Business Process Mining

Lecture 5: Conformance Checking

Gamal Elkoumy

Junior Research Fellow @ University of Tartu
Course Outline

Introduction

Process performance measurement & dashboards

Process mining
  • Automated process discovery
  • Conformance checking
  • Performance mining
  • Variant analysis

Simulation and “what-if” process mining

Data extraction & preparation for process mining

Predictive process monitoring

Business case analysis & project management for process mining

Trends: prescriptive monitoring, causal process mining, robotic process mining
Process Mining

- Automated Process Discovery
- Performance Mining
- Variants Analysis
- Conformance Checking
- Business rules / normative model
- Enhanced process model
- Discovered process model
- Event log
- Event log'
- Difference diagnostics
- Tropical
- Moderate to subtropical
- Cool-temperate
- Chilled
- Normal
- Give juice
- Give syrup
- Oranges
- Apples
- Come to Europe from Chile
Typical Process Mining Analysis Workflow

<table>
<thead>
<tr>
<th>Start with automated process discovery (exploratory phase)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perform sanity checks to determine if we have the right dataset (e.g. are there are incomplete cases?)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Depending on the emphasis of the project, use either:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compliance checking to find undesired deviations and exceptions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use variant analysis to dig deeper. Depending on the emphasis of the project:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use variant analysis to analyze non-compliant versus compliant cases</td>
</tr>
</tbody>
</table>
Conformance checking

Given an event log and a set of **business rules** or a **process model**, find, describe, and measure the impact of **differences** between the rules/model and the log.
Types of Conformance Checking

1. Rule-Based Conformance Checking
   • Checking that a process follows certain rules

2. Model-Based Conformance Checking
   • Comparing the behaviour captured in the process model against the one in the event log

3. Exceptional Behavior Analysis
   • Spotting anomalous behavior
Rule-Based Conformance Checking

Checking whether regulations and control measures embedded in processes are followed in practice.

Common types of compliance rules:

**Flow constraints**
- Certain activities may be mandatory e.g. a specific approval to proceed with an application.

**SLA constraints**
- Customer agreements on service quality define timeframes at the case, milestone or activity transition level.

**Resource constraints**
- Certain activities must be executed by different employees, e.g. creation and approval of an invoice.
Conformance Checking Levels

Accordingly, we can identify the following conformance checking levels:

• **Flow**: have all mandatory activities been performed? Have activities been performed in the right order and with the right frequency?

• **Temporal**: has a given timeframe between activities, process milestones or for the end-to-end process been adhered to?

• **Resource**: did everyone obey the segregation of duties?

• **Exception**: explore highly-infrequent behavior to identify possible deviations from the standard process

**Procedure:**

1. Define compliance rules (based on attribute-value pairs, activity orders or performance targets)
2. Apply one or more filters to keep cases that violate the rules in question
3. If needed, switch perspective to highlight non-compliance issues (e.g. show resources that violate a segregation of duties principle)
4. Assess impact on process performance
## Conformance Checking: Analysis Template

<table>
<thead>
<tr>
<th>What?</th>
<th>How?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Flow compliance checking</strong></td>
<td></td>
</tr>
<tr>
<td>1. Check skipped mandatory activities</td>
<td>Identify skipped activities in the process map or BPMN model</td>
</tr>
<tr>
<td>2. Check forbidden repetitions.</td>
<td>Visualize process map with the min/max frequency metrics</td>
</tr>
<tr>
<td>3. Check activity exclusion or co-occurrence relations</td>
<td>Visualize min/max statistics in Activity tab of the dashboard (EE)</td>
</tr>
<tr>
<td></td>
<td>Use Path filter (directly-follows or eventually-follows)</td>
</tr>
<tr>
<td><strong>Temporal compliance checking</strong></td>
<td>Use performance filter (to check for end-to-end SLA violations)</td>
</tr>
<tr>
<td>Check temporal constraints violations</td>
<td>Use Path filter (eventually-follows) with a duration constraint to identify violations of temporal constraints between two activities</td>
</tr>
<tr>
<td><strong>Resource compliance checking</strong></td>
<td>Use Path filter (eventually-follows) and require the same Resource (or different resources)</td>
</tr>
<tr>
<td>Check four-eyes principle or same-case-handler principle</td>
<td>Use the Attribute filter and search with “Activity” as the primary attribute and “Resource” as secondary attribute to check if a resource performed an activity they were not supposed to do.</td>
</tr>
<tr>
<td>Check if an activity has been performed by</td>
<td></td>
</tr>
<tr>
<td>a resource who should not do so.</td>
<td></td>
</tr>
<tr>
<td><strong>Model-to-log conformance checking</strong></td>
<td>Use a “model-to-log” conformance checker</td>
</tr>
<tr>
<td>Identify deviations between the event log</td>
<td>Or discover a model from a log and compare it manually with the manually designed process model</td>
</tr>
<tr>
<td>and the process model and analyze their</td>
<td></td>
</tr>
<tr>
<td>frequency and impact.</td>
<td></td>
</tr>
<tr>
<td><strong>Exception analysis</strong></td>
<td>Use abstraction slider and invert the order in which arcs/nodes are removed</td>
</tr>
<tr>
<td>Analyze highly infrequent behavior</td>
<td>(remove most frequent arcs/nodes)</td>
</tr>
</tbody>
</table>
**Flow Compliance Checking**

1. **Mandatory tasks:** Have all mandatory tasks been completed?

To check whether all your tasks have been executed at least once, change the frequency view to “Min.” in the visualization settings of your Process Discoverer.

The numbers in your activities now indicate the minimum amount of times it has been completed in all process instances.
Prescribed time frame: Has the given time frame been adhered to?

Example:

Compliance rule:
Waiting time between “Inform User” and “Repair (Simple)” should not exceed 1 hours.

“Time Interval” Filter:
> 1 hour, so that only those cases are retained that do not comply with the compliance rule.
Temporal Compliance Checking

1. Prescribed time frame: Has the given time frame been adhered to?

Performance filter can help to check for end-to-end SLA violations.

For example, retain only those cases that gets completed in at most 1 hour.
Segregation of duties: Did everyone obey the segregation of duties?

To check whether the segregation of duties has been breached, select a “Path” filter. Now, set another filter which indicates that the same resource must have completed the tasks (e.g. creating & approving an invoice).
Resource Compliance Checking

1. Unexpected resource assignments

User the Attribute filter with two attributes (primary and secondary) to retrieve cases where a given activity was performed by a given resource (or any other combination of attributes).
Model-Based Conformance Checking

Unfitting behaviour:
• Task C is *optional* (i.e. may be skipped) in the log

Additional behavior:
• The *cycle* including IGDF is not observed in the log
Exception Analysis

1. Analyze highly infrequent behavior

Invert the “abstraction” order to see outliers. By reversing the ordering, it is no longer the most frequent cases that are displayed first, but the rarest cases. This shows how much the less frequent process deviated from the standard process.
Exception Analysis

1. Analyze highly infrequent behavior – extract most infrequent case variants

Select the most infrequent case variants in the case variants filter. You will note here a lot of the exceptions: cases aborted prematurely, rework loops, etc.
Process Mining

Automated Process Discovery

Performance Mining

Conformance Checking

Variants Analysis

Enhanced process model

event log

Business rules / normative model

Discovered process model

event log'

Difference diagnostics

✓ / ✗
## Variant Analysis Template

<table>
<thead>
<tr>
<th>What?</th>
<th>How?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow comparison</td>
<td>• Discover a BPMN process model from the log of each variant. Visualize them side-by-side. Look for the gateways around each task to detect for example situations where two tasks are in parallel in one variant, but sequential in the other.</td>
</tr>
<tr>
<td>Identify notable diffs in activity flow</td>
<td></td>
</tr>
<tr>
<td>Frequency &amp; rework comparison</td>
<td>• Compare the maps of the process variants side-by-side (as above), but using the frequency overlay instead of the duration overlay. This allows us to detect the most frequent transitions between tasks. Look also for “thick” loop-backs.</td>
</tr>
<tr>
<td>Compare activity or handoff frequency and rework</td>
<td>• Use a multi-low dashboard to compare frequencies of activities, resources or other attributes.</td>
</tr>
<tr>
<td>Bottleneck comparison</td>
<td>• Use filtering interface to retain all traces that fulfill the condition to belong to a variant (then repeat for the other).</td>
</tr>
<tr>
<td>Compare the location and magnitude of bottlenecks across variants</td>
<td>• Open the maps of the process variants side-by-side and compare the dependencies (arcs) with the highest waiting times. If the maps are too complex, use the abstraction slider with the abstraction metric “average duration” and the ordering from slow (right) to fast (left). Use the arc slider to retain the slowest dependencies.</td>
</tr>
<tr>
<td></td>
<td>• Compare activity durations using side-by-side comparison of process maps or a multi-log dashboard (open the logs of 2+ variants in a single dashboard).</td>
</tr>
<tr>
<td></td>
<td>• Consider comparing the two variants using the “resource” perspective to identify handoffs between resources.</td>
</tr>
<tr>
<td></td>
<td>• Use a multi-log animation to observe the build up of bottlenecks.</td>
</tr>
</tbody>
</table>
Variant Analysis

Find differences in performance and control-flow between multiple variants of a process

A **process variant** is a subset of the executions of a process corresponding to a given product, customer, etc.
Variant Analysis

To start a detailed variant analysis, the event log has to be divided by a criterion.

The separation criterion can be selected using the filter functionality in Apromore. Apply at least two contrasting or differentiating filters after another. Example: divide by case duration:
Variant 1: Short case duration (Fast cases)
Variant 2: Long case duration (Slow cases)

After the desired filter has been applied, you may each variant as a filtered log.
After you have saved your filtered logs, the actual analysis can start. There are three options to carry out a variant analysis in Apromore.

1. **Visual comparison of the variants**

Open both filtered logs in the Process Discoverer. After that, place them side by side in your preferred browser.

This enables a first visual comparison. Our example shows that the underlying process of the slow and fast process instances clearly differs.
Variant Analysis

After you have saved your filtered logs, the actual analysis can start. There are three options to carry out a variant analysis in Apromore: comparison in dashboard, via animation, or side-by-side process map comparison.

2. Comparison in the dashboard

Select the filtered logs and simultaneously open them in the performance dashboard.

All dashboard functionalities are also available when comparing multiple variants.
Variant Analysis

You can animate two or more logs together to compare their dynamics over time.

Comparison in the animation

Select your filtered logs as well as a BPMN model and start the animation. The animation enables you to visually identify the different movements of cases through the model.
Special case: Temporal variant analysis

In order to get detailed information on the temporal performance of processes, a filter has to be set in Apromore.

Select the “Performance” filter in Apromore and take a detailed look at those cases that differ significantly from the average duration.

You can find the average cycle time under “Case duration” in the Process Discoverer and in the dashboard.

Save the above- and below-average cases in a separate event log or separate process models each to compare them and to identify differences using the Variant Analysis template.
Next Week

Automated Process Discovery

Performance Mining

Event log

Variants Analysis

Conformance Checking

Enhanced process model

Business rules / normative model

Discovered process model

Difference diagnostics

Performance Mining

Event log

Variants Analysis

Conformance Checking

Business rules / normative model

Difference diagnostics