Analysis of Volvo IT's Closed Problem Management Processes By Using Process Mining Software ProM and Disco

Eyüp Akçetin | Department of Maritime Business Administration, Balıkesir University, Turkey, e.akcetin@gmail.com
Ufuk Çelik | Department of Computer Technologies, Bandırma Onyedi Eylül University, Turkey, ucelik001@gmail.com
Abdulkadir Yaldır | Department of Management Information System, Pamukkale University, Turkey, akyaldir@pau.edu.tr

ABSTRACT	In this study, the event logs for closed problems which was created by Volvo IT Belgium for the Business Process Intelligence Challenge in 2013 are evaluated by using Process Mining software. One of these software (ProM) is an open source program and the other (Disco) is a commercial product. Commercial application is available as a full version with academic license. Observed event logs record consist of 6660 units. This case study described the logic of the process of mining programs and issues in problem solving stage of the Volvo Company are illustrated. In this analysis, ping-pong behavior in the processes, the product which has the most frequency and the employee who has the most duration for the processes have been revealed.
Keywords:	Data mining, business intelligence, process mining, ProM, Disco

ProM ve Disco Süreç Madenciliği Yazılımları Kullanılarak Volvo IT'nin Sonlandırılmış Problem Yönetimi Süreçlerinin Analizi

ÖZET	Bu çalışmada 2013 yılında İş Süreci Zekası yarışması için Volvo IT Belçika tarafından sağlanan çözümü gerçekleştirilmiş problemlerin olay günlükleri Süreç Madenciliği için kullanılan yazılımlar ile değerlendirilmiştir. Bu yazılımlardan birisi (ProM) açık kaynak kodlu, diğeri (Disco) ise ticari bir uygulamadır. Ticari uygulama akademik lisans ile tam versiyon şeklinde kullanılabilir. Çalışmada değerlendirilen olay günlükleri 6660 adet kayıttan oluşmaktadır. Bu örnek inceleme ile süreç madenciliği programların çalışma mantığı anlatılmış ve Volvo firmasının problem çözümü aşamasındaki sorunlar örneklendirilmiştir. Bu analizde, süreçlerdeki ping-pong davranışı, en uzun süreçli ürün, en fazla yoğunluğa sahip çalışan ortaya konulmuştur.
Anahtar Kelimeler:	Veri madenciliği, İş Zekası, Süreç Madenciliği, ProM, Disco

1. Introduction

Process mining is the new area of Computer Science which provides some tools and techniques to extract useful information from event logs in order to discover, monitor and improve business processes. In this manner process mining requires all event logs for a process and an event log requires a structure which is shown in Figure 1 (Kumaraguru, 2013, Van der Aalst et al., 2004).
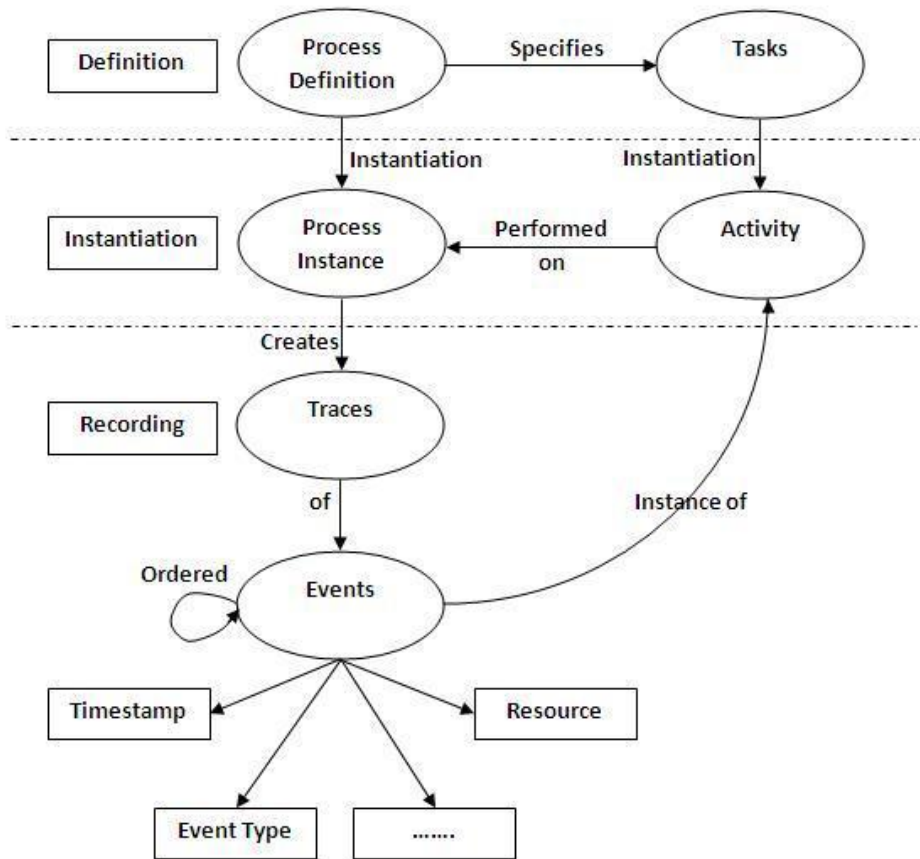

Figure 1. The structure of an event log
Reference: Kumaraguru, P. V. (2013). Machine learning approach for model discovery and process enhancement using process mining techniques Ph.D. Thesis, Dr. M.G.R. Page: 82

Process Mining is the concretization of process intelligence using event logs as a starting point and process intelligence is the combination of business intelligence and business process mining which is shown in Figure 2 (Aufaure and Zimányi, 2013, van der Aalst et al., 2012).

Figure 1. The term of Process Mining

Reference: van der Aalst, W., A. Adriansyah, A. K. A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. van den Brand, R. Brandtjen, J. Buijs, A. Burattin, J. Carmona, M. Castellanos, J. Claes, J. Cook, N. Costantini, F. Curbera, E. Damiani, M. de Leoni, P. Delias, B. F. van Dongen, M. Dumas, S. Dustdar, D. Fahland, D. R. Ferreira, W. Gaaloul, F. van Geffen, S. Goel, C. Günther, A. Guzzo, P. Harmon, A. ter Hofstede, J. Hoogland, J. E. Ingvaldsen, K. Kato, R. Kuhn, A. Kumar, M. La Rosa, F. Maggi, D. Malerba, R. S. Mans, A. Manuel, M. McCreesh, P. Mello, J. Mendling, M. Montali, H. R. Motahari-Nezhad, M. zur Muehlen, J. Munoz-Gama, L. Pontieri, J. Ribeiro, A. Rozinat, H. Seguel Pérez, R. Seguel Pérez, M. Sepúlveda, J. Sinur, P. Soffer, M. Song, A. Sperduti, G. Stilo, C. Stoel, K. Swenson, M. Talamo, W. Tan, C. Turner, J. Vanthienen, G. Varvaressos, E. Verbeek, M. Verdonk, R. Vigo, J. Wang, B. Weber, M. Weidlich, T. Weijters, L. Wen, M. Westergaard and M. Wynn (2012). Process Mining Manifesto. Business Process Management Workshops: BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I. F. Daniel, K. Barkaoui and S. Dustdar: 169-194. Berlin, Heidelberg, Springer Berlin Heidelberg. Page 13.

Process mining provides the missing link between on the one hand process model anlaysis and data-oriented analysis and on the other hand performance and coformance which is shown in Figure 3 (van der Aalst, 2014).
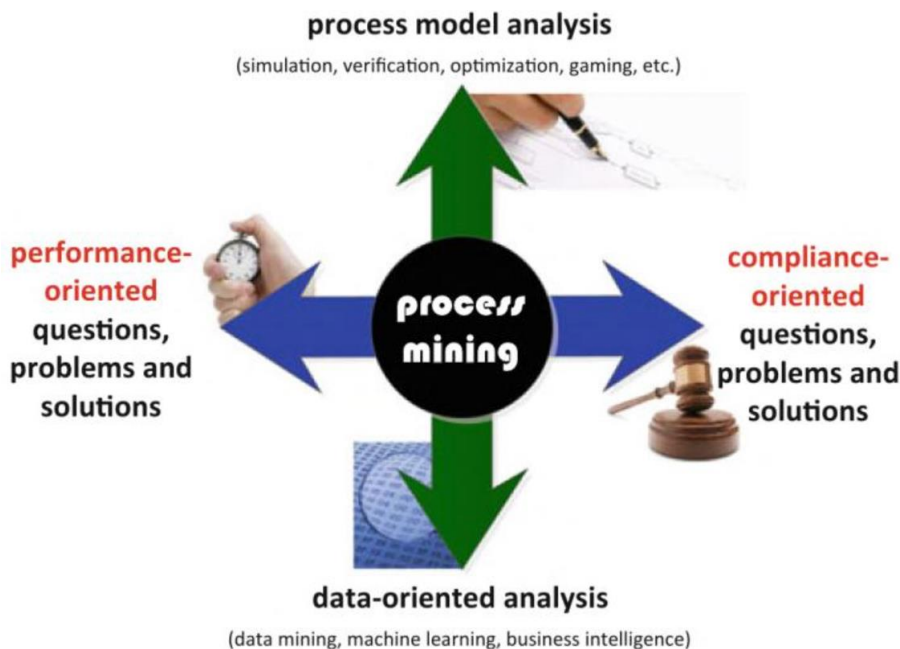

Figure 2. Process Mining stages

Reference: van der Aalst, W. M. P. (2014). Process Mining in the Large: A Tutorial. Business Intelligence: Third European Summer School, eBISS 2013, Dagstuhl Castle, Germany, July 7-12, 2013, Tutorial Lectures. E. Zimányi: 33-76. Cham, Springer International Publishing. Page: 34

Akçetin, Çelik & Yaldır | Analysis of Volvo IT's Closed Problem Management Processes By Using Process Mining Software ProM and Disco

174

2. Literature Review

According to structure an event originates a case i.e. process instance (e.g. buy an ecommerce product) which combines an activity or task (e.g. login to website) with a timestamp (e.g. time at login) by an originator (a person who makes task) (van der Aalst et al., 2005). After the collecting of event logs, process mining is used to discover a model for example by constructing a Petri net (Petri, 1962) which can recreate observed process (van der Aalst et al., 2005). Then process mining performs a conformance checking if modeled process is compatible with observed model (van der Aalst et al., 2007). Finally, it extends the model onto a new initial model (Van der Aalst et al., 2003). Three stages of process mining are shown in figure 4 (Günther, 2009).
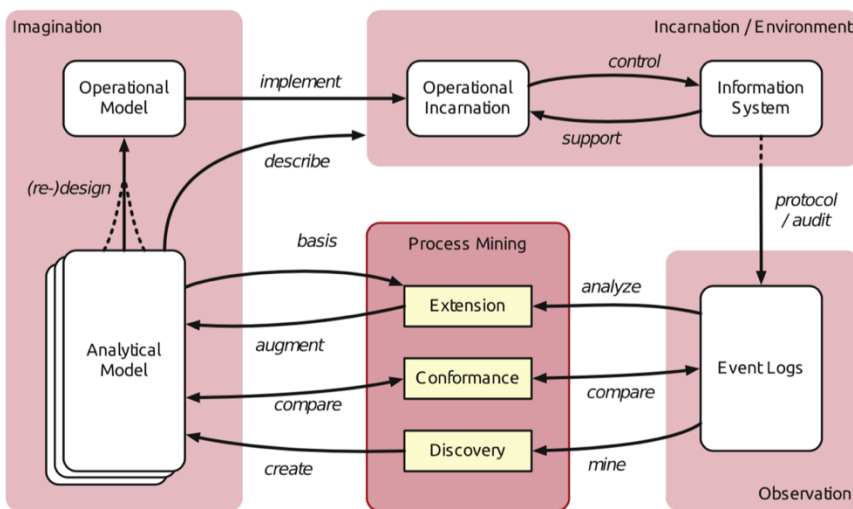

Figure 3. Process mining stages
Reference: Günther, C. W. (2009). Process Mining in Flexible Environments Ph.D. Thesis, Technische Universiteit Eindhoven. Page: 82

Process mining data can be obtained from various software such as Enterprise Resource Planning systems, Business Process Management systems, Product Data Management systems, Electronic Health Record systems or any database systems which stores event logs for process. Discovery starts with a model creation by using any process model such as Petri net (Petri, 1962) from event logs analysis. This model is actually a control-flow system and it may explain the steps of a process as well as organization and perspective (Van Der Aalst, 2011). Conformance makes a comparison between the analytical model and event logs (Van der Aalst et al., 2012). Extension is the improvements of a model by using extracted information from event logs (Burattin, 2015).

3. Materials and Methods

3.1. Aim of This Study

In industrial or organizational psychology, proactive behavior is to act and take control in advance of a future situation rather than being reactive. In this purpose process mining is the major point for being proactive in the context of business intelligence. Therefore, we learned process mining software and evaluated a real life example (it is event logs of Volvo IT closed problems in this study) and criticized the results.

3.2. Tools

We used ProM version 6.5.1 and Disco version 1.9.5 and a spreadsheet application (Excel version 2016) for the process mining.

ProM (which is short for Process Mining framework) is an Open Source framework for process mining algorithms written in JAVA (Aalst et al., 2009). Disco (which is short for Discover your process) is a commercial application for process mining but can be used as full version with an academic license (Rozinat and Günther, 2012).

3.3. Data

In this study, a real-life event log data is used which is provided by Volvo IT Belgium and published in Business Process Intelligence 2013 as a challenge (Volvo, 2013). Data contains the problem management log for closed problems. Data has several attributes and first 10 records is shown in table 1.

Problem Number	Problem Date+Time	Sub Status	Status	Impact	Product	Country	Owner
1-109135791	11.01.2006 15:49	Awaiting Assignment	Queued	High	PROD191	INDIA	Minnie
1-109135791	15.03.2012 11:53	In Progress	Accepted	High	PROD191	INDIA	Minnie
1-109135791	15.03.2012 11:56	Assigned	Accepted	High	PROD191	INDIA	Minnie
1-109135791	15.03.2012 12:09	In Progress	Accepted	High	PROD191	INDIA	Minnie
1-109135791	15.03.2012 12:11	Closed	Completed	High	PROD191	INDIA	Minnie
1-147898401	7.11.2006 10:00	In Progress	Accepted	Medium	PROD753	Sweden	Tomas
1-147898401	7.11.2006 13:05	In Progress	Accepted	Medium	PROD753	Sweden	Tomas
1-147898401	2.12.2009 14:24	Wait	Accepted	Medium	PROD753	Sweden	Tomas
1-147898401	3.09.2011 06:09	In Progress	Accepted	Medium	PROD753	Sweden	Tomas
...	...	...	...	...	...	...	...

Table 1. Volvo problem management event logs
Reference: Volvo. (2013). "Volvo IT Incident", 9th International Workshop on Business Process Intelligence, doi:10.4121/c2c3b154-ab26-4b31-a0e8-8f2350ddac11

Through this analysis 'Problem Number' attribute is used for Case ID, 'Problem Date+Time' attribute is used for Timestamp, 'Problem Status and Sub Status' attributes are used for Activity, 'Product' attribute is used for other and 'Owner' attribute (which is problem involved action owner first name) is used for Resource.

3.4. Analysis and Findings

After importing data in Disco according to those assignments, process analysis created a map for process frequency which is shown in figure 5. The numbers in figure 5 refer to case frequencies and arrows show the direction of process with the thickness according to frequency. The process from 'In Progress//Accepted' to 'Closed//Completed' has the highest density with the 1266 cases.
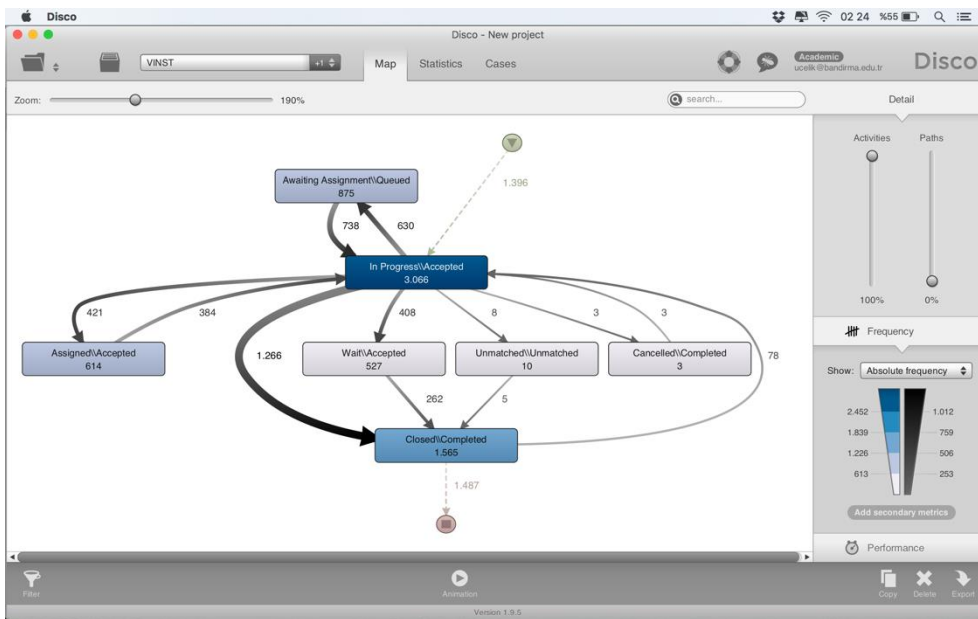
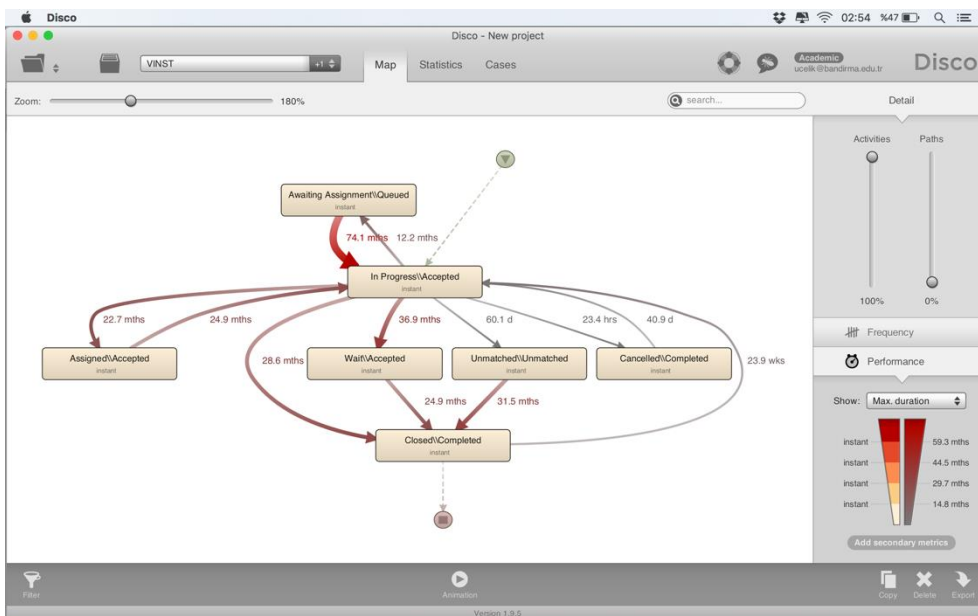**Figure 4.** Process frequency map in Disco software

Figure 5. Performance graph of the process

Process performance presented in figure 6 can be obtained from the 'Performance' section which is on the right bottom location under the 'Frequency' bar which shows absolute frequency in the combo box selection. According to performance graph we gather the total, median, mean, maximum and minimum durations by using combo box selection. From this graph, it is understood that there is the longest duration from 'Awaiting Assignment//Queued' to 'In Progress//Accepted' which makes a bottleneck in the overall process.

Statistical information of processes can be seen on the 'Statistics' tab which is shown in figure 7.
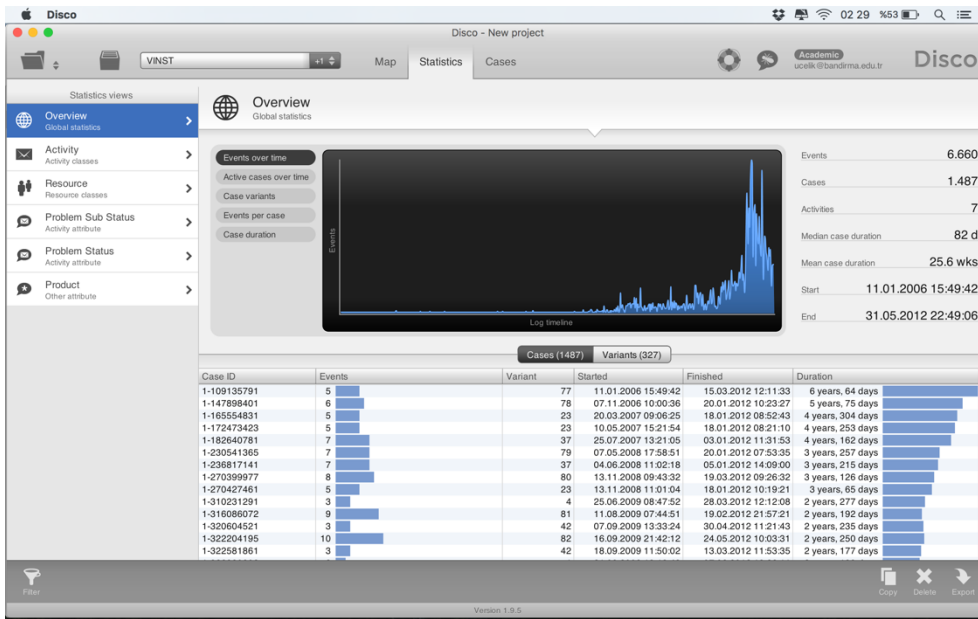
Figure 6. Statistics of process

This section gives an overall information about the process data with the activities, resources, status and product ordering to frequencies. We can gather the longest or shortest activity or status, who used how much time for their process and frequencies of each product. We can also filter the data by using filter option which is on the left bottom corner of the software according to timeframe, variations, performance, endpoints, attributes and endpoints. According to statistics there are 6660 events within 1487 cases with the 7 activities. In the table below the graph events in the cases can be seen with the start and finish time as well as duration. From this observation we learned the Carolyn used the longest time for process. Also the product PROD97 has the longest frequency.

Finally, we can observe all events as variants from 'Cases' tab. This section separates all cases into variants and variants provide a simple and sequential view on the process which is shown in figure 8.

From the created process map in figure 5, we can observe a loop between a problem 'In Progress//Accepted' and 'Assigned//Accepted' dominantly. However, we cannot observe how a single case operates through the process or how many cases loop twice or more. In order to understand this situations, we must check the variants.
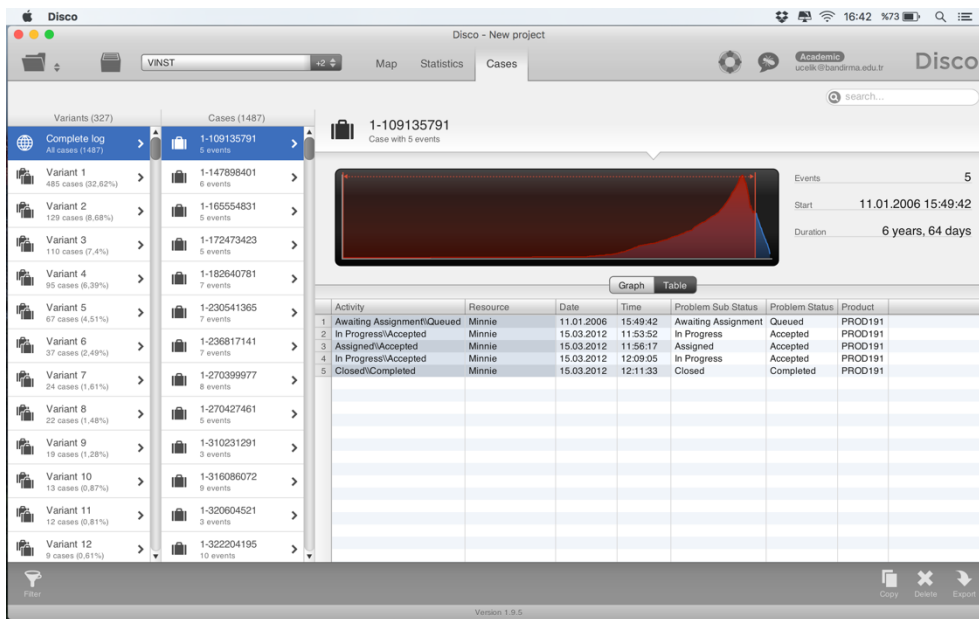
Akçetin, Çelik & Yaldır | Analysis of Volvo IT's Closed Problem Management Processes By Using Process Mining Software ProM and Disco

178

Figure 7. Cases into variants tab

There are 327 variants and 1487 cases in total which is shown in figure 8. For example, we can understand that there is a loop between 'In Progress//Accepted' and 'Assigned//Accepted' for the resource Minnie related to PROD191.

Disco also provides various export options for data to be able evaluate in another program. In process mining most important point to create a standard data format for event logs. So far, this standard has been provided by using MXML format but it has some restrictions. In order to solve these problems a new format defined named XES. It is an XML-based event format and its name come from eXtensible Event Stream (Günther and Verbeek, 2014).

After exporting VINST data from Disco as XES file format it is imported into ProM software. ProM has several algorithms for process mining that Disco cannot provide. For example, we can create a timeline of cases and observe them by using Gantt chart which is shown in figure 9.
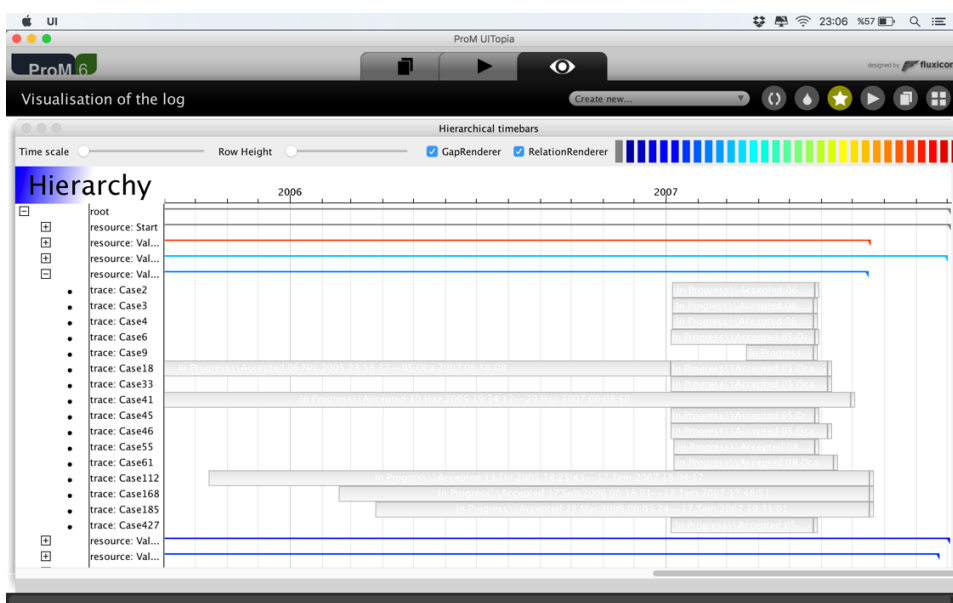

Figure 8. ProM timeline graph of process

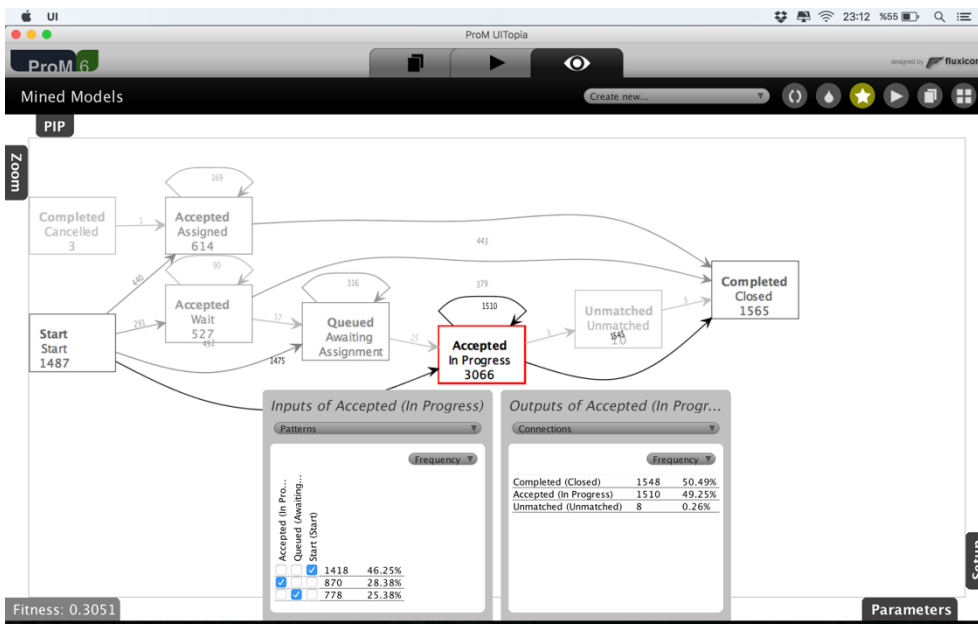ProM also provide Heuristic Miner algorithm which is shown figure 10.

Figure 9. ProM Heuristic Miner graph

If there is a complex and unstructured log data or if a simplified model is needed fuzzy miner algorithm is better to use for evaluation which is shown in figure 11.
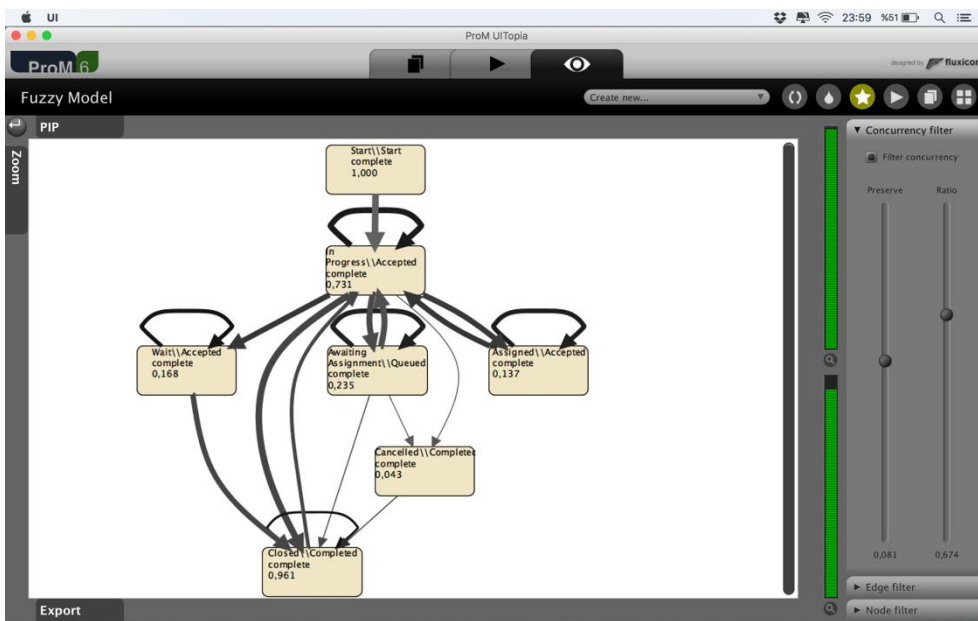
Figure 10. Fuzzy Miner in ProM

4. Results

Preferably, a problem should be closed in a short time with minimum participation of not too many people. But problems were sent repeatedly each other of support teams in Volvo. This circumstance is called 'ping-pong behavior' which actually causes to extend the total duration of a problem solving. For example, event log which is given in table 2 shows a ping-pong behavior for the product PROD425.

Akçetin, Çelik & Yaldır | Analysis of Volvo IT's Closed Problem Management Processes By Using Process Mining Software ProM and Disco

180

Problem Date+Time	Problem Sub Status	Problem Status	Problem Involved ST Function Div	Problem Involved Owner ST	Problem Involved Action Owner Name
11.08.2009 07:44	In Progress	Accepted	E_10	G260 2nd	Per
11.08.2009 07:55	Awaiting Assignment	Queued	E_10	G260 2nd	Per
11.08.2009 11:01	Assigned	Accepted	E_10	G260 2nd	Per
18.05.2010 11:35	Awaiting Assignment	Queued	E_10	G260 2nd	Per
16.06.2010 11:00	In Progress	Accepted	C_6	G152 3rd	Tony
1.10.2011 21:40	Wait	Accepted	C_6	G152 3rd	Tony
7.12.2011 12:41	Awaiting Assignment	Queued	E_10	G260 2nd	Tony
15.02.2012 12:01	In Progress	Accepted	E_10	G150 2nd	Reza
19.02.2012 21:57	Closed	Completed	E_10	G150 2nd	Reza

Table 2. Ping-Pong Behavior of a Problem

For this problem's action owners were changed through three person involved with the same product in some same activities. The person Per has queued the problem two times for the awaiting assignment. Also problem's support team were changed from E_10 to C_6 and it was changed to E_10 again thus causing the involved person changing as well as problem duration.

Involved ST	Organization	Function Div	Involved ST	Organization	Function Div
G138 2nd	Org line C	E_10	G263 2nd	Org line	E_10
G141 3rd	Org line C	C_6	G271 2nd	Org line	E_10
G152 end	Org line C	E_7	G273 3rd	Org line	C_6
G157 2nd	Org line C	E_7	G288 2nd	Org line	E_10
G165 2nd	Org line C	E_7	G290 3rd	Org line	C_6
G167 2nd	Org line A2	A2_2	G55 2nd	Org line	E_1
G181 2nd	Org line C	E_8	G56 3rd	Org line	C_1
G186 2nd	Org line C	E_8	G88 2nd	Org line	E_4
G21 2nd	Org line A2	A2_2	G92	Org line	E_5
G230 2nd	Org line B	E_10	M1 2nd	Org line	A2_2

Table 3. Responsibilities of the most of the ping pong behaviors

Reference: Kang, C. J., Y. S. Kang, Y. S. Lee, S. Noh, H. C. Kim, W. C. Lim, J. Kim and R. Hong (2013). Process Mining-based Understanding and Analysis of Volvo IT's Incident and Problem Management Processes. BPIC@ BPM. Page 13.

Support teams, organizations, and function divisions which have responsibility to most of the ping pong behaviors are given in table 3 (Kang et al., 2013).

From the observation we obtained the organizations for 'Wait//Accepted' problems which is given in table 4.

Value	Frequency	Relative frequency
Org line A2	235	44,59%
Org line C	214	40,61%
Org line G4	50	9,49%
Org line B	10	1,90%
Org line V2	7	1,33%
Org line G3	4	0,76%
Org line F	3	0,57%
Org line G1	2	0,38%
Org line V11	1	0,19%
Org line D	1	0,19%

Table 4. Organizations for 'Wait//Accepted' Problems

5. CONCLUSION

In this study, we observed the real life event logs supported by Volvo IT's VINST system for Business Process Intelligence Challenge 2013. The log includes 6660 cases for the closed problems.

We used two software (ProM and Disco) for the process mining and explained the overall usage of those applications for mining as well as data import and export. We evaluated the logs in the context of ping-pong behavior and who is responsible for problem 'Wait//Accepted' as well as bottleneck of the process performance. We found the longest duration activity that is from 'Awaiting Assignment//Queued' to 'In Progress//Accepted' which makes a bottleneck in the overall process. We gathered another bottleneck which is a person named Carolyn who involved the problem 'Awaiting Assignment'. In this manner the person who involved with the problem is the major factor in process analysis because that person is responsible for any incidents in the organizations. Sometimes people's identifications are kept secret because of the privacy. But it causes misinformation for the process. For example, Ministry of the Health started a new project in Turkey (Health, 2014). It is about the error notification system for health employments in government hospital. But they decided to keep employments' identification because of the person privacy (HaberTurk, 2016). It is a remarkable project for medical process but keeping secret of the personal's identification avoid to gather most important data for the process mining. We also identified the product which has the highest frequency.

To sum up, it is understood that process mining is definitely a key factor for the Business Intelligence. Also ProM and Disco quite powerful applications for the process mining which generate many useful graphics and statistical information. In the future, we are going to evaluate a real life example from Turkey and share the results within another publication.

References

Aalst, W. v. d., P. v. d. Brand, B. v. Dongen, C. W. Günther and E. Verbeek. (2009). "ProM (Process Mining Framework)", http://www.processmining.org/prom/start,(2016

Aufaure, M.-A. and E. Zimányi (2013). Business Intelligence: Second European Summer School, EBISS 2012, Brussels, Belgium, July 15-21, 2012, Tutorial Lectures, Springer.

Burattin, A. (2015). Process Mining Techniques in Business Environments, Springer.

Günther, C. W. (2009). Process Mining in Flexible Environments Ph.D. Thesis, Technische Universiteit Eindhoven.

Günther, C. W. and E. Verbeek. (2014). "XES Standard Definiton Version 2.0", http://www.xes-standard.org/_media/xes/xesstandarddefinition-2.0.pdf,(2016

HaberTurk. (2016). "Sağlık Bakanlığı, hata bildirim sistemi kurdu", http://www.haberturk.com/saglik/haber/1217197-saglik-bakanligi-hata-bildirim-sistemi-kurdu,(2016

Health, T. M. o. (2014). "Hata Sınıflandırma Sistemleri Hakkında", https://kalite.saglik.gov.tr/index.php?lang=tr&page=46&newsCat=1&newsID=1128,(2016

Kang, C. J., Y. S. Kang, Y. S. Lee, S. Noh, H. C. Kim, W. C. Lim, J. Kim and R. Hong (2013). Process Mining-based Understanding and Analysis of Volvo IT's Incident and Problem Management Processes. BPIC@ BPM.

Akçetin, Çelik & Yaldır | Analysis of Volvo IT's Closed Problem Management Processes By Using Process Mining Software ProM and Disco

182

Kumaraguru, P. V. (2013). Machine learning approach for model discovery and process enhancement using process mining techniques Ph.D. Thesis, Dr. M.G.R. Educational and Research Institute.

Petri, C. A. (1962). Kommunikation mit Automaten, Rhein.-Westfäl. Inst. f. Instrumentelle Mathematik an der Univ. Bonn.

Rozinat, A. and C. W. Günther. (2012). "Disco", https://fluxicon.com/disco/,(2016

Van Der Aalst, W. (2011). Process mining: discovery, conformance and enhancement of business processes, Springer Science & Business Media.

Van der Aalst, W., A. Adriansyah, A. K. A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. van den Brand, R. Brandtjen, J. Buijs, A. Burattin, J. Carmona, M. Castellanos, J. Claes, J. Cook, N. Costantini, F. Curbera, E. Damiani, M. de Leoni, P. Delias, B. F. van Dongen, M. Dumas, S. Dustdar, D. Fahland, D. R. Ferreira, W. Gaaloul, F. van Geffen, S. Goel, C. Günther, A. Guzzo, P. Harmon, A. ter Hofstede, J. Hoogland, J. E. Ingvaldsen, K. Kato, R. Kuhn, A. Kumar, M. La Rosa, F. Maggi, D. Malerba, R. S. Mans, A. Manuel, M. McCreesh, P. Mello, J. Mendling, M. Montali, H. R. Motahari-Nezhad, M. zur Muehlen, J. Munoz-Gama, L. Pontieri, J. Ribeiro, A. Rozinat, H. Seguel Pérez, R. Seguel Pérez, M. Sepúlveda, J. Sinur, P. Soffer, M. Song, A. Sperduti, G. Stilo, C. Stoel, K. Swenson, M. Talamo, W. Tan, C. Turner, J. Vanthienen, G. Varvaressos, E. Verbeek, M. Verdonk, R. Vigo, J. Wang, B. Weber, M. Weidlich, T. Weijters, L. Wen, M. Westergaard and M. Wynn (2012). Process Mining Manifesto. Business Process Management Workshops: BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I. F. Daniel, K. Barkaoui and S. Dustdar: 169-194. Berlin, Heidelberg, Springer Berlin Heidelberg.

Van der Aalst, W., A. Adriansyah and B. van Dongen (2012). Replaying history on process models for conformance checking and performance analysis. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2(2): 182-192.

Van der Aalst, W., T. Weijters and L. Maruster (2004). Workflow mining: Discovering process models from event logs. Knowledge and Data Engineering, IEEE Transactions on 16(9): 1128-1142.

Van der Aalst, W. M., B. F. van Dongen, J. Herbst, L. Maruster, G. Schimm and A. J. Weijters (2003). Workflow mining: a survey of issues and approaches. Data & knowledge engineering 47(2): 237-267.

Van der Aalst, W. M. P. (2014). Process Mining in the Large: A Tutorial. Business Intelligence: Third European Summer School, eBISS 2013, Dagstuhl Castle, Germany, July 7-12, 2013, Tutorial Lectures. E. Zimányi: 33-76. Cham, Springer International Publishing.

Van der Aalst, W. M. P., H. T. de Beer and B. F. van Dongen (2005). Process Mining and Verification of Properties: An Approach Based on Temporal Logic. On the Move to Meaningful Internet Systems 2005: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2005, Agia Napa, Cyprus, October 31 - November 4, 2005, Proceedings, Part I. R. Meersman and Z. Tari: 130-147. Berlin, Heidelberg, Springer Berlin Heidelberg.

Van der Aalst, W. M. P., A. K. A. de Medeiros and A. J. M. M. Weijters (2005). Genetic Process Mining. Applications and Theory of Petri Nets 2005: 26th International Conference, ICATPN 2005, Miami, USA, June 20-25, 2005. Proceedings. G. Ciardo and P. Darondeau: 48-69. Berlin, Heidelberg, Springer Berlin Heidelberg.

Van der Aalst, W. M. P., H. A. Reijers, A. J. M. M. Weijters, B. F. van Dongen, A. K. Alves de Medeiros, M. Song and H. M. W. Verbeek (2007). Business process mining: An industrial application. Information Systems 32(5): 713-732.

Volvo. (2013). "Volvo IT Incident" 9th International Workshop on Business Process Intelligence, http://www.win.tue.nl/bpi/doku.php?id=2013:challenge,(2016