
Common Process Extraction Pada Scalable Model Proses Bisnis

Arif Wahyu Prasetya¹, Muhammad Ainul Yaqin², Syahiduz Zaman³

^{1,2,3} Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang

¹arifwahyuprasetya@gmail.com, ²yaqinov@ti.uin-malang.ac.id, ³syahid@ti.uin-malang.ac.id

Abstrak

Ekstraksi *common process* dilakukan untuk mendapatkan rangkaian proses yang selalu ada dalam setiap model proses bisnis. *Common process* didapatkan melalui serangkaian proses berdasarkan pada kemiripan model proses bisnis. Model proses bisnis dapat diukur kemiripannya dalam tiga aspek, yaitu kemiripan semantik label, kemiripan struktural, dan kemiripan perilaku. Ketiga aspek kemiripan tersebut digabungkan untuk mendapatkan tingkat kemiripan keseluruhan dengan memberikan bobot pada setiap aspek. Kemudian hasil kemiripan tersebut digunakan untuk proses pengelompokan (*clustering*). Kombinasi bobot, tingkat skalabilitas, dan nilai *threshold* berpengaruh pada akurasi hasil *clustering*. Akurasi hasil *clustering* yang didapatkan adalah 100% dengan *threshold* 0.48 dengan kombinasi bobot kemiripan semantik, struktural dan perilaku (0.5, 0.3, 0.2) dan *threshold* 0.50 dengan kombinasi bobot (0.5, 0.4, 0.1). Dari setiap *cluster* yang telah terbentuk dicari *common process* dengan metode *graph partition approach*. Akhirnya *common process* yang didapatkan merupakan representasi dari suatu proses bisnis.

Kata kunci : kemiripan model proses bisnis, *scalable model proses bisnis*, *clustering*, *common process*.

1. Pendahuluan

Proses bisnis pada sistem *Enterprise Resource Planning* (ERP) terdiri dari penjualan dan distribusi, perencanaan produksi, manajemen material, eksekusi manufaktur, akuntansi dan keuangan, dan audit [1]. Dalam menjalankan proses bisnisnya, setiap perusahaan menerapkan prosedur yang bervariasi. Ada kalanya prosedur yang digunakan antara satu perusahaan dan perusahaan yang lain sama, namun juga tidak jarang setiap perusahaan menggunakan prosedur yang berbeda. Sehubungan dengan penerapan prosedur ERP dalam setiap perusahaan, sistem yang *survive* dibutuhkan untuk dapat melayani setiap aktivitas yang terdapat di setiap proses bisnis [2]. Sistem yang *survive* memiliki kemampuan untuk menyesuaikan diri dengan situasi yang ada. Penyesuaian terhadap situasi yang ada dilakukan dengan adanya fleksibilitas dalam sistem ERP. Dengan adanya fleksibilitas, *common process* dari suatu proses bisnis dapat divariasikan sesuai kebutuhan serta dapat juga ditambahkan proses-proses bisnis yang lain sebagai variasi. *Common process* didapatkan dengan cara menemukan model-model proses bisnis yang mirip secara label semantiknya, strukturnya, dan perilakunya [3]. Model proses bisnis secara struktural dan perilaku diukur kemiripannya dengan menggunakan formula *Jaccard* [4], *Dice* [5], *Cosine* [6], dan *Overlap Coefficient* [7]. Sedangkan kemiripan model proses bisnis secara kesesuaian label diukur dengan formula *Levenshtein* [8].

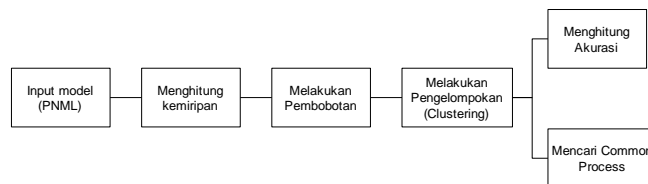
Model proses bisnis adalah representasi terurut dari semua fungsi yang terkait dengan suatu aktivitas bisnis [9]. Proses bisnis dapat dimodelkan dengan berbagai cara, yaitu UML, BPMN, EPC, dan *Petri Net*. *Petri Net* terdiri dari sintaks XML yang menggambarkan rangkaian aktivitas [9]. Salah satu cara untuk menganalisis model *Petri Net* tersebut adalah dengan menghitung kemiripannya. Kemiripan model proses bisnis ditandai dengan adanya kesamaan model proses bisnis dari suatu organisasi. Persamaan dan perbedaan model proses bisnis menjadi sebab adanya *common process* dan variasi model proses bisnis. Pengukuran kemiripan model proses bisnis diukur dalam aspek kesesuaian label (*semantic*), struktural, perilaku. Operasi perhitungan kemiripan secara *semantic* dapat dilakukan dengan membandingkan jarak *levenshtein* dengan nilai *string* terpanjang dari teks. Nilai kemiripan perilaku didapatkan dari relasi antar aktivitas yang ada. Relasi aktivitas ini juga dapat diartikan sebagai urutan eksekusi yang mungkin terjadi. Dalam istilahnya disebut sebagai *Transition Adjacent Relations* (TARs) [10]. Kemiripan struktur merupakan nilai kesamaan antar model proses bisnis dengan melihat dari struktur atau bentuknya. Menurut Remco Dijkman [11] untuk menghitung kesamaan dua proses, harus ditemukan pemetaan yang menginduksi kesamaan maksimal. Struktur dilihat dari setiap bagian model yang ada. Sehubungan dengan model yang dipakai adalah *Petri Net*, maka bagian model terdiri dari *transition*, *place*, dan *arc* [12]. Dari perhitungan yang dilakukan akan dihasilkan kelompok hasil kemiripan yang tersusun pada suatu *cluster* tertentu yang kemudian dapat dicari *common process* dari

cluster tersebut. Hasil yang didapatkan tersebut dipengaruhi pula dari faktor model proses bisnis yang bersifat *scalable*.

Ukuran skalabilitas proses bisnis dilakukan dengan membandingkan kemiripan beberapa proses bisnis dan skalanya. Metode pengukuran skala model proses bisnis menggunakan jumlah elemen model proses bisnisnya dan *Control Flow Complexity* (CFC) [13]. Kemudian nilai skala dan kemiripan model proses bisnis dihitung dengan formulasi yang telah ditentukan. Hasilnya semakin besar nilai skalabilitas, maka potensi pertumbuhan proses bisnis tersebut juga besar. Berdasarkan latar belakang di atas maka tujuan dari riset ini adalah mencari *common process* dari beberapa proses bisnis. Sehingga ketika *common process* dapat ditemukan, *common process* tersebut dapat digunakan sebagai dasar proses bisnis yang akan digunakan dalam suatu organisasi

2. Metode Penelitian

Data yang digunakan dalam riset ini berasal dari model proses bisnis dari beberapa Pondok Pesantren yang telah dibangun pada riset sebelumnya. Blok diagram proses *common process extraction* sebagaimana pada **Error! Reference source not found**.



Gambar 1 Prosedur penelitian

2.1 Input Model (PNML)

Data yang telah didapat dibentuk dalam format *.pnml. Setiap elemen yang terdapat pada *Petri Net* memiliki label yang mewakili alur dari proses yang berlangsung. Kemudian setelah model proses bisnis yang diinginkan telah didapatkan maka format yang tadinya *.pnml dikonversi menjadi format *.xml.

Tabel 1 Keterangan model : A=Model *Petri Net*, B=Keterangan label, C=*Petri Net* dalam bentuk XML

	<table border="1"> <tr><td>t1</td><td>Register in boarding schools</td></tr> <tr><td>t2</td><td>Get the receipt proof of boarding schools</td></tr> <tr><td>t3</td><td>Input the registration data</td></tr> <tr><td>t4</td><td>Submit the form and receive a registration file</td></tr> <tr><td>t5</td><td>The Committee conducts the registration data entry</td></tr> <tr><td>t6</td><td>Committee print mark registration</td></tr> <tr><td>t7</td><td>Pay to financial section</td></tr> <tr><td>t8</td><td>Orientation for new students</td></tr> <tr><td>t9</td><td>Submit the form and attaching a other file registrations</td></tr> <tr><td>t10</td><td>Selection against based on the average value of national exam</td></tr> <tr><td>t11</td><td>See the announcement escaped selection</td></tr> <tr><td>t12</td><td>Re-registration to the committee</td></tr> <tr><td>t13</td><td>Selection against based on the average value of the test</td></tr> <tr><td>t14</td><td>Verify data by committee</td></tr> <tr><td>t15</td><td>Quran-academic test, and Interview</td></tr> <tr><td>t16</td><td>Listing new student room</td></tr> </table>	t1	Register in boarding schools	t2	Get the receipt proof of boarding schools	t3	Input the registration data	t4	Submit the form and receive a registration file	t5	The Committee conducts the registration data entry	t6	Committee print mark registration	t7	Pay to financial section	t8	Orientation for new students	t9	Submit the form and attaching a other file registrations	t10	Selection against based on the average value of national exam	t11	See the announcement escaped selection	t12	Re-registration to the committee	t13	Selection against based on the average value of the test	t14	Verify data by committee	t15	Quran-academic test, and Interview	t16	Listing new student room	<pre> <?xml version="1.0" encoding="UTF-8"?> <!--PLEASE DO NOT EDIT THIS FILE Created with Workflow PetriNet Designer Version 3.2.0 (woped.org) <pnml><net type="http://www.informatik.hu-berlin.de/top/pntd/ptNetb" id="noID"> <place id="p1"> <name> <text>start</text> <graphics> <offset x="90" y="200"/> </graphics></name> <graphics> <position x="90" y="160"/> <dimension x="40" y="40"/> </graphics> </place> </pre>
t1	Register in boarding schools																																	
t2	Get the receipt proof of boarding schools																																	
t3	Input the registration data																																	
t4	Submit the form and receive a registration file																																	
t5	The Committee conducts the registration data entry																																	
t6	Committee print mark registration																																	
t7	Pay to financial section																																	
t8	Orientation for new students																																	
t9	Submit the form and attaching a other file registrations																																	
t10	Selection against based on the average value of national exam																																	
t11	See the announcement escaped selection																																	
t12	Re-registration to the committee																																	
t13	Selection against based on the average value of the test																																	
t14	Verify data by committee																																	
t15	Quran-academic test, and Interview																																	
t16	Listing new student room																																	

Model yang digunakan dalam riset ini berjumlah empat tipe model yang setiap modelnya memiliki variasi empat model proses bisnis tersendiri. Sehingga total objek model yang digunakan berjumlah 16 model.

2.2 Menghitung Kemiripan dan Pembobotan

Pengukuran kemiripan model proses bisnis dilakukan pada seluruh model proses bisnis yang ada. Kemiripan yang dihitung berdasarkan struktur, perilaku, dan kecocokan semantik. Pengukuran kemiripan struktural dan perilaku dilakukan dengan cara berikut ini :

Model pada A1= p1, p2, p3, p4, p5, p6, p7, p8, p9, p14, t5, t6, t7, t8, t9, t14, t15, t1, t2, t3, t5p6, p5t5, p9t9, t9p14, t7p8, p6t6, p7t7, t6p7, t8p9, p8t8, p1t2, p1t1, t1p2, p2t14, t14p3, p3t3, t3p4, p4t15, t15p5, t2p2 (didapatkan 40 elemen)

Model pada A2= p1, p2, p4, p5, p8, p9, p10, p12, p11, p14, t6, t7, t8, t10, t12, t11, t13, t1, t2, t3, t7p8, p12t7, p8t8, p1t2, t11p10, t8p9, t12p11, p9t10, p10t3, t6p5, p11t6, t10p14, p1t1, t2p2, t1p2, p2t11, t3p4, p4t12, p5t13, t13p12 (didapatkan 40 elemen).

Kemudian angka-angka tersebut dimasukkan ke dalam formula perhitungan kemiripan sebagai berikut:

Tabel 2 Perhitungan kemiripan

Perhitungan kemiripan struktur	Perhitungan kemiripan perilaku
$Jaccard = \frac{19}{(40+40)-19} = 0.356$	$Jaccard = \frac{2}{(9+9)-2} = 0.059$
$Dice = 2x \frac{19}{(40+40)} = 0.475$	$Dice = 2x \frac{2}{(9+9)} = 0.222$
$Cosine = \frac{19}{\sqrt{40+\sqrt{40}}} = 0.475$	$Cosine = \frac{2}{\sqrt{9+\sqrt{9}}} = 0.222$
$Overlap = \frac{19}{\min(40, 40)} = 0.475$	$Overlap = \frac{2}{\min(9, 9)} = 0.222$

Pengukuran kecocokan semantik label dilakukan dengan cara berikut :

Salah satu elemen pada model pada gambar A_1 = *Register In Boarding School*

Salah satu elemen pada model pada gambar A_2 = *Register In Boarding School*

$$L(A, B) = 1 - \left(\frac{0}{4}\right) = 1$$

Menggunakan cara yang sama, setiap kata dalam satu proses dihitung kecocokan semantik labelnya. Kemudian, semua kata yang terdapat dalam model proses bisnis tersebut dicari rata-rata *Levenstein Distance*-nya. Setelah semua perhitungan selesai dilakukan, kemudian diberikan bobot pada masing-masing hasil kemiripan mulai dari struktural, perilaku, dan semantik sehingga akan didapatkan akumulasi kombinasi dari setiap perhitungan. Berikut hasil perhitungan kemiripan struktur, perilaku, dan semantik secara keseluruhan, serta salah satu hasil pengukuran yang telah dilakukan dengan nilai bobot $w_1 = 0.3$, $w_2 = 0.2$, $w_3 = 0.5$

Tabel 3 Hasil perhitungan kemiripan dan akumulasi pembobotan

A. Perhitungan kemiripan struktur																B. Perhitungan kemiripan perilaku															
A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4
1.00	0.36	0.55	0.50	0.23	0.28	0.24	0.27	0.29	0.21	0.27	0.29	0.16	0.12	0.15	0.16	1.00	0.06	0.00	0.25	0.00	0.13	0.00	0.00	0.10	0.11	0.10	0.00	0.00	0.10	0.00	
0.36	1.00	0.38	0.53	0.20	0.43	0.41	0.32	0.30	0.19	0.10	0.09	0.19	0.17	0.18	0.19	0.06	1.00	0.00	0.20	0.00	0.11	0.13	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	
0.55	0.38	1.00	0.67	0.19	0.28	0.20	0.23	0.26	0.21	0.24	0.23	0.16	0.14	0.15	0.16	0.00	0.00	1.00	0.25	0.00	0.13	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	
0.50	0.53	0.67	1.00	0.18	0.23	0.19	0.21	0.26	0.15	0.16	0.18	0.13	0.10	0.11	0.13	0.25	0.25	1.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.23	0.20	0.19	0.18	1.00	0.25	0.55	0.42	0.21	0.22	0.14	0.26	0.29	0.16	0.27	0.29	0.00	0.00	0.00	1.00	0.00	0.22	0.43	0.07	0.08	0.08	0.15	0.00	0.08	0.15	0.00	
0.28	0.43	0.28	0.23	0.25	1.00	0.44	0.41	0.40	0.42	0.20	0.17	0.28	0.23	0.28	0.29	0.00	0.00	0.00	0.00	1.00	0.10	0.00	0.07	0.17	0.08	0.08	0.00	0.00	0.17	0.07	
0.24	0.41	0.20	0.19	0.55	0.44	1.00	0.39	0.33	0.26	0.17	0.19	0.35	0.23	0.33	0.35	0.00	0.00	0.00	0.00	0.22	1.00	0.29	0.17	0.08	0.09	0.08	0.08	0.18	0.08		
0.27	0.32	0.23	0.21	0.42	0.41	0.39	1.00	0.23	0.28	0.11	0.21	0.25	0.20	0.27	0.25	0.00	0.00	0.00	0.00	0.43	0.29	1.00	0.08	0.09	0.10	0.09	0.08	0.00	0.09	0.08	
0.29	0.30	0.26	0.26	0.21	0.40	0.33	0.23	1.00	0.40	0.38	0.31	0.33	0.31	0.33	0.33	0.00	0.08	0.20	0.10	0.07	0.07	0.17	0.08	1.00	0.06	0.07	0.06	0.20	0.06	0.31	0.29
0.21	0.19	0.21	0.15	0.22	0.42	0.26	0.28	0.40	1.00	0.28	0.27	0.35	0.23	0.35	0.37	0.00	0.10	0.00	0.00	0.08	0.17	0.08	0.09	0.06	1.00	0.25	0.23	0.06	0.00	0.23	0.13
0.27	0.10	0.24	0.16	0.14	0.20	0.17	0.11	0.38	0.28	1.00	0.35	0.21	0.19	0.22	0.21	0.10	0.11	0.00	0.00	0.08	0.08	0.09	0.10	0.07	0.25	1.00	0.15	0.07	0.00	0.15	0.07
0.29	0.09	0.23	0.18	0.26	0.17	0.19	0.21	0.31	0.27	0.35	1.00	0.20	0.22	0.24	0.20	0.10	0.00	0.00	0.00	0.08	0.08	0.08	0.09	0.06	0.23	0.15	1.00	0.06	0.00	0.14	0.06
0.16	0.19	0.16	0.13	0.29	0.28	0.35	0.25	0.33	0.35	0.21	0.20	1.00	0.36	0.55	0.50	0.00	0.00	0.00	0.00	0.15	0.00	0.08	0.08	0.20	0.06	0.07	0.06	1.00	0.00	0.62	0.39
0.12	0.17	0.14	0.10	0.16	0.23	0.23	0.20	0.31	0.23	0.19	0.22	0.36	1.00	0.38	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.13	0.06
0.15	0.18	0.15	0.11	0.27	0.28	0.33	0.27	0.33	0.35	0.22	0.24	0.55	0.38	1.00	0.67	0.00	0.00	0.00	0.00	0.08	0.17	0.18	0.09	0.31	0.23	0.15	0.14	0.21	0.13	1.00	0.42
0.16	0.19	0.16	0.13	0.29	0.29	0.35	0.25	0.33	0.37	0.21	0.20	0.50	0.53	0.67	1.00	0.00	0.00	0.00	0.00	0.15	0.07	0.08	0.08	0.29	0.13	0.07	0.06	0.39	0.06	0.42	1.00

C. Perhitungan kemiripan semantik																D. Hasil pembobotan penghitungan kemiripan																	
A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4		
1.00	0.81	0.91	0.86	0.55	0.59	0.56	0.58	0.54	0.55	0.50	0.52	0.51	0.47	0.48	0.51	1.00	0.50	0.62	0.63	0.34	0.40	0.35	0.37	0.36	0.36	0.36	0.37	0.30	0.27	0.31	0.30		
0.81	1.00	0.91	0.91	0.54	0.56	0.55	0.56	0.50	0.51	0.48	0.48	0.48	0.46	0.48	0.48	0.50	1.00	0.57	0.65	0.33	0.43	0.42	0.38	0.36	0.31	0.27	0.27	0.30	0.28	0.29	0.30		
0.91	0.91	1.00	0.86	0.55	0.58	0.56	0.58	0.53	0.54	0.50	0.51	0.51	0.47	0.49	0.51	0.57	1.00	0.68	0.33	0.40	0.34	0.36	0.38	0.33	0.32	0.32	0.31	0.28	0.29	0.31			
0.86	0.91	0.86	1.00	0.52	0.54	0.52	0.54	0.50	0.51	0.46	0.48	0.47	0.43	0.45	0.47	0.68	0.68	1.00	0.31	0.34	0.32	0.33	0.35	0.30	0.28	0.29	0.28	0.25	0.26	0.28			
0.55	0.54	0.55	0.52	1.00	0.82	0.94	0.92	0.57	0.59	0.55	0.56	0.58	0.59	0.56	0.58	0.31	0.33	0.33	1.00	0.49	0.68	0.68	0.36	0.38	0.33	0.37	0.41	0.34	0.37	0.41			
0.59	0.56	0.58	0.54	0.82	1.00	0.89	0.84	0.54	0.56	0.49	0.53	0.56	0.57	0.52	0.56	0.49	0.43	0.40	0.49	1.00	0.60	0.54	0.41	0.44	0.32	0.33	0.36	0.35	0.38	0.38			
0.56	0.55	0.56	0.52	0.94	0.89	1.00	0.87	0.56	0.58	0.52	0.55	0.55	0.56	0.53	0.55	0.60	0.61	0.60	0.61	1.00	0.61	0.41	0.38	0.33	0.35	0.40	0.36	0.40	0.40	0.40			
0.58	0.56	0.58	0.54	0.92	0.84	0.87	1.00	0.54	0.56	0.52	0.53	0.55	0.58	0.53	0.55	0.61	0.62	0.61	0.62	0.61	1.00	0.36	0.38	0.31	0.34	0.37	0.35	0.36	0.37	0.41			
0.54	0.50	0.53	0.50	0.57	0.54	0.56	0.54	1.00	0.87	0.88	0.78	0.60	0.59	0.61	0.60	0.65	0.66	0.65	0.66	0.65	0.65	1.00	0.57	0.57	0.49	0.44	0.40	0.46	0.46	0.46			
0.55	0.51	0.54	0.51	0.59	0.56	0.58	0.56	0.87	1.00	0.81	0.78	0.65	0.57	0.63	0.65	0.66	0.67	0.66	0.67	0.66	0.66	0.66	1.00	0.54	0.52	0.44	0.35	0.47	0.46	0.46			
0.50	0.48	0.50	0.46	0.55	0.49	0.52	0.52	0.88	0.81	1.00	0.78	0.58	0.59	0.61	0.58	0.65	0.66	0.65	0.66	0.65	0.65	0.65	0.65	1.00	0.52	0.37	0.35	0.40	0.37	0.40	0.37		
0.52	0.48	0.51	0.48	0.56	0.53	0.55	0.53	0.78	0.78	0.78	1.00	0.59	0.59	0.61	0.59	0.65	0.66	0.65	0.66	0.65	0.65	0.65	0.65	0.65	1.00	0.52	0.37	0.36	0.40	0.37	0.40		
0.51	0.48	0.51	0.47	0.58	0.56	0.55	0.55	0.60	0.65	0.58	0.59	1.00	0.79	0.89	0.99	0.66	0.67	0.66	0.67	0.66	0.66	0.66	0.66	0.66	0.66	1.00	0.51	0.51	0.65	0.72	0.72		
0.47	0.46	0.47	0.43	0.59	0.57	0.56	0.58	0.59	0.57	0.59	0.59	0.79	1.00	0.83	0.79	0.66	0.67	0.66	0.67	0.66	0.66	0.66	0.66	0.66	0.66	0.66	1.00	0.55	0.55	0.56	0.56		
0.48	0.48	0.49	0.45	0.56	0.52	0.53	0.53	0.61	0.63	0.61	0.61	0.89	0.83	1.00	0.96	0.66	0.67	0.66	0.67	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	1.00	0.55	0.55	0.56	0.56	
0.51	0.48	0.51	0.47	0.58	0.56	0.55	0.55	0.60	0.65	0.58	0.59	0.99	0.79	0.96	1.00	0.66	0.67	0.66	0.67	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	1.00	0.56	0.56	0.56	0.56

2.3 Melakukan Pengelompokan (*Clustering*)

Tabel 4 Hasil *clustering* (*Dice* 0.5, 0.3, 0.2) *threshold* 0.48

	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4	
A1	1.000	0.589	0.668	0.711													
A2	0.589	1.000	0.617	0.727	0.502	0.493											
A3	0.668	0.617	1.000	0.749													
A4	0.711	0.727	0.749	1.000													
B1					1.000	0.532	0.753	0.761									
B2		0.502			0.532	1.000	0.666	0.591									
B3		0.493			0.753	0.666	1.000	0.692									
B4					0.761	0.591	0.692	1.000									
C1									1.000	0.633	0.628	0.552	0.516			0.547	0.538
C2									0.633	1.000	0.618	0.591	0.503			0.546	0.534
C3									0.628	0.618	1.000	0.597					
C4									0.552	0.591	0.597	1.000					
D1																	

dari pembobotan pengukuran kemiripan model proses bisnis yang telah dilakukan sebelumnya. Hasil *clustering* ditunjukkan pada Tabel 4

2.4 Menghitung Akurasi

Setelah semua kombinasi diujicobakan pada setiap algoritma kemiripan, langkah selanjutnya menghitung akurasi dari setiap hasil proses *clustering* yang telah dilakukan. Tujuan dari proses perhitungan akurasi dimaksudkan untuk mengetahui kombinasi yang tepat, algoritma apa yang digunakan, dan nilai *threshold* berapa sehingga dapat menemukan *common process* yang akurat. Metode yang digunakan untuk mengukur akurasi adalah ROC. Hasil akurasi didapat dengan rumus :

Rumus ROC	
$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$	
TP : True Positive	TN : True Negative
FP : False Positive	FN : False Negative

Persamaan 1 Persamaan akurasi ROC

Akurasi dari pengelompokan algoritma *Dice* dengan pembobotan 0.5, 0.3, 0.2 :

$$Accuracy = \frac{64+176}{64+176+16+0} \times 100\%$$

$$Accuracy = 93.75\%$$

2.5 Mencari Common Process

Penentuan *common process* adalah penentuan *fragment* yang sama dengan mengacu pada hasil kluster dengan menggunakan pendekatan partisi graf (*graph partition approach*) [14]. Hasil kluster berupa nilai *similarity* yang terdapat pada domain yang telah ditentukan pada tahap *clustering*. Nilai domain dari proses *clustering* tersebut merupakan nilai yang lolos *threshold*. Sehingga berdasarkan nilai *similarity* yang terkluster, maka diambil nilai per-kluster. Dalam menentukan *common process* dilakukan dengan membandingkan nilai setiap kluster dengan pada kluster lainnya. Sebagai contoh, kluster 1 = {t5} {t6} {t8} {t11} {t3} kluster 2 = {t4} {t5} {t6} {t8} {t7} dari kedua kluster tersebut, dihasilkan *common process* pada kluster 1 dan kluster 2 {t5} – {t6} – {t8}.

3. Hasil Dan Pembahasan

Model Petri Net yang berformat *.pnml diparsing kedalam format *.xml untuk memudahkan pengelompokan elemen yang terdapat pada model proses bisnis. Dalam proses ini data yang diambil akan disusun sesuai dengan jenis elemennya. Elemen yang dimaksud adalah *place*, *transition*, dan *arc*. Setelah dilakukan pemarsingan pada kedua model yang akan dihitung kemiripannya, hasil dari proses *parsing* tersebut dihitung dengan membandingkan elemen yang terdapat pada model proses bisnis.

- I. Sumber data uji menggunakan model proses bisnis yang berasal dari beberapa yang meliputi proses akademik, penerimaan santri baru, kesarifan, dan sarana prasarana. pengukuran dilakukan secara keseluruhan sesuai dengan jumlah model yang tersedia. Model tersedia adalah empat proses bisnis akademik, empat proses bisnis empat proses bisnis baru, empat proses bisnis kesarifan, dan empat proses bisnis sarana prasarana. Sehingga keempat tipe model proses bisnis tersebut didapatkan sebanyak 256 hasil perhitungan. dilakukan menggunakan empat metode pada kemiripan perilaku-struktur dan satu metode semantik. Hasil dari berbagai metode tersebut menghasilkan nilai yang bervariasi sehingga perhitungan kemiripan, dilakukan pembobotan pada masing-masing hasil kemiripan. tersebut dilakukan dengan menggunakan beberapa kombinasi bobot yang terdiri dari rata-kemiripan^{II} struktur, perilaku dan semantik. Selain rata-rata, bobot yang digunakan dengan struktur, perilaku dan semantik adalah 0.5-0.4-0.1, 0.5-0.3-0.2, 0.4-0.4-0.2, 0.4-0.3-0.3. hasil perhitungan kemiripan tersebut dibobotkan, kemudian dilakukan *threshold* untuk

Tabel 4. Kemiripan model proses bisnis per bobot yang digunakan pada proses *clustering* dengan memasukkan nilai *threshold* sebesar 0.48.

Model yang telah dihitung dalam riset ini merupakan model yang bersifat *scalable*, hal ini dapat dilihat dari perbandingan yang dilakukan antar model dengan menggunakan rumus skalabilitas untuk mencari nilai skala. Nilai skala didapatkan dengan mengalikan jumlah elemen model proses bisnis dengan nilai *Control Flow Complexity* (CFC) sebagaimana persamaan $skala(A) = E(A) * CFC(A)$ [15]. Kemudian nilai skala tersebut dibandingkan dengan nilai skala pada model proses bisnis yang lain, sehingga $skala(A,B) = skala(A) / skala(B)$. Kemudian untuk mendapatkan nilai skalabilitas, maka nilai skala diperbandingkan antara skala model proses A dan B yang kemudian dikalikan dengan nilai kemiripan, sehingga dapat diformulasikan sebagai berikut $skalability(A,B) = 1 - (skala(A,B) * sim(A,B))$.

Tabel 5 Hasil clustering (0.5, 0.4, 0.1) threshold 0.48

Jaccard																Dice																	
A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4		
A1	1.00	0.50	0.68	0.69												A1	1.00	0.59	0.74	0.77													
A2	0.50	1.00	0.61	0.68												A2	0.59	1.00	0.68	0.76													
A3	0.68	0.61	1.00	0.75												A3	0.74	0.68	1.00	0.82													
A4	0.69	0.68	0.75	1.00												A4	0.77	0.76	0.82	1.00													
B1				1.00	0.49	0.70	0.68									B1					1.00	0.55	0.78	0.76									
B2				0.49	1.00	0.59	0.53									B2					0.55	1.00	0.67	0.60									
B3				0.70	0.59	1.00	0.61									B3					0.78	0.67	1.00	0.70									
B4				0.68	0.53	0.61	1.00									B4					0.76	0.60	0.70	1.00									
C1								1.00	0.60	0.61	0.53					C1									1.00	0.68	0.69	0.60					
C2								0.60	1.00	0.58	0.56					C2									0.68	1.00	0.58	0.64					
C3								0.61	0.58	1.00	0.55					C3									0.69	0.65	1.00	0.62					
C4								0.53	0.56	0.55	1.00					C4									0.60	0.64	0.62	1.00					
D1												1.00	0.54	0.69	0.73		D1													1.00	0.61	0.79	0.92
D2												0.54	1.00	0.57	0.61		D2													0.61	1.00	0.66	0.62
D3												0.69	0.57	1.00	0.75		D3													0.79	0.66	1.00	0.82
D4												0.73	0.61	0.75	1.00		D4													0.92	0.62	0.82	1.00

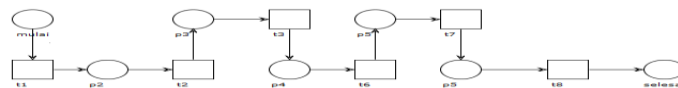
Cosine																Overlap																	
A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4		
A1	1.00	0.56	0.71	0.85												A1	1.00	0.60	0.74	0.82													
A2	0.56	1.00	0.62	0.63												A2	0.60	1.00	0.69	0.81					0.54	0.53			0.49		0.49	0.50	
A3	0.71	0.62	1.00	0.76												A3	0.74	0.69	1.00	0.88									0.51				
A4	0.85	0.63	0.76	1.00												A4	0.82	0.81	0.88	1.00													
B1				1.00	0.55	0.78	0.76									B1					1.00	0.55	0.79	0.81									
B2				0.55	1.00	0.67	0.60									B2					0.55	1.00	0.67	0.62	0.49	0.52							
B3				0.78	0.67	1.00	0.70									B3					0.79	0.67	1.00	0.72		0.51			0.49		0.48	0.49	
B4				0.76	0.60	0.70	1.00									B4					0.81	0.62	0.72	1.00									
C1								1.00	0.68	0.69	0.60					C1									1.00	0.69	0.70	0.60					
C2								0.68	1.00	0.65	0.64					C2									0.52	0.51			0.69	1.00	0.67	0.65	
C3								0.69	0.65	1.00	0.62					C3									0.70	0.67	1.00	0.63					
C4								0.60	0.64	0.62	1.00					C4									0.60	0.65	0.63	1.00					
D1												1.00	0.61	0.79	0.92		D1													1.00	0.61	0.81	0.92
D2												0.61	1.00	0.66	0.62		D2													0.61	1.00	0.67	0.62
D3												0.79	0.66	1.00	0.82		D3													0.81	0.67	1.00	0.84
D4												0.92	0.62	0.82	1.00		D4													0.92	0.62	0.84	1.00

Setelah dilakukan perhitungan terhadap setiap model proses bisnis, kemudian dicari seberapa besar akurasi metode yang digunakan untuk menemukan *common process* yang nantinya dapat digunakan sebagai *fragment* utama penyusun model proses bisnis yang baru. Pencarian akurasi menggunakan metode ROC dengan membandingkan penjumlahan kondisi *true positive* dan *true negative* dengan jumlah populasi data. Dari hasil pengukuran akurasi tersebut didapatkan hasil pada **Error! Reference source not found.**

Tabel 6 Hasil Uji Coba Perbandingan Nilai Skalabilitas

	A1	A2	A3
A1	1	0.758	0.911
A2	-	1	0.750
A3	-	-	1

Gambar 2 Salah Satu Hasil Common Process



Uji coba menunjukkan bahwa algoritma *Jaccard Coefficient* yang digunakan untuk mengukur kemiripan secara perilaku dan struktur yang dikalkulasikan dengan kemiripan semantik yang menggunakan algoritma *Levenshtein* menghasilkan nilai akurasi yang lebih tinggi daripada *Dice*, *Cosine* dan *Overlap*. Selanjutnya pencarian *common process* dapat dilakukan dengan mencari *fragment* yang terdapat pada *cluster* yang memiliki akurasi tinggi. Pada **Error! Reference source not found.** menunjukkan penemuan *common process* yang berada dalam sebuah *cluster* yang tinggi sebesar 100% dengan perpaduan kemiripan *semantic* menggunakan *Levenshtein* dan algoritma kemiripan struktur dan perilaku menggunakan *Jaccard Coefficient*.

Tabel 7 Hasil Pengukuran Akurasi

Kombinasi	Threshold											
	Jaccard			Dice			Cosine			Overlap		
	0.48	0.50	0.52	0.48	0.50	0.52	0.48	0.50	0.52	0.48	0.50	0.52
Average	89.06%	88.28%	87.50%	83.75%	82.97%	82.97%	82.19%	82.19%	82.19%	82.81%	86.72%	82.19%
5-4-1	99.22%	100.00%	96.88%	86.72%	82.19%	83.75%	83.59%	88.28%	82.97%	83.28%	72.66%	78.13%

5-3-2	100.00%	97.66%	96.09%	93.75%	94.53%	96.88%	90.63%	93.75%	96.88%	75.00%	81.25%	86.72%
4-4-2	97.66%	93.75%	92.97%	93.75%	94.53%	96.09%	92.19%	92.97%	96.88%	75.78%	78.91%	85.16%
4-3-3	94.53%	89.84%	89.06%	95.31%	96.09%	93.75%	94.53%	95.31%	91.41%	82.81%	86.72%	92.19%

4. Simpulan

1. Berdasarkan riset yang telah dilakukan, didapatkan hasil bahwa algoritma *Jaccard Coefficient* merupakan metode yang terbaik untuk menghitung kemiripan perilaku dan struktur dibandingkan dengan algoritma *Dice*, *Cosine*, dan *Overlap*.
2. *Clustering* pada beberapa model proses bisnis dilakukan dengan menggunakan dua tahap, yaitu *clustering domain* dan *clustering pattern*. Akurasi yang didapatkan dari proses *clustering* untuk mendapatkan kelompok model proses bisnis sesuai dengan domainnya adalah 100% dengan kombinasi bobot kemiripan semantik, struktur, perilaku (0.5, 0.4, 0.3) dengan *threshold* 0.50 dan kombinasi bobot kemiripan semantik, struktur, perilaku (0.5, 0.3, 0.2) dengan *threshold* 0.48. kemiripan semantik menggunakan algoritma *Levenshtein* dan struktur-perilaku menggunakan *Jaccard Coefficient*.
3. *Common proses* dapat ditemukan dengan menggunakan pendekatan partisi graf (*graph partition approach*). *Common process* tersebut merupakan *fragment* yang bersifat umum yang dapat diterapkan oleh beberapa organisasi sesuai kondisi yang diinginkan tanpa mengubah satu persatu proses bisnis yang bersifat personal.

Daftar Pustaka

- [1] S. R. Magal dan J. Word, *Integrated business processes with ERP systems*, Wiley Publishing, 2011.
- [2] M. L. Rosa, W. M. V. d. Aalst, M. Dumas dan F. P. Milani, "Business process variability modeling: A survey," 2013.
- [3] R. Sarno, E. W. Pamungkas dan D. Sunaryono, "Workflow common fragments extraction based on WSDL similarity and graph dependency," dalam *2015 International Seminar on Intelligent Technology and Its Applications (ISITIA)*.
- [4] S. Niwattanakul, J. Singthongchai, E. Naenudorn dan S. Wanapu, "Using of Jaccard Coefficient for Keywords Similarity," *Proceedings of the International MultiConference of Engineers and Computer Scientists 2013*, vol. Vol I, pp. 1-5, 2013.
- [5] E. Gallagher, *COMPAH Documentation*, Boston: University of Massachusetts, 1999.
- [6] C. D. Manning, P. Raghavan, H. Schutze dan others, *Introduction of Information Retrieval*, New York: Cambridge University Press, 2008.
- [7] G. N. D. Adwitya, *Analisis dan Implementasi Algoritma Overlapping Cover Coefficient-Based Clustering Method (OC3M) Pada Dokumen Teks Berbahasa Indonesia*, Bandung: Universitas Telkom, 2015.
- [8] A. W. Prasetya, *Aplikasi Pengolahan Proses Bisnis Menggunakan Metode Analisis Kemiripan Behavioral, Structural, dan Semantic untuk Meningkatkan Akurasi dalam Penentuan Common Fragment Workflow pada ERP Pesantren*, vol. I, Malang: UIN Maulana Malik Ibrahim Malang, 2017, pp. 1-7.
- [9] J. Billington, S. Christensen, K. Van Hee, E. Kindler, O. Kummer, L. Petrucci, R. Post, C. Stehno dan M. Weber, "The Petri net markup language: concepts, technology, and tools," dalam *International Conference on Application and Theory of Petri Nets*, Springer, 2003, pp. 483--505.
- [10] A. C. Fauzan, *Sistem Cerdas Untuk Menemukan Behavioral Similarity Pada Model Proses Bisnis Menggunakan Algoritma TArS*, Malang: UIN Maulana Malik Ibrahim Malang, 2012.
- [11] R. Dijkman, M. Dumas dan L. Garcia-Banuelos, "Graph Matching Algorithms for Business Process Model Similarity Search," *Business Process Management*, vol. Volume 5701 of the series Lecture Notes in Computer Science, pp. pp 48-63, 2009.
- [12] R. Anggrainingsih, "Analisis Dan Verifikasi Workflow Menggunakan Petri (Studi kasus; Proses Bisnis di Universitas Sebelas Maret)," *Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2014 (Semantik 2014)*, 2014.