Process Mining: Extension Mining Algorithms

Ana Karla Alves de Medeiros

Eindhoven University of Technology
Department of Information Systems
a.k.medeiros@tue.nl
Process Mining

• Short Recap
• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets
• Summary
• Announcements
• Presentation Futura Technology
Process Mining

• Short Recap

• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets

• Summary

• Announcements

• Presentation Futura Technology
Types of Algorithms

“world”
- business processes
- people
- machines
- components
- organizations

models
analyzes

supports/
controls

specifies
configures
implements

discovery
conformance
extension

information
system

records
events, e.g.,
messages,
transactions,
etc.

analyzer

(process)
model

event
logs

Process Mining Tools
Types of Algorithms

- Organizational Miner
- Social Network Miner
- Analyze Social Network

“world”
- business processes
- people
- machines
- components
- organizations

supports/controls

information system

models analyzes

specifies configures implements

analyzes

records events, e.g., messages, transactions, etc.

Process Model

Organizational Model

Social Network

event logs

/faculteit technologie management
Types of Algorithms

Compliance Process Model

Auditing/Security

Conformance Checker

LTL Checker
Main Points Lecture 4

• Organizational mining plug-ins can discover
  – Roles/Teams in organizations
  – Social networks for originators
• Some metrics of social networks are based on ordering relations (e.g., the ordering relations used by the Alpha algorithm)
• Conformance Checker assesses how much a process model matches process instances
• LTL Checker uses logics to verify properties in event logs
Process Mining

• Short Recap
• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets
• Summary
• Announcements
• Presentation Futura Technology
Process Mining

• Short Recap
• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets
• Summary
• Announcements
• Presentation Futura Technology
Types of Algorithms

“world”
- business processes
- people
- machines
- components
- organizations

models
analyzes
(specify configures implements)
supports/controls
analyzes

information system

records events, e.g., messages, transactions, etc.

(process) model

(event) logs

Process Mining Tools

Discovery
Conformance
Extension

Performance Analysis

Bottlenecks/Business Rules/Process Model

/ faculteit technologie management
Process Mining

• Short Recap

• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets

• Summary

• Announcements

• Presentation Futura Technology
Process Mining

• Short Recap
• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets
• Summary
• Announcements
• Presentation Futura Technology
Decision Point Analysis: Main Idea

• Detection of data dependencies that affect the routing the routing of process instances

Which conditions influence the choice between a full check and a policy only one?
Decision Point Analysis: Motivation

- Make tacit knowledge explicit
- Better understand the process model
Decision Point Analysis: Motivation

(amount > 500) AND (policyType = normal)

(amount <= 500) OR (policyType = premium)

status = approved

status = rejected

amount = R
clientID = String
policyType = normal | premium
status = approved | rejected

data modification
data dependency
Decision Point Analysis: Algorithm's Main Steps

1. Read a log + model
2. Identify the decision points in a model
3. Find out which alternative branch has been taken for a given process instance and decision point
4. Discover the rules for each decision point
5. Return the enhanced model with the discovered rules
Decision Point Analysis: Algorithm's Main Steps

1. Read a log + model
2. Identify the decision points in a model
3. Find out which alternative branch has been taken for a given process instance and decision point
4. Discover the rules for each decision point
5. Return the enhanced model with the discovered rules

How can we spot the decision points in a Petri net?
Decision Point Analysis:
Algorithm's Main Steps

1. Read a log + model
2. Identify the decision points in a model
3. Find out which alternative branch has been taken for a given process instance and decision point
4. Discover the rules for each decision point
5. Return the enhanced model with the discovered rules
Quick Recap Lecture 1: Decision Trees

**Illustration (10 learning examples):**

<table>
<thead>
<tr>
<th>Hair</th>
<th>Length</th>
<th>Weight</th>
<th>Suntan cream</th>
<th>Burned</th>
</tr>
</thead>
<tbody>
<tr>
<td>blond</td>
<td>medium</td>
<td>light</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>blond</td>
<td>medium</td>
<td>light</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>blond</td>
<td>long</td>
<td>medium</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>brown</td>
<td>long</td>
<td>light</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>red</td>
<td>small</td>
<td>heavy</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>brown</td>
<td>long</td>
<td>light</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>blond</td>
<td>medium</td>
<td>heavy</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>brown</td>
<td>small</td>
<td>heavy</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

**Attributes**

<table>
<thead>
<tr>
<th>Hair</th>
<th>Length</th>
<th>Weight</th>
<th>Suntan cream</th>
</tr>
</thead>
<tbody>
<tr>
<td>blond</td>
<td>long</td>
<td>heavy</td>
<td>no</td>
</tr>
<tr>
<td>brown</td>
<td>long</td>
<td>light</td>
<td>yes</td>
</tr>
<tr>
<td>brown</td>
<td>small</td>
<td>light</td>
<td>yes</td>
</tr>
</tbody>
</table>

**Classes: Yes/No**

**New (test) examples:**

<table>
<thead>
<tr>
<th>Hair</th>
<th>Length</th>
<th>Weight</th>
<th>Suntan cream</th>
<th>Burned</th>
</tr>
</thead>
<tbody>
<tr>
<td>red</td>
<td>medium</td>
<td>light</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>blond</td>
<td>medium</td>
<td>medium</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>brown</td>
<td>small</td>
<td>light</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>
Decision Point Analysis: Algorithm's Main Steps

1. Read a log + model
2. Identify the decision points in a model
3. Find out which alternative branch has been taken for a given process instance and decision point
4. Discover the rules for each decision point
5. Return the enhanced model with the discovered rules

Which elements are the classes and which are the attributes?
Step 4

Training examples for decision point "p0"

<table>
<thead>
<tr>
<th>amount</th>
<th>clientID</th>
<th>policyType</th>
<th>class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>C567894938</td>
<td>premium</td>
<td>C</td>
</tr>
<tr>
<td>700</td>
<td>C938609223</td>
<td>normal</td>
<td>B</td>
</tr>
<tr>
<td>550</td>
<td>C135697567</td>
<td>normal</td>
<td>B</td>
</tr>
<tr>
<td>500</td>
<td>C568120443</td>
<td>normal</td>
<td>C</td>
</tr>
<tr>
<td>50</td>
<td>C493823084</td>
<td>normal</td>
<td>C</td>
</tr>
<tr>
<td>200</td>
<td>C945675110</td>
<td>premium</td>
<td>C</td>
</tr>
</tbody>
</table>

Discovered decision tree for point "p0"
Decision Point Analysis: Example in ProM
Decision Point Analysis: Example in ProM

[Diagram showing a decision tree with nodes for Amount and PolicyType, branch decisions based on conditions.]
Decision Point Analysis

(process) model

discovery
conformance
extension

event logs

ProM

Model
Attributes
Log
Algorithm
Decision Tree/Rules
Evaluation
Result

Decision points
Choice 4 "p9"
Choice 5 "p2"
Choice 6 "p3"

Register Claim complete

PolicyType

Amount

(#Amount data > 500) and also (#PolicyType data = normal)

(#Amount data <= 500)

ORElse

(#Amount data > 500) and also (#PolicyType data = premium)

Check all complete

Check policy only complete

Update results

faculteit technologie management
Process Mining

• Short Recap
• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets
• Summary
• Announcements
• Presentation Futura Technology
Process Mining

• Short Recap
• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets
• Summary
• Announcements
• Presentation Futura Technology
Performance Analysis with Petri Nets

• Motivation
  – Provide different Key Performance Indicators (KPIs) relating to the execution of processes

• Main idea
  – Replay the log in a model and detect
    • Bottlenecks
    • Throughput times
    • Execution times
    • Waiting times
    • Synchronization times
    • Path probabilities etc
Bottlenecks – Waiting Times and Execution Times

How can we spot the difference between waiting and execution times?
Bottlenecks – Throughput Times

Performance information of the selected transitions:

<table>
<thead>
<tr>
<th>Frequency (cases)</th>
<th>Time in between (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>avg</td>
<td>66.59</td>
</tr>
<tr>
<td>min</td>
<td>29.0</td>
</tr>
<tr>
<td>max</td>
<td>150.0</td>
</tr>
<tr>
<td>std</td>
<td>19.45</td>
</tr>
<tr>
<td>fast25</td>
<td>45.7</td>
</tr>
<tr>
<td>low25</td>
<td>33.87</td>
</tr>
<tr>
<td>norma33</td>
<td>63.39</td>
</tr>
</tbody>
</table>

Throughput time (minutes):

- avg: 86.59
- min: 29.0
- max: 150.0
- std: 19.45
- fast25: 45.7
- low25: 33.87
- norma33: 63.39

Change

Percentages

Export

Time-Metrics

Selected:

- Transition - Register complete
- Transition - Archive Repair complete
Bottlenecks – Synchronization Times
Bottlenecks – Synchronization Times

What are these average synchronization times telling us?

1.3 minutes

20.8 minutes
Bottlenecks – Path Probabilities

What are these path probabilities telling us?
Performance Analysis with Petri Nets
Process Mining

• Short Recap
• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets
• Summary
• Announcements
• Presentation Futura Technology
Process Mining

• Short Recap
• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets
• Summary
• Announcements
• Presentation Futura Technology
Summary

• Extension techniques enhance existing models with information discovered from event logs

• The *Decision Point Analysis* plug-in can discover the “business rules” for the moments of choice in a process model

• The *Performance Analysis with Petri Nets* plug-in provides various KPIs w.r.t. the execution of processes

• The results of both techniques can be used to create simulation models for CPN Tools
Process Mining

- Short Recap
- Extension Techniques
  - Decision Miner
  - Performance Analysis with Petri Nets
- Summary
- Announcements
- Presentation Futura Technology
Process Mining

• Short Recap
• Extension Techniques
  – Decision Miner
  – Performance Analysis with Petri Nets
• Summary
• Announcements
• Presentation Futura Technology
Announcements

• Assignment 5
  – Individual assignment
  – Q&A session during Instruction 5
  – Posting of Report with Answers
    • Digital version at StudyWeb (folder Assignment 5)
    • Printed version to be delivered at secretary’s office of IS group (room Pav D3)
      – There will be a box on the desk
    • Deadline: March 14th, 2008 at 6pm

• Invited talk after the break!