Business Process Management Journal, 2012, Vol. 18 Iss: 3, pp.493 - 512

Turner CJ, Tiwari A, Olaiya R, Xu Y

Business Process Mining: From Theory to Practice

Abstract

Purpose - This paper presents a comparison of a number of business process mining tools currently available in the UK market. An outline of the practice of business process mining is given along with an analysis of the main techniques developed by academia and commercial entities. This paper also acts as a primer for the acceptance and further use of process mining in industry suggesting future directions for this practice.

Design/methodology/approach –Secondary research has been completed to establish the main commercial business process mining tool vendors for the market. A literature survey has also been undertaken into the latest theoretical techniques being developed in the field of business process mining.

Findings – The authors have identified a number of existing commercially available business process mining tools and have listed their capabilities within a comparative analysis table. All commercially available business process mining tools included in this paper are capable of process comparison and at least 40% of the tools claim to deal with noise in process data.

Originality/value - The contribution of this paper is to provide a state of the art review of a number of commercial business process mining tools available within the UK. This paper also presents a summary of the latest research being undertaken in academia in this subject area and future directions for the practice of business process mining.

Keywords – Business process mining, Business process management

Paper type – General review

1. Introduction

One of the major drivers behind the rise of business process mining is the need for companies to learn more about how their processes operate in the real world (Tiwari et al. 2008). Business process mining can be described in various ways. For instance, business process mining can be seen as Business Process Intelligence (BPI), Business Activity Monitoring (BAM), Business Process Management (BPM), Business Process Analysis (BPA), Automated Business Process Discovery (ABPD) and Workflow Mining (Alves de Medeiros et al., 2004, Chang, 2006, van der Aalst et al. 2007, Gartner 2008, Schimm, 2004). The practice of business process mining attempts to reconstruct complete process models from data logs containing real process execution data (Tiwari et al., 2008).

In many cases business process mining practice can be traced back to techniques used for data mining practice (Tiwari et al., 2008). Weijters et al. (2007) describes business process mining as a technique that attempts to extract non trivial and useful process information from event logs. Moreover, van der Aalst (2007) suggests that the aim of business process mining is to extract information in the form of process models from event logs. Alves de Medeiros et al. (2004) also states that the goal of business process mining is to discover process models from event logs, i.e. events logged by some information systems are used to extract information about activities and their casual relations. According to Lassen et al. (2007), the aim of business process mining is to extract information about processes from transaction logs, in order to discover the control flow of a process model. Furthermore, van der Aalst and Weijters (2004) stated that business process mining aims to extract information from event logs to capture the business process as it is being executed. An additional commentary on business process practice may be found in Tiwari et al. (2008).

This paper presents the current state of business process mining tool availability and use in the UK market. An outline of the practice of business process mining is given along with an analysis of the main techniques developed by academia and commercial entities. This paper marks the transition of process mining from an academic research topic to a practice accepted by industry and acts as a primer for the acceptance and further use of process mining suggesting future directions for this practice.

2. Latest theoretical developments in the mining of business processes

Business process mining can be used as a tool to find out how people enact processes in the real world, and for delta analysis (the comparison of the mined process to a pre defined ‘ideal’ template) (Wen et al., 2009).

2.1 Business process mining perspectives

Dustdar et al. (2005) distinguishes three different perspectives in business process mining: (1) process perspective, (2) organisational perspective and (3) case perspective.

The process perspective focuses on the ordering of activities (the control-flow of the process). The aim is to find an acceptable representation of all possible paths within the process. These paths can then be expressed in terms of a process model (e.g., Petri net or Event-driven Process Chain (EPC)) (Dustdar et al., 2005).

The organisational perspective focuses on the originators within a process, i.e., which people and roles are involved and how are they related. This approach can be used to portray the roles and relationships between individuals in a process in terms of a social network (Dustdar et al., 2005).

The case perspective takes into account the properties of cases, that is, the attributes that can differentiate one path through a process (case) from another (Dustdar et al., 2005). An example is a process that traces out paths followed by call centre workers in handling a customer complaint. With such a process it may be beneficial to differentiate the paths through that process my knowing attributes such as the individual call centre operator and the type of complaint.

2.2 Business process mining types

Other than the three perspectives identified by Dustdar et al. (2005), van der Aalst et al. (2010) describe three basic types of business process mining: (1) Process discovery (2) Conformance checking (3) Extension.

Discovery: No a-priori model exists. Mining is based on event logs without the use of a template of the complete process. Van der Aalst et al. (2010) give the example of their α algorithm where a process model can be discovered from an event log without prior knowledge of the process itself.

Conformance Checking: An a-priori model exists for a process. This model is used to check if the mined process conforms to real life (i.e. the a-priori model acts a template for the complete process) (van der Aalst et al. 2010). For example, there may be a process model indicating the standard procedure for a customer to obtain a credit card. In conformance checking, this model will show if the standard procedures in obtaining a credit card have been followed for a particular customer. Conformance checking may be used to detect, locate and explain these deviations, and to measure the severity of the process disparity.

Extension: An a-priori model exists for a process. This model is enhanced by the addition of extra parameters detailing a new aspect or perspective (van der Aalst et al. 2010). Attributes could include data on the cost, performance or importance of different tasks within a process.

2.3 Business process mining techniques

One of the most recent developments in the field of business process mining is the rise of commercial tools capable of performing process mining. The Gartner Group (Gartner, 2008; 2009) has highlighted process mining (in the form of Business Process Discovery) in a number of recent reports. Most notably Gartner name a number of Business Process Management Suite vendors who now provide automated methods for process discovery (Gartner 2009). Such commercial tools will be discussed later on in the comparative tool analysis section of this paper

(section 3). This heightened commercial awareness of business process mining reinforces the need to also explore the latest research in this area.

(1) Transition systems and regions

Notable contributions to this area include the work of van der Aalst et al. (2010) which utilises transition systems and regions to discover process models from event logs. This technique is said to mine models that offer a balance between over fitting (only modelling process transitions observed in the event log) and under fitting (modelling process transitions that may potentially be inferred from the event log) the event log, so providing a mined model that does not over generalise the event log it is based on (van der Aalst et al., 2010).

(2) Clustering techniques

Many techniques highlight the possibility of combining a number of process mining approaches to mine more challenging event logs, such as those that contain noise (event log data that contains incorrect or incomplete information about a process). Most notably, clustering techniques have been used in this way. The work of Greco et al. (2006) provided such a technique utilising the κ-means algorithm to provide hierarchical views of a mined process model. The technique of Greco et al. (2006) can utilise the Heuristic approach of Weijters and van der Aalst (2003) to successfully mine noisy event logs. Ferreira (2009) has experimented with the use of sequence clustering for use in process mining. In this research the work processes of a software development team are clustered to reveal a set of behavioural patterns which are then related to the set processes being followed (Ferreira, 2009). The use of clustering for process mining has been further investigated in the work of Alves de Mederios et al. (2007) addressing the mining of complex ‘spaghetti-like’ process models. Such ‘spaghetti’ models have also been addressed by Veiga and Ferreira (2009) in their work utilising sequence clustering. An interesting new direction for clustering is its use with the Fuzzy Process Mining approach of Gunther and van der Aalst (2007) put forward by van Dongen and Adriansyah (2009). This work concentrates on the performance analysis of business processes and the identification of bottlenecks within processes.

(3) Heuristic approach

One of the most long standing process mining techniques is the heuristic approach of Weijters and van der Aalst (2001, 2003). In the work of Weijters and van der Aalst (2001, 2003) a set of rules are used to determine the precedence between tasks and overall task sequences. Further research in the field of heuristic process mining has been made by Burattin and Sperduti (2010a) who, with their Heuristics++ miner, adapt the approach of Weijters and van der Aalst (2003) to allow the duration of process activities to become part of the process mining parameter set. Burattin and Sperduti also explore the area of parameter setting with their Heuristics++ miner and

formulate an approach potentially applicable to a wide range of process mining approaches (Burattin and Sperduti, 2010b).

(4) Evolutionary techniques

The use of evolutionary techniques for process mining is another new research direction. The approach of Alves de Mederios et al. (2005) utilises a Genetic Algorithm (GA) to mine a variety of event logs containing noise and duplicate tasks. This technique allows for process patterns to be represented as chromosome strings, and successive populations of potential process mining solutions to be generated. Each individual process mining solution in a population is a mined process which is given a fitness score based on how accurately it describes the event log on which it is based. The use of a GA allows for a global search of the event log allowing for the possibility of mining process constructs that span the entire length of the process (a construct that may be difficult to mine if using a mining approach that utilised just a local search strategy. The work of Turner et al. (2008) develops further the concept of evolutionary process mining with the Genetic Programming (GP) process miner. This approach uses a representation based on a graph construct negating the need to translate to and from a binary string (necessary with the GA approach). The GP approach is suited to mining complex highly parallel processes.

(5) Declarative mining approach

The declarative mining approach, featured in the work of Goedertier et al. (2009), allows for the introduction of artificially generated negative events; that is events that show that a particular path through a process (or activity recorded in a process trace) is not possible. It is argued in this work that process logs rarely contain such negative traces and their introduction can aid the task of process discovery. Also of note in this area is the incremental declarative approach put forward by Cattafi et al. (2010) that recognises that process can change over time and is able to revise the mined model taking into account newer process traces and the possible deviations they may bring. This technique utilises Inductive Logic Techniques in order to revise the process models being mined. Both of the aforementioned approaches point to the contribution of Ferreira and Ferreira (2006), on the use of Inductive Programming Logic for the discovery of workflow activities, as influences on the research directions that they took. One area that has received limited attention in process mining research is that of business policy mining within business processes. Research by Li et al. (2010) introduces the concept of using a text mining approach to extract process related business policies from policy documents. The authors of this approach point to the potential commercial value of being able to extract business processes from, sometimes, poorly structured policy documents (Li et al., 2010).

2.4 Application of business process mining techniques

Table 1 presents a list of common mining problems and shows which techniques can be used to address them. For further details on the process mining problem types list in the first column of Table 1 please see Tiwari et al. (2008). It is interesting to note from Table 1 that soft computing approaches to process mining (encompassing Evolutionary, Fuzzy and Neural Network techniques) are now starting to form a focus of research in recent years. It may also be noted that Markovian approaches to process mining are still being explored (Li and Feng, 2007).

Table 1: Mapping process-mining challenges to techniques in current research (adapted and updated from Tiwari et al. (2008))

	Data Mining Based	Heuristic Approach	Soft Computing Algorithms	Markovian Approach	Other Approaches
Noise	Agrawal et al. (1998) Hwang and Yang (2002)	Weijters and van der Aalst (2003) Weijters and van der Aalst (2001) Burattin and Sperduti (2010a)	Alves de Medeiros, et al. (2005) Cook and Wolf (1998a)	Cook and Wolf (1998a) Li and Feng (2007)	Gaaloul and Godart (2005) Cook et al.. (2004) van der Aalst and Alves de Medeiros (2005) Greco et al. (2006) Ferreira et al. (2007) Alves de Medeiros et al. (2007) Ferreira (2009) Goedertier et al. (2009)
Hidden Tasks			Alves de Medeiros, et al. (2005)		Rozinat and van der Aalst (2006) Wen et al. (2007)
Duplicate Tasks	Herbst and Karagiannis (2004) Agrawal et al. (1998)		Alves de Medeiros, et al. (2005)		van Dongen and van der Aalst (2005a) Rozinat and van der Aalst (2006) van der Aalst et al. (2010) Goedertier et al. (2009)
Non Free Choice Constructs		Weijters and van der Aalst (2001) Burattin and Sperduti (2010a)	Alves de Medeiros, et al. (2005) van Dongen and Adriansyah (2009)		van der Aalst et al. (2002) Greco et al. (2006) van der Aalst et al. (2010) Goedertier et al. (2009) Gaaloul et al. (2009) Wen et al. (2007)
Mining Loops	Herbst and Karagiannis (2004) Schimm (2003) Schimm (2004)	Weijters and van der Aalst (2001) Weijters and van der Aalst (2003) Burattin and Sperduti (2010a)	Alves de Medeiros, et al. (2005) Gunther and van der Aalst (2007)		Alves de Medeiros et al. (2004) van der Aalst and Alves de Medeiros (2005) Lamma et al. (2007) Greco et al. (2006) van der Aalst et al. (2010) Alves de Medeiros et al. (2007) Cattafi et al. (2010) Wen et al. (2007) Wen et al. (2009) Ferreira (2009) Goedertier et al. (2009) Gaaloul et al. (2009)
Different Perspectives					van der Aalst and Song, 2004 Rozinat and van der Aalst, 2006
Visualising Results	Herbst and Karagiannis (2004) Hammori et al. (2004)		Alves de Medeiros, et al. (2005)		Dustdar et al. (2004) Alves de Medeiros et al. (2004) van Dongen and van der Aalst (2005) van Dongen et al. (2005) Lamma et al. (2007)

Concurrent processes	Schimm (2004)	Weijters and van der Aalst (2001) Weijters and van der Aalst (2003) Burattin and Sperduti (2010a)	Alves de Medeiros, et al. (2005) Gunther and van der Aalst (2007) van Dongen and Adriansyah (2009) Turner et al. (2008)	Li and Feng (2007)	Golani and Pinter (2003) Alves de Medeiros et al. (2004) Lamma et al. (2007) Greco et al. (2006) Nikovski and Baba (2007) Wen et al. (2007) van der Aalst et al. (2010) Alves de Medeiros et al. (2007) Cattafi et al. (2010) Wen et al. (2009) Ferreira (2009) Goedertier et al. (2009) Gaaloul et al. (2009)

As demonstrated in this section, the range and sophistication of process mining techniques has seen a substantial increase in just the last 3-4 years. Note must be made of the business process mining framework ProM (2010), which many of the techniques, in Table 1, utilise as a hosting platform. ProM is 'a generic open-source framework' for the support of process mining techniques (the framework provides event log handling functionality along with other utility and visualisation capabilities required for process mining) (ProM, 2010). With such maturity it is perhaps no surprise that commercial business process mining products are starting to emerge onto the market. The following section of this paper compares the leading products currently available on the market utilising a comparative methodology devised by the authors of this paper.

3. Analysis of Business Process Mining Practice

While there has been much innovation within, and debate surrounding, the field of process mining, within academia the results of this activity are only now filtering through into the commercial offerings of a number of business process management software vendors. So far there has been little discussion on the actual market for business process mining tools or how to compare them. In this section of the paper a methodology for the comparative analysis of commercial business process mining tools is put forward along with a comparison table of a number of leading tools identified from the authors' research.

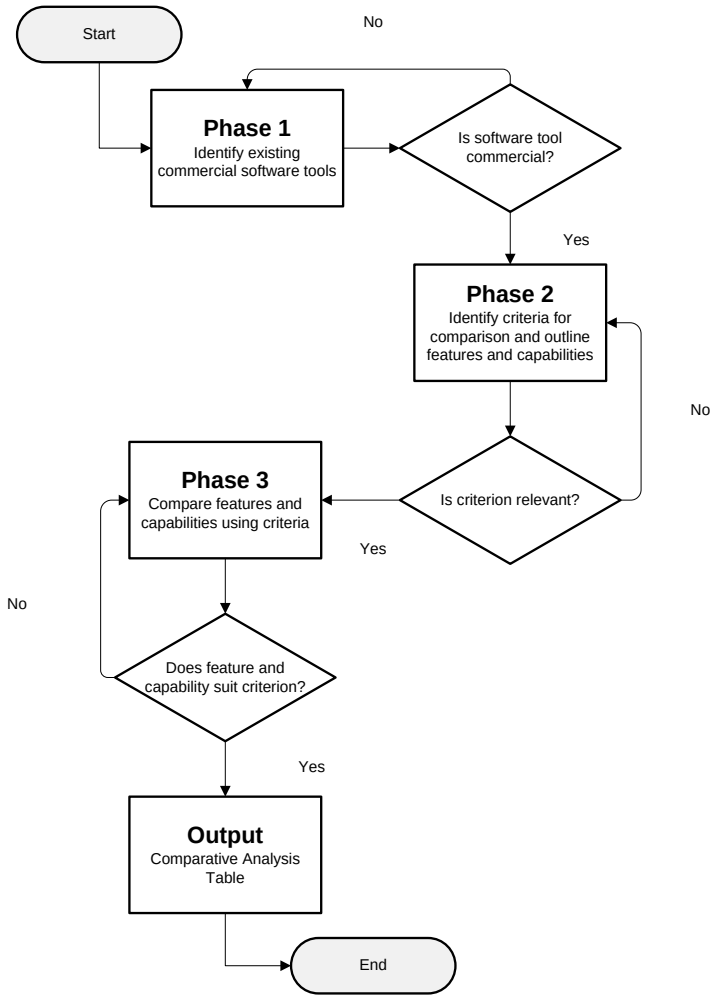

Figure 1: Methodology for comparative analysis of business process mining tools

3.1. Methodology for Comparative Analysis of Business Process Mining Tools

The methodology used for the comparative analysis of business process mining is illustrated in Figure 1. The methodology embodies a degree of freedom (i.e., a previous research phase can be revisited when necessary), indicated by the back loops in Figure 1. In fact, this methodology can also be used for a comparative analysis of disciplines outside the field of business process mining. The methodology is composed of three phases (shown in Figure 1).

Phase 1: of the comparative analysis identified existing commercially available software tools 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. Five tools were

identified at this stage, Futura Reflect, Fluxicon, ARIS Process Performance Manger, BPM One, and Fujitsu Automated Process Discovery Service (APDS).

Phase 2: 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. After gaining sufficient knowledge about each identified business process mining software tool, the criteria for comparison of business process mining tools were developed by locating and using the common characteristics of business process mining identified from the literature review. At this stage an initial list of criteria for comparison of the process mining tools was established comprising of the three headings ‘Algorithm’, ‘Process mining problem’ and Type of business process mining’. These criteria were then reviewed and refined after consultation with experts (sought from the literature review and personal recommendation) in the field of business process mining through a focus group held in phase 3.

Phase 3: involved the convening of a focus group on business process mining tools. This focus group consisted of 10 experts drawn from industry and academia (from companies such as Santander, Enginsoft and Mergis) and was convened to gain a better understanding of the requirements from industry for business process mining tools. The experts had a minimum of 5 years experience with a number of participants having over 20 years experience. This ensured the depth, quality and reliability of the findings. The backgrounds of the participants varied from business process management to general management and consultancy. Two moderators were used for this focus group (a dual moderator focus group). One moderator ensured the smooth progress of the focus group (facilitating a constructive discussion between members in the group) and the other made sure that all the topics were covered (addressing members when the discussion was moving off topic). Topics discussed by the focus group included, the use of business process mining in industry, knowledge of business process mining and business process mining functionality. The introduction of topics in the focus group followed a funnel approach; where a broad discussion of a topic was narrowed down to obtain detailed responses on particular issues.

When asked about the criteria for comparison for business process mining tools the focus group was able at add two additional criteria, Process model (what type of process modelling notation is used to display the mining results) and Additional features (what additional features are available through the process mining tool). This focus group was also asked to suggest additional process mining tools for inclusion in the survey and the Iontas and Comprehend tools were offered as examples by focus group members. This phase then compared the features and capabilities of each of the identified business process mining software tools, including the two tools suggested by the focus group; using the criteria set designed in phase 2 and the two additional criteria suggested by the focus group members.

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.

3.2. Business process mining software tools

This section briefly describes the business process mining software tools that are capable of constructing a process model or a flowchart (describing a process instance) from event logs, existing process pre-templates, or user applications found in an enterprise information system. As mentioned in section 3.1 a focus group was convened which added additional criteria and tools to the comparison. Out of all the software tools found, the following software tools were identified to be capable of performing business process mining:

- Futura Reflect
- Fluxicon
- Comprehend
- ARIS Process Performance Manager (PPM)
- BPM one
- Iontas Focus Suite
- Fujitsu Automated Process Discovery Service (APDS)

Futura Reflect

The Futura Reflect tool is offered as part of a Software as a Service (SaaS) package, providing process discovery, social network discovery, process animation, and performance analysis capabilities. Futura Reflect can automatically construct a process model without any prior knowledge of the process (such as a template) directly from log files provided by enterprise information systems (process discovery). The tool is offered as a suite paid for on a per use basis. The company also offers a consultant based service, where logs are mined by the company and the results returned to the client. The software tool was developed by a company called Futura Process Intelligence. This company was founded in 2006, and has subsequently been recognised by Gartner (2009) as one of most innovative companies in the field of business process management (Futura Process Intelligence, 2010).

Fluxicon

Fluxicon is also offered as part of a Software as a Service (SaaS) package that provides automatic process discovery, social network discovery, detailed descriptions of process executions, and identification of bottlenecks within a process. The control flow of a process can be visualised using this tool. In addition Fluxicon can utilise process documentation to establish the distance between process goals and an actual process in situ. Like the Futura Reflect tool Fluxicon offers process discovery capabilities. In addition to offering a software tool a process mining

consultancy service (similar in some ways to that offered by Futura) is also provided. This company has a partnership with Eindhoven University of Technology (Fluxicon, 2010).

Comprehend

Comprehend is a software tool that offers information capture, process identification, process centric business intelligence, and process variation discovery. Processes can be filtered by attributes in order to gain new insights and understand variations between process instances. The Comprehend tool is designed to complement other business intelligence tools in the market collecting data from various sources in order to construct a quantified probabilistic process model. This software tool was created by OpenConnect, an American company providing business process analysis software. This company has a partnership with the University of Leuven in Belgium (OpenConnect, 2010).

ARIS Process Performance Manager (PPM)

ARIS PPM is a tool developed by IDS Scheer, a market leader in business process management (BPM) software, solutions, and services for corporations and public organisations worldwide. ARIS PPM provides facilities such as quantitative measurement of objectives and visualisation of process instances. In addition an aggregate process view can be compiled from a variety of data sources and realised as a process graph (using the Event-Driven Process Chain (EPC) notation) or a Gantt chart (in order to better understand the sequence and overlap of the activities in a process). Process instances may be annotated with Key Performance Indicator (KPI) information and performance bottlenecks can be discovered using built in interactive analysis. (IDS Scheer, 2010).

BPM one

BPM one is a tool that offers process discovery, reconstructing a process model from activities recorded in an enterprise information systems event log. Process models are mined in an iterative fashion, improving with each iteration and converging on a final model of the process data. Parts of the information systems log can be filtered for incomplete or exceptional process instances. BPM one was created by Pallas Athena, a leading European supplier of Business Process Management (BPM) solutions, software and services (Pallas Athena, 2010). This company has also been recognised by Gartner (2009) as a leading light in BPM field (Pallas Athena, 2010).

Iontas Focus Suite

Iontas Focus suite is a software tool that analyses event based data within an organisation. The Focus Suite can assemble a graphical business process view from data situated in multiple applications so the individual tasks (or steps) of the process may be examined. The Iontas Focus suite provides facilities such as automatic discovery and visualisation of business processes, identification of inefficient process

flows and best practices, access to individual process details and suggestions for process improvements. One way in which the Focus Suite differs from other vendors products is the way in which it gathers process data. Such data is derived from user's desktops rather than event logs and the overall process may be visualised in the form of a Microsoft Visio flowchart (Iontas, 2010).

Fujitsu Automated Process Discovery Service (APDS)

Fujitsu APDS automatically discovers and visualises business processes, identifies business process bottlenecks and highlights inefficiencies, and compares mined processes to pre-existing templates. This tool can be integrated with process modelling tools such as Interstage Business Process Manager. As with other vendors mentioned here, processes may mined from data provided by a range of enterprise systems. A range of analytical functions are provided by APDS to improve processes, such as a process comparison utility (Fujitsu, 2010).

3.3. Comparison of Business Process Mining Tools

Table 2 highlights the key functionality of the commercial business process mining tools featured in this paper. From Table 2 it becomes clear that each tool can provide a graphical representation of a process though the method and notation used in each representation may differ. A number of vendors such as ARIS (IDS Scheer) and Fluxicon, provide a choice in the graph representation used. It is possible that an end user may select a process mining tool partly on the basis that it fits with the process modelling notation currently in use within their organisation (most likely the notation of the process modelling tools currently in use in an organisation). The integration with Visio used by the Iontas Focus Suite is therefore an interesting development in this direction. Integration with other business process management software will be another consideration for the user. Tools such as Fujitsu APDS allow for integration with the Interstage Business Process Manager Studio environment for business process modelling functionality. Both BPMOne and ARIS are Business Process Management suites.

Consultancy is used as an alternative way of delivering business process mining functionality to the end user. Fluxicon's process mining functionality is backed up on this basis, with consultants who can carry out the process mining and provide results back to clients. The Futura Reflect tool is also complimented by a consultancy service. Many vendors, such as BPMOne and Comprehend, offer process discovery as a form of process mining. As mentioned earlier in this paper process discovery allows for the mining of process models from data without the need for templates. This feature can potentially offer a greater level of flexibility when mining complex and infrequent process patterns.

All vendors offer a form of conformance checking whereby a mined process can be checked against a template (or ideal process model). A number of tools go further than this and also highlight process variants, low frequency process patter that depart from the overall mined process flow. Comprehend is one such tool that offers this insight along with Futura Reflect.

The type of data being mined can vary from organisation to organisation. The completeness of event logs and the amount of errors that they contain can also present a challenge to a process mining tool. Mining tools that offer to tackle a wider range of process mining problems may be suited to organisations that know that their process data is incomplete or contains challenging constructs (such as non free choice constructs).

Table 2: Comparative analysis table of business process mining

	Software tool	Type of business process mining	Process model	Process mining problem	Algorithm	Additional Features
1	**Fujitsu Automated Process Discovery Service (APDS)**	• Extension • Conformance checking • Process discovery	• Interstage BPM studio (similar to Business Process Modelling Notation (BPMN))	• Duplicate task • Mining loop • Delta analysis • Visualising results • Mining different perspectives	• Custom algorithm	• Process performance metrics • Snapshots of visualised business processes including typical routes, repeats, process variance, exception routes and loop-backs • Findings, impacts and recommendations for process improvement • Discovers business processes
2	**Futura Reflect**	• Process discovery • Conformance checking • Extension	• Workflow Net	• Noise • Duplicate task • Non-free choice construct • Mining loops • Delta analysis • Mining different perspectives • Visualising results • Concurrent processes	• Genetic Algorithm • Geodesic paths	• Advanced visual reporting facility that has been specially designed with a business process mining perspective • Displays extra process information • Provides an in-depth overview of executed business processes • Visualizes different types of stream e.g. a document stream in an editorial process • Eradicates exceptional cases found in a process • Consultancy service offered
3	**Fluxicon**	• Conformance checking • Extension • Process discovery	• Yet Another Workflow Language (YAWL) • Petri-Net • Event-driven Process Chain (EPC) • Mining Extensible	• Noise • Duplicate task • Non-free choice construct • Mining loops • Delta analysis • Mining different perspectives	• Genetic Algorithm • Fuzzy logic • Decision trees • Geodesic paths • Heuristic nets	• Provides a variety of algorithms • Visualise activities of a process with multiple levels of abstraction • Measures distance, and detects deviations, from existing business process objectives • Consultancy service offered

			Markup Language (MXML)	• Visualising results • Heterogeneous results • Concurrent processes			
4	**ARIS Process Performance Manager (PPM)**	• Conformance checking • Extension	• Event- driven Process Chain (EPC) • Business Process Modelling Notation (BPMN) • Unified Modelling Language (UML) • Business Process Execution (BPEL) • Web Service Description Language (WSDL) • Gant chart	• Duplicate task • Mining loops • Delta analysis • Mining different perspectives • Visualising results • Concurrent processes	• Aggregation graphs	• Performance Dashboard-uses process maps and associated KPI's to monitor the current performance of a process • Instance Controlling-captures problems encountered in live processes • Advanced Analysis-quickly searches for strange patterns and transactions (with high optimization) in a process • Organisational Analysis-viewing actual processes from the perspective of organisational unit, teams and departments
5	**Iontas Focus Suite**	• Conformance checking • Process discovery • Extension	• Microsoft Visio 2007	• Mining different perspectives • Visualising results	• Custom algorithm	• Provides macro and micro views of business processes • Captures timing of events and particular values that take place during a process • Tamper proof • Data integration – sends data from one application to many applications without complex integration issues
6	**Comprehend**	• Conformance checking • Extension • Process discovery	• Workflow NET	• Duplicate task • Mining loops • Delta analysis • Mining different perspectives • Visualising	• Hidden Markov models	• Process duration statistics • Data aware process variations • Probabilistic event transition information • Drill down into transition data correlations

				results • Current processes		
7	**BPM one**	• Process discovery • Conformance checking • Extension	• Workflow NET	• Noise • Duplicate task • Non-free choice construct • Mining loops • Delta analysis • Mining different perspectives • Visualising results • Concurrent processes	• Genetic Algorithm • Geodesic paths	• Risk, quality and process analysis • Modelling and simulation of business processes • Constant measuring and monitoring of business processes • Extended monitoring and analysing of executed processes • Real-time insight and control of business processes • Equips the user to actively intervene to avoid potential problems in business processes

4. Discussion

Most of the software tools in Table 2 perform process discovery (as defined earlier in the literature review section), mining complete process models without the need for a template. In addition all of the identified software tools offer the service of conformance checking. However, the authors were unable to assess the quality of conformance checking performed by each of the software tools.

The Fluxicon tool is offered on the basis of Software as a Service (SaaS) administered by process consultants. In addition to their software suite Futura Reflect is also able to offer a consultant backed process mining service for customers.

The ARIS tool offers the possibility to monitor processes against a set of Key Performance Indicators, which is a common requirement especially in larger organisations. ARIS along with Fluxicon are the only tools to output results in the popular EPC graph format.

The Futura Reflect, BPMone and the Fluxicon software tools address a wide range of business process mining problems (e.g. noise, duplicate tasks, non-free choice constructs etc.). It was also found that noise (i.e., incorrect logged data) is the most common problem experienced in business process mining. From Table 2 it can be seen that the Fujitsu APDS tool, Iontas Focus suite, ARIS PPM tool and the Comprehend tool, while potentially being able to mine noisy data, do not explicitly claim to deal with noise.

BPMOne, in addition to its standard business process analysis functionality, is able to simulate a business process, showing changes in the process in the form of an animation. This type of functionality is also offered by the Futura Reflect tool; concentrating in particular on the identification of slow running or abnormal processes. The analysis tools of Iontas can also compare business processes against best practice and suggest improvements. The analysis tools provided with Comprehend allow the user to discover correlations in parameter data within business processes, providing the user with a finer grain of detail on process variations.

All of the commercial business process mining software tools reviewed here can be used to identify bottlenecks in business processes. However, the authors were unable to ascertain the quality of bottleneck identification for each of the commercial tools compared in this paper. From academic literature Petri Nets are the most commonly used process modelling notation in business process mining research (Tiwari, et al. 2008). However, the authors uncovered that only Fluxicon is capable of constructing Petri Net process models (among other graph visualisation forms). This

shows that only one commercially available business process mining software tool can construct process models in the form of a Petri-Net.

5. Conclusions

This paper sets out a number of existing commercially available business process mining tools and their capabilities and provides a comparative analysis table for these software tools. From the comparative analysis table, the authors discovered that more than 85% of the business process mining tools perform process discovery, and approximately 40% of the tools claim to deal with noise.

In addition, the authors combined the knowledge of the present state of research in business process mining with current industry perceptions to identify areas where further research would be helpful in enhancing the discipline. The authors suggest that further research should be made into the use of process modelling techniques more commonly used by companies, such as the EPC notation. The analysis presented in this paper has enabled the authors to make recommendations on the future of business process mining for industry. For instance, existing business process mining vendors may consider offering their tools in the form of an online service. With the advent of web service connected applications and the increasing acceptance of ‘cloud’ based data repositories the market for consultant backed software offerings may increase. Such an offering could potentially be integrated with in house business process management systems.

While the process mining tools featured in this paper will all perform business process mining across a wide range of process data sets it is possible that certain tools may be better suited to a particular organisations process mining challenges. It is hoped that the comparison table and discourse in this paper will help the prospective user of business process mining to better understand the tools on offer and perhaps experiment with such tools (featured in this paper or others available in the market) in their organisation.

It is important to point out that research is required to identify the full market potential of business process mining. For instance, the authors suggest that a market analysis should be conducted to investigate the current state of business process mining problems that exist in industry. From this work could come a better understanding in the research community of the business process mining functionality that is required from industry. A more quantitative study of the performance of each mining tool with a number of representative process data sets is beyond the scope of this paper. It is likely that such a study could focus on a particular sector or case study organisation.

It is clear that the latest developments in business process mining research are filtering through into commercial products. It must also be noted that business process mining is just one part of the wider field of Business Process Management (BPM). Within this field a new research area has recently gained attention, that of

business process optimisation. This research is exemplified by the work of Vergidis et al. (2007) who propose an evolutionary approach to the optimisation of business processes. This approach puts forward a multi-objective optimisation framework that allows for a mined process to be optimised in terms of its task attributes such as duration and cost. This work points to the potential for business processes to be optimised and then enacted in an automated fashion by future intelligent BPM software systems. Real-time reconfiguration of business processes could provide advantages in terms of both time and cost savings and enable the streamlined design and enactment of a new generation of 'just-in-time business processes'.

References

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.

Alves de Medeiros, A. K., van Dongen, B. F., van der Aalst, W. M. P. and Weijters, A. J. M. M. (2004), "Process mining: Extending the a-algorithm to mine short loops", BETA Working Paper Series, WP 113, Eindhoven University of Technology, Eindhoven.

Alves de Medeiros, A. K., Weijters, A.J.M.M. and van der Aalst, W. M. P. (2005). "Genetic process mining: A basic approach and its challenges" In: *Business Process Management Workshops (BPM),* Nancy, France, September 5, 2005.

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.

Burratin, A. and Sperduti, A. (2010a), "Heuristics miner for time intervals", *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, Bruges, Belgium, 28-30 April.

Burratin, A. and Sperduti, A. (2010b), "Automatic determination of parameters values for heuristics miner++", *World Congress on Computational Intelligence (WCCI) incorporating Congress on Evolutionary Computation (CEC) 2010*, Barcelona, Spain, 18-23 July.

Cattafi, M., Lamma, E., Riguzzi, F. And Storari, S. (2010), "Incremental declarative process mining", in Szczerbicki, E. et al. (Eds) *Smart Information and knowledge management*, SCI 260, Springer, Berlin, pp. 103-127.

Chang, J.F. (2006), "*Business process management systems: Strategies and implementation*", Vol. 19, Auerbach Publications, Taylor and Francis Group, Florida, USA.

Cook, J. E., Du, Z., Liu, C. and Wolf, A. L. (2004), “Discovering models of behaviour for concurrent workflows”, *Computers in Industry*, Vol. 53, No. 3, pp. 297-319.

Cook, J. E. and Wolf, A. L. (1998a), “Discovering models of software processes from event-based data”, *ACM transactions on software engineering and methodology*, Vol. 7, No. 3, pp. 215-249.

Dustdar, S., Hoffmann, T. and van der Aalst, W. (2004), “Mining of ad-hoc business processes with teamlog”, Technical University of Vienna, Vienna.

Dustdar, S, Hoffmann, T, and van der Aalst, W. (2005), “Mining of ad-hoc business Processes with TeamLog”, *Data Knowledge and Engineering*, Vol. 55, No. 2, pp. 129-158.

Ferreira, D. R., (2009),
“Applied Sequence Clustering Techniques for Process Mining”, in: Cardoso, J., van der Aalst, W.M.P. (Eds), *Handbook of Research on Business Process Modelling*, IGI Global, pp. 492-513.

Ferreira, H. and Ferreira, D. R., (2006), “An integrated life cycle for workflow management based on learning and planning”. *International Journal of Cooperative Information Systems*, Vol. 15, No. 4, pp. 485–505.

Ferreira, D., Zacarias, M., Malheiros, M. and Ferreira, P. (2007), “Approaching process mining with sequence clustering: Experiments and findings”, in: Alonso, G., Dadam, P. and Rosemann, M. (Eds), *Business Process Management (BPM 2007)*, Springer Verlag, Heidelberg, pp. 360-374.

Fluxicon Process Laboratories (2010), “Process Mining Technology”, available at: http://fluxicon.com/technology/. Accessed on (23/06/2010).

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

Gaaloul, W., Gaaloul, K., Bhiri, S., Haller, A. and Hauswirth, M. (2009), “Log-based transactional workflow mining”, *Distributed Parallel Databases*, Vol. 25, pp. 193-240.

Gaaloul, W. and Godart, C. (2005), “Mining workflow recovery from event based logs”, in: van der Aalst, W. M. P., Benatallah, B., Casati, F., et al. (Eds), *Business Process Management (BPM 2005),* Springer Verlag, Heidelberg, pp. 169-185.

Gabriel M. Veiga and Diogo R. Ferreira. (2009), “Understanding spaghetti models with sequence clustering for ProM”, in: Rinderle-Ma, S. et al. (Eds), *BPM 2009 Workshops, LNBIP 43*, Springer-Verlag, Berlin.

Gartner (2009), “Cool vendors in business process management”, Gartner Inc. ID:G00165554

Gartner (2008), “Automated business process discovery improves BPM outcomes”, Gartner Inc. ID:G00164422

Goedertier, S., Martens, D., Vanthienen, J. and Baesens, B. (2009), “Robust process discovery with artificial negative events”, *Journal of Machine Learning Research*, Vol. 10, pp. 1305-1340.

Golani, M. and Pinter, S. (2003), “Generating a process model from a process audit log”, in: van der Aalst, W. M. P., ter Hofstede, A. H. M. and Weske, M. (Eds), *Business Process Management (BPM 2003),* Springer Verlag, Heidelberg, pp. 1020.

Greco, G, Guzzo, A, and Pontieri, L. (2006), “Discovering Expressive Process Models by Clustering Log Traces”, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 18, pp.1010-1027.

Gunther, C. W. and van der Aalst, W. M. P. (2007), “Fuzzy mining: Adaptive process simplification based on multi perspective metrics”, in: Alonso, G., Dadam, P. and Rosemann, M. (Eds), *Business Process Management (BPM 2007),* Springer Verlag, Heidelberg, pp. 328-343.

Hammori, M., Herbst, J. and Kleiner, N. (2004), “Interactive workflow mining”. in: Desel, B. P. and Weske, M. (Eds), *Business Process Management (BPM 2004)*, Springer Verlag, Heidelberg, pp. 211.

Herbst, J. and Karagiannis, D., (2004), “Workflow mining with InWoLvE”, *Computers in Industry*, Vol. 53, No. 3, pp. 245-264.

Hwang, S. Y. and Yang, W. S. (2002), “On the discovery of process models from their instances”, *Decision Support Systems*, Vol. 34, No. 1, pp. 41-57.

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).

Iontas (2010), “Iontas Focus Software Suite”, available at: http://www.iontas.com/pages/products/products.php. Accessed on (23/06/2010).

Li, J., Wang, H. J., Zhang, Z., Zhao, J. L. (2010). “A policy-based process mining framework: Mining business policy texts for discovering process models”, *Information Systems and E-Business Management*, Vol. 8, pp. 169-188.

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.

Lassen, K.B, van Dongen, B.F, and van der Aalst W.M.P. (2007), “Translating Message Sequence Charts to other Process Languages Using Process Mining”, Beta Working Paper WP207, Eindhoven University of Technology, Eindhoven.

Lamma, E., Mello, P., Riguzzi, F. and Storari, S. (2007), “Applying inductive logic programming to process mining”, in: Blockeel, H., Ramon, J., Shavlik, J., et al. (Eds) *Inductive Logic Programming*, Springer Verlag, Heidelberg.

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

OpenConnect (2010), “Comprehend”, available at: http://www.oc.com/technology/index.jsp. Accessed on (23/06/2010).

Pallas Athena (2010), “Process Mining”, available at: http://www.pallas-athena.com/process-mining. Accessed on (23/06/2010).

ProM (2010) “ProM Process Mining Framework”, available at: http://prom.win.tue.nl/research/wiki/prom/start. Accessed (23/06/2010).

Rozinat, A. and van der Aalst, W. M. P. (2006), “Decision mining in ProM”, in: Dustdar, S., Fiadeiro, J. L. and Sheth, A. (Eds), *Business Process Management (BPM 2006)*, Springer Verlag, Heidelberg, pp. 420-425.

Schimm, G. (2003). “Mining most specific workflow models from event-based data”, in: van der Aalst, W. M. P., ter Hofstede, A. H. M. and Weske, M. (Eds), *Business Process Management (BPM) 2003,* Springer Verlag, Heidelberg, pp. 1021.

Schimm, G. (2004). “Mining exact models of concurrent workflows”, *Computers in Industry*, Vol. 53, No. 3, pp. 265-281.

Tiwari, A., Turner, C. J. and Majeed, B. (2008). “A review of business process mining: State of the art and future trends”, *Business Process Management Journal (BPMJ)*, Vol. 14, No. 1, pp. 5-22.

Turner, C.J, and Tiwari, A. Mehnen, J, (2008), “A genetic programming approach to business process mining”, in: *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, 12-16 July 2008, Atlanta, USA, pp 1307-1314.

van Dongen, B. F. and van der Aalst, W. M. P. (2005), “Multi-Phase process mining: Aggregating instance graphs into EPCs and petri nets”, *In: 2nd International Workshop on Applications of Petri Nets to Coordination, Workflow and Business Process Management (PNCWB) at the ICATPN 2005*, pp. 35-58.

van Dongen, B. F., Alves de Medeiros, A. K., Verbeek, H. M. W., Weijters, A. J. M. M. and van der Aalst, W. M. P. (2005), “The ProM Framework: A new era in process mining tool support”, in: Ciardo, G. and Darondeau, P. (Eds), *26th International Conference on Applications and Theory of Petri Nets (ICATPN 2005)*, Springer Verlag, Heidelberg, pp. 444-454.

van Dongen, B.F. and Adriansyah, A., (2009), “Process mining: Fuzzy clustering and performance visualisation”, in: Rinderle-Ma, S. et al. (Eds): *BPM 2009 Workshops, LNBIP 43*, Springer-Verlag Berlin pp. 158–169,

van der Aalst, W.M.P. (2007), “Trends in business process analysis”, in: *Proceedings of the 9th International Conference on Enterprise Information Systems (ICEIS) 2007*, Madeira, Institute for Systems and Technologies of Information, Control and Communication, INSTICC, Portugal, p. 12-22.

van der Aalst, W. M. P., Weijters, A. J. M. M. and Maruster, L. (2002), “Workflow mining: Which processes can be rediscovered?”, Beta Working Paper WP74, Eindhoven University of Technology, Eindhoven.

van der Aalst, W. M. P. and Song, M. (2004), “Mining social networks: Uncovering interaction patterns in business processes”, in: Desel, J., Pernici, B. and Weske, M. (Eds), *Business Process Management (BPM 2004),* Springer Verlag, Berlin, pp. 244-260.

van der Aalst, W.M.P, and Weijters, A.J.M.M. (2004), “Process Mining: a Research Agenda”, *Computers in Industry*, Vol. 53, pp. 231-244.

van der Aalst, W. M. P., Rubin, V., Verbeek, H. M. W., van Dongen, B. F., Kindler, E. and Gunther, C. W. (2010), “Process mining: A two step approach to balance between underfitting and overfitting”, *Software Systems Modelling*, Vol. 9, pp. 87-111.

van der Aalst, W. M. P. and Alves de Medeiros, A. K. (2005), “Process mining and security: Detecting process executions and checking process conformance”, *Electronic Notes in Theoretical Computer Science*, Vol. 121, pp. 3-21.

Vergidis, K., Tiwari, A., Majeed, B. and Roy, R. (2007), “Optimisation of business process designs: An algorithmic approach with multiple objectives”, *International Journal of Production Economics*, Vol. 109, pp. 105-121.

Wen, L., Wang, J., van der Aalst, W.M.P., Wang, Z., Sun, J. (2009), “A novel approach for process mining based on event types”, *Journal of Intelligent Information Systems*, Vol. 32, pp. 163-190.

Weijters, A.J.M.M, and van der Aalst, W.M.P. (2003), “Rediscovering workflow models from event-based data using little thumb”, *Integrated Computer-Aided Engineering*, Vol. 10, pp.151–162.

Weijters, A.J.M.M, and van der Aalst, W.M.P. (2001), “Process mining: Discovering workflow models from event-based data”, in: Krose, B, de Rijke, M, Schreiber, G, van Someren, M. (Eds), *Proceedings of the 13th Belgium-Netherlands Conference on Artificial Intelligence (BNAIC 2001),* pp. 283–290.

Weijters, A.J.M.M, van der Aalst, W.M.P, van Dongen, B, Günther, C, Mans, R, Alves de Medeiros , A.K, Rozinat, A, Song M, and Verbeek E. (2007), “Process Mining with ProM”, Dastani, M, and de Jong, E. (Eds), in: *Proceedings of the 19th Belgium-Netherlands Conference on Artificial Intelligence (BNAIC).*

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.

https://dspace.org/

Business process mining: from theory to practice

Turner, Christopher J.; Tiwari, Ashutosh; Olaiya, Richard; Xu, Yuchun

2012-06-01T00:00:00Z
Emerald Group Publishing Limited
http://dx.doi.org/10.1108/14637151211232669
Turner CJ, Tiwari A, Olaiya R, Xu Y. (2012) Process mining: from theory to practice. Business Process Management Journal, Volume 18, Issue 3, pp. 493 - 512

Downloaded from a DSpace repository. https://dspace.org/