Prediction and Scenario-based Traffic Management (POST) – Clustering, Classification and Prediction

David Gundlegård (LiU), **Matej Cebecauer (KTH)**, Erik Jenelius (KTH), **Clas Rydergren (LiU)**, Rasmus Ringdahl (LiU), Wilco Burghout (KTH), Anna Danielsson (LiU)

Trafik Stockholm, TrV STRESS
Sweco, UC Berkeley
Aim and purpose

• Means for evaluating traffic management controls
  – Results of towing directly or from waiting to after peak hour traffic
  – Effects of early information to travelers of severe incidents (i.e. do not use car)
• Predict demand and route choice for scenario evaluation and action ranking
  – Offline processes for demand prediction and scenario evaluation
  – Online processes for classification of traffic situation and choice of control measure
Research Target

1. Scenario evaluation for historical events
2. Scenario evaluation using real-time data
Research Developments

• Past: Traffic flow analysis using traffic flow models and data analytics for travel time estimation

• Current: Analytics of data sources for travel demand prediction. Data analytics of data from: Inrix travel times on road segments, Trip data (GPS tracks) from Inrix, Congestion charging portal data (flow and "origin"), Mobile network data, Motorway control system (MCS) data

• Outcome: Integrated data analytics and model based scenario analysis
Overview of computational modules

• Route use analysis
  – Route flows (planned destination) for incident link
  – Route choice during incidents
• Demand prediction
  – Link flow destination distribution
  – Local upstream prediction
  – OD estimation and prediction
• Scenario evaluation
  – Traffic flow model for selected incidents
Module overview

Supply
(Network, road works, incidents, control)

Traffic State Estimation

Observations

Demand
(OD, boundary flows)

Traffic modeling
Route set generation
Route flow modeling
Network loading

Traffic State Prediction

Pattern matching
Incident detection

Offline processing
Clustering
Offline Traffic modeling
Action ranking
Archiving

Supply & Incident Management

Knowledge, actions

Automatic Traffic Control
(RM, VSL)

Traffic Management
Scenario Evaluation

Data-driven Analytics

Demand Management

Information, incentives

Traffic model generation
Route set generation
Route flow modeling
Network loading

Assimilation and fusion

Network loading
Route set generation
Route flow modeling
Network loading

CTR

KTH

Lund University
Prediction and Scenario-based Traffic management (POST) – Clustering, Classification and Prediction

David Gundlegård, Matej Cebecauer (KTH), Erik Jenelius (KTH), Clas Rydergren (LiU), Rasmus Ringdahl (LiU), Wilco Burghout (KTH), Anna Danielsson (LiU)

Trafik Stockholm, TrV STRESS
Sweco, UC Berkeley
Clustering, Classification and Prediction for scenario-based traffic management

Subprolems
- Spatio-temporal partitioning
- Day clustering
- “Typical” days inference
- New day classification
- Prediction

What we search for?
- Best performance with reasonable resources

Tradeoff of:
- Costs (training time, calibration time, pc memory)
- Easy implementation and scaling in practice
Clustering, Classification and Prediction for scenario-based traffic management

14 sensors at highway

93 highway routes

11,071 OD pairs in metro and rail public transport network
Subprolems

• **Spatio-temporal partitioning**
• Day clustering
• “Typical” days inference
• New day classification
• Prediction

For large-scale heterogeneous areas

• What is the best approach and which method?

• What is the most appropriate number of clusters?
Subproblems

- Spatio-temporal partitioning
- **Day clustering**
- “Typical” days inference
- New day classification
- Prediction

- What is the best method?
- How to measure similarity between days?
- What is the most appropriated number of clusters?
Subprolems

- Spatio-temporal partitioning
- Day clustering
- "Typical" days inference
- New day classification
- Prediction
Subproblems

- Spatio-temporal partitioning
- Day clustering
- “Typical” days inference
- New day classification
- Prediction
Subproblems
- Spatio-temporal partitioning
- Day clustering
- “Typical” days inference
- New day classification
- Prediction

What is the best method?

- **Euclidean distance**
  - No training or calibration,
  - Fast and easy to implement in practice
  - Sum the distance in network-time
    - may not reflect the structural difference

Machine learning methods

- **Random Forest**
- **Decision trees**
- **Naïve Bayes**
  - Can reflect structural differences
  - Costs (time & pc memory)
Subproblems

- Spatio-temporal partitioning
- Day clustering
- “Typical” days inference
- New day classification
- Prediction

- Application of “typical” days.
- Important for real-time scenario-based traffic management
- Validation tool to above subproblems and could help reveals the best performing tradeoff of:
  - Costs
  - Implementation in practice
  - Performance
Short-term prediction (15 minutes to the future)

- Training (all days 2017)
- Prediction (all days 2018)
Short-term prediction (15 minutes to the future)

- network-wide smoothing model
Short-term prediction (15 minutes to the future)

- network-wide smoothing model & spatio-temporal zoning
Conclusions

The prediction performance analysis shows that:

- **Day clusters**
  - reveal recurring patterns with its “typical” days
  - could be a reasonable input to scenario-based traffic management

- **Classification**
  - Euclidean Distance seems to give best performance with reasonable resources

- **Prediction**
  - Tool for validating revealed most “typical” days
  - Adding some smoothing or “neighborhoods” zones, could help to boost prediction performance
Prediction and Scenario-based Traffic management (POST) – Clustering, Classification and Prediction

David Gundlegård, **Matej Cebecauer (KTH)**, Erik Jenelius (KTH), **Clas Rydergren (LiU)**, Rasmus Ringdahl (LiU), Wilco Burghout (KTH), Anna Danielsson (LiU)

matej.cebecauer@abe.kth.se