LTAT.05.015

Business Process Mining

Week 14:

Recap and Advanced Topics in Process Mining

Marlon Dumas

Professor of Information Systems @ University of Tartu
Co-founder @ Apromore
Process Mining Recap

historical data

live data

Enterprise systems, IoTs...

process mining algorithms

process insights

process model
Automated Process Discovery

Process Map (directly follows graph)

BPMN process model

<table>
<thead>
<tr>
<th>CID</th>
<th>Task</th>
<th>Time Stamp</th>
<th>…</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>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>
Process Mining Recap

- Historical data
- Live data
- Process mining algorithms
- Process insights

Enterprise systems, IoTs...

Process model
Performance measurements
Performance Mining

Process map with duration overlay

Process performance dashboards
Process Mining Recap

Enterprise systems, IoTs...

historical data

live data

process mining algorithms

process insights

process model

performance measurements

conformance report
Conformance Checking

Modeled Process
(Expected: 8 Hours)

Actual Process
(In reality: 18 Hours)
Process Mining Recap

- Historical data
- Live data
- Process mining algorithms
- Process insights
- Process model
- Performance measurements
- Conformance report
- Differences, root causes

Enterprise systems, IoTs...
Case: Process Mining @ Nordic financial company

**Context:**
Mid-sized European payment systems provider operating in multiple countries

**Goal:**
Analysis of customer onboarding and customer support processes (B2B sales)

**Questions:**
Why are we performing in terms of customer satisfaction and resolution times better in some countries and for some customer segments and not for others?

**Data sources:**
SAP CRM and ServiceNow, centralized via a data warehouse solution

**Timeframe:**
8 weeks of data extraction & analysis, continued use afterwards
Case: Process Mining @ Nordic financial company

We found a dozen patterns including:
- performing some activities earlier in the process lead to better customer feedback
- Certain bounce-backs between teams negatively affected feedback.

Outcomes (after ca. 6 months)
• Process changes leading to reductions in customer onboarding time of several days in lower-performing countries
• Changes leading to reductions in rework loops, increase in NPS
• Analysts are able to perform regular review of the process in days, instead weeks (more than 3 x speed-up)
From Process Mining to Augmented BPM

Descriptive Process Mining
• Automated Process Discovery & Analysis

Augmented BPM
• Predictive & Prescriptive Process Monitoring
• Digital Twins & Automated Process Improvement

Prescriptive Analytics
Predictive Analytics
Diagnostic Analytics
Descriptive Analytics
## Augmented BPM: Opportunities

<table>
<thead>
<tr>
<th>Operational Level</th>
<th>Predictive Process Monitoring</th>
<th>Prescriptive process monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Predicting future states, outcomes, or properties of a process instance or group of process instances</td>
<td>Recommending actions to optimize one or more performance indicators, taking into account their causal effect</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tactical Level</th>
<th>Digital Twins for What-If Process Analysis</th>
<th>Automated Process Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Discovering accurate and reliable simulation scenarios to evaluate improvement options</td>
<td>Automatically discovering optimal sets of improvement opportunities from data to optimize one or more performance indicators</td>
</tr>
</tbody>
</table>
Predictive Process Monitoring

Running case: A B C ...

Current state

Time

Completion time

Predict Process Outcome
Is this loan offer going to be rejected?

Predict Process Performance
Will this claim take more than 5 days to be handled?

Predict Future Events
What activity is likely to be executed next?
And after that?
Predictive Process Monitoring

Enterprise System

Database

Event log

Predictive models

Event stream

Detailed predictive dashboards
Challenges in Predictive Process Monitoring

Explaining predictions
- Helping managers to understand the reasons behind a prediction

Turning predictions into actions
- Determining which actions can help to prevent a negative outcome effectively

Turning predictions into gain
- Triggering actions that improve a KPI in a cost and resource-aware manner
Prescriptive Process Monitoring

- Event log (completed traces)
- Predictive model(s)
- Cost model
- Ongoing case
- Apply
- Prediction
- Warning + Recommendation
- Recommender System
- P(X)
- +/- -
- - +
## Augmented BPM: Opportunities

<table>
<thead>
<tr>
<th>Operational Level</th>
<th>Predictive Process Monitoring</th>
<th>Prescriptive process monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Predicting future states, outcomes, or properties of a process instance or group of process instances</td>
<td>Recommending actions to optimize one or more performance indicators, taking into account their causal effect</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tactical Level</th>
<th>Digital Twins for What-If Process Analysis</th>
<th>Automated Process Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Discovering accurate and reliable simulation scenarios to evaluate improvement options</td>
<td>Automatically discovering optimal sets of improvement opportunities from data to optimize one or more performance indicators</td>
</tr>
</tbody>
</table>
“What-If” Process Mining

How to determine if a given intervention/change will improve a business process, and by how much?
What-If Process Mining

Enterprise System → Process Constraints or Process Model → Simulation Model Discoverer → Simulation Model → Simulation Engine → Predicted Performance Profile & Reliability Estimate

https://github.com/AutomatedProcessImprovement/Simod
SIMOD: Simulation Model Discovery from Event Logs

Event log → Process Model Discovery → Model enhancement → Simulator

- Control Flow Discovery
- Model-to-trace alignment & repair

Simulation parameters extraction
BPS model assembly

Simulation

Tuning

Hyperparameter optimizer

Accuracy assessment

Generated

Vs.

Ground truth

https://github.com/AutomatedProcessImprovement/Simod
## Hyper-Parameter Tuning

<table>
<thead>
<tr>
<th>Phase</th>
<th>Category</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control flow discovery</td>
<td></td>
<td>Parallelism threshold (ε)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Filtering threshold (η)</td>
</tr>
<tr>
<td></td>
<td>Parameters for log repair</td>
<td></td>
</tr>
<tr>
<td>Simulation parameters discovery</td>
<td></td>
<td>Thresholds for resource pool discovery</td>
</tr>
<tr>
<td></td>
<td>Parameters for fitting temporal distributions</td>
<td></td>
</tr>
</tbody>
</table>

Paired Damerau-Levenshtein (DL) distance with penalty for temporal mismatch

Test Fold of Event-Log

Simulated Log
SIMOD: Empirical Evaluation Procedure

Particle 1

Training

<table>
<thead>
<tr>
<th>σ₁</th>
<th>σ₂</th>
<th>σ₃</th>
<th>σ₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ₅</td>
<td>σ₆</td>
<td>σ₇</td>
<td>σ₈</td>
</tr>
<tr>
<td>σ₉</td>
<td>σ₁₀</td>
<td>σ₁₁</td>
<td>σ₁₂</td>
</tr>
<tr>
<td>σ₁₃</td>
<td>σ₁₄</td>
<td>σ₁₅</td>
<td>σ₁₆</td>
</tr>
</tbody>
</table>

Generated log

<table>
<thead>
<tr>
<th>σ'₁₂</th>
<th>σ'₁₃</th>
<th>σ'₁₄</th>
<th>σ'₁₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ'₁₆</td>
<td>σ'₁₇</td>
<td>σ'₁₈</td>
<td>σ'₁₉</td>
</tr>
<tr>
<td>σ'₂₀</td>
<td>σ'₂₁</td>
<td>σ'₂₂</td>
<td>σ'₂₃</td>
</tr>
<tr>
<td>σ'₂₄</td>
<td>σ'₂₅</td>
<td>σ'₂₆</td>
<td>σ'₂₇</td>
</tr>
<tr>
<td>σ'₂₈</td>
<td>σ'₂₉</td>
<td>σ'₃₀</td>
<td>σ'₃₁</td>
</tr>
</tbody>
</table>

Partition 2 (Testing)

<table>
<thead>
<tr>
<th>σ₂₁</th>
<th>σ₂₂</th>
<th>σ₂₃</th>
<th>σ₂₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ₂₅</td>
<td>σ₂₆</td>
<td>σ₂₇</td>
<td>σ₂₈</td>
</tr>
<tr>
<td>σ₂₉</td>
<td>σ₃₀</td>
<td>σ₃₁</td>
<td>σ₃₂</td>
</tr>
</tbody>
</table>

Timeline

- Delete(A₁)
- Substitute(A₂)
- Insert(A₄)
- Transpose(A₈ ↔ A₇)

Damerau-Levenshtein distance (DL)
Control-Flow Similarity (CLFS)
Business Process Trace Distance (BPTD)
Event Log Similarity (ELS)

Error in waiting and processing time
No penalization if parallel activities
Transposing (A₈ ↔ A₇)
Transposing (A₆ ↔ A₇)
## SIMOD: Evaluation Results

**Discover simulation model**  
**Simulate model 10 times**  
**Evaluate Similarity (mean string-edit distance & timed-string edit distance)**

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Control-Flow Similarity (string-edit distance)</th>
<th>Temporal Similarity (timed-string edit distance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call centre</td>
<td>0.37</td>
<td>0.41</td>
</tr>
<tr>
<td>Pharmacy customer service</td>
<td>0.29</td>
<td>0.30</td>
</tr>
<tr>
<td>Purchase-to-Pay</td>
<td>0.55</td>
<td>0.57</td>
</tr>
<tr>
<td>Make-to-order manufacturing</td>
<td>0.65</td>
<td>0.69</td>
</tr>
<tr>
<td>Academic credentials recognition</td>
<td>0.32</td>
<td>0.29</td>
</tr>
<tr>
<td>Insurance claims handling</td>
<td>0.39</td>
<td>0.43</td>
</tr>
<tr>
<td>Loan Origination</td>
<td>0.41</td>
<td>0.42</td>
</tr>
</tbody>
</table>
Skip credit history check when customer has previous loans with bank.

Allocate an additional clerk on Monday-Tuesdays, one less officer on Fridays.

This task can be automated with an RPA script.

If loan-to-annual-income ratio > 1.5, allocate a senior officer.

If credit rating is C or D, do not wait for appeal.

For consumer loans, check credit history before income sources.

Make credit offer

Notify rejection

Receive customer feedback

Assess application

Check credit history

Check income sources

Credit application received

entry point

exit point

entry point

exit point

exit point

entry point

exit point
Automated Process Improvement

Enterprise System

Domain Knowledge

Search-Based Process Optimizer

Cycle time vs. Resource cost

Improvement Opportunities