Adarules: Learning rules for real-time road-traffic prediction

Rafael Mena-Yedra\textsuperscript{1,2}  
Ricard Gavaldà\textsuperscript{2}  
Jordi Casas\textsuperscript{1}

\textsuperscript{1} TSS-Transport Simulation Systems, Spain  
\textsuperscript{2} Universitat Politècnica de Catalunya, Spain

20th EURO Working Group on Transportation Meeting (EWGT 2017)  
4-6 September 2017, Budapest, Hungary
Traffic (flow) prediction
How and what for?
Traffic prediction research

“Traffic flow prediction”

Article count

<table>
<thead>
<tr>
<th>Decade</th>
<th>Article Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970s</td>
<td>1</td>
</tr>
<tr>
<td>1980s</td>
<td>4</td>
</tr>
<tr>
<td>1990s</td>
<td>20</td>
</tr>
<tr>
<td>2000-2005</td>
<td>41</td>
</tr>
<tr>
<td>2005-2010</td>
<td>93</td>
</tr>
<tr>
<td>2010s</td>
<td>187</td>
</tr>
</tbody>
</table>
Why traffic prediction

- Traveler Information Services
- Active Traffic Management

- Beneficial impact on the network performance in terms of throughput, congestion length and average network speeds.

- Decision support systems for real-time traffic management.
  - Example: Aimsun Online

- Valuable input for other processes: trend to merge both approaches, purely data-driven methods and simulation models.
Motivation
Case study: San Diego (I-15)

Data source: California Department of Transportation (Caltrans) Performance Measurement System (PeMS). State of California.
Case study: San Diego (I-15)

Data source: California Department of Transportation (Caltrans) Performance Measurement System (PeMS). State of California.
Identified issues

- Diversity (kind of network, or even within the same network)
- Sudden change
- Gradual change (drift)
- Missing data observations
- Dependence on the data scientist or traffic engineer criteria for each case
Our approach: learning adaptive rules

"Adarules"
Adarules

Ruleset (Gama, 2010)

Default rule

Rule #1

Antecedent
if ‘weekday’ is [Sunday]
& ‘time’ is [7 - 9]
& ‘detector.x.flow’ > 1000

Consequent
Prediction Model #1
Prediction Model #n

Rule #n

Antecedent
if ‘season’ is [Summer]
& ‘detector.x.occupancy’ > 10
& ‘detector.x.flow’ > 1000

Consequent
Prediction Model #1
Prediction Model #n

Prediction Model #1
Prediction Model #n

Prediction Model #1
Prediction Model #n
Expanding rules

• To further specialize a current rule after observing enough data
  ○ Select n combinations (random, smart guess...) of attributes/splitpoints
  ○ Calculate entropy (measuring the randomness of data) on the outcome distribution
  ○ Hoeffding bound (as in Gama, 2010); statistical test to decide if the best scored split significantly reduces the metric

★ Non-parametric approach (finding spatiotemporal patterns in the network)
★ Minimum number of assumptions (i.e. maximizing the outcome probability)
★ Better interpretability than black-box models
Online learning: Sudden change

- Concept drift detection. Algorithm used based on the Page-Hinkley test.
- It starts to monitor the rule’s mean error when a new rule is built. Rule mean error should be located at 0.
- When a change is detected, the rule is removed from the ruleset.
  - Other approaches could be considered: changing the ruleset structure, merging rules...
- This kind of (sudden) change is handled at rule level
Rule prediction models

- Weighted (historical) mean (in the scope of the rule)
- LASSO: Sparse linear regression to capture the spatial dependencies in the network:

\[
\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{N} \| y - X\beta \|_2^2 + \lambda \| \beta \|_1 \right\}
\]

- High-dimensional problem (San Diego district 11 has +1500 detection stations)

\[ n \ll p \]
Online learning: Gradual change

- Seasonality, traffic demand growth...
- This kind of gradual change is handled at rule predictor level.
- Specific solution for each rule predictor
  - Weighted historical mean: age decaying factor
  - LASSO: coordinate-wise descent with soft-thresholding
Adarules

Real-time

Streaming data

Network state

Weekday

Time

Weather

(...)

Predictive system

Rule

Ruleset

Variable selection

Rule prediction model(s)

Context information

Change detection

Split evaluation

Anomaly detection

Forecasting output

Prediction point-estimate

Error prediction interval

15 / 22
Results
60-min traffic flow prediction

- Dataset: 2013/01 to 2015/12
- Tested approaches
  - Adarules (real-time)
  - Lassos for each 15-min interval trained in batch mode
    - 1 year train data set (2013/01 to 2013/12)
    - 6 month train data set (2013/01 to 2013/06)
  - Lassos for each 15-min interval retrained (blindly) every month
    - Using the last 6 month as training data
    - Using the last 1 month as training data
  - Lassos for each 15-min interval retrained (blindly) every week
    - Using the last 6 month as training data
    - Using the last 1 month as training data
60-min traffic flow prediction

Framework
- Adarules
- Batch
- Blind Adapt - monthly
- Blind Adapt - weekly

TrainSize
- ---
- ▲ 1 month
- ■ 1 year
- + 6 months
60-min traffic flow prediction

Number of 'valid' rules: 48
60-min traffic flow prediction

Number of ‘valid’ rules: 21

Framework
- Adarules
- Batch
- Blind Adapt - monthly
- Blind Adapt - weekly

TrainSize
- ---
- 1 month
- 1 year
- 6 months

San Diego - 1116415
2015/01 - 2015/12
Conclusions & Future work
Conclusions

➔ Fast adaption to change
➔ Autonomy to decide the best decisions with more data
➔ Interpretable spatiotemporal patterns for traffic managers
➔ Prediction accuracy is important, but not the only criteria (Karlaftis and Vlahogianni, 2011; Kirby et al., 1997). Autonomy, maintenance and adaptation, interpretability

Future work

➔ Multi-task learning
➔ Incident management
➔ Improving real-time efficiency
Adarules: Learning rules for real-time road-traffic prediction

Rafael Mena-Yedra¹,²
Ricard Gavaldà²
Jordi Casas¹

¹ TSS-Transport Simulation Systems, Spain
² Universitat Politècnica de Catalunya, Spain

20th EURO Working Group on Transportation Meeting (EWGT 2017)
4-6 September 2017, Budapest, Hungary
References