CROWDEDNESS

Gemeente Amsterdam

Universiteit van Amsterdam

Using varied data sources to predict crowdedness and congestion levels across the city. Methods employed range from agent based simulation to machine learning. Insights gathered from models and simulations developed can be adopted to make strategic urban and transport planning decisions.

Agent-based Modeling of Public Transport

Sophie Ensing

RESEARCH QUESTION

How can an **agent-based model** be used to assess the effects of changes in **public transport** on passenger flows?

Figure 1. The subset of the public transport network that was used.

An agent-based model was created to simulate passenger flows through the public transport network in Amsterdam. First, a baseline model was created with the help of GVB data, literature and domain knowledge. After that, the model was used to assess the effects of changes in the public transport network. Only tram 24 and 12 and metro 50, 52 and 53 were used. Two scenarios were simulated. A regular scenario and a scenario where tram 12 experiences problems.

THE MODEL

Class 1: people who prefer the shortest trip

Class 2: people who prefer minimal transfers

Class 3: people who prefer minimal waiting time

Three different types of agents were created to simulate a population. For every person an origin and destination was assigned based on a probability distribution for all stations. Based on the class type, their optimal route was calculated. It was not possible to compare the complete routes with GVB data, but parts of the trip could be evaluated. This resulted in a Mean Average Error of 13 to 36 people per trip.

SCENARIO 1 vs. SCENARIO 2

(Results between 08:00 - 09:00 on weekdays. Metro 50 was not affected so is not included.)

Fig 2. Tram 12 occupancy

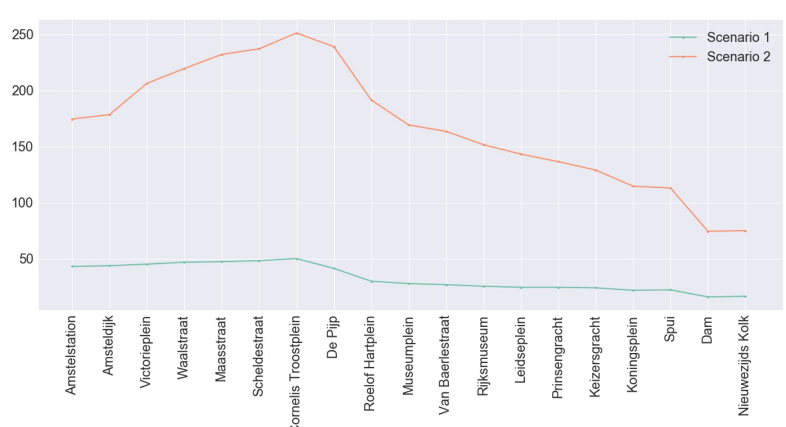


Fig 3. Metro 53 occupancy

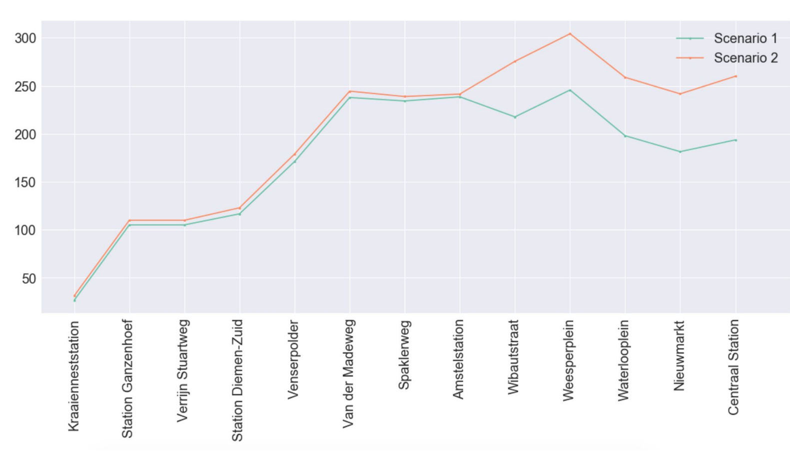


Fig 4. Metro 52 occupancy

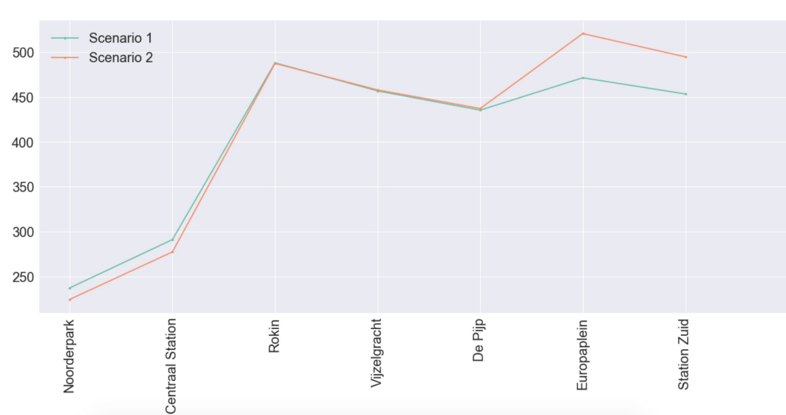
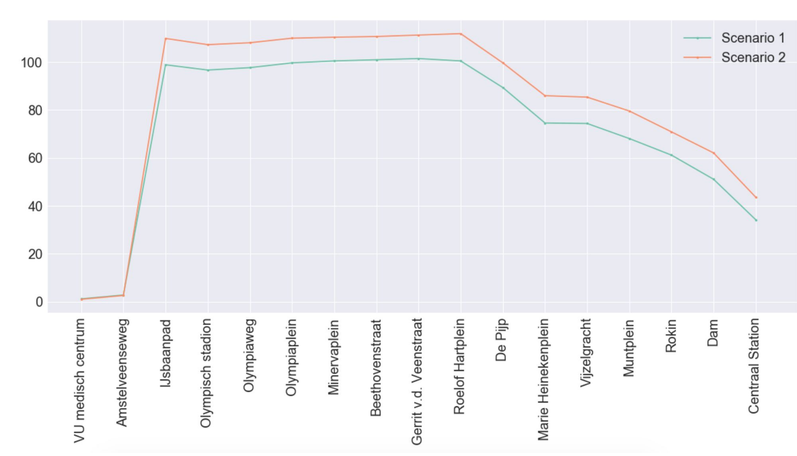


Figure 5. Tram 24 occupancy



RESULTS AND CONCLUSION

- The results show that the occupancy of all lines except metro 50 are affected during/after the malfunctions. There is either a peak in a different place during the malfunction because people choose a different route or after the malfunction because people were not able to reach their transfer.
- All effects were found for all simulations at different times/days.
- With the use of GVB data, domain knowledge and literature an agent-based model was created to assess the effects of changes in public transport.
- Future work should be aimed at implementing the model for the complete network of Amsterdam and increasing the complexity of the agents.

Crowdedness Prediction Model

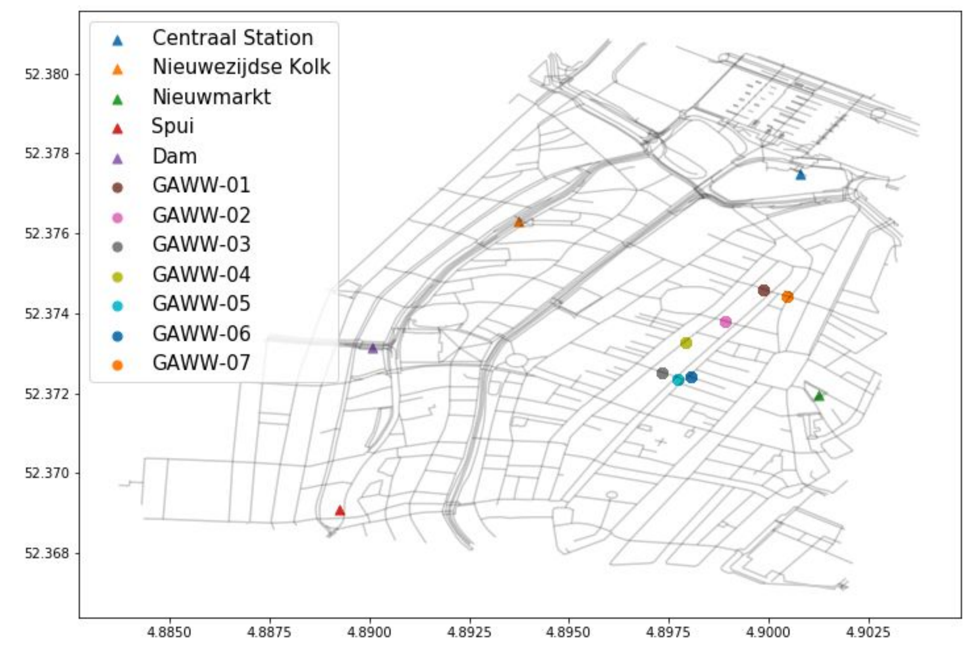
Don de Lange

Research Question

How can a **prediction** of the level of **crowdedness** within the city of Amsterdam be given, based on input from city-wide available **data sources**?

Data

Fig 1 Map Sensors & Stations



Single Dataset:

1. Public Transport Data:
2. Sensor Crowdedness Counts
3. Event Dates

Transformations:

1. Circular Time
2. Scale Longitude & Latitude
3. Compute Distances from sensor to station
4. Transform Passenger counts

Prediction Models

80% **Train using Cross-Validation** 20% **Evaluation**

Results

Regression		
Model	**R^2**	**RMSE**
Baseline	49.2%	718.15
Random Forest	83.3%	409.19
XGBoost	*85.2%*	*389.34*

Classification			
Model	**Accuracy**	**Precision**	**Recall**
Baseline	24.1%	-	-
Random Forest	84.4%	84.4%	84.4%
XGBoost	*85.8%*	*85.8%*	*85.8%*

Fig 2 Plot predictions

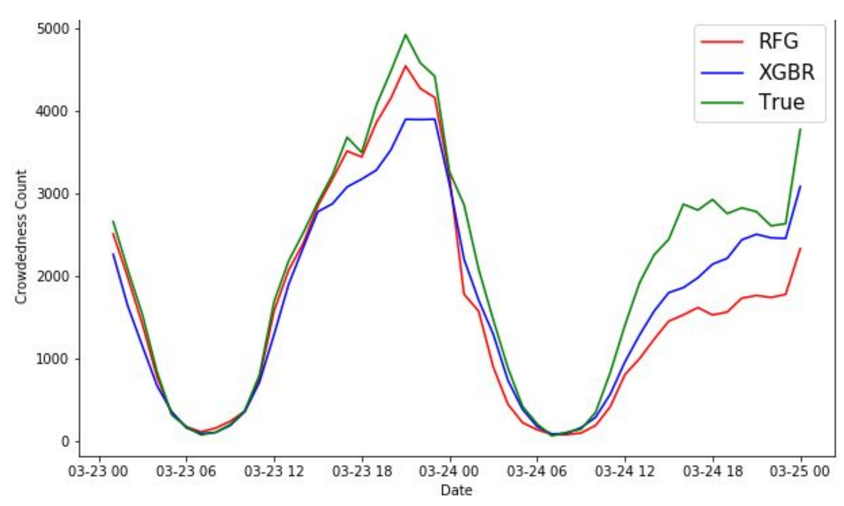
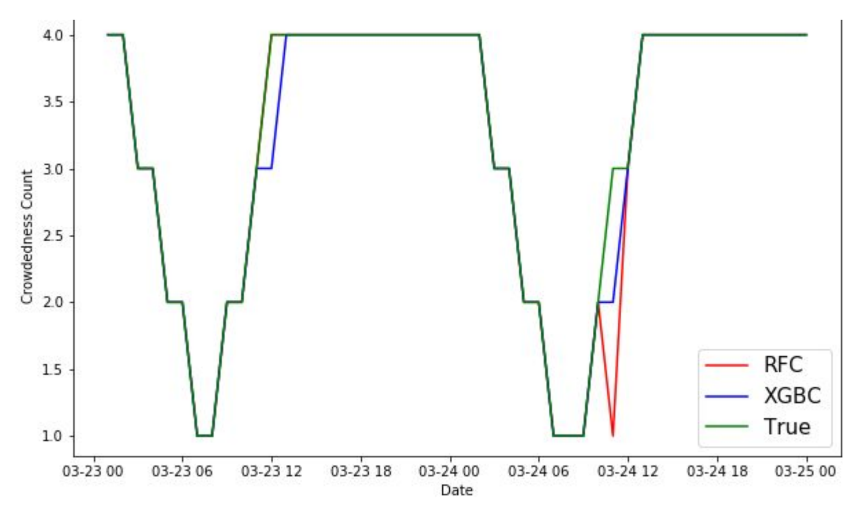


Fig 3 Plot predictions



Conclusion

- With the use of **GVB** and **CMSA** data, the needed **input** data can be formed
- **Random Forest** and **XGBoost** provide accurate **predictions** of crowdedness, with **XGBoost** outperforming Random Forest