Process mining for organizational AGILITY

BY MATHIAS KIRCHMER, FRANCISCO GUTIÉRREZ AND SIGIFREDO LAENGLE

EXECUTIVE SUMMARY

Industrial engineers have techniques at their disposal that make it possible to discover the model behind an information technology-based process and explain the determinants of its performance level. These methods are known generically as process mining. In this article, the authors describe how a Chilean company named Spevi was able to apply process mining to implement a continuous improvement cycle that increased its organizational flexibility and strengthened its sustainability in a highly competitive and continuously changing business environment.

The ability to make rapid changes in business processes is an increasingly important competitive advantage for companies that operate in highly volatile and competitive environments. Such firms must be able to design, implement, execute and monitor their processes quickly if they are to satisfy customers who demand both high product quality and swift adaptation to their changing needs.

In 2008 and 2009, for example, Chile’s residential mortgage lending market had to compete on approval times as well as interest rates. Early in 2008, the average time for approving a mortgage loan was close to 90 days; in the following year many banks had reduced the cycle time to 45 days. In other words, competitive advantage in this market depends not just on the current speed of the approval process but the ability to increase that speed. This requires the organizational ability to meet the needs of customers. Rapidly changing a business process means being able to

Firms must be able to design, implement, execute and monitor their processes quickly.
(re)design, implement, execute and monitor that process in an agile and efficient manner. In effect, a company must be capable of executing the continuous improvement cycle shown in Figure 1 at high speed.

This is no simple task. Often there are two teams, one business and one technical, working on continuous improvement projects. Each has its own techniques and ideas, and often these will conflict. In other cases, the challenges of managing the changes and the employee learning process they inevitably involve also will have to be considered. This article looks at just one of the phases of the continuous improvement cycle that is particularly important for the industrial engineer: process monitoring.

Within the monitoring phase, process mining has become more important. Process mining is a technique for discovering the model describing an IT-based process using event logs or audit trails and explaining its behavior. The three guiding questions of the technique are What, How and Why. Thus, a company might want to know how many nonstandard orders were processed (the What), which was the process executed (the How) and why it occurred (the Why). Not only do we need to know what occurred, but also how it occurred and the reasons behind it.

Process monitoring has become a key factor in the ability to change rapidly and ensure organizational agility, especially for complex but highly structured processes. By employing process mining, industrial engineers can provide a solid basis for launching the continuous improvement cycle illustrated in Figure 1 rather than relying on imprecise (or even wrong) assumptions.

Process mining can be seen in the broader context of business process intelligence (BPI) and business activity monitoring (BAM). Recently, in his book *High Performance through Process Excellence — From Strategy to Operations*, Mathias Kirchmer clarified concepts of process performance manager (PPM)/process mining. BPI tools typically do not allow for process discovery (the What) and offer relatively simple performance (the How and the Why) analysis tools that depend on a correct a priori process model. Instead, we want to emphasize the discovery and analysis together regardless of the (a priori) process model available.

**The solution**

To demonstrate the power of process mining, we applied it to the real
To demonstrate the power of process mining, we applied it to the real problem of processing a company’s purchase orders. The last two steps were supported by ProM, a free software developed at Eindhoven Technical University in the Netherlands. In what follows, we take a closer look at the procedure (Figure 2).

The company in question, known as Spevi, was founded in Chile around 18 years ago to install equipment. Today it offers a wide range of related products and services. In the last five years, business has literally exploded due to the boom in construction of large buildings in the country as well as in neighboring Argentina and Peru. Spevi’s annual billings now average close to $8 million, and it receives some 300 purchase orders per month that range from the import or construction of specialized products to large equipment installation projects. The variety and complexity of the services it provides made handling work orders increasingly difficult to manage, finally convincing the firm that this process was ready for an upgrade.

A plan was designed for improving the firm’s order process that involved four steps:

1. Install “connectors” at different points in the process to capture process instances and the corresponding tasks executed.
2. Record the process instances during execution time in an event log file.
3. Process the event log file using a process mining algorithm to discover the model behind the executed process.
4. Establish the determinants of the process performance level.

The company was founded in Chile around 18 years ago to install equipment. Today it offers a wide range of related products and services. In the last five years, business has literally exploded due to the boom in construction of large buildings in the country as well as in neighboring Argentina and Peru. Spevi’s annual billings now average close to $8 million, and it receives some 300 purchase orders per month that range from the import or construction of specialized products to large equipment installation projects. The variety and complexity of the services it provides made handling work orders increasingly difficult to manage, finally convincing the firm that this process was ready for an upgrade.

A plan was designed for improving the firm’s order process that involved four steps:

1. Install “connectors” at different points in the process to capture process instances and the corresponding tasks executed.
2. Record the process instances during execution time in an event log file.
3. Process the event log file using a process mining algorithm to discover the model behind the executed process.
4. Establish the determinants of the process performance level.

The last two steps were supported by ProM, a free software developed at Eindhoven Technical University in the Netherlands. In what follows, we take a closer look at the procedure (Figure 2).

In the first step, which was the installation of connectors, we implemented the computer code for recording the execution of the process. At a minimum, this would include the activities performed in each process instance, but in our case, we also recorded the state of each activity, the employee or system performing the corresponding task, and a timestamp and links to the data processed during the task, such as purchase or work orders. The process is computationally supported by the Order Processing System, a software developed ad-hoc by Spevi on an Oracle database. Since all tasks are carried out semimanually, it was not complex to install

![Diagram illustrating process mining steps](image-url)

**Figure 2.** Process mining is carried out in four steps: installation of connectors in the business process, recording of execution instances in the event log files, discovery of the process model and establishing the determinants of the performance results.
connectors that allow an event log record with the required information.

The second step was to generate the event log file during execution time. This file contained all of the data on the process instances that had been executed. An example of an event log file is shown in Figure 3.

The third step was the discovery of the process model and the determination of performance levels. For this we used the ProM α+ algorithm, recommended for relatively well-constructed event log files with low noise levels. Over a period of five months from January 2009 to May 2009, a total of 1,330 process instances were recorded. The model discovered, slightly simplified for ease of presentation, is shown in Figure 4.

The model describes the entry of an order by a customer via a Web portal. A company employee examines the order manually and applies a heuristic to decide whether it can be processed. If the order can be fulfilled, the same employee makes a tentative offer that is sent to the customer by e-mail or regular mail. The customer must then confirm the order within a defined time limit or the process terminates. If the order cannot be fulfilled, the employee informs the customer to that effect, also by e-mail or regular mail. A third option arises after the heuristic analysis if it is found that clarification of the order is required due to incomplete information. This is done by telephone and results in either of the two previous options — that the order can or cannot be filled.

Various performance indicators can be obtained from the event log records, such as cycle times and processing costs. In our case, we measured performance in terms of the percentage of orders accepted. In this process, an accepted order may arrive by either of two paths, with or without telephone clarification. As shown in Figure 4, the proportion of instances ending in an accepted order without telephone clarification, which we will call Path 1 (P1), is 50 percent. Path 2 (P2) orders are those accepted after telephone clarification and account for 16 percent of total instances (20 percent of 80 percent). Path 3 (P3) orders are those rejected after telephone clarification and are 4 percent of the total (20 percent of 20 percent). Path 4 (P4) orders are those that are rejected...
Let us investigate further.

As a statistical study, it was necessary to explain a dependent variable (in this case “performance”) based on independent variables (“determinants”) in order to make assumptions about the independent variables and measure and analyze the relationships between them. Thanks to the knowledge of the business managers, we identified two sources of performance determinants measured using the ProM software.

To begin with, the order system offered a catalogue of products that did not always match the production possibilities and was often out of date or marred by errors. This impacted the clarity of the orders entered by customers and reflected a deficient order entry system design. The quality of telephone order clarification is very sensitive to the experience levels of the telephone operators, a factor of particular importance given the high staff turnover typical of this type of work.

The combination of these performance determinants was identified as occurrences of a random variable. If in a given instance of the order process the ordered item was not precisely defined and the telephone clarification was performed by an inexperienced operator, the event was denoted \( x_1 \). If the ordered item was imprecisely defined and the operator was experienced, the event was called \( x_2 \). If the ordered item was precisely defined but the operator was inexperienced, the event was named \( x_3 \). And if the ordered item was precisely defined and the operator was experienced, the event was denominated \( x_4 \).

The findings

We applied process mining techniques to find the relationship between the results of the process (accept or reject orders) and the determinants of the performance results \( (x_1, x_2, x_3 \text{ and } x_4) \). To accomplish this, however, we had to establish the relationship between performance (accept or reject orders) without telephone clarification and make up 30 percent of instances.

The question that arose when the results were reviewed was why the number of accepted orders was relatively low (66 percent of the total) compared to the rate expected by the company, which was 80 percent or higher. The causes were multiple and their analysis was likely to be highly complex. Was it because the order entry system was poorly designed or because customers did not know how to use it properly? Were the rules comprising the order acceptance analysis heuristic badly defined? Or was the telephone clarification step being executed incorrectly?

Let us investigate further.

As a statistical study, it was necessary to explain a dependent variable (in this case “performance”) based on independent variables (“determinants”) in order to make assumptions about the independent variables and measure and analyze the relationships between them. Thanks to the knowledge of the business managers, we identified two sources of performance determinants measured using the ProM software.

To begin with, the order system offered a catalogue of products that did not always match the production possibilities and was often out of date or marred by errors. This impacted the clarity of the orders entered by customers and reflected a deficient order entry system design. The quality of telephone order clarification is very sensitive to the experience levels of the telephone operators, a factor of particular importance given the high staff turnover typical of this type of work.

The combination of these performance determinants was identified as occurrences of a random variable. If in a given instance of the order process the ordered item was not precisely defined and the telephone clarification was performed by an inexperienced operator, the event was denoted \( x_1 \). If the ordered item was imprecisely defined and the operator was experienced, the event was called \( x_2 \). If the ordered item was precisely defined but the operator was inexperienced, the event was named \( x_3 \). And if the ordered item was precisely defined and the operator was experienced, the event was denominated \( x_4 \).

The findings

We applied process mining techniques to find the relationship between the results of the process (accept or reject orders) and the determinants of the performance results \( (x_1, x_2, x_3 \text{ and } x_4) \). To accomplish this, however, we had to establish the relationship between performance (accept or reject orders)
orders were associated with the inexperience of the operator handling the telephone clarifications ($x_1$ and $x_3$). Operators with more experience would thus bring about a reduction in order rejections.

The conclusions and perspectives
There are three main conclusions of this study. First, process monitoring is a fundamental tool in the management of organizations and the basis for ensuring organizational flexibility. It identifies what process has been executed, its level of performance and the determinants of that level. Process mining techniques have been developed to carry out these operations.

Second, process performance can be explained once the flow of task execution has been determined, which is only possible using process discovery techniques. Other approaches based solely on searching for patterns or relationships between variables without taking into account process flows are unable to solve the type of problem discussed in this study.

Third, the case analyzed here demonstrated that process mining techniques can be particularly useful with high-volume, structured and highly complex processes. And the processes have to be IT-based, thus supported through application systems that create the logs.

Industrial engineers today have a very powerful tool at their disposal. Process mining makes it possible to discover IT-based executed processes and accurately explain the determinants of their performance levels. These techniques provide a secure basis for supporting continuous improvement projects in organizations looking to improve organizational flexibility as they battle to survive in fiercely competitive and highly changeable environments.