Identifikasi Pola Fraud dalam Transaksi Online

Solichul Huda[1)], Heru Agus Santoso[2)]
Universitas Dian Nuswantoro
Jl. Imam Bonjol 155 Semarang
e-mail: solichul.huda@dinus.ac.id

Abstrak

Kuantitas fraud (penipuan) pada transaksi online meningkat dari waktu ke waktu. Beberapa penelitian sebelumnya telah mengusulkan metode deteksi fraud; namun metode tersebut tidak dapat mendeteksi fraud dalam transaksi online dengan baik. Hal itu disebabkan oleh indikator fraud tidak dapat mengidentifikasi dengan tepat pelanggaran Standard Operating Sistem (SOP) pada transaksi online. Penelitian ini mengusulkan atribut/indikator fraud dan pola fraud untuk menganalisis proses bisnis pada transaksi online. Penentuan fraud atau penipuan dilakukan dengan, pertama menganalisis proses bisnis yang melanggar SOP. Selanjutnya, proses bisnis yang melanggar SOP diidentifikasi atribut/indikator fraud. Kemudian, dilakukan pembobotan atribut/indikator berdasarkan jumlah indikator/atribut yang teridentifikasi. Terakhir, melakukan uji similarity untuk menentukan pelanggaran SOP tersebut merupakan fraud atau bukan. Dari eksperimen yang dilakukan menunjukkan bahwa metode yang diusulkan ini mampu mendeteksi fraud pada transaksi online dengan akurasi, sensitivitas dan spesifisitas masing-masing 0.97, 1 dan 0.97.

Kata kunci: *identifikasi, fraud, penipuan, transaksi, online, pola.*

1. Pendahuluan

Dewasa ini pemerintah daerah atau pemerintah kota berlomba meningkatkan pelayanan mereka kepada masyarakat dengan memujudkan *smart city*. Layanan tersebut meliputi bidang kesehatan, bidang pembangunan fisik infrastruktur jaringan komunikasi, termasuk bidang ekonomi. Dalam bidang ekonomi misalnya, pemerintah mendorong pemanfaatan teknologi informasi oleh Usaha Kecil Menengah (UKM) sehingga mendukung terwujudnya *smart city*. Pemanfaatan teknologi informasi dapat diwujudkan dalam bentuk sistem informasi transaksi *online*.

Transaksi *online* merupakan implementasi teknologi informasi untuk transaksi jual beli. Teknologi tersebut meliputi teknologi perangkat keras komputer, telepon seluler, dan teknologi perangkat lunak. Perkembangan teknologi informasi ini membuat transaksi jual beli tidak dibatasi oleh ruang dan waktu. Penjualan *online* ini sejalan dengan program pemerintah dalam menggerakkan masyarakat untuk melakukan transaksi lewat transaksi *online*. Namun disisi lain, pelaku penipuan (*fraud*) memandang transaksi *online* merupakan area baru memperoleh keuntungan [1].

Dalam dua tahun terakhir ini penipuan *online* (*cyber crime*) dalam bentuk penipuan transaksi *online* jumlahnya semakin meningkat. Modus penipuan/fraud tersebut berubah-ubah dengan metode atau teknik yang dikembangkan. Oleh karena itu pemerintah dan peneliti perlu kerja keras untuk mengembangkan metode deteksi fraud / penipuan agar transaksi *online* aman.

Beberapa penelitian sebelumnya sudah pernah mengusulkan metode deteksi fraud/penipuan dalam transaksi online [2],[3],[4]. Namun, metode tersebut hanya dapat mendeteksi *fraud* setelah muncul kerugian. Penelitian ini mengusulkan metode deteksi fraud berbasis *process mining* yang akan menganalisis proses yang sedang dijalankan, sehingga *fraud* dapat dideteksi lebih dini. Penelitian deteksi *fraud* berbasis *process mining* sudah pernah dilakukan oleh beberapa penelitian sebelumnya[5],[6],[7],[8],[9], namun metode yang diusulkan tersebut kurang akurat untuk mendeteksi fraud/penipuan dalam transaksi *online*. Penelitian ini akan mengembangkan metode deteksi fraud pada transaksi *online* berbasis proses transaksi yang sedang dilakukan. Analisis proses bisnis tersebut diharapkan dapat menunjukkan indikasi terjadinya *fraud*/penipuan, sehingga *fraud* terdeteksi lebih dini sebelum kerugian terjadi.

Dalam penjualan *online*, penjual dan pembeli dapat mengaktualisasikan bentuk bisninya dengan berbagai media mulai dari desain tampilan sampai umpan balik pembeli, yang umumnya berupa testimoni. Selain itu mereka dapat memilih barang yang akan diperjualbelikan dengan mudah dan cepat tanpa dibatasi tempat dan waktu. Model ini yang membuat masyarakat tertarik melakukan transaksi *online*.

Fraud/penipuan yang terjadi dalam transaksi *online* dapat dilakukan oleh penjual atau pembeli. Penipuan terhadap penjual, umumnya berupa tidak diterimanya uang dari pembeli sesuai dengan aturan yang mereka disepakati. Sedangkan penipuan terhadap pembeli, umumnya berujud ketidaksesuaian antara barang pesanan dengan barang yang dikirim, atau barang sama sekali tidak diterima.

Penipuan *online* mudah dilakukan karena ada celah keamanan yang berupa tidak adanya pertemuan antara penjual dengan pembeli. Saat transaksi, penjual maupun pembeli kesulitan untuk melakukan validasi dan verifikasi pisik tentang identitas penjual atau pembeli, seperti alamat usaha atau kondisi barang yang perjualbelikan. Selain itu, pembeli atau penjual tidak mengetahui perilaku diantara mereka.

Teknologi informasi dalam transaksi *online*, sebetulnya dapat menunjukkan lokasi pelaku transaksi. Kemudian, pola penjual dan pembeli dalam transaksi juga dapat dipelajari dari *event logs* atau data transaksi yang ada. *Event logs* berfungsi untuk menyimpan proses transaksi yang dilakukan [10],[11]. Dalam *event logs* ini minimal tersimpan informasi mengenahi nama *event*, lama melalukan *event*, dan *user* (originator) yang menjalankan *event* [12]. Contoh *event* dalam transaksi *online* adalah pesan_barang. Selama ini, *event logs* hanya dipakai untuk menganalisis proses setelah terjadi fraud. Penelitian ini akan mengunakan *event logs* untuk menganalisis proses transaksi yang sedang berlangsung sehingga indikasi *fraud*/penipuan teridentifikasi lebih dini. Penelitian ini akan mengidentifikasi *fraud*/penipuan berdasarkan proses transaksi yang terjadi. Untuk menganalisis data transaksi menggunakan teknik *data mining*, sedangkan untuk menganalisis proses transaksi menggunakan teknik *process mining* [13]. Studi ini akan menggunakan pendekatan *process mining* dan *data mining* untuk mengembangkan metode deteksi fraud/penipuan berdasarkan data transaksi dan *event logs*.

Fraud dalam aplikasi kredit kemungkinan sedikit berbeda dengan *fraud* dalam transaksi *online*, sehingga kemungkinan indikator atau atributnya juga berbeda. Oleh karena itu, penelitian ini akan mengidentifikasi indikator /atribut *fraud* pada transaksi *online*. Begitupun pola penipuannya kemungkinan berbeda dengan *fraud* pada aplikasi kredit; penelitian ini akan mengidentifikasi pola *fraud* dalam transaksi *online*. Hipotesis penelitian ini, metode yang diusulkan mampu mendeteksi penipuan pada transaksi *online* dengan akurat.

2. Metode Penelitian

2.1 Data Penelitian

Dalam melakukan penelitian ini, data sumber utama berupa data primer dan data sekunder. Data primer berupa *event logs* transaksi *online* yang diambil dari tiga Usaha Kecil Menengah (UKM) yang melakukan transaksi *online* dalam tiga tahun terakhir ini. Sedangkan data sekunder berupa data tambahan yang diperoleh dari buku panduan atau *standard operating system* (SOP).

Event logs transaksi ini berjumlah 4.025 *case* atau *process instance* dengan 12.000 *event*/*record*. Data tersebut dibagi menjadi data *trining* dan data *testing* masing-masing 2.415 *case* dengan 7.200 *record* dan 1.610 case dengan 4.800 *record*.

2.2. Penentuan Indikator

Indikasi penipuan yang terjadi diidentifikasi menggunakan metode analisis proses dan analisis data. Penelitian ini menggabungkan pendekatan *process mining* dengan *data mining*. Pertama, analisis proses bisnis menggunakan metode analisis *skip*, analisis *throughput time* dan analisis *wrong pattern*[8]. Dan selanjutnya, menganalisis data untuk memperoleh pembedaan antara *case* dengan SOP.

2.2.1. Klasifikasi pelanggaran

Penelitian ini akan menganalisis proses bisnis semua *case* dalam data *training*. Selanjutnya, hasil analisis tersebut dicluster sesuai dengan pola proses bisnis masing-masing *case*. Analisis ini menghasilkan delapan *cluster*, dimana satu *cluster* berisi *case* yang sesuai SOP, sedangkan tujuh *cluster* lainnya melanggar SOP. Metode cluster yang kami gunakan sesuai metode yang digunakan dalam [9]. Kami menganalisis tujuh *cluster* tersebut dan menunjukkan ada tujuh jenis pelanggaran SOP, yaitu *throughput time, quantity, same location, wrong pattern, skip, map* dan *relationship*. Indikator *throughput time, wrong pattern,*dan *skip* sudah pernah diidentifikasi dalam [7],[8],[9]. Selanjutnya, tujuh jenis pelanggaran SOP tersebut dikenal dengan istilah atribut atau indikator *penipuan*/fraud.

2.2.2. Penentuan indikator *fraud.*

Penelitian ini mengidentifikasi tujuh pelanggaran SOP. Penelitian ini menguji bobot korelasi antara tujuh atribut/indikator *fraud* tersebut dengan bobot *fraud* [8]. Dari uji korelasi menunjukkan bahwa enam indikator/atribut memiliki korelasi yang signifikan dengan *fraud*/penipuan. Dengan demikian, enam atribut tersebut ditentukan sebagai atribut/indikator *fraud*. Enam indikator atau atribut tersebut adalah *throughput time, quantity, same location, wrong pattern, skip* dan *relationship*. Indikator/atribut *fraud* tersebut ditunjukkan dalam Tabel 1.

Tabel 1. Indikator/atribut *Fraud* Dalam Transaksi *Online*

No.	Nama indikator	Keterangan	Penjelasan
1	*Throughput time*	Waktu menjalankan *event* yang lebih kecil dibanding dengan waktu standar menjalankan *event*	Contoh *event* masukkan_pesanan memerlukan waktu 30 menit, padahal waktu standard masukkan_pesanan 15 menit, maka *event* ini terindikasi *throughput time* karena melebihi waktu standard event masukkan_pesanan.
2	*Quantity*	Jumlah pembelian yang terlalu besar	Misalnya dari data *trining* bahwa rata-rata pembelian 5 item. Ada case yang melakukan pemesanan sejumlah 15 item, maka terindikasi *quantity* karena jumlah pemesanannya melebihi jumlah umum pemesanan
3	*Same location*	Tempat pembeli dengan lokasi penjual satu lokasi	Misalnya alamat pemesan jl. Diponegoro no. 158 semarang, sedangkan alamat penjual online Jl. Diponegoro no. 80 Semarang. Disebabkan lokasi pembeli dan penjual dalam nama jalan dan kota yang sama maka case terindikasi *same location*
4	*Wrong pattern*	Urutan proses bisnis berbeda dengan pola SOP	Urutan proses dalam SOP seharusnya *event* A, *event* B kemudian *event* C. Dalam sebuah case urutannya *event* A, *event* C selanjutnya *event* B, maka *case* ini terindikasi *wrong pattern*
5	*Skip*	Melompati *event* dibanding dengan SOP	Urutan proses nya seharusnya *event* A, *event* B kemudian *event* C. Dalam proses bisnis sebuah case *event* A, *event* C kemudian *event* D, karena melompati event B, case ini terindikasi *Skip*
6	*Relationship*	Pelanggan tetap	Jika kustomer sudah membeli tiga kali atau lebih, maka *relationship* .t.

2.3 Penentuan Pola *Fraud*

Pola *fraud* digunakan untuk merekam berbagai pola *fraud*/penipuan yang yang diperoleh dari data *training*. Pola ini menunjukkan bobot atribut/indikator fraud dari *case* yang terindentifikasi *fraud*. Pola ini digunakan sebagai rujukan penentuan sebuah *case* yang melanggar SOP merupakan penipuan/*fraud* atau bukan.

2.3.1. Identifikasi Penipuan

Indikator penipuan yang dari sebuah proses bisnis diidentifikasi dari proses bisnis yang melanggar SOP. Sebuah pelanggaran SOP diidentifikasi sebagai atribut/indikator penipuan. Selanjutnya penentuan pelanggaran SOP tersebut sebagai fraud atau bukan dilakukan dengan uji *similarity* dengan pola fraud.. Hasilnya *case* tersebut *fraud* atau bukan.

2.3.2. Penentuan Pola *Fraud*

Penelitian ini akan menganalisis pola dari *case* yang terindikasi melanggar SOP. Pola tersebut menunjukkan bobot atribut/indikator *fraud.* Pola fraud yang diperoleh akan menjadi dasar untuk uji *similarity*. Contoh pola penipuan atau *fraud* ditunjukkan dalam Tabel 2.

Tabel 2. Contoh Pola *Fraud* Pada Transaksi *Online*

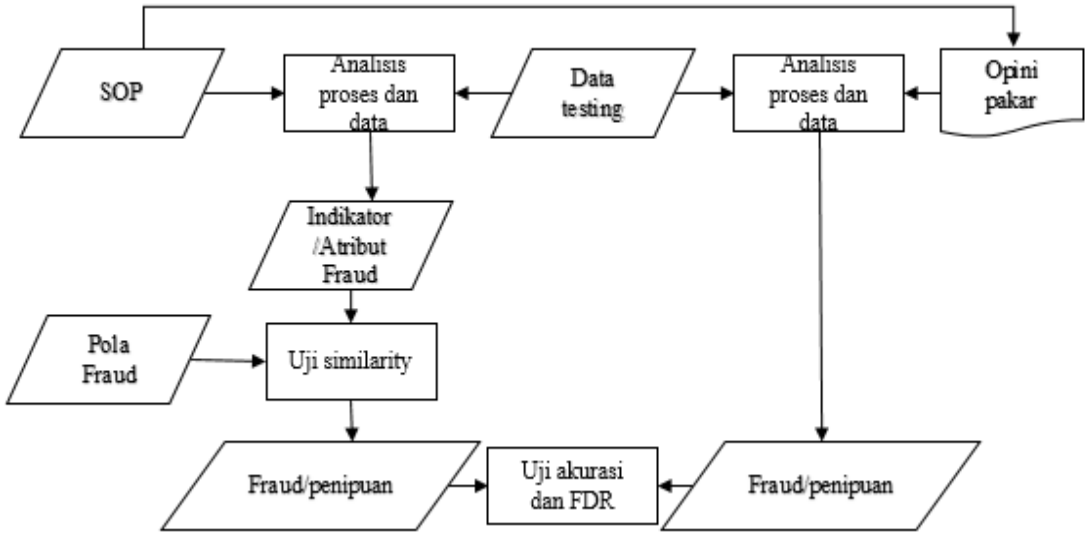
Pola Fraud	Throughput Time	Wrong pattern	Skip	Same location	Quantity	Relationship
1	Low	Low	Low	f	f	F
2	Middle	Low	Low	f	f	F
3	High	Low	Low	f	f	F
4	Low	Middle	Low	t	t	T
5	Middle	Middle	Low	t	t	T
6	High	Middle	Low	t	t	T
7	Low	High	Low	f	t	T
8	Middle	High	Low	f	t	T
9	High	High	Low	f	f	F

2.4. Penentuan *Fraud*

Indikasi penipuan yang teridentifikasi dalam sebuah *case* menjadi kunci untuk menentukan *fraud*. Berdasarkan indikator/atribut yang diperoleh, penelitian ini akan melakukan cek *similarity* dibanding dengan pola yang ada dalam database pola penipuan. *Case* yang memperoleh nilai *similarity* diatas nilai *threshold* ditentukan sebagai *fraud*.

3. Hasil dan Pembahasan

Uji akurasi dilakukan untuk mengukur kemampuan metode yang diusulkan dalam mendeteksi *fraud* dalam transaksi *online*. Proses evaluasi ditunjukkan dalam Gambar 1.



Gambar 1. Proses Evaluasi

Dalam eksperimen ini, peneliti mengumpulkan data *event logs* dari tiga Usaha Kecil Menengah (UKM) dalam periode 2014-2016. Data tersebut dikelompokkan dalam data *training* dan data *testing*, masing-masing 2.415 *case* (7.200 *event/record*) dan 1.610 *case* (4.800 *event/record*).

Kami menganalisis data dan proses dalam data *training* untuk memperoleh *case* yang melanggar SOP. Selanjutnya penelitian ini melakukan pembobotan indikator/atribut yang diperoleh oleh *case*. Metode pembobotan masing-masing indikator/atribut penipuan seperti metode pembobotan yang digunakan dalam[8]. Pola *case* yang ditetapkan sebagai fraud oleh pakar disimpan dalam pola fraud/penipuan.

Analisis terhadap data *test* menghasilkan 243 *case* melanggar SOP. *Case* 1021 memiliki tiga atribut yaitu *throughput time*, *wrong pattern* dan *quantity*, masing-masing 1,1, dan 't'. Sedangkan dalam case ID 8772 memiliki 2 atribut, yaitu *throughput time* dan *wrong pattern*, masing-masing 1. Contoh *case*

yang teridentifikasi *fraud* ditunjukkan dalam Tabel 3. Metode analisis proses bisnis yang dimplementasikan dalam penelitian ini seperti metode yang digunakan dalam [8]. Selanjutnya melakukan pembobotan indikator/atribut pada *case* yang teridentifikasi melanggar SOP. Terakhir, melakukan uji *similarity* untuk menentukan *case* tersebut *fraud* atau bukan.

Metode deteksi *fraud* yang diusulkan dalam paper ini untuk memperoleh metode deteksi *fraud* dalam transaksi *online* yang lebih akurat dibanding metode sebelumnya [8],[9]. Evaluasi ini dilakukan dengan menganalisis dan membobot atribut/indikator *fraud* dalam data *testing* dan uji *similarity* dengan pola *fraud*. Disisi lain, pakar menganalisis data *testing* menggunakan metode mereka. Evaluasi akurasi, sensitivitas dan spesifisitas terhadap metode ini dilakukan untuk melihat kelebihan metode yang diusulkan ini. Uji sensitivitas dan spesifisitas dilakukan karena ketidakseimbangnya jumlah case yang *fraud* dan bukan *fraud*. Rumus (1), Rumus(2) dan Rumus (3) masing-masing digunakan untuk menghitung akurasi, sensitivitas dan spesifisitas. Metode evaluasi ini seperti yang dilakukan dalam penelitian [7],[8],[9].

Metode *receiver operating characteristic* (ROC) digunakan untuk mengukur akurasi metode deteksi *fraud* dalam transaksi *online*. *Framework* ini mengukur akurasi dengan mempertimbangkan *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN). TP berarti pakar dan metode ini sama menentukan bahwa *case* tersebut *fraud* atau penipuan. TN juga mengangap bahwa pakar dan metode menentukan bahwa *case* tersebut bukan *fraud*. Jika pakar menentukan *fraud* sedangkan metode bukan fraud, berarti FN. Jika pakar memutuskan bukan *fraud* sedangkan metode menentukan *fraud*, berarti FP.

Tabel 3. Contoh *Case* Yang Terindikasi *Fraud*

ID Case	Throughput Time	Wrong pattern	Skip	Same location	Quantity	Relationship
1021	1	1	-	f	t	t
3324	1	1	-	f	f	t
3581	1	-	1	f	f	t
6567	2	-	-	f	f	t
8521	2	-	-	f	f	t
8533	2	-	-	f	t	f
8645	1	-	-	f	f	t
8700	2	1	-	t	f	t
8772	1	1	-	f	f	f
8905	1	-	1	f	f	f

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}. \quad (1)$$

$$Sensitivitas = \frac{TP}{TP+FN} \quad (2)$$

$$Spesifisitas = \frac{TN}{TN+FP} \quad (3)$$

Hasil evaluasi terhadap data *testing* menggunakan metode yang diusulkan menghasilkan 243 *case* yang melanggar SOP. Metode ini menentukan 243 *case* tersebut merupakan *fraud*. Disisi lain pakar juga menganalisis data *testing* menggunakan metode mereka. Pakar menentukan hanya 201 *case* yang dianggap *fraud* atau penipuan. Dengan demikian ada perbedaan antara hasil analisis metode yang diusulkan dengan hasil analisis oleh pakar. Hasil diskusi dengan pakar tersebut menunjukkan bahwa menggunakan metode yang diusulkan 201 *case* diidentifikasi sebagai *true positive*, artinya bahwa pakar dan metode ini sama mengidentifikasi bahwa 201 case tersebut *fraud*. Kemudian, 42 *case* sebagai *false positive*, berarti ada 42 *case* yang menurut metode ini terindikasi fraud, sedangkan menurut pakar bukan fraud. Dan 1.367 case sebagai *true negative*, dimana pakar dan metode ini sama menentukan 1.367 case bukan *fraud*. Menggunakan Rumus (1), Rumus (2) dan Rumus (3), metode ini memperoleh akurasi 0,97, sensitivitas 1dan spesifisitas 0,97. Hasil evaluasi metode yang diusulkan ditunjukkan dalam Tabel 4.

Tabel 4. Hasil Evaluasi Metode

Variabel ROC				Akurasi	Sensitivitas	Spesifisitas
True Positive	*False Positive*	*False negative*	*True negative*			
201	42	0	1.367	0,97	1	0,97

4. Simpulan

Fraud dalam transaksi *online* dapat diidentifikasi dengan akurat dengan mengembangkan atribut/indikator *fraud* dan pola *fraud*. Ada enam atribut/indikator *fraud* dalam transaksi *online* yang teridentifikasi dalam penelitian ini yaitu *same location, relationship,quantity, througput time, skip* dan *wrong pattern*. Tiga atribut fraud yaitu *througput time, skip* dan *wrong pattern* sudah diusulkan oleh penelitian sebelumnya [8],[9], sedangkan tiga atribut berikutnya *same location, relationship*, dan *quantity* diusulkan penelitian ini. Penelitian ini membuktikan bahwa tiga atribut *same location, relationship*, dan *quantity* dapat meningkatkan akurasi metode deteksi *fraud* atau penipuan ini pada transaksi *online*. Untuk memperoleh akurasi deteksi *fraud*, pola yang ada *fraud* diperbaiki sesuai dengan pola baru *fraud*. Metode deteksi *fraud* yang diusulkan ini mampu mendeteksi *fraud* pada transaksi *online* dengan akurat.

Daftar Pustaka

[1] I. Amara, A. B. Amar dan A. Jarboui. Detection of Fraud in Financial Statements: French Companies as a Case Study. *International Journal of Academic Research in Accounting, Finance and Management Sciences.* 2013: 3(3), 44-55.

[2] H. Jung-Woo, P. Hyuna dan K. Jeonghee. *Large-Scale Item Categorization in e-Commerce Using Multiple Recurrent Neural Network*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco. 2016;107-115.

[3] C. Khyati dan M. Bhawna, Credit Card Fraud: Bang in E-Commerce. *International Journal Of Computational Engineering Research* . 2012: 3(2), 935-941.

[4] C. Evandro , B. Gabriel dan P. Adriano C. M. Fraud Analysis and Prevention in e-Commerce Transactions. *IEEE*. 2014;42-49.

[5] M. Jans, M. J. van der Werf, N. Lybaert dan K. Vanhoof. A Business Process Mining Application for Internal Transaction Fraud Mitigation. *Expert Systems with Applications.* 2011. 38(10). 13351-13359.

[6] J. J. Stoop. Process Mining and Fraud Detection. *Thesis*. Netherlands:Twente University; 2012.

[7] R. Sarno, D. R. Dewandono, T. Ahmad, M. F. Naufal dan F. Sinaga. Hybrid Association Rule Learning and Process Mining for Fraud Detection. *IAENG International Journal of Computer Science*. 2015:42(2).59-72.

[8] S. Huda, R. Sarno dan T. Ahmad. Fuzzy MADM approach for Rating of Process-based Fraud. *Journal ICT. Research Application.* 2015: 9(2). 111-128.

[9] S. Huda, R. Sarno dan T. Ahmad. Increasing accuracy of Process-based Fraud Using Behavior Models, *International Journal of Software Engineering and Its Applications*.2016. 10(5). 175-188.

[10] W. M. P. van der Aalst. Discovery, Conformance dan Enhancement of Business Processes. *Springer*. 2010: 7-8.

[11] R. Sarno, P. L. I. Sari. H. Ginardi, D. Sunaryono, dan I. Mukhlash. Decision Mining For Multi Choice Workflow Patterns, *International conference on Computer, Control, and Its Application.* 2013.19-21.

[12] R. Sarno, A.B. Sanjoyo, I. Mukhlash dan M.H. Astuti. Petri Net Model of ERP Business Process Variations for Small and Medium Enterprises. *Journal of Theoretical and Applied Information Technology*. 2013. 54(1). 31-38.

[13] W. M. P. van der Aalst, H.A. Reijers dan M. Song. Discovering Social Networks from Event Logs. *Computer Supported Cooperative Work*. 2005. 14. 549-593.