Data-Driven Network Performance Model (NPM) for Urban Rail Systems

Baichuan Mo
Zhenliang Ma, Haris N. Koutsopoulos, Jinhua Zhao

Jan, 2020
Motivation

- Monitoring network performance (online/offline) is crucial
  - Understand system
  - Improve service attractiveness
  - Assist planning and operations

- Objective
  - Develop a self-calibrated data-driven monitoring & decision support platform
    - Performance monitoring
    - Operations planning
Network Performance Model (NPM)

- Timetable/AVL
- OD Demand/AFC
- Network
- Route Choice
- Train Capacity

NPM Engine

- Station Crowding
- Pax Waiting Time
- Link/Train Load
- Pax Journey time
- OD Exit Flow

Self-calibration/Optimization

AFC: Automated Fare Collection
AVL: Automated vehicle location
Network Performance Model (NPM)

Self-calibration/Optimization
NPM Engine

• Event-based simulation:
  • First-come-first-serve
  • Strick capacity constraints

- Train arrival:
  • Offload passengers
- Train departure:
  • Load passengers
  • Update states of train and platform
Train Capacity

• **Effective train capacity:**
  • Number of passengers in the train when it departures while there are left-behind passengers on the platform.

• Train capacity may vary by station and crowding levels
  • Platform geometry and access impact passenger distribution along the platform, and hence, load distribution on trains.
  • Different crowding levels affect passenger willingness to board

• Expectation:
  • High train load $\rightarrow$ high effective capacity [Liu et al. 2018]
  • High crowding on the platform $\rightarrow$ high effective capacity
Effective Capacity Model

• Effective capacity of a train at platform $i$ ($EC_i$) is:

$$EC_i = \begin{cases} C + \beta_1 L_i + \beta_2 Q_i & \text{if platform } i \text{ is in the list of congested stations} \\ C & \text{otherwise} \end{cases}$$

where $C_i$: base capacity

$L_i$: train load and

$Q_i$: number of passengers waiting at platform
Estimation of Capacity Model Parameters

• Simulation-Based Optimization

• Minimize the error between
  • observed OD exit flow and model-derived OD exit flow
  • observed journey time distribution (JTD) and model-derived JTD

\[
\begin{align*}
\min_{\beta_1, \beta_2} & \quad w_1 \sum_{i,j,t} (q^{i,j,t} - \tilde{q}^{i,j,t})^2 + w_2 \sum_{i,j,t} D_{KL}(p_{i,j,t}(x) \| \tilde{p}_{i,j,t}(x)) \\
\text{s.t.} & \quad q^{i,j,t}, p_{i,j,t}(x) = \text{NPM}(\beta_1, \beta_2) \quad \forall i, j, t \\
D_{KL}(p_{i,j,t}(x) \| \tilde{p}_{i,j,t}(x)) & = \int_x p_{i,j,t}(x) \cdot \log \frac{p_{i,j,t}(x)}{\tilde{p}_{i,j,t}(x)} \, dx.
\end{align*}
\]
Applications

• Hong Kong MTR network
• Demand on March 16th, 2017. Evening Peak
• Path choice from survey
• Validation
  – OD exit flow by time
  – Left behind survey at key stations
Path Choice

• Path choices are modeled using a C-logit model from survey data

\[ p_{i,m,j}^r = \frac{\exp(\beta_X \cdot X_{r,m} + \beta_{CF} \cdot CF_r)}{\sum_{r' \in R(i,j)} \exp(\beta_X \cdot X_{r',m} + \beta_{CF} \cdot CF_{r'})} \]

**TABLE 3: Route Choice Model Estimation Results**

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-value</th>
<th>**</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-vehicle time</td>
<td>-0.147</td>
<td>0.011</td>
<td>-13.64</td>
<td>***</td>
</tr>
<tr>
<td>Relative walking time</td>
<td>-1.271</td>
<td>0.278</td>
<td>-4.56</td>
<td>***</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.573</td>
<td>0.084</td>
<td>-6.18</td>
<td>***</td>
</tr>
<tr>
<td>CF</td>
<td>-3.679</td>
<td>1.273</td>
<td>-2.89</td>
<td>**</td>
</tr>
</tbody>
</table>

\[ \rho^2 = 0.54 \]

***: p<0.01; **: p<0.05.

• AFC-data based path choice estimation [Poster session A139]
Results: Effective Capacity

- Parameter estimation (Bayesian Optimization Algorithm)

\[ EC_i = \begin{cases} 
C + \beta_1 L_i + \beta_2 Q_i \\
C 
\end{cases} \]

\( C = 230 \text{ pax/car} \times \text{Num of cars in a train (fixed)} \)

Optimal Solution:
\( \beta_1 = 0.0904 \)
\( \beta_2 = 0.0718 \)
Results: Effective Capacity

- Train load (pax/train) at Admiralty station

Train load

Peak period

![Graph showing train load over time with peak period highlighted.](image-url)
Validation: OD Exit Flow Estimates

18:00 – 18:30
Validation: Left Behind Estimates

**Number of Boarding Passengers**
- NPM
- Ground truth

**Number of Arrival Passengers**
- NPM
- Ground truth

**Left Behind Times**
- NPM
- Ground Truth

**Graphs**
- Time vs. Number of Boarding/Pax
- Time vs. Number of Arrival/Pax
Applications

- [History] **Monitor crowding patterns**: train load, left behind, waiting time, ...
- [History] **Diagnose crowding sources**: where does the congestion come from?
- [History] **Evaluate network resilience**: how does system change if link disruption happens?
- [Future] **Operations planning**: time table evaluation, dispatching strategies, ...

Performance Monitoring  Operational Planning

History  Future
Dispatching strategies evaluation

• Impact of dispatching an empty train from upstream to relieve the crowding in the platform at Admiralty station
Interactive Visualization

• What is happening?
• What is the problem?
• Why it happens?
• What will happen
  – if nothing change?
  – if things change?
  – if actions taken?
Conclusion

• Data-driven NPM platform:
  • Performance monitoring (what was/is...)
  • Operations control and strategic planning (what if...)
• Effective train capacity formulation
• Effective train calibration using AFC data
• Future work
  • Simultaneous calibration of route choice and train capacity
Thanks

Q&A