region of the target model based on the behavioral profiles and simulated behavioral profiles relations in the source model.

Input: the relations of the transitions in the source model $N_1 = (P_1, T_1, F_1)$, the relations of the transitions in the target model $N_2 = (P_2, T_2, F_2)$.

Output: the change region of the target model.

(1) According to the definition about the behavioral profiles and simulated behavioral profiles relations, we can derive the suspicious areas $R' = \{t'_1, t'_2, \cdots, t'_m\}$ of the target model with the comparison and analysis to the source model. The corresponding area in the source model is $R = \{t_1, t_2, \cdots, t_n\}$.

(2) $\forall t_k, t_l \in R$, corresponding to $t'_i, t'_j \in R'$, if $t_k \rightarrow t_l$ ($t_k \Rightarrow t_l$, $t_k + t_l$, $t_k \Updownarrow t_l$, $t_k // t_l$, $t_k \Leftrightarrow t_l$, $t_k \leftarrow t_l$, $t_k \Leftarrow t_l$), observe t'_i and t'_j whether or not satisfy the same relation, if they satisfy it, then the suspicious area is $R' - \{t'_i, t'_j\}$, if not, then the suspicious area is $\{t'_i, t'_j\}$.

(3) Then analyze the relations of all the transitions, and find out the transitions in N_2 which do not satisfy the corresponding transition relations in N_1. The region set composed of them named C_R is the change region of the target model.

5. **Experimental Evaluation.** According to Figure 1, Figure 2 and Definition 3.4, then we can derive the relations of all activities from the source and target models. According to Algorithm 1, then the strict order similarity degree is: $sim_{\rightarrow}(N_1, N_2) \approx 0.788$, the simulated strict order similarity degree is: $sim_{\Rightarrow}(N_1, N_2) \approx 0.807$, the exclusiveness similarity degree is: $sim_{+}(N_1, N_2) \approx 0.667$, the simulated exclusiveness similarity degree is: $sim_{\Updownarrow}(N_1, N_2) \approx 0.732$, the interleaving order similarity degree is: $sim_{//}(N_1, N_2) \approx 0.571$, the simulated interleaving order similarity degree is: $sim_{\Leftrightarrow}(N_1, N_2) = 0.5$, the inverse strict order similarity degree is: $sim_{\leftarrow}(N_1, N_2) \approx 0.788$, and the inverse simulated strict order similarity degree is: $sim_{\Leftarrow}(N_1, N_2) \approx 0.807$.

We assign $w_1 = w_3 = w_5 = w_7 = 0.15$ and $w_2 = w_4 = w_6 = w_8 = 0.1$, and then the degree of behavior similarity between N_1 and N_2 is: $sim(N_1, N_2) \approx 0.7067$.

Due to that the degree of behavior similarity is below than 0.8, then we use Algorithm 2 to find the change region of the target model. Through analysis of N_1 and N_2, we can derive the suspicious area in Figure 1 and Figure 2, as shown in the dotted place. According to Algorithm 2, we know that in N_1, $M \Leftrightarrow K$, $S//T$, but in N_2, $M \rightarrow K$, $S + T$. The relations between the other corresponding transitions in the suspicious area are the same, so we can obtain the change region of the target model is $\{P_{12}MP_{13}KP_{14}\} \cup \{P_{20}STP_{21}\}$.

6. **Conclusions.** On the basis of existing research, we consider the direct behavior relations and indirect behavior relations for all of the transitions in source and target models. Then calculate eight kinds of the similarity degree, and assign each of them with a weight to calculate the degree of behavior similarity of the models. By looking for similarities and differences between two models, conclude their suspicious areas. Based on the behavioral profiles and simulated behavioral profiles relations of corresponding transitions, we can derive the final change region of the target model.

In the future, we want to explore the similarities and differences between the two models from the perspective of the structure and the behavior. Then identify the key areas of the various relations and take effective methods to analyze the key areas to make it meet the needs of business process modeling in the commercial fields.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China under Grant No. 61272153, No. 61402011, No. 61340003 and No. 6157

2035, Anhui Provincial Soft Science Foundation (12020503031), the Natural Science Foundation of Educational Government of Anhui Province of China (KJ2012A073, KJ2014A067), Anhui Provincial Natural Science Foundation (1508085MF111), the youth academic backbone foundation of AUST, the Academic and Technology Leader Foundation of Anhui Province. We also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] M. Weidlich, J. Mendling and M. Weske, Efficient consistency measurement based on behavioral profiles of process models, *2011 IEEE Trans. Software Engineering*, vol.37, no.3, pp.410-429, 2011.

[2] P. Zezula, V. Dohnal and G. Amato, *Similarity Search: The Metric Space Approach*, Springer-Verlag, New York, 2005.

[3] M. Kunze, M. Weidlich and M. Weske, m^3 – A behavioral similarity metric for business processes, *CEUR-WS*, pp.89-95, 2011.

[4] M. Dumas, L. G. Banuelos and R. Dijkman, Similarity search of business process models, *IEEE Data Eng.*, vol.32, no.3, pp.23-28, 2009.

[5] R. Dijkman, M. Dumas, B. van Dongen, R. Kaarik and J. Mendling, Similarity of business process models: Metrics and evaluation, *Information System*, vol.36, no.2, pp.498-516, 2011.

[6] C. Li, M. Reichert and A. Wombacher, On measuring process model similarity based on high-level change operations, *The 27th International Conference on Conceptual Modeling*, Barcelona, Spain, pp.248-264, 2008.

[7] M. Weidlich and J. Mendling, Perceived consistency between process models, *Information Systems*, vol.37, no.2, pp.80-98, 2012.

[8] S. Smirnov, M. Weidlich and J. Mendling, Business process model abstraction based on behavioral profiles, *Proc. of ICSOC*, San Francisco, USA, pp.1-16, 2010.