

Removing Operational Friction in Shared Services Using Process Mining

Both shared services and process mining aim at improving performance and compliance of operational processes. The key idea of shared services is to share efforts and resources for processes that are common among organizations or departments. The goal is twofold: (1) increasing efficiency and reducing costs by avoiding the replication of resources and (2) improving quality and effectiveness by the industrialization of service processes. Shared-service organizations aim to provide 'economies of scale', but many shared-service projects fail because moving the work to a central location may lead to hand-offs, rework, duplication, and ineffective communication. Fortunately, process mining can be used to address these problems. Using the event data collected in any shared-service organization, one can show the real processes and uncover inefficiencies (e.g., rework), bottlenecks, and undesired deviations.

What is process mining?

The author started to work on process mining in the late nineties when he developed the first process mining techniques to discover operational processes from event data. The main motivation for looking at event data was the low quality of process models used as input for Business Process Management (BPM) and Workflow Management (WFM) projects. Processes modeled in notations such as BPMN, UML activity diagrams, or EPCs tend to oversimplify reality. Implementing BPM/WFM systems based on these simplified process diagrams are a recipe for disaster. As an example, take the Order-to-Cash (O2C) process of a large multinational that processes over 30 million ordered items per year. These 30 million cases (i.e., instances of the O2C process) generate over 300 million events per year. Over 60 different activities may occur. Although the O2C process is fairly standard, over 900,000 process variants can be observed in one year! These variants describe different ways of executing this process. This real-life example shows that traditional process modeling cannot capture the complexity of real-life operational processes.

Input for process mining is an event log. An event log 'views' a process from a particular angle. Each event in the log refers to (1) a particular process instance (called case), (2) an activity, and (3) a timestamp. There may be additional event attributes referring to resources, people, costs, etc., but these are optional. Events logs are related to process models (discovered or hand-made). Process models can be expressed using different formalisms ranging from Directly-Follows Graphs (DFGs) and accepting automata to Petri nets, BPMN diagrams, and UML activity diagrams. Typically, four types of process mining are identified.

- **Process discovery:** learning process models from event data. A discovery technique takes an event log and produces a process model without using additional information. An example is the well-known Alpha-algorithm, which takes an event log and produces a Petri net explaining the behavior recorded in the log. Most of the commercial process mining tools first discover DFGs before conducting further analysis.
- **Conformance checking:** detecting and diagnosing both differences and commonalities between an event log and a process model. Conformance checking can be used to check if reality, as

recorded in the log, conforms to the model and vice versa. The process model used as input may be descriptive or normative. Moreover, the process model may have been made by hand or learned using process discovery.

- **Process reengineering:** improving or extending the model based on event data. Like for conformance checking, both an event log and a process model are used as input. However, now, the goal is not to diagnose differences. The goal is to change the process model. For example, it is possible to repair the model to better reflect reality. It is also possible to enrich an existing process model with additional perspectives. For example, replay techniques can be used to show bottlenecks or resource usage. Process reengineering yields updated models. These models can be used to improve the actual processes.
- **Operational support:** directly influencing the process by providing warnings, predictions, or recommendations. Conformance checking can be done 'on-the-fly' allowing people to act the moment things deviate. Based on the model and event data related to running process instances, one can predict the remaining flow time, the likelihood of meeting the legal deadline, the associated costs, the probability that a case will be rejected, etc. The process is not improved by changing the model, but by directly providing data-driven support in the form of warnings, predictions, and/or recommendations.

All techniques start from the so-called control-flow perspective, which focuses on the ordering of activities. Then the time perspective (bottlenecks, delays, and frequencies), the data perspective (understanding decisions), and the resource and organization perspective (social networks, roles, and authorizations) are added.

Until 2010 there were only a few commercial process mining tools (Futura Reflect by Futura Process Intelligence, Disco by Fluxicon, and Interstage Automated Business Process Discovery by Fujitsu were notable exceptions). Since 2010 there has been a rapid increase in the number of tools and their maturity. For example, Celonis Process Mining (Celonis) was introduced in 2011, minit (Gradient ECM) was introduced in 2014, and ProcessGold Enterprise Platform (ProcessGold) was introduced in 2016. Currently, there are over 25 commercial tools available. These tools can easily deal with event logs having millions of events.

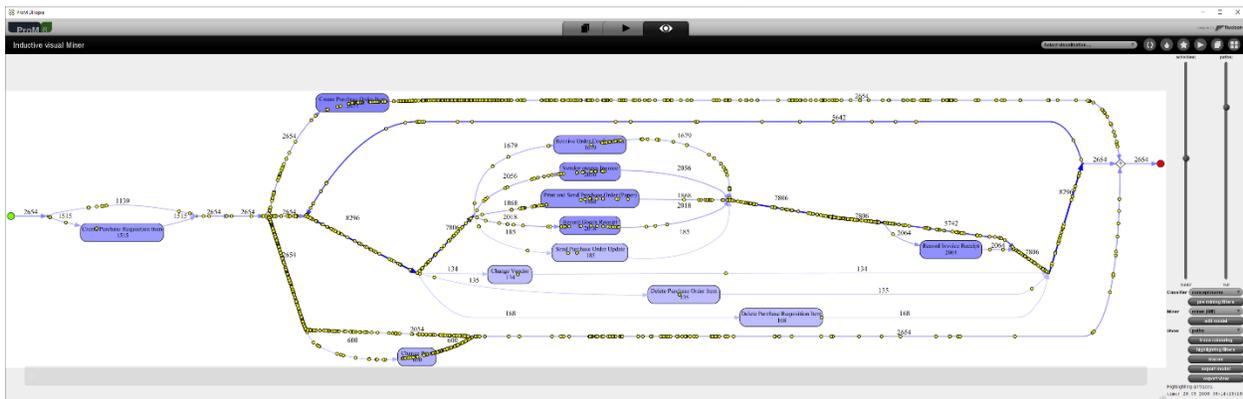


Figure 1: A process model discovered by ProM based on SAP data. The process model shows the dominant behavior in the Purchase-to-Pay (P2P) process. The numbers indicate frequencies, and the yellow dots refer to actual purchase orders.

How to remove operational friction?

Shared-service organizations aim to streamline processes and benefit from economies of scale. However, as our earlier O2C example already showed, also standard processes tend to have many variants. Thirty million cases may generate 900,000 different process variants. A process variant is a sequence of activities, also called a trace, followed by at least one case. The most frequent process variant occurred over 3 million times, but there are also variants that are rare. Typically, activities and traces (i.e., process variants) follow a Pareto distribution (also known as the "80-20 rule" or "power law"). Often, a small percentage of activities accounts for most of the events and a small percentage of trace variants accounts for most of the traces. 20% of all variants may be able to explain 80% of all cases. However, the remaining 20% of cases account for 80% of the variants. Many of these infrequent variants involve rework, passing the buck (leaving a difficult problem for someone else to deal with), communication errors, and repair actions. Some of these process variants make sense when dealing with exceptional cases. However, most deviations from the so-called 'happy path' represent 'operational friction'. Process discovery and conformance checking can reveal such operational frictions. It is possible to identify cases (1) that deviate from a normative process or that can be considered as outliers and (2) that have a poor performance (e.g., taking too long or inducing high costs). This information can be used to improve processes. After identifying sources of friction, process mining can be used in a continuous manner providing actionable information.

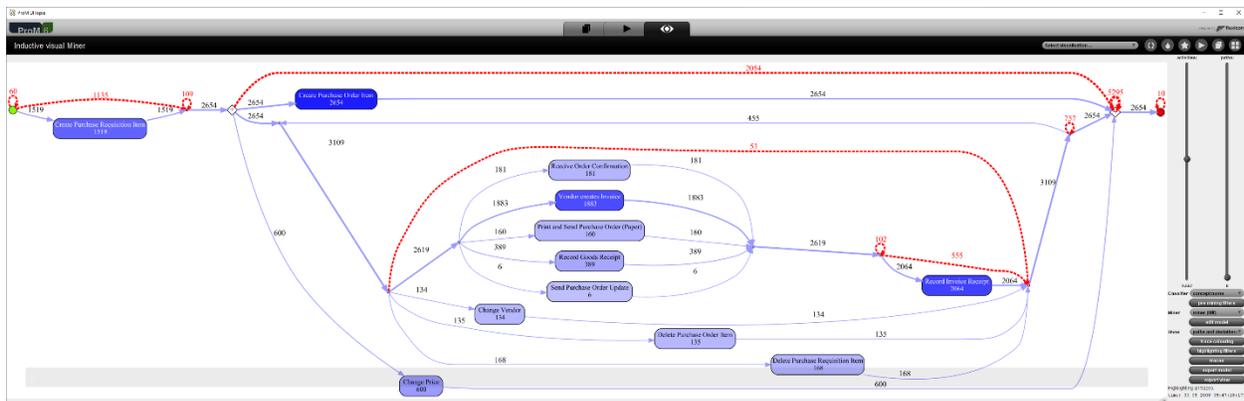


Figure 2: Conformance diagnostics provided by ProM for the Purchase-to-Pay (P2P) process using SAP data. The red arcs show deviations from the mainstream process indicated in red. It is possible to drill-down on the cases exhibiting particular deviations.

How can I start?

Next to the commercial process mining systems that are generally easy to use, one can also start with open-source software like ProM. ProM provides over 1500 plug-ins supporting process discovery, conformance checking, process reengineering, and operational support. Event data can be loaded from databases. However, it is often easier to start with a simple event log stored in CSV format or XES format. In a CSV file each row refers to an event and the columns refer to case, activity, timestamp, etc. XES is the IEEE standard for storing event data (see www.xes-standard.org) supported by tools such as ProM, Celonis, Disco, ProcessGold, Minit, QPR, and myInvenio. Several repositories provide publicly available XES data, see for example https://data.4tu.nl/repository/collection:event_logs_real.

How about RPA?

Most process mining projects do not involve Robotic Process Automation (RPA). The scope of process mining is much broader than RPA. Process mining often results in organizational and managerial changes without automation or the introduction of new IT systems. However, process mining may play a key role in successful RPA projects. RPA aims to replace people by automation done in an 'outside-in' manner. This differs from the classical 'inside-out' approach to improve information systems. Unlike traditional workflow technology, the information system remains unchanged and the robots use the same interface as the humans they are replacing or supporting. Process mining can be used to automatically visualize and select processes with the highest automation potential, and subsequently, build, test, and deploy RPA robots driven by the discovered process models.

About the first International Process Mining Conference (ICPM)

The International Conference on Process Mining (ICPM) is the first conference devoted to the rapidly growing process mining discipline. The conference will take place in Aachen, Germany from 24-26 June 2019. All thought leaders working on process mining will be present at ICPM 2019. The program includes a fully packed industry day next to exciting scientific talks by the leading scientists in the field. The conference is supported by all the main vendors providing process-mining software. However, the conference is neutral and provides insights that are tool independent. This makes ICPM 2019 'the place to be' for anyone working on process mining. Moreover, it provides a great stepping-stone of those that are new to the topic. Next to a range of excellent speakers, there will be the possibility to see the state-of-the-art in action (e.g., dozens of tool demonstrations) and discuss real-life process mining experiences. Next to the leading academics, speakers from Gartner, Siemens, Deloitte, Ernst & Young, DHL, Merck, Metronic, and many other organizations will provide novel insights in this new and exciting technology. Visit <https://icpmconference.org> to register.

About the author

Prof.dr.ir. Wil van der Aalst is a full professor at RWTH Aachen University leading the Process and Data Science (PADS) group. He is also part-time affiliated with the Fraunhofer-Institut für Angewandte Informationstechnik (FIT) where he leads FIT's Process Mining group. His research interests include process mining, Petri nets, business process management, workflow management, process modeling, and process analysis. Wil van der Aalst has published over 220 journal papers, 20 books (as author or editor), 500 refereed conference/workshop publications, and 75 book chapters. Next to serving on the editorial boards of over ten scientific journals, he is also playing an advisory role for several companies, including Fluxicon, Celonis, ProcessGold, and Bright Cape. Van der Aalst received honorary degrees from the Moscow Higher School of Economics (Prof. h.c.), Tsinghua University, and Hasselt University (Dr. h.c.). He is also an elected member of the Royal Netherlands Academy of Arts and Sciences, the Royal Holland Society of Sciences and Humanities, and the Academy of Europe. In 2018, he was awarded an Alexander-von-Humboldt Professorship (the most valuable German research award).

For the photo use http://www.padsweb.rwth-aachen.de/wvdaalst/about_me/WvdA-BvO-24059.jpg (photo is license-free, credits go to Bart van Overbeeke Photography).

