Activation Patterns
and Event Correlation
Peter Teufl, Udo Payer, Reinhard Fellner
TOC

• Event Correlation, Knowledge/Data Mining/Extraction
• AI and Machine Learning methods
• Activation Patterns
• Examples
  • Cars
  • IDS
Event Correlation

- Events reported by multiple sensors
- Analysis
- Knowledge Mining
Event correlation - an overview/example

- **Protocol Sensor**
  - HTTP
  - HTTPS
  - SSH

- **Direction**
  - in-in
  - in-out
  - out-in

- **Stats Sensor**
  - Connection duration: 10 sec
  - Packet Loss: 90%
  - Connection duration: 12 sec
  - Packet Loss: 10%
  - Connection duration: 110 sec
  - Packet Loss: 95%

- **Sensor Table**
  - HTTP: 12 sec, 10%, in-out
  - HTTP: 100 sec, 90%, in-in
  - SSH: 12 sec, 10%, out-in

- **Event Correlator**

- **Analysis Methods**
  - Unsupervised Clustering
  - Anomaly Detection
  - Semantic Search Queries
  - Knowledge Mining
  - Feature Relevance
  - Event Relations
  - Fingerprinting

**Example Events**

- HTTP: 10 sec, 10%
- HTTPS: 100 sec, 90%
- SSH: 12 sec, 10%

**Statistics**

- Packet Loss: 10%
Thoughts

Arbitrary sensors?
Thoughts

Arbitrary sensors?

Symbolic features?
Thoughts

Arbitrary sensors?

Distance-based features?

Symbolic features?
Thoughts

- Arbitrary sensors?
- Distance-based features?
- Combine symbolic and distance-based features?
- Symbolic features?
Thoughts

- Arbitrary sensors?
- Distance-based features?
- Combine symbolic and distance-based features?
- Value ranges?
- Symbolic features?
Thoughts

- Arbitrary sensors?
- Distance-based features?
- Combine symbolic and distance-based features?
- Value ranges?
- Symbolic features?
- Normalization?
Thoughts

Arbitrary sensors?

Normalization?

Distance-based features?

Analysis?

Combine symbolic and distance-based features?

Value ranges?

Symbolic features?
Thoughts

- Arbitrary sensors?
- Distance-based features?
- Value ranges?
- Symbolic features?
- Combine symbolic and distance-based features?
- Normalization?
- Analysis?
- Machine Learning?
Thoughts
AI and Machine Learning

- Spreading Activation
- Euclidean
- Cosine
- Similarity Measures
- Unsupervised Learning
- Neural Gas Family
- SOM Family
- HAC
- Supervised Learning
- Decision Trees
- Neural Networks
- Support Vector Machines
- Semantic Networks
- Associative Networks
Activation Patterns

- **L1**: Feature Extraction
- **L2**: Node Gen.
- **L3**: Network Gen.
- **L4**: SA
- **L5**: Analysis
Activation Patterns

- L1: Feature Extraction
- L2: Node Gen.
- L4: SA
- L5: Analysis

Event Correlation
Activation Patterns

- L1: Feature Extraction
- L2: Node Gen.
- L4: SA
- L5: Analysis

Analysis of Activation Patterns

Spreading Activation

Associative Network

Generation of nodes for the associative network

Nominal Features
Nominal Features
NG
NG
Distance-Based Features
Distance-Based Features

Event Correlation

e-Participation
Text analysis

Friday, February 19, 2010
Activation Patterns

- L1: Feature Extraction
- L2: Node Gen.
- L4: SA
- L5: Analysis

Analysis of Activation Patterns

Spreading Activation

Associative Network

Generation of nodes for the associative network

Nominal Features
Nominal Features
NG
NG

Distance-Based Features
Distance-Based Features

e-Participation
Text analysis

Twitter Mining

Event Correlation

Friday, February 19, 2010
Activation Patterns

- L1: Feature Extraction
- L2: Node Gen.
- L4: SA
- L5: Analysis

Analysis of Activation Patterns

Spreading Activation

Associative Network

Generation of nodes for the associative network

Nominal Features

Nominal Features

NG

NG

Distance-Based Features

Distance-Based Features

Twitter Mining

Event Correlation

e-Participation

Text analysis

User Tracking in WIFI Networks

Friday, February 19, 2010
Activation Patterns

- L1: Feature Extraction
- L2: Node Gen.
- L4: SA
- L5: Analysis

• Spreading Activation
• Associative Network
• Generation of nodes for the associative network
• Nominal Features
• Nominal Features
• NG
• NG
• Distance-Based Features
• Distance-Based Features

Twitter Mining

Event Correlation

RDF data
CIA World Factbook

e-Participation
Text analysis

User Tracking in WIFI Networks

Friday, February 19, 2010
Activation Patterns

- L1: Feature Extraction
- L2: Node Gen.
- L4: SA
- L5: Analysis

Malware Analysis

Generation of nodes for the associative network

Nominal Features

Distance-Based Features

NG

NG

Analysis of Activation Patterns

Spreading Activation

Associative Network

Twitter Mining

Event Correlation

RDF data

CIA World Factbook

e-Participation

Text analysis

User Tracking in WIFI Networks

Friday, February 19, 2010
Activation Patterns: L5

- Supervised learning
- Unsupervised learning
- Anomaly detection
- Feature relevance
- Semantic search
- Activation energy
- Interconnections
- Relations
- Connections between nodes
- Search for related Patterns

Activate nodes → Spread activation → Search query

Instance from data set → All instances (clustering) → Single feature → New instance
Examples

• Cars
  • Semantic Search
  • Relations
  • Clustering
• IDS
  • Clustering
Cars

- UCI data set
- 1985 Model Import Car and Truck Specifications
- 26 Features (categorical, integer, real)
- Make, length, horsepower, MPG-1, MPG-2...
Cars

• Relations:
  Porsche, VW, 40 MPG (highway) and 200
Cars

• Relations:
Porsche, VW, 40 MPG (highway) and 200

<table>
<thead>
<tr>
<th>Relations</th>
<th>Porsche</th>
<th>Porsche</th>
<th>167 cm</th>
<th>205 HP</th>
<th>16 mpg</th>
<th>24 mpg</th>
<th>RWD</th>
<th>gas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porsche</td>
<td>6</td>
<td>167 cm</td>
<td>205 HP</td>
<td>16 mpg</td>
<td>24 mpg</td>
<td>RWD</td>
<td>gas</td>
<td>5961</td>
</tr>
<tr>
<td>VW</td>
<td>4</td>
<td>172 cm</td>
<td>92 HP</td>
<td>27 mpg</td>
<td>32 mpg</td>
<td>FWD</td>
<td>gas</td>
<td>5364</td>
</tr>
<tr>
<td>40</td>
<td>Chevrolet</td>
<td>4</td>
<td>157 cm</td>
<td>70 HP</td>
<td>37 mpg</td>
<td>42 mpg</td>
<td>FWD</td>
<td>gas</td>
</tr>
<tr>
<td>200</td>
<td>Porsche</td>
<td>6</td>
<td>167 cm</td>
<td>205 HP</td>
<td>16 mpg</td>
<td>24 mpg</td>
<td>RWD</td>
<td>gas</td>
</tr>
</tbody>
</table>

Ranking M1 (test-e.csv)
Cars

- Search for VW with moderate horsepower and high MPG rates
Search for VW with moderate horsepower and high MPG rates

<table>
<thead>
<tr>
<th>Rank</th>
<th>Car</th>
<th>Length (cm)</th>
<th>Horsepower (HP)</th>
<th>MPG (miles per gallon)</th>
<th>Drive</th>
<th>Fuel Type</th>
<th>Price (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Toyota</td>
<td>4 172 cm</td>
<td>58</td>
<td>37 mpg</td>
<td>FWD</td>
<td>diesel</td>
<td>4319</td>
</tr>
<tr>
<td>6</td>
<td>VW</td>
<td>4 167 cm</td>
<td>92</td>
<td>27 mpg</td>
<td>FWD</td>
<td>gas</td>
<td>5364</td>
</tr>
<tr>
<td>11</td>
<td>Mazda</td>
<td>4 177 cm</td>
<td>70</td>
<td>37 mpg</td>
<td>FWD</td>
<td>diesel</td>
<td>4874</td>
</tr>
<tr>
<td>204</td>
<td>Jaguar</td>
<td>12 188 cm</td>
<td>260</td>
<td>16 mpg</td>
<td>RWD</td>
<td>gas</td>
<td>4874</td>
</tr>
<tr>
<td>205</td>
<td>Mercedes</td>
<td>8 200 cm</td>
<td>180</td>
<td>16 mpg</td>
<td>RWD</td>
<td>gas</td>
<td>4319</td>
</tr>
</tbody>
</table>
Cars

• Search for powerful BMW
Cars

- Search for powerful BMW

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Make</th>
<th>Length (cm)</th>
<th>Power (HP)</th>
<th>Combined MPG</th>
<th>Highway MPG</th>
<th>Drive Configuration</th>
<th>Fuel Type</th>
<th>Price (in $)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>BMW</td>
<td>200</td>
<td>180</td>
<td>16</td>
<td>18</td>
<td>RWD</td>
<td>gas</td>
<td>5364</td>
</tr>
<tr>
<td>11</td>
<td>Peugeot</td>
<td>200</td>
<td>92</td>
<td>19</td>
<td>24</td>
<td>RWD</td>
<td>gas</td>
<td>4874</td>
</tr>
<tr>
<td>15</td>
<td>Audi</td>
<td>188</td>
<td>152</td>
<td>16</td>
<td>18</td>
<td>FWD</td>
<td>gas</td>
<td>5364</td>
</tr>
<tr>
<td>204</td>
<td>Honda</td>
<td>147</td>
<td>70</td>
<td>30</td>
<td>37</td>
<td>FWD</td>
<td>gas</td>
<td>5961</td>
</tr>
<tr>
<td>205</td>
<td>Honda</td>
<td>147</td>
<td>58</td>
<td>46</td>
<td>48</td>
<td>FWD</td>
<td>gas</td>
<td>4874</td>
</tr>
</tbody>
</table>
Cars

- Activate Porsche and search for related patterns
Cars

- Activate Porsche and search for related patterns

<table>
<thead>
<tr>
<th>Rank</th>
<th>Car</th>
<th>Model</th>
<th>HP</th>
<th>MPG</th>
<th>Transmission</th>
<th>Fuel Type</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Porsche</td>
<td>6</td>
<td>205</td>
<td>16</td>
<td>24</td>
<td>RWD</td>
<td>5961</td>
</tr>
<tr>
<td>6</td>
<td>Nissan</td>
<td>6</td>
<td>205</td>
<td>16</td>
<td>24</td>
<td>RWD</td>
<td>5364</td>
</tr>
<tr>
<td>8</td>
<td>Mazda</td>
<td>6</td>
<td>205</td>
<td>16</td>
<td>24</td>
<td>RWD</td>
<td>5364</td>
</tr>
<tr>
<td>203</td>
<td>VW</td>
<td>4</td>
<td>70</td>
<td>37</td>
<td>42</td>
<td>FWD</td>
<td>4319</td>
</tr>
<tr>
<td>205</td>
<td>Toyota</td>
<td>4</td>
<td>70</td>
<td>30</td>
<td>32</td>
<td>FWD</td>
<td>4874</td>
</tr>
</tbody>
</table>
Cars

- Activate VW and search for related patterns
Cars

- Activate VW and search for related patterns

<table>
<thead>
<tr>
<th>Rank</th>
<th>Car</th>
<th>Engine</th>
<th>HP</th>
<th>MPG</th>
<th>MPG</th>
<th>Drive</th>
<th>Fuel</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VW</td>
<td>4</td>
<td>172</td>
<td>70</td>
<td>37</td>
<td>42</td>
<td>FWD</td>
<td>diesel</td>
</tr>
<tr>
<td>13</td>
<td>Mazda</td>
<td>4</td>
<td>176</td>
<td>70</td>
<td>37</td>
<td>42</td>
<td>FWD</td>
<td>diesel</td>
</tr>
<tr>
<td>14</td>
<td>Toyota</td>
<td>4</td>
<td>167</td>
<td>58</td>
<td>37</td>
<td>42</td>
<td>FWD</td>
<td>diesel</td>
</tr>
<tr>
<td>203</td>
<td>Porsche</td>
<td>6</td>
<td>167</td>
<td>205</td>
<td>16</td>
<td>24</td>
<td>RWD</td>
<td>gas</td>
</tr>
<tr>
<td>205</td>
<td>Jaguar</td>
<td>6</td>
<td>200</td>
<td>180</td>
<td>16</td>
<td>18</td>
<td>RWD</td>
<td>gas</td>
</tr>
</tbody>
</table>
Cars

• Clustering
Cars

- Clustering

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Make</th>
<th>Model</th>
<th>Length (cm)</th>
<th>Horsepower (HP)</th>
<th>Fuel Efficiency (mpg)</th>
<th>Wheel Drive</th>
<th>Fuel Type</th>
<th>Fuel Consumption (g/km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>Toyota</td>
<td>4</td>
<td>188 cm</td>
<td>92 HP</td>
<td>19 mpg</td>
<td>24 mpg</td>
<td>RWD</td>
<td>5364</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>Toyota</td>
<td>4</td>
<td>167 cm</td>
<td>69 HP</td>
<td>39 mpg</td>
<td>37 mpg</td>
<td>FWD</td>
<td>5364</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>Toyota</td>
<td>4</td>
<td>176 cm</td>
<td>69 HP</td>
<td>30 mpg</td>
<td>32 mpg</td>
<td>FWD</td>
<td>4874</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>Mazda</td>
<td>4</td>
<td>167 cm</td>
<td>92 HP</td>
<td>16 mpg</td>
<td>24 mpg</td>
<td>RWD</td>
<td>5916</td>
</tr>
<tr>
<td>Cluster 5</td>
<td>Toyota</td>
<td>4</td>
<td>172 cm</td>
<td>92 HP</td>
<td>23 mpg</td>
<td>32 mpg</td>
<td>FWD</td>
<td>5364</td>
</tr>
<tr>
<td>Cluster 6</td>
<td>Toyota</td>
<td>4</td>
<td>167 cm</td>
<td>69 HP</td>
<td>30 mpg</td>
<td>32 mpg</td>
<td>FWD</td>
<td>5364</td>
</tr>
<tr>
<td>Cluster 7</td>
<td>Toyota</td>
<td>4</td>
<td>176 cm</td>
<td>92 HP</td>
<td>23 mpg</td>
<td>32 mpg</td>
<td>FWD</td>
<td>4874</td>
</tr>
<tr>
<td>Cluster 8</td>
<td>Mercedes</td>
<td>4</td>
<td>188 cm</td>
<td>180 HP</td>
<td>16 mpg</td>
<td>24 mpg</td>
<td>RWD</td>
<td>4874</td>
</tr>
<tr>
<td>Cluster 9</td>
<td>Subaru</td>
<td>4</td>
<td>172 cm</td>
<td>92 HP</td>
<td>23 mpg</td>
<td>32 mpg</td>
<td>FWD</td>
<td>4874</td>
</tr>
<tr>
<td>Cluster 10</td>
<td>Toyota</td>
<td>4</td>
<td>176 cm</td>
<td>114 HP</td>
<td>23 mpg</td>
<td>28 mpg</td>
<td>FWD</td>
<td>4874</td>
</tr>
</tbody>
</table>
IDS

- KDD Dataset
- Generation of Activation Patterns
- Unsupervised Analysis (RGNG)
- Visualization with SOMs
<table>
<thead>
<tr>
<th>Detected Cluster</th>
<th>Attack 1</th>
<th>Attack 2</th>
<th>Attack 3</th>
<th>Attack 4</th>
<th>Attack 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2L (27)</td>
<td>Warezcl.-32</td>
<td>Norm.-27</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Probe (3)</td>
<td>Portsw.-93</td>
<td>NMAP-11</td>
<td>Norm.-2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DOS (8)</td>
<td>Neptune-24</td>
<td>NMAP-14</td>
<td>Portsw.-8</td>
<td>Norm.-3</td>
<td>Land-2</td>
</tr>
<tr>
<td>DOS (7)</td>
<td>Back-208</td>
<td>Norm.-119</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DOS (18)</td>
<td>Smurf-28</td>
<td>Norm.-5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Probe (28)</td>
<td>IPSweep-116</td>
<td>Pod-16</td>
<td>NMAP-6</td>
<td>Norm.-1</td>
<td>-</td>
</tr>
<tr>
<td>Probe (12)</td>
<td>Satan-138</td>
<td>Portsw.-3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DOS (19)</td>
<td>Neptune-148</td>
<td>Satan-10</td>
<td>Portsw.-2</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Cluster 3/19

<table>
<thead>
<tr>
<th>Detected Cluster</th>
<th>Attack 1</th>
<th>Attack 2</th>
<th>Attack 3</th>
<th>Attack 4</th>
<th>Attack 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2L (27)</td>
<td>Warezcl.-32</td>
<td>Norm.-27</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Probe (3)</td>
<td>Portsw.-93</td>
<td>NMAP-11</td>
<td>Norm.-2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DOS (8)</td>
<td>Neptune-24</td>
<td>NMAP-14</td>
<td>Portsw.-8</td>
<td>Norm.-3</td>
<td>Land-2</td>
</tr>
<tr>
<td>DOS (7)</td>
<td>Back-208</td>
<td>Norm.-119</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DOS (18)</td>
<td>Smurf-28</td>
<td>Norm.-5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Probe (28)</td>
<td>IPsweep-116</td>
<td>Pod-16</td>
<td>NMAP-6</td>
<td>Norm.-1</td>
<td>-</td>
</tr>
<tr>
<td>Probe (12)</td>
<td>Satan-138</td>
<td>Portsw.-3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DOS (19)</td>
<td>Neptune-148</td>
<td>Satan-10</td>
<td>Portsw.-2</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Why???

• One model for arbitrary analysis procedures applied to all kind of data
• Enables us to understand relations between unknown data
• Correlation of arbitrary sensor data
• Semantic network offers quite much flexibility
Outlook

- Visualization with GWT
- Matlab Framework almost done
- Java Framework in progress (include RDF support, improve SA algorithms)
- Apply to other data sets
- Utilize different relations (e.g. time)
Thank You!

• http://www.iaik.tugraz.at
• peter.teufl@iaik.tugraz.at
• udo.payer@campus02.at
Thank You!

Looking forward to answering your questions!

- http://www.iaik.tugraz.at
- peter.teufl@iaik.tugraz.at
- udo.payer@campus02.at