Runtime Integration of Machine Learning and Simulation for Business Processes

Authors: Francesca Meneghello, Fondazione Bruno Kessler, Trento, Italy & Sapienza University, Rome, Italy
Chiara Di Francescomarino, University of Trento, Trento, Italy
Chiara Ghidini, Fondazione Bruno Kessler, Trento, Italy
Context: Business Process Simulation

- Optimization
- What-if scenarios
- Analysis
- Monitoring
- Process improvement
- Redesign process
Typical Approach: BPSimulation model
Typical Approach: BPSimulation model

Resources:
- Operator Man: 2 resources
- Stock Man: 2 resources
- Deliver Man: 1 resource

Case Arrival Rate: Exp(5)

Role: Operator Man
- Processing Time: TN(2,1)
- noOfPackage: Uniform(2,6)

Role: Stock Man
- Processing Time: TN(10,2)
- Waiting Time: Uniform(10,12)

Role: Delivery Man
- Processing Time: Exp(15)

Error in charging packages?
- Yes: 5%
- No: 95%
Typical Approach: BPSimulation model

Resources:
- Operator Man: 2 resources
- Stock Man: 2 resources
- Deliver Man: 1 resource

Role: Operator Man
- Processing Time: TN(2,1)
- noOfPackage: Uniform(2,6)

Role: Stock Man
- Processing Time: TN(10,2)
- Waiting Time: Uniform(10,12)

Role: Delivery Man
- Processing Time: Exp(15)

Case Arrival Rate: Exp(5)

Error in charging packages?
- Yes: 5%
- No: 95%

Easy and intuitive for implementing what-if scenarios
- Leveraging Process Mining Techniques
- Monitoring the simulation
Many unrealistic or oversimplifying assumptions

Typical Approach: BPSimulation model

Resources:
- Operator Man: 2 resources
- Stock Man: 2 resources
- Deliver Man: 1 resource

Case Arrival Rate: Exp(5)

Role: Operator Man
- Processing Time: TN(2,1)
- noOfPackage: Uniform(2,6)

Role: Stock Man
- Processing Time: TN(10,2)
- Waiting Time: Uniform(10,12)

Role: Delivery Man
- Processing Time: Exp(15)

Easy and intuitive for implementing what-if scenarios
Leveraging Process Mining Techniques
Monitoring the simulation

Error in charging packages?
- Yes: 5%
- No: 95%

Context/scenario is not considered in determining process paths or processing time
Alternative approach: Generative models

- LSTM networks used to predict the activity of the next event in a case, its timestamp, and the resource/role associated to the event.
- The approach has the ability to generate complete sequences from scratch by repeatedly predicting the subsequent event.

1 Discovering generative models from event logs: data-driven simulation vs deep learning, Manuel Camargo, Marlon Dumas, Oscar González-Rojas. (PeerJ Computer Science 2021)
Alternative approach: Generative models

- LSTM networks used to predict the activity of the next event in a case, its timestamp, and the resource/role associated to the event.
- The approach has the ability to generate complete sequences from scratch by repeatedly predicting the subsequent event.

- Able to consider the features of the context/scenario for prediction
- Not suitable for what-if scenarios
- Not global view, only on the single trace

---

1 Discovering generative models from event logs: data-driven simulation vs deep learning, Manuel Camargo, Marlon Dumas, Oscar González-Rojas. (PeerJ Computer Science 2021)
Hybrid Simulation: the Idea

Simulation model

Generative model
Hybrid Simulation: the Idea

Simulation model

Generative model
Hybrid Simulation: the Idea

- Easy and intuitive for implementing what-if scenarios
- Able to consider the context/scenario
- Monitoring the simulation
- Leveraging Process Mining Techniques

Simulation model

Generative model
Hybrid Simulation: State of the Art

Example of Post-integration
Example of Post-integration

Simulated Log

CHECK ORDER  CHARGE PACK  DELIVER  ID_Order: 1
CHECK ORDER  CHARGE PACK  DELIVER  ID_Order: 2
Example of Post-integration

Simulated Log

CHECK ORDER  CHARGE PACK  DELIVER  ID_Order: 1
CHECK ORDER  CHARGE PACK  DELIVER  ID_Order: 2

Enriched Simulated Log
Example of Post-integration

Simulated Log

- ID_Order: 1
  - CHECK ORDER
  - CHARGE PACK
  - DELIVER

- ID_Order: 2
  - CHECK ORDER
  - CHARGE PACK
  - DELIVER

Enriched Simulated Log

- CHECK ORDER
  - 8.15

- CHARGE PACK
  - 8.30

- DELIVER
  - 8.40
Example of Post-integration

Simulated Log

ID_Order: 1
CHECK ORDER
CHARGE PACK
DELIVER
CHECK ORDER
CHARGE PACK
DELIVER

ID_Order: 2
CHECK ORDER
CHARGE PACK
DELIVER
CHECK ORDER
CHARGE PACK
DELIVER

Enriched Simulated Log

CHECK ORDER
CHECK ORDER
CHARGE PACK
CHARGE PACK
DELIVER
DELIVER

8.15 8.20 8.25 8.30 8.37 8.40
Example of Post-integration

Simulated Log

Enriched Simulated Log

Resource Queue

1 Delivery Man
Problems with Post-integration

Simulated Log
- ID_Order: 1
  - CHECK ORDER
  - CHARGE PACK
  - DELIVER

ID_Order: 2
- CHECK ORDER
- CHARGE PACK
- DELIVER

Enriched Simulated Log
- CHECK ORDER
  - 8.15
- CHARGE PACK
  - 8.30
- DELIVER
  - 8.40
Problems with Post-integration

Simulated Log

Enriched Simulated Log

ID_Order: 1

ID_Order: 2

CHECK ORDER

CHECK ORDER

CHECK ORDER

CHECK ORDER

CHARGE PACK

CHARGE PACK

CHARGE PACK

CHARGE PACK

DELIVER

DELIVER

8.15

8.20

8.25

8.30

8.37

8.40
Problems with Post-integration

Simulated Log

<table>
<thead>
<tr>
<th>ID_Order: 1</th>
<th>ID_Order: 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHECK ORDER</td>
<td>CHECK ORDER</td>
</tr>
<tr>
<td>CHARGE PACK</td>
<td>CHARGE PACK</td>
</tr>
<tr>
<td>DELIVER</td>
<td>DELIVER</td>
</tr>
</tbody>
</table>

Enriched Simulated Log

<table>
<thead>
<tr>
<th>CHECK ORDER</th>
<th>CHECK ORDER</th>
<th>CHARGE PACK</th>
<th>CHARGE PACK</th>
<th>DELIVER</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.15</td>
<td>8.20</td>
<td>8.25</td>
<td>8.30</td>
<td>8.37</td>
</tr>
<tr>
<td>8.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A single delivery man

![Simulated Log](image1)

![Enriched Simulated Log](image2)
Problems with Post-integration

Simulated Log

ID_Order: 1
- CHECK ORDER
- CHARGE PACK
- DELIVER

ID_Order: 2
- CHECK ORDER
- CHARGE PACK
- DELIVER

Enriched Simulated Log

A single delivery man

8.15
- CHECK ORDER
- CHECK ORDER
- CHARGE PACK
- CHARGE PACK
- DELIVER
- DELIVER

8.20
8.25
8.30
8.37
8.40
Problems with Post-integration

Simulated Log

Enriched Simulated Log

A single delivery man

Adapted Resource Queue
Problems with Post-integration

Simulated Log

- ID_Order: 1
  - CHECK ORDER: 8.15
  - CHECK ORDER: 8.20
  - CHARGE PACK: 8.25
  - DELIVER

- ID_Order: 2
  - CHECK ORDER
  - CHECK ORDER
  - CHARGE PACK
  - CHARGE PACK
  - DELIVER
  - DELIVER

Enriched Simulated Log

- A single delivery man

Predicted Resource Queue

Adapted Resource Queue
Our solution: Runtime-Integration

Simulation model

Predictive model
Our solution: Runtime-Integration

Simulation model

CHECK ORDER → CHARGE PACK → DELIVER

Predictive model
Our solution: Runtime-Integration

Simulation model

CHECK ORDER → CHARGE PACK → DELIVER

Predictive model
Our solution: Runtime-Integration

Simulation model

Predictive model

8.15
Our solution: Runtime-Integration
Our solution: Runtime-Integration

- Simulation model
- Predictive model
RIMS: Runtime Integration of Machine Learning and Simulation for Business Processes

Event Log → BPS Simulation Model → Data Driven Process Simulation (DDPS) approach

Predictive models for time perspective
1. Definition of the DDPS elements
2. Training of the predictive models
3. Integration and run simulation
1. **Definition of the DDPS elements**

**Resources:**
- Operator Man: 2 resources
- Stock Man: 2 resources
- Deliver Man: 1 resource

**Roles:**
- Operator Man
- Stock Man
- Delivery Man

**Yes:** 5%

**No:** 95%
Training predictive models

1. Time series to predict the trace start times

Predictive model for Waiting time:
1) Next activity
2) End timestamp of the current activity
3) Intercase features: WIP(work-in-progress) and resources’ occupation

Predictive model for Processing Time:
1) Current activity
2) Start timestamp of the current activity
3) Intercase features: WIP(work-in-progress) and resources’ occupation
Integration and run simulation

1. Predict new arrival
2. Predict Waiting time
3. Predict Processing time
Integration and run simulation

1. Predict new arrival
2. Predict waiting time
3. Predict processing time
Integration and run simulation

1. Predict new arrival
2. Predict Waiting time
3. Predict Processing time
Integration and run simulation

1. Predict new arrival
2. Predict waiting time
3. Predict processing time
Integration and run simulation

1. Predict new arrival
2. Predict waiting time
3. Predict processing time
How does RIMS perform in terms of simulation quality compared to other techniques?
Evaluation: Procedure

- 25 simulations for each simulation approach
- MAE (Mean Absolute Error) of cycle times
- EMD (Earth Mover’s Distance) of the normalized histograms of activity timestamps grouped by day and hour

Metrics: MAE e EMD
## Evaluation: Datasets

<table>
<thead>
<tr>
<th>Log</th>
<th>Type</th>
<th>#Traces</th>
<th>#Events</th>
<th>#Activity</th>
<th>Avg. Trace length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidential1000</td>
<td>Syn</td>
<td>800</td>
<td>21221</td>
<td>29</td>
<td>26.53</td>
</tr>
<tr>
<td>Confidential2000</td>
<td>Syn</td>
<td>1670</td>
<td>44373</td>
<td>29</td>
<td>26.57</td>
</tr>
<tr>
<td>Cvs Pharmacy</td>
<td>Syn</td>
<td>10000</td>
<td>103906</td>
<td>15</td>
<td>10.39</td>
</tr>
<tr>
<td>Purchasing Example</td>
<td>Syn</td>
<td>608</td>
<td>9119</td>
<td>21</td>
<td>15</td>
</tr>
<tr>
<td>SynLoan</td>
<td>Syn</td>
<td>2000</td>
<td>43164</td>
<td>25</td>
<td>39</td>
</tr>
<tr>
<td>Production</td>
<td>Real</td>
<td>225</td>
<td>4503</td>
<td>24</td>
<td>20</td>
</tr>
<tr>
<td>ConsultaDataMining</td>
<td>Real</td>
<td>954</td>
<td>4962</td>
<td>16</td>
<td>5.2</td>
</tr>
<tr>
<td>BPI12W</td>
<td>Real</td>
<td>8616</td>
<td>59302</td>
<td>6</td>
<td>6.88</td>
</tr>
<tr>
<td>BPI17W</td>
<td>Real</td>
<td>30276</td>
<td>240854</td>
<td>8</td>
<td>7.96</td>
</tr>
</tbody>
</table>
# Evaluation RIMS: MAE metric

<table>
<thead>
<tr>
<th></th>
<th>DDPS</th>
<th></th>
<th>LSTM</th>
<th></th>
<th>Dsim</th>
<th></th>
<th>RIMS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>CI&lt;sub&gt;l&lt;/sub&gt;</td>
<td>CI&lt;sub&gt;u&lt;/sub&gt;</td>
<td>Mean</td>
<td>CI&lt;sub&gt;l&lt;/sub&gt;</td>
<td>CI&lt;sub&gt;u&lt;/sub&gt;</td>
<td>Mean</td>
<td>CI&lt;sub&gt;l&lt;/sub&gt;</td>
</tr>
<tr>
<td><strong>Confidential 1000</strong></td>
<td>55477</td>
<td>51509</td>
<td>59444</td>
<td>32754</td>
<td>32377</td>
<td>33132</td>
<td>36337</td>
<td>35796</td>
</tr>
<tr>
<td><strong>Confidential 2000</strong></td>
<td>234520</td>
<td>179570</td>
<td>289471</td>
<td>37344</td>
<td>37063</td>
<td>37625</td>
<td>8095</td>
<td>7309</td>
</tr>
<tr>
<td><strong>Cvs Pharmacy</strong></td>
<td>1061957</td>
<td>1061300</td>
<td>1062613</td>
<td>638665</td>
<td>637616</td>
<td>639714</td>
<td>137636</td>
<td>136104</td>
</tr>
<tr>
<td><strong>PurchasingExample</strong></td>
<td>2127361</td>
<td>2109665</td>
<td>2145058</td>
<td>2299619</td>
<td>2294319</td>
<td>2304919</td>
<td>2280717</td>
<td>2273954</td>
</tr>
<tr>
<td><strong>SynLoan</strong></td>
<td>2295797</td>
<td>2295579</td>
<td>2296015</td>
<td>1446177</td>
<td>1446177</td>
<td>1463812</td>
<td>1675036</td>
<td>1635297</td>
</tr>
<tr>
<td><strong>Production</strong></td>
<td>312153</td>
<td>307036</td>
<td>317270</td>
<td>128942</td>
<td>117516</td>
<td>140368</td>
<td><strong>83700</strong></td>
<td>73572</td>
</tr>
<tr>
<td><strong>ConsultaDataMining</strong></td>
<td>274786</td>
<td>271536</td>
<td>278035</td>
<td>343520</td>
<td>343422</td>
<td>343618</td>
<td>230921</td>
<td>228349</td>
</tr>
<tr>
<td><strong>BPI12W</strong></td>
<td>667139</td>
<td>665889</td>
<td>668389</td>
<td><strong>316642</strong></td>
<td><strong>309766</strong></td>
<td><strong>323517</strong></td>
<td>597090</td>
<td>594860</td>
</tr>
<tr>
<td><strong>BPI17W</strong></td>
<td>952502</td>
<td>949997</td>
<td>955007</td>
<td>757742</td>
<td>653898</td>
<td>861585</td>
<td>417641</td>
<td>416631</td>
</tr>
</tbody>
</table>
## Evaluation RIMS: MAE metric

<table>
<thead>
<tr>
<th></th>
<th>DDPS</th>
<th>LSTM</th>
<th>Dsim</th>
<th>RIMS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>CI&lt;sub&gt;l&lt;/sub&gt;</td>
<td>CI&lt;sub&gt;u&lt;/sub&gt;</td>
<td>Mean</td>
</tr>
<tr>
<td><strong>Confidential 1000</strong></td>
<td>55477</td>
<td>51509</td>
<td>59444</td>
<td>32754</td>
</tr>
<tr>
<td><strong>Confidential 2000</strong></td>
<td>234520</td>
<td>179570</td>
<td>289471</td>
<td>37344</td>
</tr>
<tr>
<td><strong>Cvs Pharmacy</strong></td>
<td>1061957</td>
<td>1061300</td>
<td>1062613</td>
<td>638665</td>
</tr>
<tr>
<td><strong>PurchasingExample</strong></td>
<td>2127361</td>
<td>2109665</td>
<td>2145058</td>
<td>2299619</td>
</tr>
<tr>
<td><strong>SynLoan</strong></td>
<td>2295797</td>
<td>2295579</td>
<td>2296015</td>
<td>1446177</td>
</tr>
<tr>
<td><strong>Production</strong></td>
<td>312153</td>
<td>307036</td>
<td>317270</td>
<td>128942</td>
</tr>
<tr>
<td><strong>ConsultaDataMining</strong></td>
<td>274786</td>
<td>271536</td>
<td>278035</td>
<td>343320</td>
</tr>
<tr>
<td><strong>BPI12W</strong></td>
<td>667139</td>
<td>665889</td>
<td>668389</td>
<td><strong>316642</strong></td>
</tr>
<tr>
<td><strong>BPI17W</strong></td>
<td>952502</td>
<td>949997</td>
<td>955007</td>
<td>757742</td>
</tr>
</tbody>
</table>
## Evaluation RIMS: MAE metric

<table>
<thead>
<tr>
<th></th>
<th>DDPS</th>
<th>LSTM</th>
<th>Dsim</th>
<th>RIMS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>CI_l</td>
<td>CI_u</td>
<td>Mean</td>
</tr>
<tr>
<td>Confidential 1000</td>
<td>55477</td>
<td>51509</td>
<td>59444</td>
<td>32754</td>
</tr>
<tr>
<td>Confidential 2000</td>
<td>234520</td>
<td>179570</td>
<td>289471</td>
<td>37344</td>
</tr>
<tr>
<td>Cvs Pharmacy</td>
<td>1061957</td>
<td>1061300</td>
<td>1062613</td>
<td>638665</td>
</tr>
<tr>
<td>PurchasingExample</td>
<td>2127361</td>
<td>2109665</td>
<td>2145058</td>
<td>2299619</td>
</tr>
<tr>
<td>SynLoan</td>
<td>2295797</td>
<td>2295579</td>
<td>2296015</td>
<td>1446177</td>
</tr>
<tr>
<td>Production</td>
<td>312153</td>
<td>307036</td>
<td>317270</td>
<td>128942</td>
</tr>
<tr>
<td>ConsultaDataMining</td>
<td>274786</td>
<td>271536</td>
<td>278035</td>
<td>343520</td>
</tr>
<tr>
<td>BPI12W</td>
<td>667139</td>
<td>665889</td>
<td>668389</td>
<td>316642</td>
</tr>
<tr>
<td>BPI12W*</td>
<td>694558</td>
<td>694083</td>
<td>695033</td>
<td>316642</td>
</tr>
<tr>
<td>BPI17W</td>
<td>952502</td>
<td>949997</td>
<td>955007</td>
<td>757742</td>
</tr>
</tbody>
</table>
## Evaluation RIMS: EMD metric

<table>
<thead>
<tr>
<th></th>
<th>DDPS</th>
<th>LSTM</th>
<th>Dsim</th>
<th>RIMS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  CI&lt;sub&gt;L&lt;/sub&gt;  CI&lt;sub&gt;U&lt;/sub&gt;</td>
<td>Mean  CI&lt;sub&gt;L&lt;/sub&gt;  CI&lt;sub&gt;U&lt;/sub&gt;</td>
<td>Mean  CI&lt;sub&gt;L&lt;/sub&gt;  CI&lt;sub&gt;U&lt;/sub&gt;</td>
<td>Mean  CI&lt;sub&gt;L&lt;/sub&gt;  CI&lt;sub&gt;U&lt;/sub&gt;</td>
</tr>
<tr>
<td>Confidential 1000</td>
<td>0.00115 0.00103 0.00127</td>
<td>0.01435 0.01434 0.01437</td>
<td>0.01269 0.01263 0.01274</td>
<td>0.01264 0.01257 0.01270</td>
</tr>
<tr>
<td>Confidential 2000</td>
<td>0.00109 0.00105 0.00113</td>
<td>0.00992 0.00991 0.00993</td>
<td>0.00897 0.00895 0.00899</td>
<td>0.00907 0.00905 0.00909</td>
</tr>
<tr>
<td>CVS Pharmacy</td>
<td>0.00093 0.00092 0.00094</td>
<td>0.00237 0.00237 0.00237</td>
<td>0.00659 0.00658 0.00659</td>
<td>0.00672 0.00672 0.00673</td>
</tr>
<tr>
<td>Purchasing Example</td>
<td>0.00702 0.00699 0.00706</td>
<td>0.00525 0.00500 0.00550</td>
<td>0.00645 0.00634 0.00656</td>
<td>0.00856 0.00831 0.00880</td>
</tr>
<tr>
<td>SynLoan</td>
<td>0.00799 0.00798 0.00800</td>
<td>0.01926 0.01924 0.01928</td>
<td>0.00473 0.00469 0.00477</td>
<td>0.00321 0.00317 0.00325</td>
</tr>
<tr>
<td>Production</td>
<td>0.01486 0.01438 0.01534</td>
<td>0.00715 0.00647 0.00784</td>
<td>0.00635 0.00589 0.00680</td>
<td>0.00523 0.00487 0.00559</td>
</tr>
<tr>
<td>ConsultaDataMining</td>
<td>0.01573 0.01555 0.01591</td>
<td>0.03830 0.03820 0.03840</td>
<td>0.01233 0.01202 0.01265</td>
<td>0.00777 0.00739 0.00816</td>
</tr>
<tr>
<td>BPI12W</td>
<td>0.00642 0.00640 0.00644</td>
<td>0.00766 0.00765 0.00768</td>
<td>0.00807 0.00806 0.00808</td>
<td>0.00805 0.00804 0.00806</td>
</tr>
<tr>
<td>BPI12W*</td>
<td>0.00468 0.00465 0.00470</td>
<td>0.00766 0.00765 0.00768</td>
<td>0.00799 0.00798 0.00799</td>
<td>0.00798 0.00798 0.00799</td>
</tr>
</tbody>
</table>
## Evaluation RIMS: EMD metric

<table>
<thead>
<tr>
<th></th>
<th>DDPS</th>
<th></th>
<th>LSTM</th>
<th></th>
<th>Dsim</th>
<th></th>
<th>RIMS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>(CI_l)</td>
<td>Mean</td>
<td>(CI_l)</td>
<td>Mean</td>
<td>(CI_l)</td>
<td>Mean</td>
<td>(CI_u)</td>
</tr>
<tr>
<td>Confidential 1000</td>
<td>0.00115</td>
<td>0.00103</td>
<td>0.00127</td>
<td>0.01435</td>
<td>0.01434</td>
<td>0.01437</td>
<td>0.01269</td>
<td>0.01263</td>
</tr>
<tr>
<td>Confidential 2000</td>
<td>0.00109</td>
<td>0.00105</td>
<td>0.00113</td>
<td>0.00992</td>
<td>0.00991</td>
<td>0.00993</td>
<td>0.00897</td>
<td>0.00895</td>
</tr>
<tr>
<td>Cvs Pharmacy</td>
<td>0.00093</td>
<td>0.00092</td>
<td>0.00094</td>
<td>0.00237</td>
<td>0.00237</td>
<td>0.00237</td>
<td>0.00659</td>
<td>0.00658</td>
</tr>
<tr>
<td>PurchasingExample</td>
<td>0.00702</td>
<td>0.00699</td>
<td>0.00706</td>
<td>0.00525</td>
<td>0.00500</td>
<td>0.00550</td>
<td>0.00645</td>
<td>0.00634</td>
</tr>
<tr>
<td>SynLoan</td>
<td>0.00799</td>
<td>0.00798</td>
<td>0.00800</td>
<td>0.01926</td>
<td>0.01924</td>
<td>0.01928</td>
<td>0.00473</td>
<td>0.00469</td>
</tr>
<tr>
<td>Production</td>
<td>0.01486</td>
<td>0.01438</td>
<td>0.01534</td>
<td>0.00715</td>
<td>0.00647</td>
<td>0.00784</td>
<td>0.00635</td>
<td>0.00589</td>
</tr>
<tr>
<td>ConsultaDataMining</td>
<td>0.01573</td>
<td>0.01555</td>
<td>0.01591</td>
<td>0.03830</td>
<td>0.03820</td>
<td>0.03840</td>
<td>0.01233</td>
<td>0.01202</td>
</tr>
<tr>
<td>BPI12W</td>
<td>0.00642</td>
<td>0.00640</td>
<td>0.00644</td>
<td>0.00766</td>
<td>0.00765</td>
<td>0.00768</td>
<td>0.00807</td>
<td>0.00806</td>
</tr>
<tr>
<td>BPI12W*</td>
<td>0.00468</td>
<td>0.00465</td>
<td>0.00470</td>
<td>0.00766</td>
<td>0.00765</td>
<td>0.00768</td>
<td>0.00799</td>
<td>0.00798</td>
</tr>
</tbody>
</table>
## Evaluation RIMS: EMD metric

<table>
<thead>
<tr>
<th>Dataset</th>
<th>DDPS</th>
<th>LS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Mean</strong></td>
<td><strong>CI_t</strong></td>
</tr>
<tr>
<td>Confidential 1000</td>
<td>0.00115</td>
<td>0.00103</td>
</tr>
<tr>
<td>Confidential 2000</td>
<td>0.00109</td>
<td>0.00105</td>
</tr>
<tr>
<td>CVS Pharmacy</td>
<td>0.00093</td>
<td>0.00092</td>
</tr>
<tr>
<td>PurchasingExample</td>
<td>0.00702</td>
<td>0.00699</td>
</tr>
<tr>
<td>SynLoan</td>
<td>0.00799</td>
<td>0.00798</td>
</tr>
<tr>
<td>Production</td>
<td>0.01486</td>
<td>0.01438</td>
</tr>
<tr>
<td>ConsultaDataMining</td>
<td>0.01573</td>
<td>0.01555</td>
</tr>
<tr>
<td>BPI12W</td>
<td><strong>0.00642</strong></td>
<td><strong>0.00640</strong></td>
</tr>
<tr>
<td>BPI12W*</td>
<td><strong>0.00468</strong></td>
<td><strong>0.00465</strong></td>
</tr>
</tbody>
</table>
The End??
What about the queue??

Simulated Log

ID_Order: 1
- CHECK ORDER at 8.15
- CHECK ORDER at 8.20
- CHARGE PACK at 8.25
- DELIVER

ID_Order: 2
- CHECK ORDER
- CHARGE PACK
- DELIVER

Enriched Simulated Log

- CHECK ORDER at 8.20
- CHECK ORDER at 8.25
- CHARGE PACK at 8.30
- DELIVER at 8.37
- DELIVER at 8.40
What about the queue??

Simulated Log

- ID Order: 1
  - Time: 8.15
  - Action: Check Order
  - Time: 8.20
  - Action: Charge Pack
  - Time: 8.25
  - Action: Deliver

- ID Order: 2
  - Time: 8.25
  - Action: Check Order
  - Time: 8.30
  - Action: Charge Pack
  - Time: 8.37
  - Action: Deliver

Enriched Simulated Log

- ID Order: 1
  - Time: 8.15
  - Action: Check Order
  - Time: 8.20
  - Action: Charge Pack
  - Time: 8.25
  - Action: Deliver

- ID Order: 2
  - Time: 8.25
  - Action: Check Order
  - Time: 8.30
  - Action: Charge Pack
  - Time: 8.37
  - Action: Deliver

Queue of Delivery Man Requests
What about the queue??

Simulated Log

- ID_Order: 1
- ID_Order: 2

Enriched Simulated Log

- How many deliveries are waiting?
What about the queue??

Simulated Log

Enriched Simulated Log

Add the Queue feature

Queue of Delivery Man Requests

How many deliveries are waiting?

Add the Queue feature
1. Definition of the DDPS elements

2. Training of the predictive models:
   2a. Retrieve the queue feature from the log
   2b. Train the waiting time predictive model with the queue feature

3. Integration and run simulation
2a Retrieve the queue feature from the log
2a Retrieve the queue feature from the log

$e_{S6} = (B, 08:11, 08:14, \text{Role1})$

2b Train the waiting time predictive model with the queue feature

$e_{S4} = (B, 08:05, 08:11, \text{Role1})$

$e^*_{S6} = (B, 08:11, 08:14, \text{Role1, 1})$

$e^*_{S4} = (B, 08:05, 08:11, \text{Role1, 0})$
RIMS$^+$: Integration and run simulation
RIMS\textsuperscript{+}: Integration and run simulation
RIMS⁺: Integration and run simulation
How does RIMS+ improve the quality of simulation compared to approaches that do not use the queue feature?
Evaluation: Procedure

- 25 simulations for each simulation approach
- MAE (Mean Absolute Error) of cycle times
- EMD (Earth Mover’s Distance) of the normalized histograms of activity timestamps grouped by day and hour

Metrics: MAE e EMD
## Evaluation: Datasets

<table>
<thead>
<tr>
<th>Log</th>
<th>Type</th>
<th>#Traces</th>
<th>#Events</th>
<th>#Activity</th>
<th>Avg. Trace length</th>
<th>Queue mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidential1000</td>
<td>Syn</td>
<td>800</td>
<td>21221</td>
<td>29</td>
<td>26.53</td>
<td>9.91</td>
</tr>
<tr>
<td>Confidential2000</td>
<td>Syn</td>
<td>1670</td>
<td>44373</td>
<td>29</td>
<td>26.57</td>
<td>2.37</td>
</tr>
<tr>
<td>Cvs Pharmacy</td>
<td>Syn</td>
<td>10000</td>
<td>103906</td>
<td>15</td>
<td>10.39</td>
<td>59</td>
</tr>
<tr>
<td>Purchasing Example</td>
<td>Syn</td>
<td>608</td>
<td>9119</td>
<td>21</td>
<td>15</td>
<td>0.30</td>
</tr>
<tr>
<td>SynLoan</td>
<td>Syn</td>
<td>2000</td>
<td>43164</td>
<td>25</td>
<td>39</td>
<td>68.89</td>
</tr>
<tr>
<td>Production</td>
<td>Real</td>
<td>225</td>
<td>4503</td>
<td>24</td>
<td>20</td>
<td>1.17</td>
</tr>
<tr>
<td>ConsultaDataMining</td>
<td>Real</td>
<td>954</td>
<td>4962</td>
<td>16</td>
<td>5.2</td>
<td>0</td>
</tr>
<tr>
<td>BPI12W</td>
<td>Real</td>
<td>8616</td>
<td>59302</td>
<td>6</td>
<td>6.88</td>
<td>1115</td>
</tr>
<tr>
<td>BPI17W</td>
<td>Real</td>
<td>30276</td>
<td>240854</td>
<td>8</td>
<td>7.96</td>
<td>1646</td>
</tr>
</tbody>
</table>
### Evaluation: Datasets

<table>
<thead>
<tr>
<th>Log</th>
<th>Type</th>
<th>#Traces</th>
<th>#Events</th>
<th>#Activity</th>
<th>Avg. Trace length</th>
<th>Queue mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidential1000</td>
<td>Syn</td>
<td>800</td>
<td>21221</td>
<td>29</td>
<td>26.53</td>
<td>9.91</td>
</tr>
<tr>
<td>Confidential2000</td>
<td>Syn</td>
<td>1670</td>
<td>44373</td>
<td>29</td>
<td>26.57</td>
<td>2.37</td>
</tr>
<tr>
<td>Cvs Pharmacy</td>
<td>Syn</td>
<td>10000</td>
<td>103906</td>
<td>15</td>
<td>10.39</td>
<td>59</td>
</tr>
<tr>
<td>Purchasing Example</td>
<td>Syn</td>
<td>608</td>
<td>9119</td>
<td>21</td>
<td>15</td>
<td>0.30</td>
</tr>
<tr>
<td>SynLoan</td>
<td>Syn</td>
<td>2000</td>
<td>43164</td>
<td>25</td>
<td>39</td>
<td>68.89</td>
</tr>
<tr>
<td>Production</td>
<td>Real</td>
<td>225</td>
<td>4503</td>
<td>24</td>
<td>20</td>
<td>1.17</td>
</tr>
<tr>
<td>ConsultaDataMining</td>
<td>Real</td>
<td>954</td>
<td>4962</td>
<td>16</td>
<td>5.2</td>
<td>0</td>
</tr>
<tr>
<td>BPI12W</td>
<td>Real</td>
<td>8616</td>
<td>59302</td>
<td>6</td>
<td>6.88</td>
<td>1115</td>
</tr>
<tr>
<td>BPI17W</td>
<td>Real</td>
<td>30276</td>
<td>240854</td>
<td>8</td>
<td>7.96</td>
<td>1646</td>
</tr>
</tbody>
</table>
## Evaluation RIMS+ vs Best

<table>
<thead>
<tr>
<th></th>
<th>Best</th>
<th>RIMS+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>CI&lt;sub&gt;l&lt;/sub&gt;</td>
</tr>
<tr>
<td>Confidential 1000</td>
<td>26950</td>
<td>25683</td>
</tr>
<tr>
<td>Cvs Pharmacy</td>
<td>102950</td>
<td>101130</td>
</tr>
<tr>
<td>SynLoan</td>
<td>1278782</td>
<td>1271133</td>
</tr>
<tr>
<td>BPI12W</td>
<td>316642</td>
<td>309766</td>
</tr>
<tr>
<td>BPI12W*</td>
<td>310293</td>
<td>304608</td>
</tr>
<tr>
<td>BPI17W</td>
<td>393749</td>
<td>391442</td>
</tr>
</tbody>
</table>

- RIMS<sup>+</sup> outperforms the best approach in MAE
- RIMS<sup>+</sup> worsens slightly in EMD
RIMS vs RIMS+: the impact of queue

6 Synthetic logs with different levels of queuing
Conclusions

- New hybrid simulation approach
- \textbf{RIMS} and \textbf{RIMS}^+ outperform state of the art

Future work:

- Add the remaining perspectives
- Consider Resource Calendar
- Consider queue policies
THANK YOU!

Questions, doubts, thoughts?

Contact me: fmeneghello@fbk.eu
Process & Data Intelligence group website: https://pdi.fbk.eu/
THANK YOU!

Questions, doubts, thoughts?

Contact me: fmeneghello@fbk.eu
Process & Data Intelligence group website: https://pdi.fbk.eu/