Consideration of dynamic demand requirements for traffic management

Carlos Lima Azevedo, climaz@dtu.dk

https://mlsm.man.dtu.dk/

Paris, 13/05/2024

Outline

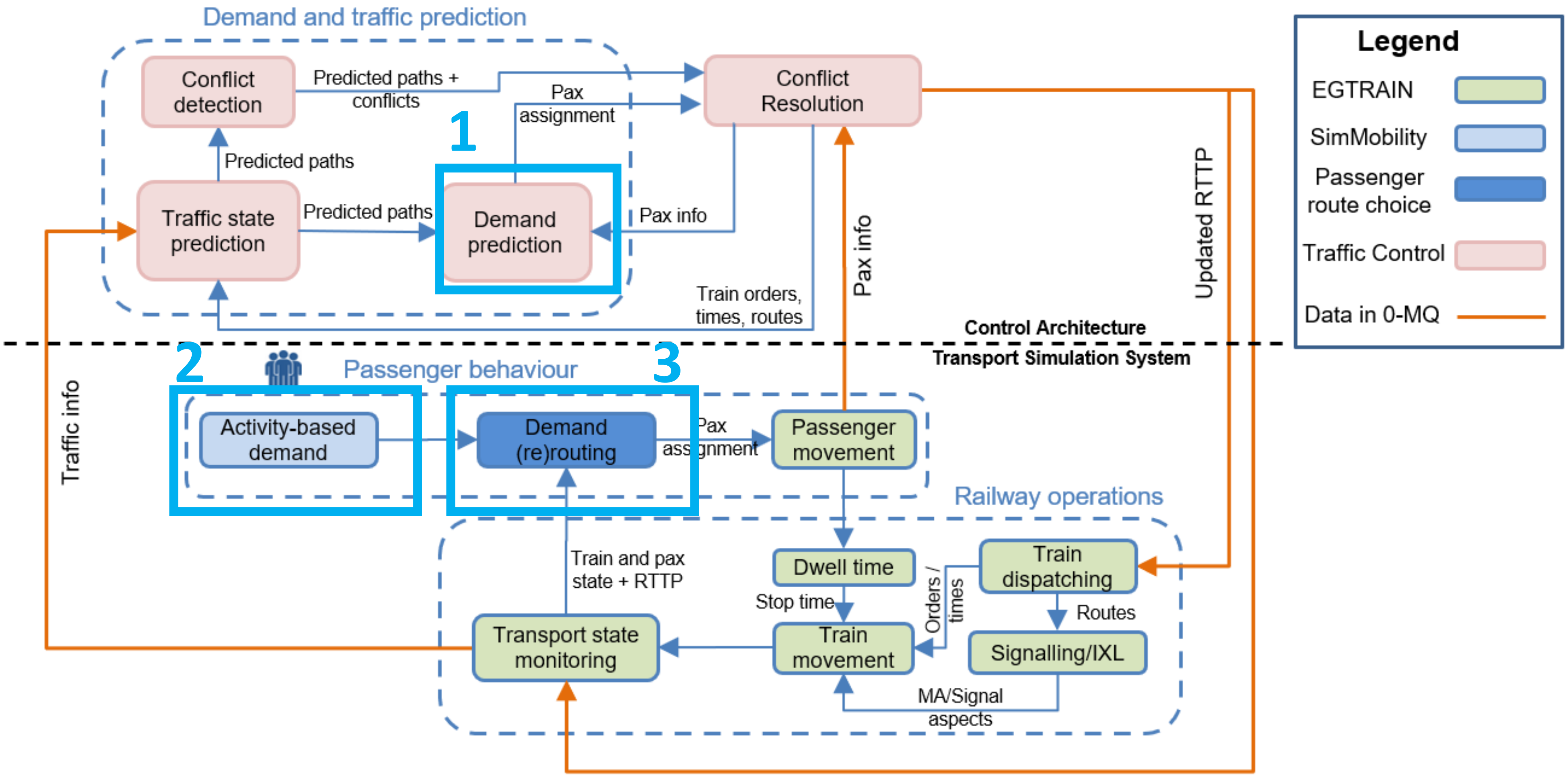
1. **Overview of Demand in Sortedmobility**
2. Demand Prediction
3. Demand generation for simulation
4. Demand (re) routing

Motivation

Why demand is essential in the design of future rail systems?

- Urban rail is part of a **multimodal system** > competition and complementarity
- Future rail may be **demand responsive** > demand is sensitive to such dynamics
- There is a trend to move towards **passenger LoS** instead of just train delays as key objective

Sortedmobility Framework



1. Demand Prediction: The problem

Let t represent a time interval. We define the OD matrix for time interval t as δ_t, where δ_{tij}, represents the **number of trips from station $\boldsymbol{i}$ to station $\boldsymbol{j}$ that are initiated within time interval $\boldsymbol{t}$**.

Problem: Based on real-time information about OD passenger numbers for historical time intervals (as well as external variables), a model is desired that can **predict OD matrices for the next time interval(s)**.

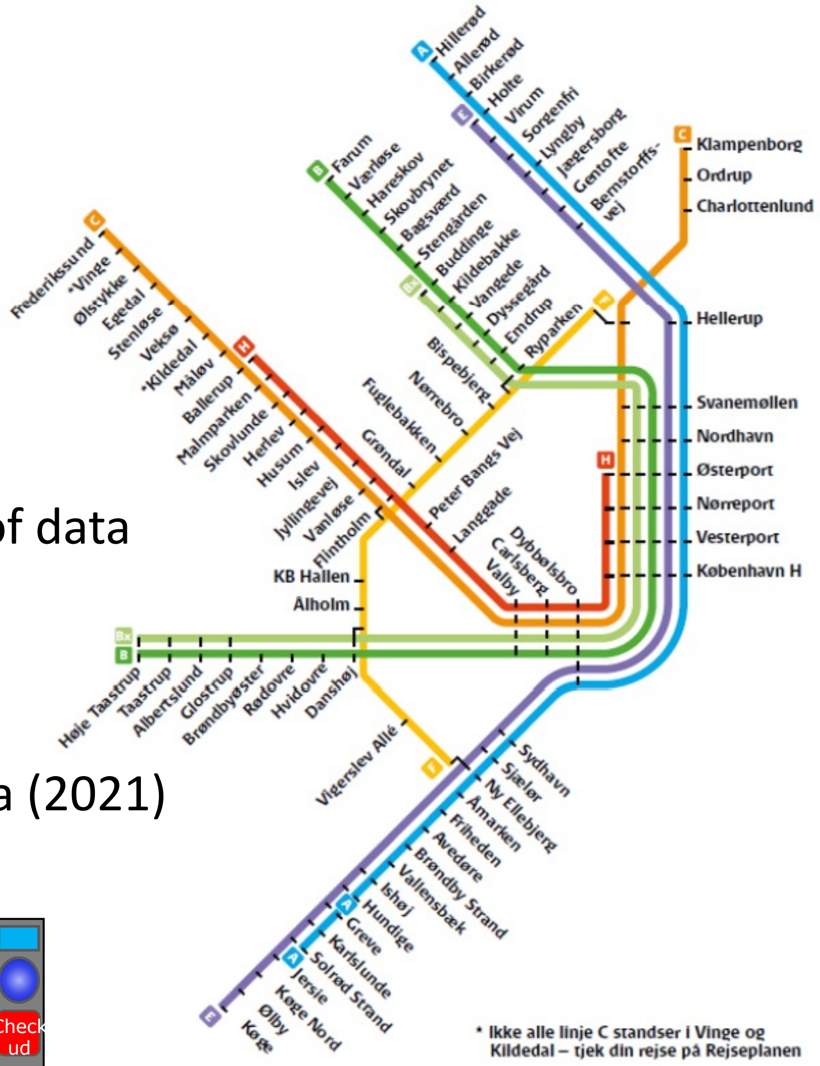
Challenges:

- delayed data
- high-dimensional data
- sparse input/output (many "zeros")

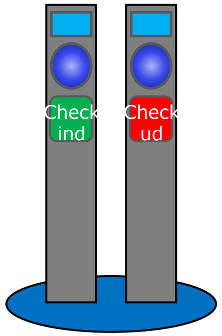
1. Demand Prediction: The Setting

Copenhagen S-bane

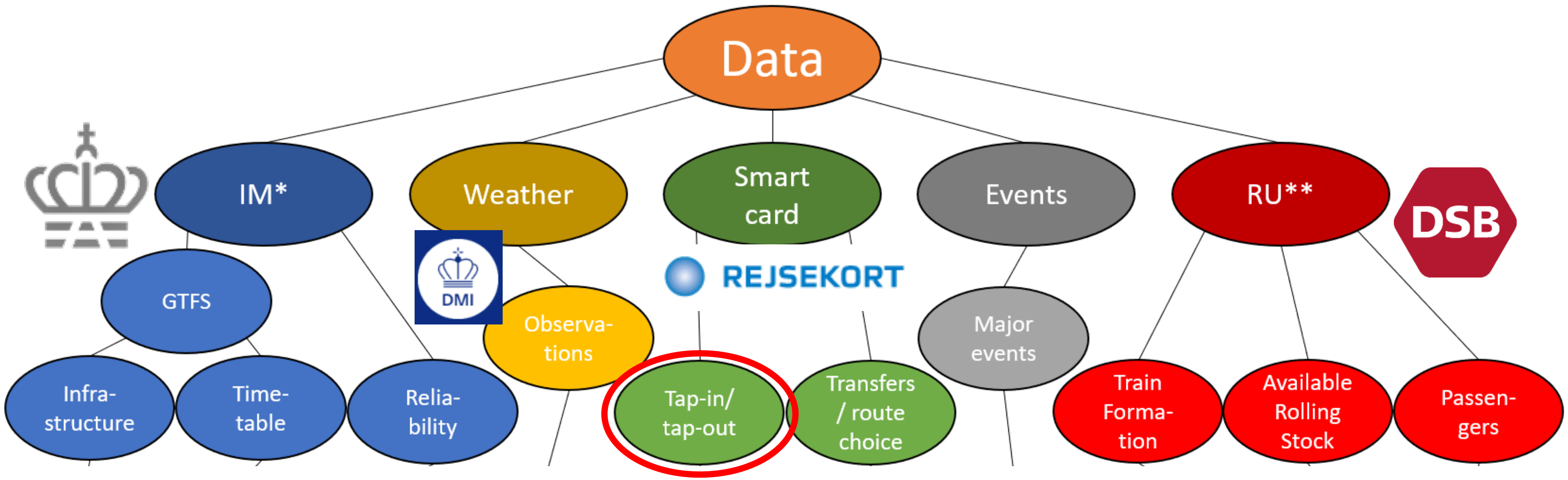
- 100% separated from the main line in daily operation
- ”5 fingers” layout with one ring line (Fingerplan)
- Heavy traffic and many commuters
- Smart cards like Rejsekort automatically collect large amounts of data
  - Large share of travelers use Rejsekort (*)
  - 74.21 million passengers in S-trains (2021)
  - 89.92 million journeys on Rejsekort in the Copenhagen area (2021)
- Estimation stages:
  - Formulation exploration for main 12 OD
  - Refining for Tiny CPH
  - Tollout for full CPH



(DSB, 2021)



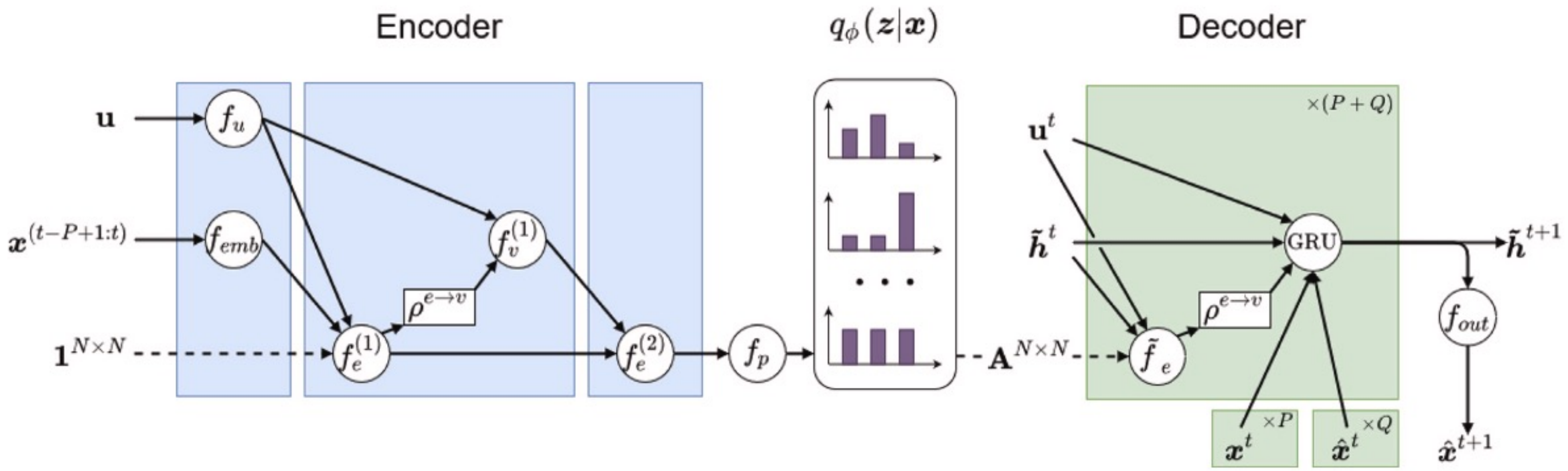
1. Demand Prediction: The Data



- Demand data from smart card (historical OD period, day, and weekly patters + tap-ins)
- Supply data from Banedanmark Regularity and Operational Statistics (RDS)
- RTTP
- Static line-based pathsets for route-choice and passenger assignment

1. Demand Prediction: The Model

- Supervised deep learning
- **Graph Convolutional Network + Neuro Relational Inference**



- Benchmarks:
  - Linear Regression
  - STZINB (Zhuang et al. 2022)

Zhuang, D., Shenhao, W., Koutsopoulos, H., & Zhao, J. (2022). Uncertainty Quantification of Sparse Travel Demand Prediction with Spatial-Temporal Graph Neural Networks. *KDD '22: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.*

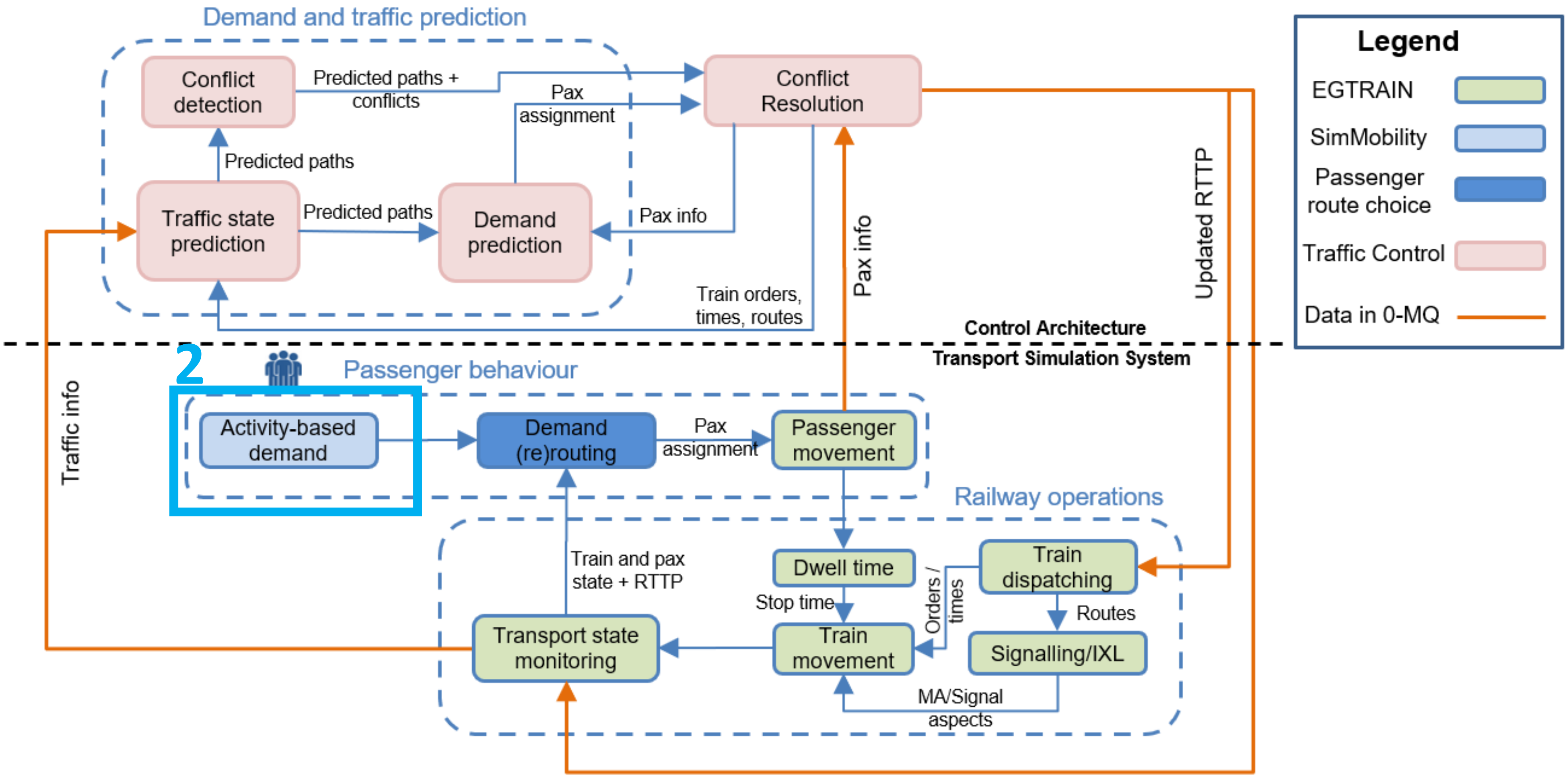
1. Demand Prediction: The Model

	RMSE	MAE
GCN - with lags only	5.006	3.458
STZINB – with lags only	5.159	3.663
NRI - with lags only	4.921	3.449
NRI - adding weather features	4.887	3.454
NRI - adding day of week and time of day features	4.853	3.413
NRI - adding node ID features	4.819	3.382
NRI - adding supply features	4.824	3.390

Table 1: Prediction error statistics for the different models considered (12 largest ODs).

- NRI improves prediction
- Including supply features as input to the model does not lead to a noticeable improvement in **overall** prediction performance (rare events)
- But, more noticeable impact in including supply features in days with disruption (cancelled trains)
- NRI not scalable for full CPH
- Assignment follows OD prediction (based on route choice)

Sortedmobility Framework



2. Activity-Based Demand Modelling Basics

- Travel demand is derived from demand for **activities**
- People face **time and space constraints** that limit their activity schedule choice
- Activity and travel scheduling decisions are made in the context of a broader framework
  - Conditioned by outcomes of **long-term processes**
  - **Interacts** with the transportation system
  - Influenced by **intra-household** interactions
  - Occurs **dynamically** with influence from past and anticipated future events

Ben-Akivai, M., Bowman, J. L., & Gopinath, D. (1996). Travel demand model system for the information era. *Transportation*, *23*, 241-266.

2. Utility Maximization

- **Agent: Decision-Maker**
  - Individual (person/household)
  - State characterized by socio-economic characteristics (e.g., age, gender, income, vehicle ownership)
- **Alternatives**
  - Decision-maker n selects one and only one alternative from a choice set $C_n = \{1, 2, \ldots, i, \ldots, J_n\}$ with J_n alternatives from the environment
- **Attributes of alternatives**
- **Decision Rule**
  - Dominance, satisfaction, utility etc.

2. Multiple Choice

- Logit Model

ε_{jn} is independent and identically distributed (i.i.d.)

$\varepsilon_{jn} \sim ExtremeValue\ (0, \mu), \forall\ j \in C_n$

$$F(\varepsilon) = \exp[-e^{-\mu\varepsilon}], \mu > 0$$

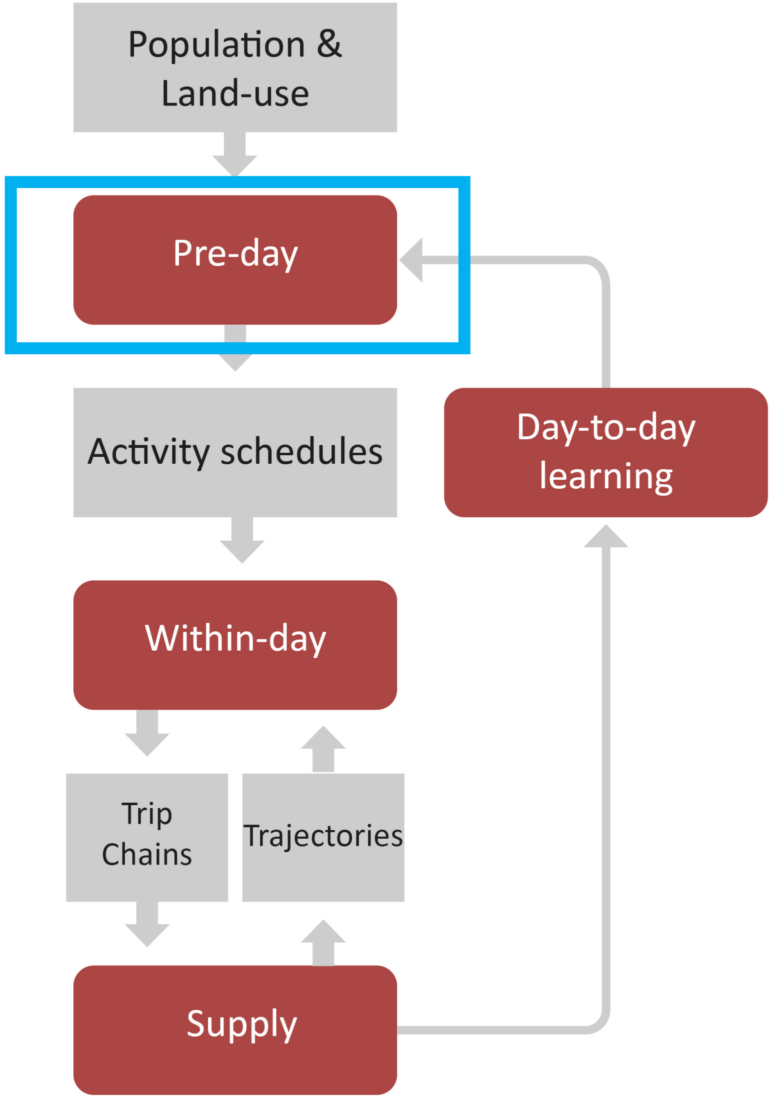
$$f(\varepsilon) = \mu e^{-\mu\varepsilon} \exp[-e^{-\mu\varepsilon}]$$

$$P(i|C_n) = \frac{e^{\mu V_{in}}}{\sum_{j\in C_n} e^{\mu V_{jn}}}$$

2. The SimMobility Framework

- A **laboratory** for analyzing future urban scenarios
- **Integrated/modular** agent-based platform
- Mobility-sensitive **behavioral** dynamic plan/action models
- Local and city-wide multimodal **networks**
- Multiple spatial-temporal **scales**
- **Open-source** started in 2010

https://github.com/smart-fm/simmobility-prod/

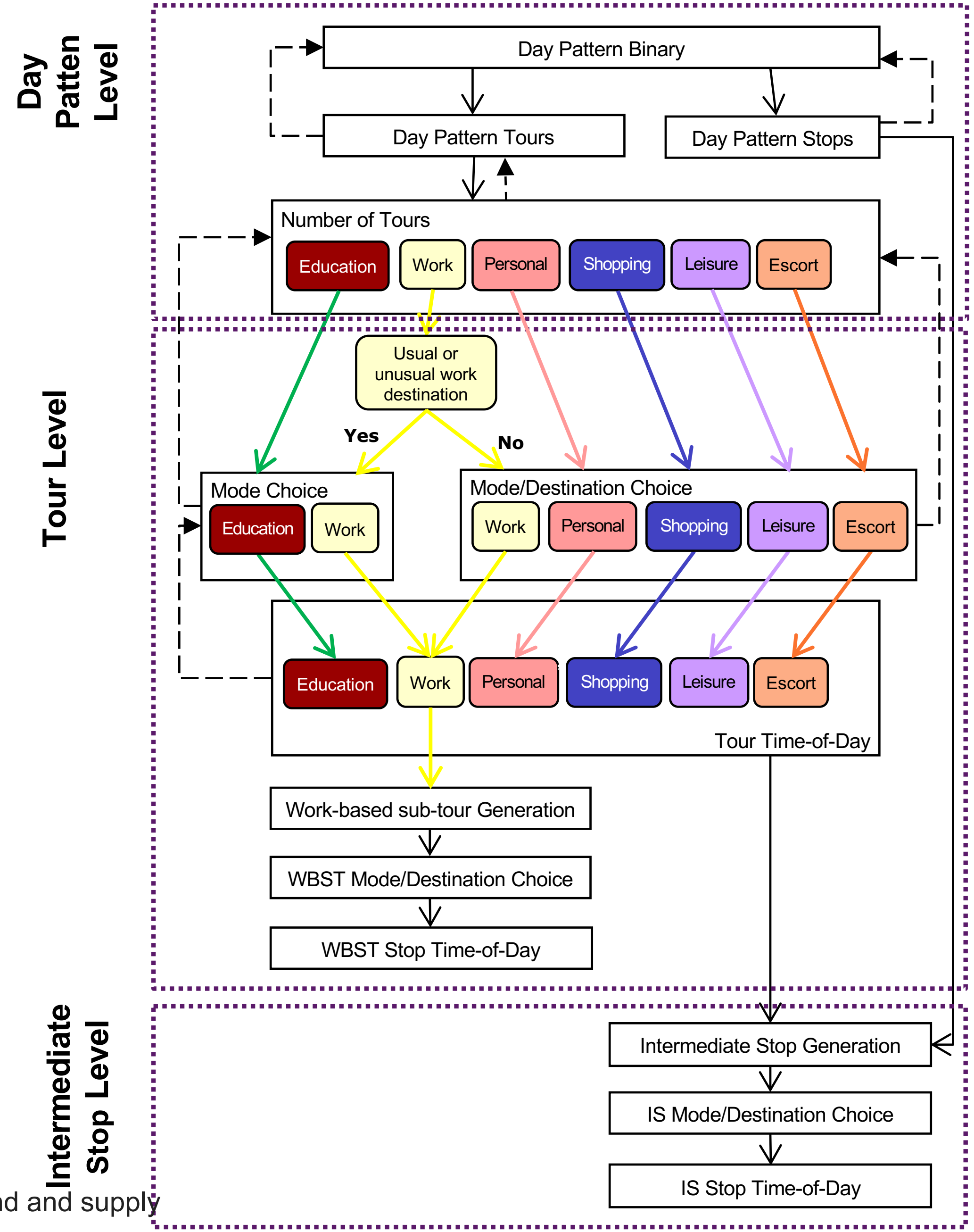


In Sortedmobility we use the SimMobility Mid-Term Pre-Day

Lu, Yang, et al. (2015) Simmobility mid-term simulator: A state of the art integrated agent based demand and supply model. In: *94th Annual Meeting of the TRB Washington, DC*. 2015.

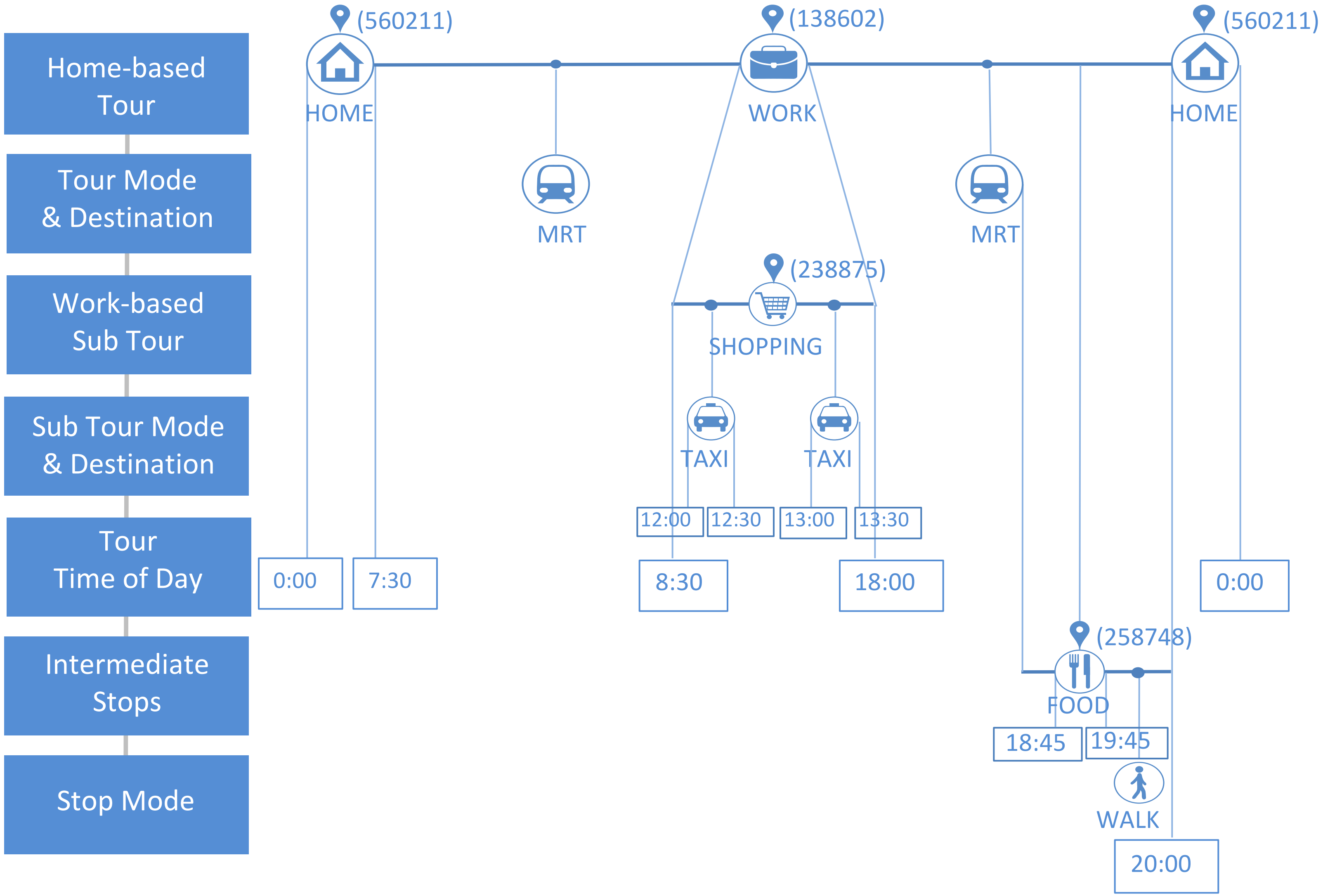
2. SimMobility Pre-day

- Day Activity-Schedule
- Predicts agents' activity-travel schedule (plan) including
  - Activity types and number
  - Activity durations and time-of-day
  - Activity locations
  - Modes (bus, rail, drive alone, shared ride, motorbike, taxi, walk)
- Choices are sensitive to level-of-services, upstream choices and accessibility gains from downstream choices



Lu, Yang, et al. (2015) Simmobility mid-term simulator: A state of the art integrated agent based demand and supply model. In: *94th Annual Meeting of the Transportation Research Board, Washington, DC*. 2015.

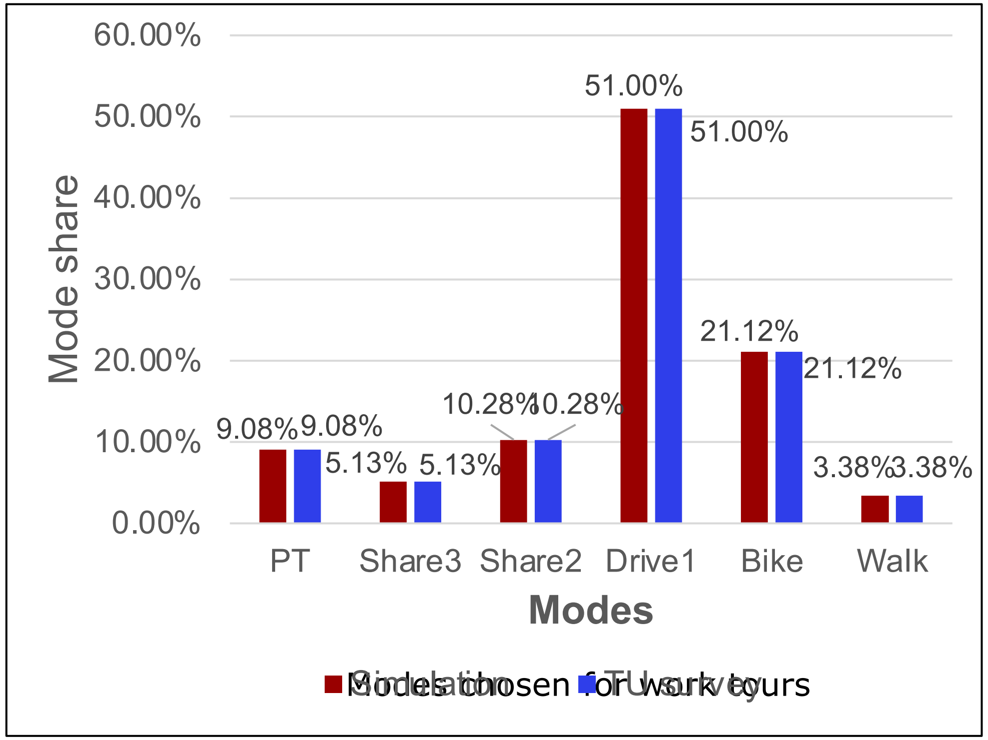
2. SimMobility Pre-day Output

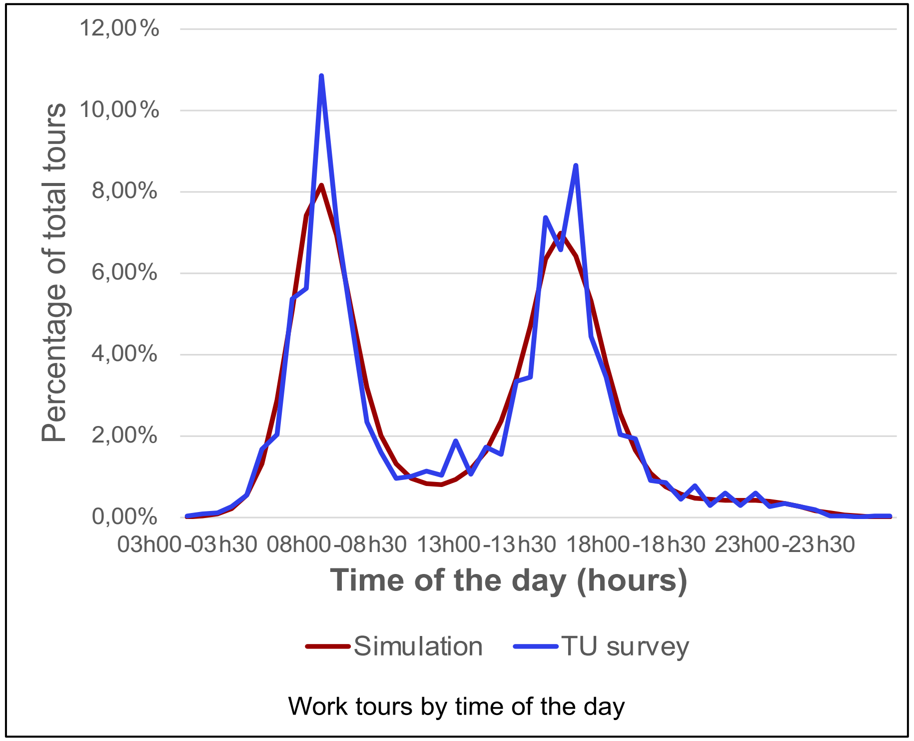


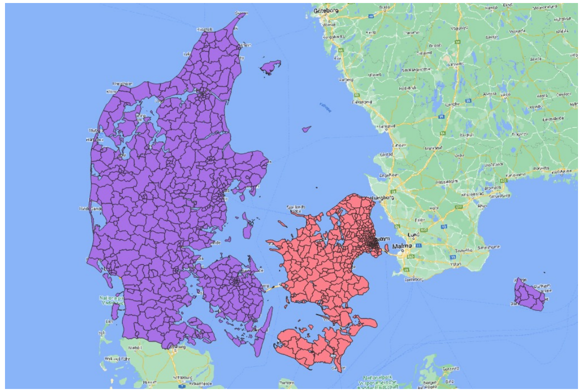
2. The model built for Denmark

- Demand: Full Denmark (5,8M)
- Supply: Zealand, Falster, Lolland, Møn
- Bus lines, metro, regional and urban trains

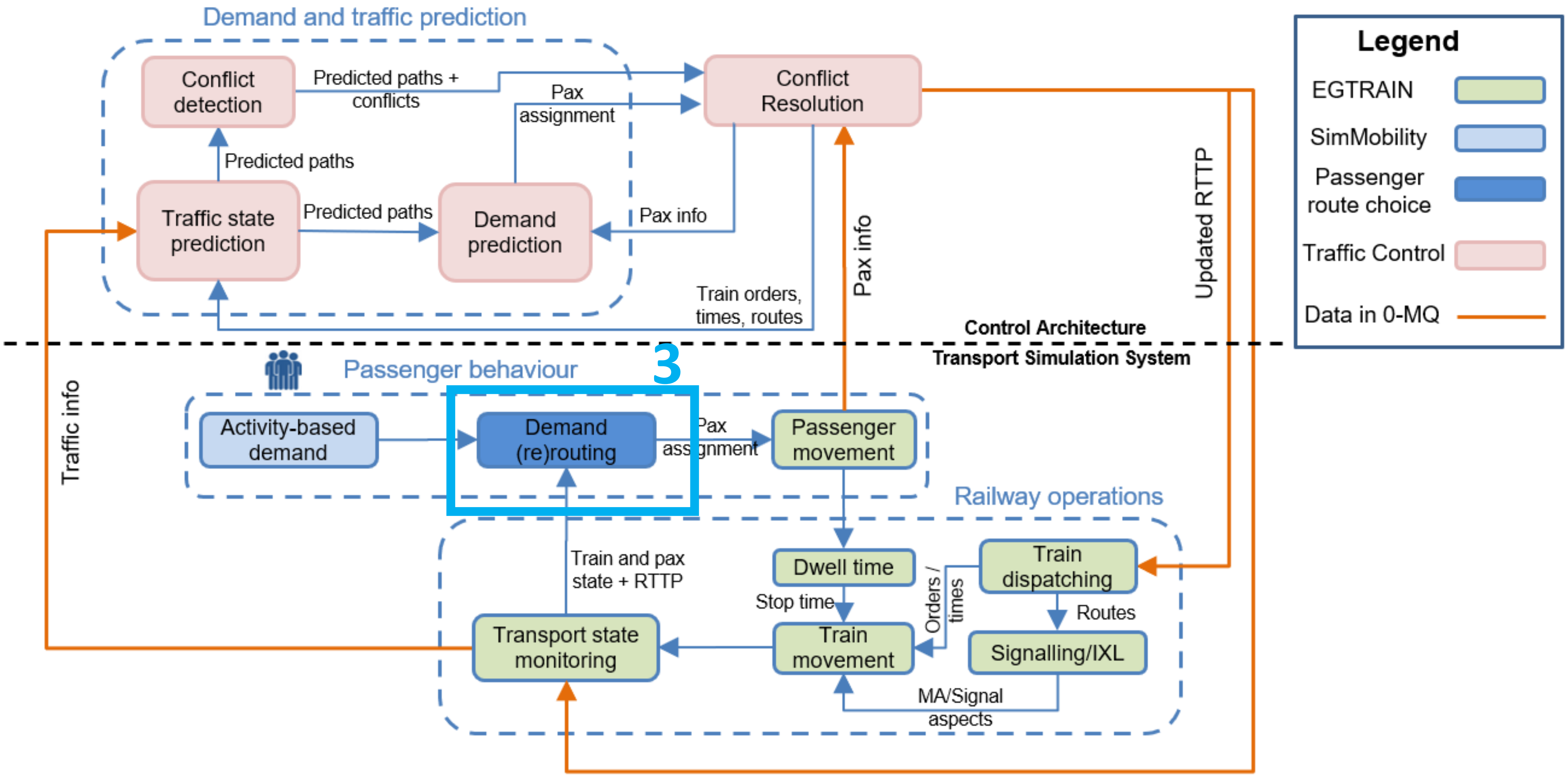
- Data from the Danish National Travel Survey (TU)
- Skim matrices from the Danish Road Directorate (VD)
- Acctivity diaries for 19,588 individuals from 2017 to 2019.



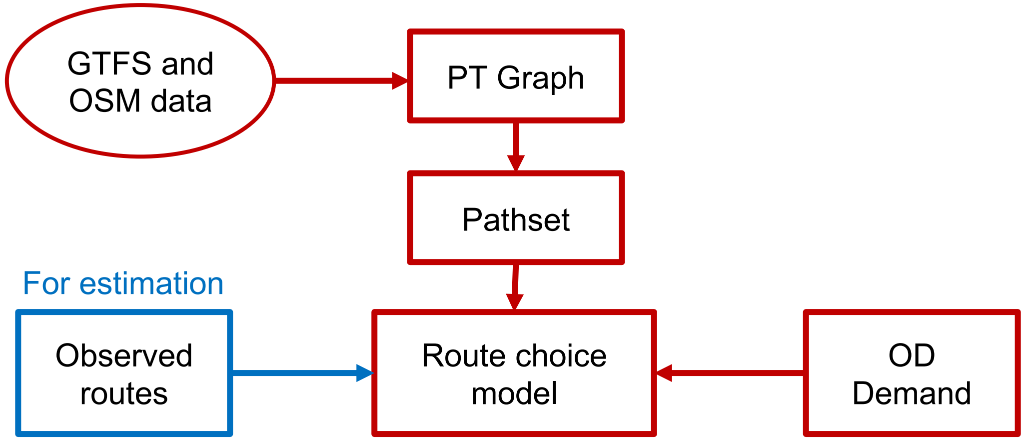




Sortedmobility Framework



3. Route Choice Model Formulation



- Formulation based on discrete choice – random utility theory
- Decision-maker ***n*** associates a utility ***U*** to each available alternative ***j*** at time period ***t***:

$$U_{njt} = V_{njt} + \varepsilon_{njt} = \beta_X X_{njt} + \varepsilon_{njt}$$

Random disturbance term

Vector of unknown parameters

Vector of attributes of alternative j including a constant

$$P_{njt} = \frac{e^{\beta_X X_{njt}}}{\sum_{j'=1}^{J} e^{\beta_X X_{nj't}}}$$

- ***J*** total number of available alternatives between OD pairs (based on the generated pathset)
- Account for correlation between alternatives

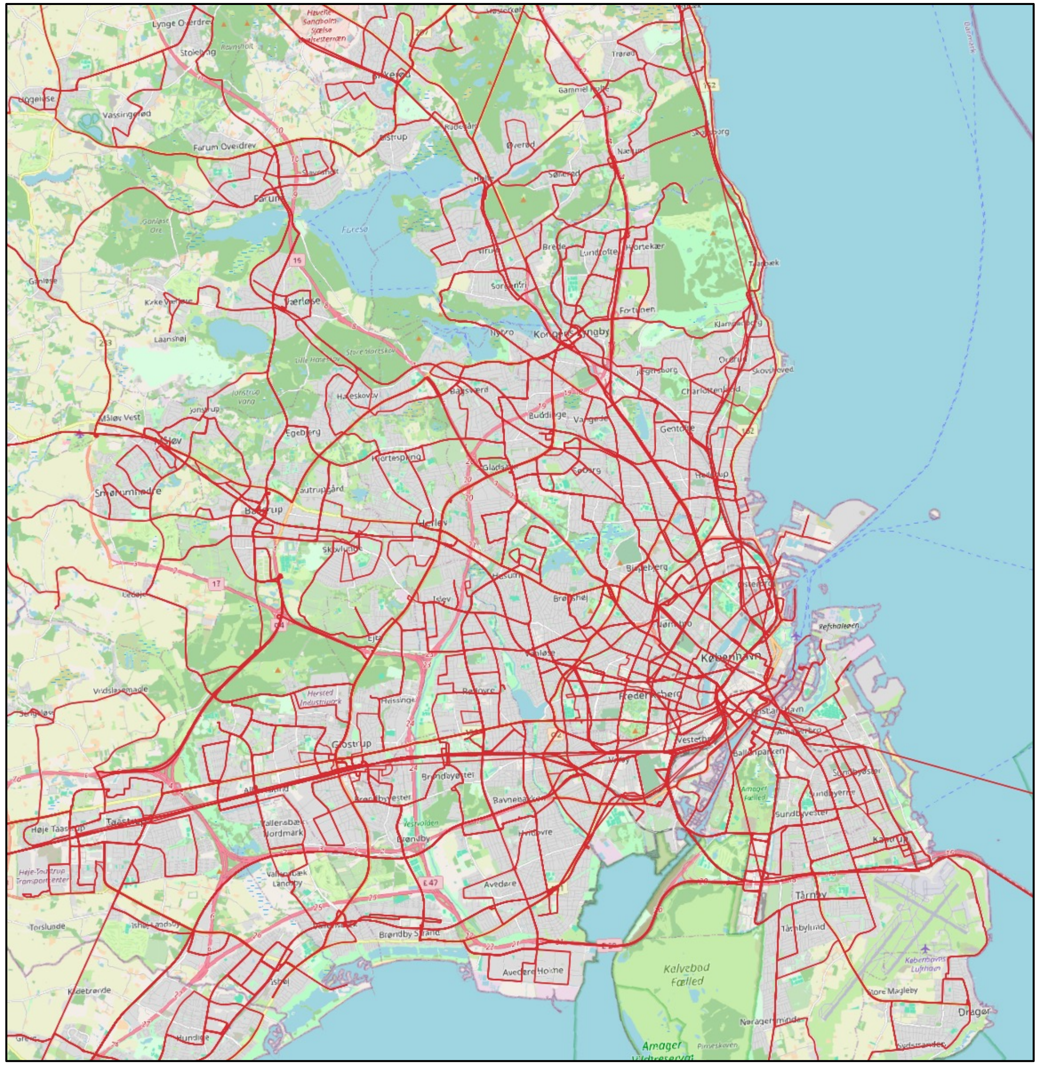
Tan, R. et al. (2015). New path size formulation in path size logit for route choice modeling in public transport networks. *Transportation Research Record*, *2538*(1), 11-18.

3. PT Graph

- A PT graph contains:
  - a set of vertices
  - a set of edges connecting pairs of those vertices
- The vertex set consists of:
  - nodes from the road network
  - bus stops and rail stations in the network
- The edge set consists of:
  - PT edge: If a bus- or rail-line serves stop a and then eventually stop b, there is an edge (a, b) in the edge set representing this line service
  - walk edge:
    - For access and egress: connecting each node vertex to each bus- or rail-stop vertex (and vice-versa) if the distance between the node and the PT stop is "walkable“
    - For transfers: connecting every pair of PT stops that are at a "walkable" distance from one-another

3. Example of a PT graph

- Area: East Great Belt
- Modes: Bus, Train, Metro
- Summary of the network:
  - 11,419 bus stops
  - 1,425 bus lines
  - 244 train and metro stops
  - 46 train and metro lines
  - 124,368 nodes (road network)
  - 784,603 road segments
- Summary of PT Graph:
  - 136,031 vertices (nodes and stops)
  - 3,513,457 edges



3. Pathset

- A traveler who wants to travel from Origin (O) to Destination (D) is required to select a route to navigate the network from O to D
- A set of route alternatives (Pathset) between all Origin-Destination pairs is needed:
  - to estimate a route choice model
  - for simulation
- A PT path is a sequence of:
  - access trip
  - PT trip
  - transfers
  - egress trip
- Generation of path alternatives on the PT graph
  - K shortest path
  - Link elimination
  - Random perturbation
  - other algorithms…

Paths for 86,128 Ods (1 month) were generated, with a coverage rate of 98.88%.

3. The Model

Variable	Value	Rob. Std err
β_{IVTT}	-0.186	0.003
$\beta_{Nb\_transfers}$	-1.398	0.022
β_{PS_j}	0.801	0.022
$\beta_{TT\_waiting}$	-0.049	0.009
$\beta_{TT\_walking}$	-0.506	0.004

Variable	Value	Rob. Std err	Rob. t-test	Rob. p-value
β_{IVTT}	-0.186	0.003	-70.70	0.000
β_{MEDIAN}	-0.106	0.008	-13.56	0.000
$\beta_{Nb\_transfers}$	-1.403	0.021	-65.46	0.000
β_{PS_j}	0.812	0.022	36.20	0.000
β_{STD_M}	0.376	0.028	13.35	0.000
$\beta_{TT\_waiting}$	-0.047	0.009	-5.24	0.000
$\beta_{TT\_walking}$	-0.506	0.004	-125.77	0.000

Parameter	Value
No. of parameters	5
Sample size	165,548
Excluded data	0
Null log likelihood	-303,476.1
Final log likelihood	-121,490.4
Likelihood ratio test (null)	363,971.3
Rho square (null)	0.6
Rho bar square (null)	0.6
Akaike Information Criterion (AIC)	242,990.8
Bayesian Information Criterion (BIC)	243,040.9

$$U_{njt} = V_{njt} + \varepsilon_{njt} = \beta_X X_{njt} + \varepsilon_{njt}$$

$$P_{njt} = \frac{e^{\beta_X X_{njt}}}{\sum_{j'=1}^{J} e^{\beta_X X_{nj't}}}$$

Summary

- Sortedmobility proposed passenger **prediction, generation and behaviour** for rail
- State-of-the-art demand models were included **in both operations and simulation** of future rail systems
- **Scalability and case-specificity** are still bottlenecks for a widespread usage
- **Smart-card** (especially tap -in and –out) makes a difference in model performances