

Process Mining Tools Comparison

Ufuk ÇELİK, Assist. Prof., Bandırma Onyedi Eylül University, O.S. Applied Science Faculty, Management Information System Department, ucelik@bandirma.edu.tr

Eyüp AKÇETİN, Assoc. Prof., Muğla Sıtkı Kocman University, Seydikemer Applied Science School, e.akcetin@gmail.com

ABSTRACT *Process-mining analysis related with business intelligence, is a new idea in the science of data mining. This analysis is made by using event logs and it provides a knowledge about a general process. It compares each process with others in the terms of time in order to discover bottlenecks. To do that it is also compares the responsible people who deal with the process. Event logs that are used in this analysis consist of large data. Because all the records that occur during short time intervals are kept in event logs. In order to examine such a big data specific programs are used. These programs generate a process map at first then bottlenecks in the process are discovered after that. In this manner, processes are improved by discovering and monitoring the all the information such as event ID, activity, time and responsible person. In this study, the tools named ProM, Disco, Celonis and My-Invenio used in process mining were examined and their performance according to usage features compared. According to the obtained results, the usefulness, performance and reporting features of the software used in a process analysis are revealed.*

Keywords: Data Mining, Business Intelligence, Process Mining.

Süreç Madenciliği Araçları Karşılaştırması

ÖZ *İş zekâsı ile ilgili süreç-madenciliği analizi, veri madenciliği biliminde yeni bir fikirdir. Bu analiz olay günlükleri kullanılarak yapılır ve genel bir süreç hakkında bilgi sağlar. Darboğazları keşfetmek için her süreci zamanla diğerleriyle karşılaştırır. Bunu yapmak için aynı zamanda süreçle ilgilenen sorumluları karşılaştırır. Bu analizde kullanılan olay günlükleri büyük verilerden oluşmaktadır. Çünkü kısa zaman aralıklarında meydana gelen tüm kayıtlar olay kayıtlarında tutulur. Bu kadar büyük bir veriyi incelemek için belirli programlar kullanılmaktadır. Bu programlar önce bir süreç haritası oluştururlar, daha sonra süreçteki darboğazlar bundan sonra keşfedilir. Bu şekilde olay kimliği, aktivite, zaman ve sorumlu kişi gibi tüm bilgileri keşfedip izleyerek süreçler geliştirilir. Bu çalışmada, ProM, Disco, Celonis ve My-Invenio adlı süreçlerde kullanılan süreç madenciliğine bakılmış ve kullanım özelliklerine göre performansları karşılaştırılmıştır. Elde edilen sonuçlara göre, bir süreç analizinde kullanılan yazılımın kullanışlılığı, performans ve raporlama özellikleri ortaya konmuştur.*

Anahtar Kelimeler: Veri Madenciliği, İş Zekası, Süreç Madenciliği.

1. INTRODUCTION

Business Intelligence (BI) is extensive collection of tools and technique that use information in order to provide decision-making. Business process intelligence (BPI) is a branch of BI focusing on business process management (BPM). BMP is a field that combines information technology and management sciences and applies both to operational business processes. Process mining can be positioned under the umbrella of BI and (BPI) that can be seen in figure 1.

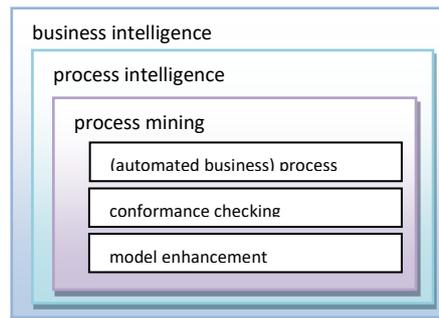


Figure 1. The term of process mining (W. Van Der Aalst et al., 2012)

Process mining is a bridge between the process science and data science, therefore it is an essential technique for ambitious and fast-growing manufacturing organizations that operated according to Industry 4.0 (W. Van Der Aalst & Damiani, 2015). It is a new type of Big Data Analytics (Rinke, 2017).

Business processes contain large amount of events recorded by information systems. These data is a collection of information such as event ID, activity, timestamp and responsible person and is called "Event Logs" (W. Van Der Aalst, 2011) and the structure of an event is shown in figure 2. Process mining is a rising discipline in Management Information Systems as well as Computer Science that is based on model-driven methods within the data mining techniques that also considers complex business processes (W. Van Der Aalst et al., 2003). It aims to understand as-is process from observed system behavior by deriving process models (Helm & Paster, 2015).

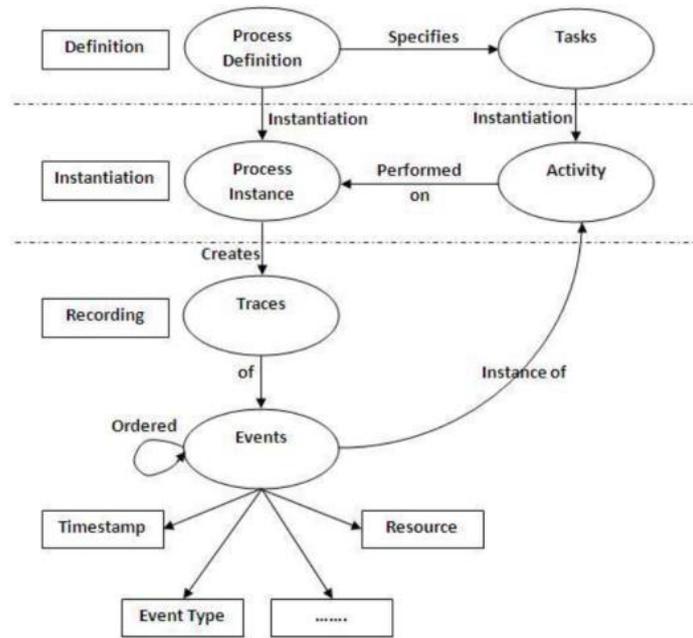


Figure 2. The structure of an event log (Kumaraguru & Rajagopalan, 2013)

An event is the origin of a case i.e. processes instance (e.g. transfer money from a bank account) which consists of an activity (e.g. login to website of a bank) within a timestamp (e.g. the duration of bank's website usage from login to logout) by an originator (a person who makes task) (W. Van Der Aalst, de Beer, & van Dongen, 2005). Process mining is used to discover a model by constructing a Petri net (C Adam Petri & Reisig, 2008; Carl Adam Petri, 1962) from observed process after the collection of all event logs (W. Van Der Aalst, de Medeiros, & Weijters, 2005). Process mining makes a conformance checking in order to proof that observed model is compatible with modeled process (W. Van Der Aalst et al., 2007). Finally, it extends the model onto new initial model. These stages are shown in figure 3.

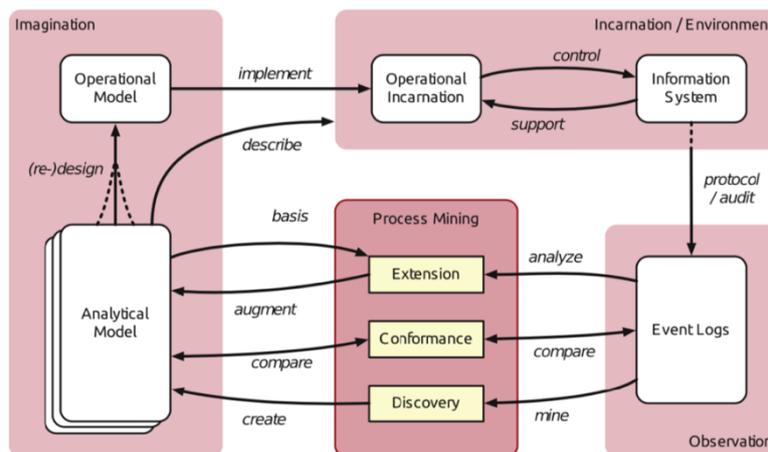


Figure 3. Process mining stages (Günther, 2009)

All these event logs can be obtained from various data sources such as databases, ERP or CRM system and log records. Finally, there are four software for process-mining namely ProM, Disco, Celonis and My-Invenio that we compared in this study. (Celonis, 2017;

Fluxicon, 2017; “My-Invenio,” 2017; “Process Mining,” 2017; “ProM 6,” 2017; Anne Rozinat & Günther, 2012)

Tablo 1. A fragment of event log

Case ID	Activity	Start Date	End Date	Product	Service Type	Resource
Case 1	Inbound Call	9.3.10 8:05	9.3.10 8:10	MacBook Pro	Referred to Servicer	Helen
Case 1	Handle Case	11.3.10 10:30	11.3.10 10:32	MacBook Pro	Referred to Servicer	Helen
Case 1	Call Outbound	11.3.10 11:45	11.3.10 11:52	MacBook Pro	Referred to Servicer	Henk
Case 2	Inbound Call	4.3.10 11:43	4.3.10 11:46	MacBook Pro	Referred to Servicer	Susi
Case 3	Inbound Call	25.3.10 9:32	25.3.10 9:33	MacBook Pro	Referred to Servicer	Mary
Case 4	Inbound Call	6.3.10 11:41	6.3.10 11:51	iPhone	Referred to Servicer	Fred
Case 5	Inbound Call	18.3.10 10:54	18.3.10 11:01	MacBook Pro	Product Assistance	Kenny
Case 6	Inbound Call	25.3.10 17:09	25.3.10 17:13	MacBook Pro	Referred to Servicer	Harold

2. PROCESS MINING TOOLS

In this study, we intent to bridge the gap between the software developer and the potential users by providing a test framework for process mining (Devi, 2006). This framework is based on a list of business process problems refers to related process mining methods. In addition, we aim to compare and analyze these process mining software (ProM, Disco, Celonis and My-Invenio) using the framework.

One of the tools that we compared is ProM and it is an open source software maintained with Java so it has platform independence. Will van der Aalst and his research group develop proM at Eindhoven University of Technology.

Disco is a commercial tool developed by Fluxicon but it has an academic license option with full support. It runs on Windows or Mac OsX.

Celonis is another commercial tool but has academic license option. It runs on Windows or Mac OsX.

My-Invenio is also commercial product with academic license option. However, it is a web-based system. So it can be accessible from any device (mobile, tablet, computer) that has browser and active internet connection.

3. FINDINGS

The capabilities of process mining tools are subject to some fields that are listed in Table 2 as features (Kebede, 2015). Process mining tools should support the filtering data in order the avoid malfunction of software. Obviously, these programs must implement the process discovery and provide some models. Therefore, models that are generated by software can be investigated for conformance checking. Conformance checking is a plug-in, which controls the process model from two point of view. First point is the exact model in real situation and second point is how to improve this model in further application (A. Rozinat & Van Der Aalst, 2008; W. Van Der Aalst, 2011; W. Van der Aalst, Adriansyah, & van Dongen, 2012). Some process mining tools discover the relationships between the resources by using

social network mining (W. Van Der Aalst et al., 2007). If a process mining tool uses decision rule mining capability, it is possible to enhance models by discovering some useful patterns between the activities (A Rozinat & van der Aalst, 2006). Process visualization is an animation of process executions in various form (Der, Van Aalst, de Leoni, & ter Hofstede, 2011). Another feature of process mining is to measure process performance according to process model elements through the some actions such as bottleneck detection, cycle or flow time analysis etc. Discriminative rule mining indicates process deviation, which cases deviate, what is wrong (Bernardi, Cimitile, Di Francescomarino, & Maggi, 2014). Trace clustering is an operation that enables to split unstructured process into homogeneous subsets in order to create a process model for each subset (Song, Günther, & van der Aalst, 2009). Delta analysis (W. M. P. van der Aalst, 2005) compares the generated model and reference model in order to answer the problems related to business alignment. In the view of these features, process mining tools we compared here were presented in Table 2.

Table 2. Comparison of process mining tools

Features	ProM (6.5.1)	Disco (1.9.5)	Celonis	My-Invenio
Import type support	MXML, XES	CSV, XLS, MXML, XES, FXL	CSV, XLS	CSV, XES
Import log size capacity	unlimited	up to 5 million events	Based on the quota of database	30 MegaByte
License	Open source	Evaluation, Academic, Commercial	Evaluation, Academic, Commercial	Evaluation, Academic, Commercial
Output model notation	BPMN, WF, Petri nets, ECPs, transition systems, heuristics	Fuzzy model	Fuzzy model and Support of charts	BPMN, SVG, CSV, XPDL, Activity Map, Social Network
Supported platform	Standalone desktop version	Standalone desktop version	On premise, Software as a service version, Web based	Web based only
Filtering data	Yes	Yes	Yes	Yes
Process discovery	Yes	Yes	Yes	Yes
Conformance checking	Yes	No	Yes	Yes
Social network mining	Yes	No	Yes	Yes
Decision rule mining	Yes	No	No	No
Process visualization	Yes	Yes	Yes	Yes
Performance reporting	Yes	Yes	Yes	Yes
Discriminative rule mining	Yes	No	No	No
Trace clustering	Yes	No	No	No
Delta analysis	Yes	Yes	Yes	Yes

According to comparison result, ProM has the wide variety of features and it is the best tools for process mining. Nevertheless, it has different user interface, which is not easy for beginners. Other tools are easy to use and they have successful user interface. Disco and My-Invenio are much more similar to each to other. On the other hand Celonis has its unique design and supports for a custom designer like My-Invenio. Disco offers a guided data import. In this manner, My-Invenio has a problem with data size. It can import up to 30 MB

of file size for data. ProM is the only one open source tool. In addition, it supports the rule mining and trace clustering whereas other tools does not. Trace clustering is a technique, which makes clusters from unstructured processes and creates process models from each cluster. Social network mining discovers relationships between resources. Conformance checking is another important feature of process mining. It is about a comparison of a process model with an event log of the same process (A. Rozinat & Van Der Aalst, 2008). This technique is supported by ProM, My-Invenio and Celonis with web version named Celonis PI. Celonis PI that is web-based edition of Celonis software uses advanced artificial intelligence and machine learning capabilities.

4. DISCUSSION AND CONCLUSIONS

The main goal of this study was to compare process-mining software such as ProM, Disco, Celonis and My-Invenio. In order to evaluate problem, we tried all the tools with some data sources. The results showed that ProM is the indispensable tools among them. Because it is open source, expandable and has all the features that you need for process mining.

Although ProM has the best solution, it has user interface problem. It is not easy to use as much as any other tools. That is why to use a software rather than ProM firstly and export the model into the ProM is the best way of the process mining. For example, Disco has very easy to use user interface and anybody easily understand how the process mining is evaluated.

My-Invenio is the web based tool and it can be reachable from anywhere but it has limitation of file size and it support upto 30 MB.

Celonis is a rapidly fast-developed tool and it has another version of process mining system named Celonis PI. These tools can be reached from a web site and uses artificial intelligence.

REFERENCES

- Bernardi, M. L., Cimitile, M., Di Francescomarino, C., & Maggi, F. M. (2014). Using Discriminative Rule Mining to Discover Declarative Process Models with Non-atomic Activities BT - Rules on the Web. From Theory to Applications. In A. Bikakis, P. Fodor, & D. Roman (Eds.) (pp. 281–295). Cham: Springer International Publishing.
- Celonis. (2017). Celonis Process Mining. Retrieved October 1, 2017, from <http://www.celonis.com/en/>
- Der, Van Aalst, W. M. P., de Leoni, M., & ter Hofstede, A. H. M. (2011). Process mining and visual analytics: Breathing life into business process models. BPM Center Report BPM-11-15, BPMcenter. Org, 17, 699–730.
- Devi, A. T. (2006). An informative and comparative study of process mining tools. *Int. J. Sci. Eng. Res*, 8(5), 8–10.
- Fluxicon. (2017). Disco. Retrieved October 1, 2017, from <https://fluxicon.com/disco/>

- Günther, C. W. (2009). *Process Mining in Flexible Environments*. Technische Universiteit Eindhoven. Technische Universiteit Eindhoven, Eindhoven University of Technology Library.
- Helm, E., & Paster, F. (2015). First Steps Towards Process Mining in Distributed Health Information Systems. *International Journal of Electronics and Telecommunications*. <https://doi.org/10.1515/eletel-2015-0017>
- Kebede, M. (2015). *Comparative Evaluation of Process Mining Tools*. University of Tartu.
- Kumaraguru, P. V., & Rajagopalan, D. S. P. (2013). Machine learning approach for model discovery and process enhancement using process mining techniques. Department of Computer Science. Dr. M.G.R. Educational and Research Institute, India.
- My-Invenio. (2017). Retrieved October 1, 2017, from <https://www.my-invenio.com>
- Petri, C. A. (1962). *Kommunikation mit Automaten*. Hamburg.
- Petri, C. A., & Reisig, W. (2008). Petri net. *Scholarpedia*, 3, 6477.
- Process Mining. (2017). Retrieved October 1, 2017, from <http://www.processmining.org>
- ProM 6. (2017). Retrieved October 1, 2017, from <http://www.promtools.org/>
- Rinke, A. (2017). Why Manufacturers Need Process Mining – A New Type Of Big Data Analytics.
- Rozinat, A., & Günther, C. W. (2012). Disco. Retrieved from <https://fluxicon.com/disco/>
- Rozinat, A., & Van Der Aalst, W. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems*, 33(1), 64–95. <https://doi.org/10.1016/j.is.2007.07.001>
- Rozinat, A., & van der Aalst, W. M. P. (2006). Decision Mining in ProM BT - Business Process Management. In S. Dustdar, J. L. Fiadeiro, & A. P. Sheth (Eds.) (pp. 420–425). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Song, M., Günther, C. W., & van der Aalst, W. M. P. (2009). Trace Clustering in Process Mining BT - Business Process Management Workshops. In D. Ardagna, M. Mecella, & J. Yang (Eds.) (pp. 109–120). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Van Der Aalst, W. (2011). *Process mining: discovery, conformance and enhancement of business processes*. Springer Science & Business Media.
- Van Der Aalst, W., Adriansyah, A., de Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., ... Wynn, M. (2012). *Process Mining Manifesto*. In F. Daniel, K. Barkaoui, & S. Dustdar (Eds.), *Business Process Management Workshops: BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I* (pp. 169–194). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-28108-2_19
- Van der Aalst, W., Adriansyah, A., & van Dongen, B. (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., & Damiani, E. (2015). Processes Meet Big Data: Connecting Data Science with Process Science. *IEEE Transactions on Services Computing*, 8(6), 810–819. <https://doi.org/10.1109/TSC.2015.2493732>
- Van Der Aalst, W., de Beer, H. T., & van Dongen, B. F. (2005). Process Mining and Verification of Properties: An Approach Based on Temporal Logic. In R. Meersman & Z. Tari (Eds.), *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* (pp. 130–147). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/11575771_11

- Van Der Aalst, W., de Medeiros, A. K. A., & Weijters, A. J. M. M. (2005). Genetic Process Mining. In G. Ciardo & P. Darondeau (Eds.), *Applications and Theory of Petri Nets 2005: 26th International Conference, ICATPN 2005, Miami, USA, June 20-25, 2005. Proceedings* (pp. 48–69). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/11494744_5
- Van Der Aalst, W., Dongen, B. van, Herbst, J., Maruster, L., Schimm, G., & Weijters, A. J. M. M. (2003). Workflow mining: a survey of issues and approaches. *Data & Knowledge Engineering*, 47(2), 237–267.
- van der Aalst, W. M. P. (2005). Business alignment: using process mining as a tool for Delta analysis and conformance testing. *Requirements Engineering*, 10, 198–211.
- Van Der Aalst, W., Reijers, H. A., Weijters, A. J. M. M., van Dongen, B. F., Alves de Medeiros, A. K., Song, M., & Verbeek, H. M. W. (2007). Business process mining: An industrial application. *Information Systems*, 32(5), 713–732. <https://doi.org/http://dx.doi.org/10.1016/j.is.2006.05.003>