MTAT.03.231 Business Process Management

Lectures 12 and 13
Process Monitoring and Process Mining

Marlon Dumas
University of Tartu
Where are we?

1. **Process identification**
   - Process architecture
   - Conformance and performance insights

2. **Process discovery**
   - As-is process model
   - Process analysis
   - Insights on weaknesses and their impact

3. **Process monitoring and controlling**
   - Executable process model
   - Process implementation
   - To-be process model

4. **Process redesign**
   - Insights on weaknesses and their impact
   - Process architecture

---

Where are we?
## Structure of a Business Process Event Log

<table>
<thead>
<tr>
<th>Case ID</th>
<th>Timestamp</th>
<th>Activity</th>
<th>Resource</th>
<th>Loan goal</th>
<th>Requested amt</th>
<th>Offered amt</th>
</tr>
</thead>
<tbody>
<tr>
<td>C001</td>
<td>18-10-2016</td>
<td>Check completeness</td>
<td>Sue</td>
<td>Mortgage</td>
<td>100 000</td>
<td>-</td>
</tr>
<tr>
<td>C001</td>
<td>19-10-2016</td>
<td>Check credit history</td>
<td>Sue</td>
<td>Mortgage</td>
<td>100 000</td>
<td>-</td>
</tr>
<tr>
<td>C001</td>
<td>19-10-2016</td>
<td>Calculate risk score</td>
<td>Bob</td>
<td>Mortgage</td>
<td>100 000</td>
<td>-</td>
</tr>
<tr>
<td>C001</td>
<td>20-10-2016</td>
<td>Make offer</td>
<td>Mike</td>
<td>Mortgage</td>
<td>100 000</td>
<td>70 000</td>
</tr>
<tr>
<td>C001</td>
<td>25-10-2016</td>
<td>Make offer</td>
<td>Mike</td>
<td>Mortgage</td>
<td>100 000</td>
<td>80 000</td>
</tr>
<tr>
<td>C002</td>
<td>20-10-2016</td>
<td>Check completeness</td>
<td>Sue</td>
<td>Car</td>
<td>15 000</td>
<td>-</td>
</tr>
<tr>
<td>C002</td>
<td>20-10-2016</td>
<td>Check credit history</td>
<td>Sue</td>
<td>Car</td>
<td>15 000</td>
<td>-</td>
</tr>
<tr>
<td>C002</td>
<td>22-10-2016</td>
<td>Calculate risk score</td>
<td>Elsa</td>
<td>Car</td>
<td>15 000</td>
<td>-</td>
</tr>
<tr>
<td>C002</td>
<td>24-10-2016</td>
<td>Reject application</td>
<td>Elsa</td>
<td>Car</td>
<td>15 000</td>
<td>-</td>
</tr>
<tr>
<td>C003</td>
<td>02-11-2016</td>
<td>Check completeness</td>
<td>Maria</td>
<td>Mortgage</td>
<td>30 000</td>
<td>-</td>
</tr>
<tr>
<td>C003</td>
<td>04-11-2016</td>
<td>Ask for additional data</td>
<td>Maria</td>
<td>Mortgage</td>
<td>30 000</td>
<td>-</td>
</tr>
<tr>
<td>C003</td>
<td>10-11-2016</td>
<td>Check credit history</td>
<td>Maria</td>
<td>Mortgage</td>
<td>30 000</td>
<td>-</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
What is a Process Performance Dashboard?
Process performance dashboard

A visualization of a collection of process performance measures to put into evidence information and patterns to help business users answer one or more business questions.
Operational process dashboards

- Aimed at process workers & operational managers
- Emphasis on monitoring (detect-and-respond), e.g.:
  - Work-in-progress
  - Problematic cases – e.g. overdue/at-risk
  - Resource load
Tactical dashboards

• Aimed at process owners / managers
• Emphasis on analysis and management
  • E.g. detecting bottlenecks
• Typical process performance indicators
  • Cycle times
  • Error rates
  • Resource utilization
Tactical Performance Dashboard @ Australian Insurer
Strategic dashboards

• Aimed at executives & managers
• Emphasis on linking process performance to strategic objectives

<table>
<thead>
<tr>
<th>Key Performance</th>
<th>Manage Unplanned Outages</th>
<th>Manage Emergencies &amp; Disasters</th>
<th>Manage Work Programming &amp; Resourcing</th>
<th>Manage Procurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Satisfaction</td>
<td>0.5</td>
<td>0.55</td>
<td>-</td>
<td>0.2</td>
</tr>
<tr>
<td>Customer Complaint</td>
<td>0.6</td>
<td>-</td>
<td>-</td>
<td>0.5</td>
</tr>
<tr>
<td>Customer Feedback</td>
<td>0.4</td>
<td>-</td>
<td>-</td>
<td>0.8</td>
</tr>
<tr>
<td>Connection Less Than Agreed Time</td>
<td>0.3</td>
<td>0.6</td>
<td>0.7</td>
<td>-</td>
</tr>
</tbody>
</table>
What does a performance dashboard consist of?
Process performance dashboards

A dashboard is a collection of widgets

Each widget displays at least one performance measure
E.g. The average case duration

Four types of widgets:
- Single-dimensional widgets
- Two-dimensional widgets
- Three-dimensional widgets
- N-dimensional widgets (tables)
Process performance dashboard widgets

- Single-dimensional widgets display one variable

- We can group multiple measures in one widget for convenience.
Process performance dashboard widgets

Two-dimensional widgets display two variables:

- Independent variable (x-axis) – variable with respect to which we want to analyze
- Dependent variable (y-axis) – the variable that we wish to analyze

Usually, a two-dimensional widget is a chart

- Bar chart, pie chart, etc.

Types of two-dimensional charts:

- Longitudinal charts: x-axis is time, y-axis is a performance measure
- Cross-sectional chart: x-axis is an attribute (e.g. activity, resource, country), y-axis is a performance measure
- Histogram: x-axis is a performance measure, y-axis is a “count”
Process performance dashboard widgets

Longitudinal chart

Activity instances over time

From 21.01.2020 03:14:46 to 24.01.2020 00:16:52
Claims_Management: 190 activities

Work-in-Progress (WIP)
Active cases over time
Process performance dashboard widgets

Cross-sectional chart
Process performance dashboard widgets

Cross-sectional chart

Case duration

from 10.7 hrs to 2.6 d
Claims_Management: 96
Three-dimensional widgets display three variables, usually:
- Two cross-sectional variables
- A performance measure

The third dimension can be coded in different ways:
- Heat map: intensity and/or color
- Bubble chart: size and/or intensity
- Scatter plot: useful when the third dimension is Boolean
Process performance dashboard widgets

- N-dimensional widgets display N dimensions, usually:
  - A cross-sectional variable (e.g. the customer, the case identifier, etc.)
  - Several performance measures

- In general, N-dimensional widgets are tables
Designing Process Performance Dashboard

1. Identify a user or users and define a clear question or set of questions that users will answer with this dashboard.
2. Identify the type of dashboard elements based on the insights required to answer the question.
3. Identify a type of visualization (e.g., type of chart) for the element.
4. Determine x-axis (independent variable): longitudinal time, cross-sectional (attribute).
5. Determine y-axis (dependent variable): performance measure or attribute and the aggregation function.
Demo Time!
Tools for performance dashboarding

Operational-Tactical – Business Activity Monitoring (BAM)

- Axway AMPLIFY
- VITRIA Operational Process Intelligence
- Oracle BAM
- SAP Operational Process Intelligence (OPI)

Tactical

- Business Intelligence (BI) tools: PowerBI, Qlikview, Tableau...
- Process Mining (dashboard modules): Apromore, Celonis, QPR, ...

Strategic Level – Balanced Scorecard tools

- BSC Designer Online, Quickscore, Sisense, etc.
Business Process Monitoring

Enterprise System → Event stream → Event log

Performance Dashboards

Process Mining
Process Mining

Automated Process Discovery

discovered process model

event log
Automated Process Discovery

<table>
<thead>
<tr>
<th>CID</th>
<th>Task</th>
<th>Time Stamp</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>13219</td>
<td>Enter Loan Application</td>
<td>2007-11-09 T 11:20:10</td>
<td></td>
</tr>
<tr>
<td>13219</td>
<td>Retrieve Applicant Data</td>
<td>2007-11-09 T 11:22:15</td>
<td></td>
</tr>
<tr>
<td>13220</td>
<td>Enter Loan Application</td>
<td>2007-11-09 T 11:22:40</td>
<td>-</td>
</tr>
<tr>
<td>13219</td>
<td>Compute Installments</td>
<td>2007-11-09 T 11:22:45</td>
<td>-</td>
</tr>
<tr>
<td>13219</td>
<td>Notify Eligibility</td>
<td>2007-11-09 T 11:23:00</td>
<td>-</td>
</tr>
<tr>
<td>13219</td>
<td>Approve Simple Application</td>
<td>2007-11-09 T 11:24:30</td>
<td>-</td>
</tr>
<tr>
<td>13220</td>
<td>Compute Installments</td>
<td>2007-11-09 T 11:24:35</td>
<td>-</td>
</tr>
</tbody>
</table>

BPMN process model

Process Map (directly follows graph)
A process map of an event log is a graph where:

- Each activity is represented by one node
- An arc from activity A to activity B means that B is directly followed by A in at least one trace in the log
Anatomy of a Process map

- **Activity** (and its frequency)
- **Directly-follows relation** between two activities (and its frequency)
- **Eventually-follows relation** between two activities (not visualized)
- **Activity self-loop** (rework)
- **Short loop** (possible rework)
- **Final activity** (multiple possible)

Process maps are easy to understand, but they can be misleading!
Example: Process Map with frequency

Event log

5: a, b, c, g, e, h
4: a, b, c, f, g, h
3: a, b, d, g, e, h
3: a, b, d, e, g, h
Example: Process Map
with frequency

Event log

5: a,b,c,g,e,h
4: a,b,c,f,g,h
3: a,b,d,g,e,h
3: a,b,d,e,g,h
Example: Process Map
with frequency

Event log

5: a,b,c,g,e,h
4: a,b,c,f,g,h
3: a,b,d,g,e,h
3: a,b,d,e,g,h
Example: Process Map
with frequency

Event log

5: a,b,c,g,e,h
4: a,b,c,f,g,h
3: a,b,d,g,e,h
3: a,b,d,e,g,h
Example: Process Map with frequency

Event log

5: a,b,c,g,e,h
4: a,b,c,f,g,h
3: a,b,d,g,e,h
3: a,b,d,e,g,h
Process Maps

• Commercial process mining tools use process maps as the main visualization technique for event logs

• These tools also provide three types of operations:
  1. **Abstract** the process map:
  2. **Filter** the traces in the event log...
Examples

Patient Treatment Process @ Hospital (Sepsis infections)

IT Incident Management @ Bank
Common process mining tools provide abstraction capabilities on top of process maps, to:

- Show only most frequent activities
- Show only most frequent arcs

Some tools offer further abstraction capabilities to:

- Show only least frequent activities
- Show only least frequent arcs
- Show only fastest/slowest activities
- Show only fastest/slowest arcs
Log filters

- **Case-Level Log Filters**
  - Retain (or remove) all cases that fulfill a given condition, for example all cases that take more than 3 days.

- **Event-Level Filter**
  - Retain or remove individual events (*activity instances*) that fulfill a given condition, e.g. all activity instances that were performed by “John Smith”

- A case-level filter either retains an entire case or it removes the whole case while an event-level filter may remove some events in a case, and remove others.
Case-Level log filters

Attribute filters

- Retain only cases such that at least one event fulfils a condition over one or more attributes, for example: at least one event where Activity = “Close order” and Resource = “Anti Alman”

Timeframe filter

- Retain traces that start, end, or are contained in a time period (e.g. all cases that started in January)

Performance filter

- Retain only cases that have a duration above or below a given value (e.g. 3 days)

Path filter

- Retain traces where there is a pair of events that fulfil a given condition (e.g. “Create invoice” eventually followed by “Create purchase order”)

Repetition filter

- Retain traces that start with or finish with an event that fulfils a given condition
Event-Level log filters

Attribute filters

• Retain only events that fulfil a condition, for example: \textit{at least one event where Activity = “Close order”}

Timeframe filter

• Retain all events that happen during a certain time period (e.g. all events in January)

Between filter

• For each case, retain all events between the (first or last) occurrence of an activity and the (first or last occurrence) of another activity (e.g. between “Create PO” and “Approve PO”
Demo Time!
Limitations of Process Maps

• Process maps over-generalize: some paths of a process map might not exist and might not make sense
  • Example: Draw the process map of \([ abc, adc, afce, afec ]\) and check which traces it can recognize, which don’t make sense given the log.

• Process maps make it difficult to distinguish conditional branching, parallelism, and loops.
  • See previous example... or a simpler one: \([abcd, acbd]\)

• Solution: automated BPMN process discovery
Process Mining Using BPMN

• In addition to discovering a process map from an event log, process mining tools allows us to discover BPMN models.

• This allows us to distinguish between conditional branching, parallelism, and loops

• To get an intuition of how this works, let’s draw the process maps of the following logs and let’s try to imagine what is the BPMN process model corresponding to these logs:
  • [abc, abd]
  • [abcd, acbd]
  • [abcd, abcbcd]
1. [abc, abd]
2. [abcd, acbd]
3. [abcd, abcbcd]
Discovery of BPMN models

Note: Apromore uses an algorithm called **Split Miner** to turn process maps into process models.
Demo Time!
Process Mining

Automated Process Discovery

Discovered process model

event log

Conformance Checking

✅ / ✗

Process model or compliance rules
Conformance Checking (model-based)

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

Event log:
- ABCDEH
- ACBDEH
- ABCDFH
- ACBDFH
- ABDEH
- ABDFH
Conformance checking (rule-based)

• Oftentimes, we don’t have a full process model available

• Instead, we have a set of compliance rules, e.g.
  • Certain tasks MUST be executed in every case
  • Certain tasks can be executed at most once
  • When a task A occurs, another task B must also occur (within e.g. 2 hours)
  • Two tasks A and B CANNOT be executed by the same person (four-eyes principle)
Demo Time!
Process Mining

Automated Process Discovery

Performance Mining

Conformance Checking

event log

discovered process model

Enhanced process model

input process model

✓ / ✗
Performance Mining

• Process mining tools typically support performance mining in three ways:
  • Performance-enhanced process maps
  • Dashboards
  • Log animation
Performance-Enhanced Process Map

Activity label (and its duration)

Create Purchase Requisition
30.4 mins

Analyze Purchase Requisition
6.5 mins

Create Request for Quotation Requester
10.4 mins

Create Request for Quotation Requester Manager
2 mins

Analyze Request for Quotation
23 mins

Amend Request for Quotation Requester
9.8 mins

Send Request for Quotation to Supplier
22.7 mins

Amend Request for Quotation Requester Manager
19.3 mins

Directly-follows relation between two labels (with waiting time)

Bottleneck node (all incoming arcs are slow)
Demo Time!
Process Mining

Automated Process Discovery

Performance Mining

Event log

Variants Analysis

Event log

Conformance Checking

Enhanced process model

Difference diagnostics

Input process model

discovered process model

✓ / ✗
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.
Case Study: Suncorp Group

• General & life insurance, banking, superannuation and investments management
• 9M customers
• 16K employees
• $85 billion in assets
End to end insurance process

Each process is varied by product & brand...

Total process variants: 3,000+

Source: Guidewire reference models
Case Study: Variants Analysis at Suncorp

Expected Performance Line

NOTE: x and y axis are logarithmically-scaled

- + (simple and quick) - expected
- □ (complex and slow) - expected
- x (simple and slow) - needs improvement
- *(complex and quick) - good
Variants Analysis via Process Map Comparison

Simple claims and quick  

Simple claims and slow

S. Suriadi et al.: Understanding Process Behaviours in a Large Insurance Company in Australia: A Case Study. CAiSE 2013
Variant Analysis

In process mining tools, such as Apromore, variant analysis can be approached as follows:

1. First, use filtering to “slice” the event log into multiple logs (one per variant).

2. Then compare the multiple logs using:
   ▪ Side-by-side comparison of process maps or process models
   ▪ A comparative multi-log dashboard
   ▪ A multi-log animation

![Image of bar charts showing resource usage over time for different complexity levels: Simple and Complex]
Demo Time!

http://apromore.cs.ut.ee
Summary: Four Process Mining Capabilities

The bulk of process mining techniques can be divided into four capabilities.

**Automated Process Discovery**

By using the automated process discovery, an actual process model is created from the event log. This represents the current sequence of work steps and does not necessarily have to correspond to the target process.

**Conformance Checking**

Conformance checking provides a comparison of the actual and the target process and thus enables the detection of differences and deviations.

**Variant Analysis**

Process outcomes often differ in practice. But why? The variant analysis enables a simple comparison of several process variants and serves as the basis for process optimization.

**Performance Mining**

By applying performance mining techniques, e.g. nodes or cycle times can be highlighted in color to simplify the subsequent analysis.