Advances in Computer Science and Information Technology (ACSIT)
Print ISSN: 2393-9907; Online ISSN: 2393-9915; Volume 1, Number 2; November, 2014 pp. 26-29

http://www.krishisanskriti.org/acsit.html

Process Mining Tools: A comparative Analysis and Review

Neha Agarwal[1] and Leena Singh[2]

[1]M.Tech(SE) SRMSCET Bareilly(U.P.)*
[2]Department of CS Astt. Prof SRMSCET Bareilly (U.P.)
E-mail: [1]neha.4eva20@gmail.com, [2]leena_12_2001@yahoo.com

Abstract: *This paper presents a comparison of a number of business process mining tools currently available. A number of commercially business process mining tools are available but which tool should be used for which purpose that is explained in this paper. Technologies such as workflow management (WFM), enterprise application integration (EAI), enterprise resource planning (ERP), and web services (WS) typically focus on the realization of IT support rather than monitoring the operational business processes. Process mining aims at extracting information from event logs to capture the business process as it is being executed. The goal of process mining is to extract process- related information, e.g., to automatically discover a process model by observing events recorded by some enterprise system. In this paper, we give a comparison of a number of commercial business process mining tools available on the basis of Type of business process mining, Process model, Process mining problem, Algorithm and Additional features.*

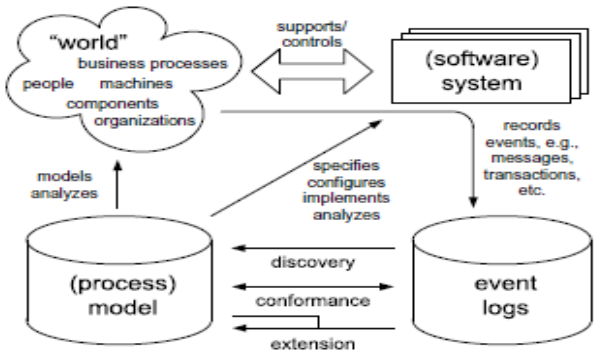
Keywords*: process mining, event logs.*

1. INTRODUCTION

Process mining is a process management technique that allows for the analysis of business processes based on event logs. The basic idea is to extract knowledge from event logs recorded by an information system. During the last decade explicit process concepts (e.g., workflow models) [2, 4, 17, 30, 32, 33, 36, 59] have been applied in many enterprise information systems. These concepts are also playing a major role in cross- organizational processes; cf. the work on web services composition languages such as BPEL4WS and BPML [13, 9]. Workflow management (WFM) systems such as Staff ware, IBM MQSeries, COSA, and etc. offer generic modeling and enactment capabilities for structured business processes. By making graphical process definitions, i.e., models describing the life-cycle of a typical case (process instance) in isolation, one can conFig. these systems to support business processes. Besides pure WFM systems many other software systems use explicit process models. Consider for example enterprise resource planning (ERP) systems such as SAP, PeopleSoft, Baan and Oracle, customer relationship management (CRM) software, etc. Although enterprise information systems are increasingly ‘‘process aware’’ and there are generic languages and tools (e.g., WFM languages and systems), little attention is devoted to process monitoring and improvement. Note that typically WFM systems do not provide functionality to diagnose the running workflow. Process mining offers a way to distill process models from event logs originating from transactional systems in logistics, banking, e-business, health-care, etc. The algorithms used for process mining are complex and in practice large logs are needed to derive a high-quality process model. To support these efforts, the process mining tools have been built.

2. PROCESS MINING

Now day’s process mining is an important phenomenon. Process mining is a technique that allows for the analysis of business process based on event logs. The basic idea behind this is to extract information from event logs recorded by an information system. It is used to fetch important and useful information from event logs. This information in the form of business process models can be extracted by process mining algorithm. Process mining is closely related to BAM (business activity monitoring), BPI (business process intelligence), BOM (business operations management) and Data/workflow mining. [1]. Process mining provides a new means to improve processes in a variety application area. In process mining technique more and more events are recorded thus providing detailed knowledge or information about the process.



Process mining is a relatively young research discipline that sits between computational intelligence and data mining on the one hand, and process modeling and analysis on the other hand. The idea of process mining is to discover, monitor and improve real processes There are three classes of process mining techniques. This classification is based on whether there is a prior model and, if so, how it is used.

(a) Discovery, (b) Conformance(c) extension [2]

- *Discovery*: There is no *a priori* model, i.e., based on an event log some model is constructed a process model can be discovered based on low-level events. For example, using the alpha algorithm, which is a didactically driven approach, where the authors state the lack of analytic capability for large event data volumes with such simple method.[3] There exist many techniques to automatically construct process models (e.g., in terms of a Petri net) based some event log.[3][4][5][6][7] Recently, process mining research also started to target the other perspectives (e.g., data, resources, time, etc.). For example, the technique described in (Aalst, Reijers, & Song, 2005)[8] can be used to construct a social network.
- *Conformance analysis*: There is an *a priori* model. This model is compared with the event log and discrepancies between the log and the model are analyzed. For example, there may be a process model indicating that purchase orders of more than 1 million euro require two checks. Another example is the checking of the so-called "four-eye" principle. Conformance checking may be used to detect deviations to enrich the model. An example is the extension of a process model with performance data, i.e., some *a priori* process model is used to project the bottlenecks on. Another example is the decision miner described in (Rozinat & Aalst, 2006b) [9] which takes an *a priori* process model and analyzes every choice in the process model. For each choice the event log is consulted to see which information is typically available the moment the choice is made. Then classical data mining techniques are used to see which data elements influence the choice. As a result, a decision tree is generated for each choice in the process.
- *Extension*: There is a prior model also. This model is extended with a new aspect or perspective, i.e., the goal is not to check conformance but to enrich the model. An example is the extension of a process model with performance data, i.e., some prior process model dynamically annotated with performance data (e.g., bottlenecks are shown by coloring parts of the process model).

Table 1: Three Classes of Process Mining Techniques

Discovery	It is used to extract models from raw event data in information system.
Conformance	It compares an existing process model with an event log of the same process.
Extension	It improves the actual process model using information about the actual process recorded in event log.

3. COMPARATIVE ANALYSIS OF BUSINESS PROCESS MINING TOOLS

The methodology used for the comparative analysis of business process mining is illustrated in Fig. 1. The methodology is composed of three phases (shown in Fig. 1).

Phase 1: this phase is used for the comparative analysis of identified existing commercially available software tools that are capable of performing business process mining. This phase of the comparative analysis was completed using the major search engines to locate software tools capable of performing business process mining along with academic citation and abstract databases such as Scopus and Science Direct. Four tools were identified at this stage Fluxicon, ProM, **EMiT,** BPM One.

Phase 2: in this phase, identified the features and capabilities of each of the business process mining software tools enabling the criteria for the comparison of the tools to be developed. This task involved reading and extracting relevant information about the identified software tools from their respective brochure/technical paper and website.

Phase 3: This phase then compared the features and capabilities of each of the identified business process mining software tools, using the criteria set designed in phase 2. The results of this analysis were formatted into a table capable of illustrating the similarities and differences between the business process mining software tools. This Comparison table formed the output of the three phases.

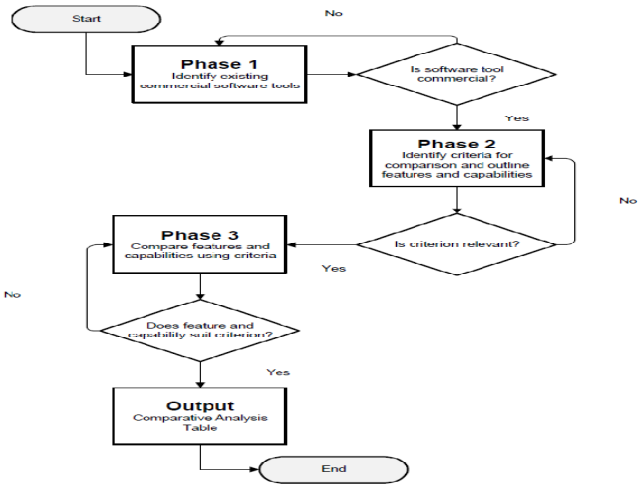


Fig. 1: Methodology for Comparative Analysis of Business Process Mining Tools

4. BUSINESS PROCESS MINING SOFTWARE TOOLS

In this section we narrates that the business process mining software tools are capable of constructing a process model from event logs , user applications found in an information system.

There are different tools which are used in process mining-

1. ProM
2. Fluxicon
3. Emit
4. ARIS PPM

In this paper the main focus in on Emit tool.

1. **ProM:** ProM stands for Process Mining framework. ProM framework i.e., an "plugable "environment for process mining. It is easy to use and extend.

There are different versions of ProM- ProM6.3, ProM6.2, ProM6.1, ProM6.0, and ProM5.2

ProM supports process mining techniques in the form of plug-ins. It is "platform independent" as it is implemented in java and can be downloaded as free of charge. The Fig. shows the overview of the ProM framework. [13]

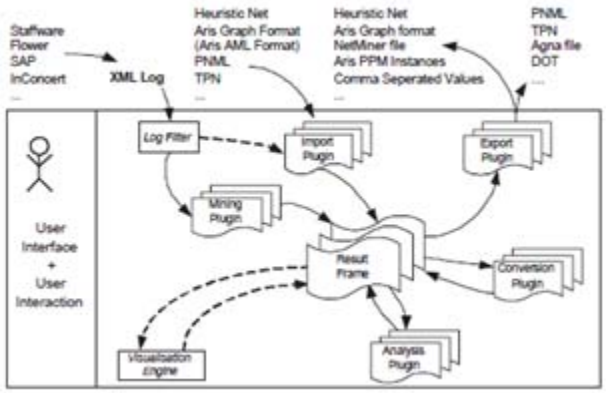


Fig. 2: The ProM framework

As shown in Fig. 2 there are five kind of plug-ins:

- **Mining plug-ins:** it is used to implement some mining algorithm, e.g., mining algorithms that constructs a Petri net based on some event log.

- **Export plug-ins:** it is used to implement some "save as" functionality for some objects. Currently there are 9 export plug - ins.

- **Import plug-ins:** it is used to implement an "open" functionality for exported objects. Currently there are 4 import plugs - ins.

- **Analysis plug-ins:** it is used to implement some property analysis based on some mining result. Currently there are 7 analysis plug - ins.

- **Conversion plug-ins:** it is used to implement conversion between different data formats. Currently there are 3 conversion plug - ins.

ProM6.3 is distributed in 2 parts-

α) ProM6.3 core

β) ProM6.3 plug-ins

ProM6.3	Currently 120 packages and more than 500 plug-ins
ProM6.2	Currently more than 100 packages and more than 400 plug-ins
ProM6.1	More than 170 plug-ins
ProM6.0	More than 170 plug-ins
ProM5.2	More than 230 plug-ins

2. **Fluuxicon:** Fluxicon is a process mining tool which is able to accord with large amount of event logs, complex models and filtering. In this tool we can show the performance metrics in the direct and intuitive manner.

Disco or fluxicon is a tool which allows you to focus on the job at the hand rather than the tool. Disco is a commercial tool. Disco is a tool which is easy to use even you're not a process mining expert, you can easily use this tool to improve your processes.

3. **Emit:** Emit is a process mining tool which stands for "Enhanced mining tool". The algorithm used for process mining are highly complex, so large logs are needed to derive a high quality process model. For this purpose we use EmiT tool.

Emit is a tool that brings event logs using a standard XML format as input. Using an extended version of the α- algorithm [3,8], it can discover the underlying process model and represent it in terms of petrinet.

The mining process consists of a number of steps, namely– Pre-processing, Processing and the Past –processing.

a) Pre-processing: In this phase the log is read into Emit and the log based ordering relations are inferred based on that log. In this thing the assumption is made that the events in the log are totally ordered. But when we are using the default settings, only the "complete" event of a task is taken into account. Basically the log on which the ordering relations will be built

is the real log. Where all entries related to event types not selected are removed.

b) Processing: Processing is the next phase, the core algorithm is called. Since the α-algorithm cannot deal with loops of length one and two, some changes are made. First of all we identified all tasks that are as a loop of length one [3, 8] are taken out of the set of tasks. Second, for all tasks that are identified as loops of length two, the relations are changed. To construct Petri net the "Make petri net (alpha)" button should be clicked. Now the core α-algorithm is called and petri net is build. When the petri net is build EmiT is ready for the post-processing phase.

c) Post processing: In post processing phase the real log is loaded into the program again. Using the real log and the petri net generated in the processing step, we can drive additional information. For the calculation of additional information of the petri net, the real log is used.

IV. ARIS PPM: ARIS PPM stands for ARIS process performance manager. It is a tool that allows organizations to observe and analyze the performance and structure of their business processes. ARIS PPM take outs run time data from the useful source systems, e.g. .ERP, workflow or legacy systems. It helps organizations to access their business processes in terms of speed, cost, quantity, quality and opportunities. ARIS PPM drives the continuous optimization of your internal and external workflows, thus making a key contribution to your business success.

5. CONCLUSION

In this paper, we put the topic of process mining into context; discuss the main issues around process mining and its related tools. There are different tools available for process mining like ProM, Fluxicon, Emit, and ARIS PPM. Basically process mining is the analysis of the information contained in an event log. Instead of starting with a process creation process mining starts by collecting information about the processes as they take place. The basic idea behind process mining is to extract knowledge from event logs recorded by an information system. In the Emit tool mining can be viewed as a three phase process: pre-processing, processing, post-processing.

6. ACKNOWLEDGEMENTS

We are really thankful to God, our family members for making things possible.

REFERENCES

[1] Agrawal, R., Gunopulos, D. and Leymann, F. (1998), "Mining process models from workflow logs", in Schek, H. J. (Ed.), *Proceedings of the 6th International Conference on Extending Database Technology: Advances in Database Technology,* Springer Verlag, Heidelberg, pp 469-490.

[2] W.M.P. van der Aalst, J. Desel, and A. Oberweis, editors. Business Process Man-agement: Models, Techniques, and Empirical Studies, volume 1806 of Lecture Notes in Computer Science. Springer-Verlag, Berlin, 2000.

[3] W.M.P. van der Aalst, J. Desel, and A. Oberweis, editors. Business Process Man-agement: Models, Techniques, and Empirical Studies, volume 1806 of Lecture Notes in Computer Science. Springer-Verlag, Berlin, 2000.

[4] Alves de Medeiros, A. K., Guzzo, A., Greco, G., van der Aalst, W. M. P., Weijters, A. J. M. M., van Dongen, B. F. and Sacca, D. (2007), "Process mining based on Clustering: A quest for precision", in: *Business Process Management Workshops 2007*, Brisbane, Australia, 25th-28th September 2007.

[5] Fujitsu (2010), "Automated Business Process Discovery", available at:http://www.fujitsu.com/global/services/software/interstage/. Accessed on (23/06/2010).

[6] IDS Scheer (2010), "ARIS Process Performance Manager", available at: http://www.ids-scheer.com/us/en/ARIS/ARIS_Platform/ARIS_Process_Performance_Manager/32600.html. Accessed on (23/06/2010).

[7] IDS Scheer (2010), "ARIS Process Performance Manager", available at: http://www.ids-scheer.com/us/en/ARIS/ARIS_Platform/ARIS_Process_Performance_Manager/32600.html. Accessed on (23/06/2010).

[8] Li, Y. and Feng, Y. (2007), "Design of automatic business process modelling method based on process logs", *Computer Integrated Manufacturing Systems*, Vol. 13, No. 1, pp. 24-30.

[9] Nikovski, D. and Baba, A. (2007), "Workflow trees for representation and mining of implicitly concurrent business processes", TR2007-072, Mitsubishi Electric Research Laboratories, Cambridge

[10] Wen, L, van der Aalst, W.M.P, Wang, J, and Sun, J. (2007), "Mining Process Models with Non Free Choice Constructs", *Data Mining and Knowledge Discovery*, Vol. 15, pp.145-180.

[11] W. M. P. van der Aalst and B.F. van Dongen . Discovering Workflow Performance Models from Timed Logs . In Y. Han, S. Tai, and D. Wikarski, editors, *International Conference on Engineering and Deployment of Cooperative Information Systems (EDCIS 2002)*, volume 2480 of *Lecture Notes in Computer Science*, pages 45–63. Springer-Verlag, Berlin, 2002.

[12] D. Grigori, F. Casati, U. Dayal, and M.C. Shan. Improving Business Process Quality through Exception Understanding, Prediction, and Prevention. In P. Apers, P. Atzeni, S. Ceri, S. Paraboschi, K. Ramamohanarao, and R. Snodgrass, editors, *Proceedings of 27th International Conference on Very Large Data Bases (VLDB'01)*, pages 159–168. Morgan Kaufmann, 2001.

[13] B.F. van Dongen, A.K.A. de Medeiros, H.M.W. Verbeek, A.J.M.M. Weijters, and W.M.P. van der Aalst" **"**The ProM framework: A new era in process mining tool support"