Real-time prediction of employee workload in digital railway control rooms

Léon Sobrie*, Marijn Verschelde*^ & Bart Roets*§

*Ghent University, ^IÉSEG School of Management, §Infrabel
Outline

1. Setting the stage
2. Research on workload
3. Implementation
4. Lessons learned
5. Future implications
1. Setting the stage
The current context

Current trends

• Increasing heterogeneity, complexity and interconnectedness of many business processes (Vasconcelos & Ramirez, 2011)
• Digitization of business processes (Davenport & Ronanki, 2018)
• The adoption of machine/deep learning in industry is still in its infancy (Kraus et al., 2020)

→ A need for data-driven decision support for management

Initiative

• European Commission: Industry 5.0
Control rooms

• The nerve center for real-time monitoring and intervention

• Manage and coordinate many environments: rail and air traffic, nuclear power plants, chemical production sites, ambulance, etc.

Characteristics

1. Real-time decision making

2. Highly variable workload

3. Safety-critical environment
All Belgian railway traffic is managed in real-time by the control rooms of Infrabel

- Dense railway network
- Huge amount of events
  - Trains passing signals (50 million/year)
  - All actions taken by operators in control rooms (150 million/year)
Control room dynamics
2. Control room operator workload
The importance of workload

- Lack of attention (Young, 2021)
- Extra performance-seeking risks (Xu et al., 2021)
- Lower well-being (Ilies et al., 2010)
- Preference for easier tasks (Kc et al., 2020)
- Human fatigue (Li et al., 2020)
- Lower safety levels (Fereira & Balfe, 2014)

The importance of balanced workload within and between operators (Inegbedion et al., 2020)
Contributions

1. Insights from a granular data structure containing all anonymized operator events

2. Empirical usefulness of the proposed model and insights into the importance of the different organizational & operational characteristics

3. Development of an application to provide decision-support for the control room manager
Input: operational and organizational characteristics

- Operational features
  - Experience
  - Railway operations
  - Time
  - Current workload

- Organizational features
  - Control room characteristics
  - Partner interactions
Output: Operational workload categories

• In line with the multi-attribute task battery for human operator workload (Comstock & Arnegard, 1992) = communication, resource management, automation, scheduling, monitoring and tracking

<table>
<thead>
<tr>
<th>category</th>
<th>content</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOVE</td>
<td>proactively monitoring of railway traffic</td>
</tr>
<tr>
<td>ADAPT</td>
<td>changing tracks and station platforms</td>
</tr>
<tr>
<td>AUT</td>
<td>changing the automation</td>
</tr>
<tr>
<td>SAFETY</td>
<td>safety interventions</td>
</tr>
<tr>
<td>PHONE</td>
<td>phone calls between operator and driver</td>
</tr>
<tr>
<td>JUSTIF</td>
<td>justification of train delays</td>
</tr>
</tbody>
</table>

Task distribution of traffic and safety controller
Model: linking characteristics with workload categories

Q1: To what extent can we predict whether the operator will have workload for a specific category?

Q2: To what extent can we predict how much workload the operator will have for a specific category?
Model: linking characteristics with workload categories

Q1: To what extent can we predict whether the operator will have workload for a specific category?

Q2: To what extent can we predict how much workload the operator will have for a specific category?

LSTM encoder-decoder

Input t-4
Input t-3
Input t-2
Input t-1
Input t

Output t+1
Output t+2
Output t+3
Output t+4
Output t+5

Random forest

sample
sample
...

(eXtreme) Gradient Boosting

data
model
error

data
model
error
Managing model risk

• Different types of risk to be managed when modeling
  1. Data
  2. Specification
  3. Development
  4. Validation
  5. Operational
  6. Security
  7. Managerial

Source: ‘Managing model risk’ by Seppe Vanden Broucke & Bart Baesens
Results: classification ability

• Q1: To what extent can we predict whether the operator will have workload for a specific category in the next 15 minutes?

![Graph showing accuracy for different workload categories]

- Perfect classification
- Random model
Results: error of prediction

• Q2: To what extent can we predict how much workload the operator will have for a specific category in the next 15 minutes?

<table>
<thead>
<tr>
<th>category</th>
<th>RF</th>
<th>XGB</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOVE</td>
<td>23s</td>
<td>22s</td>
</tr>
<tr>
<td>AUT</td>
<td>22s</td>
<td>20s</td>
</tr>
<tr>
<td>ADAPT</td>
<td>60s</td>
<td>54s</td>
</tr>
<tr>
<td>SAFETY</td>
<td>18s</td>
<td>18s</td>
</tr>
<tr>
<td>PHONE</td>
<td>70s</td>
<td>70s</td>
</tr>
<tr>
<td>JUSTIF</td>
<td>57s</td>
<td>57s</td>
</tr>
</tbody>
</table>

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{N}}
\]
Results: feature importance

Insights in the importance of features of random forest model

- **Delays**
- **Partner interactions**
- **Previous workload levels**
- **End of shift**
Results: SHAP values

- Shift
- Specific workstations
- Control room
- Time of the day
- Partner workload

The control room of the future: AI-empowered dashboards
Tool for management of workload

Predicted workload level

10:50:00
3. Implementation
Technology readiness level (TRL)

A compass for assessing how ready the technology is for the real-world (developed by NASA, originates from ’70s)

2 stage approach

→ using proofs of concepts
- Replay real-time simulation
  - Face validity
  - Flexibility
- Real-time implementation
  - R Shiny

<table>
<thead>
<tr>
<th>TRL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>System proven in operational environment</td>
</tr>
<tr>
<td>8</td>
<td>System complete &amp; qualified</td>
</tr>
<tr>
<td>7</td>
<td>Integrated pilot system demonstrated</td>
</tr>
<tr>
<td>6</td>
<td>Prototype system verified</td>
</tr>
<tr>
<td>5</td>
<td>Laboratory testing of integrated system</td>
</tr>
<tr>
<td>4</td>
<td>Laboratory testing of prototype</td>
</tr>
<tr>
<td>3</td>
<td>Proof of concept established</td>
</tr>
<tr>
<td>2</td>
<td>Technology concept/ application formulated</td>
</tr>
<tr>
<td>1</td>
<td>Basic principles are observed</td>
</tr>
</tbody>
</table>
Implementation for management
Real-time implementation

Sobrie, Verschelde, Hennebel & Roets (2022) – Capturing complexity over space and time: An application to real-time delay prediction in railways

DL model predictions

![Diagram showing DL model predictions for train delays](image_url)
4. Lessons learned
Lessons learned

- There is untapped potential for machine learning in control rooms
  - Multidisciplinary approach required
  - Close collaboration between academia and practice
  - Learning iteratively: FAIL = first attempt in learning

- A roadmap towards implementation requires
  - Focus a practical issue
  - Construction of a real-time data flow
  - Model validation by operational testing
5. Future implications
Future research avenues

• Balancing the workload within and between operators
  • Research done in the BALANCE project of the On Track Lab

• Estimating the evolution of workload thresholds within a shift
  • Research done by the System Dynamics Lab of Virginia Tech

• More granular insights on the relationship between workload, delays, human errors and fatigue
Any questions?

Stay on track at https://ontracklab.com

Contact me at leon.sobrie@ugent.be
References


• Young, M. S. (2021, November). In Search of the Redline: Perspectives on Mental Workload and the ‘Underload Problem’. In International Symposium on Human Mental Workload: Models and Applications (pp. 3-10). Springer, Cham.