USC Viterbi
School of Engineering
Integrated Media Systems Center

Data-driven time-dependent freight volume estimation

Chrysovalantis Anastasiou, *Luciano Nocera, Cyrus Shahabi*
Integrated Media Systems Center
Viterbi School of Engineering
University of Southern California

Motivation

- Urban planning
  - Introduce new lanes/roads where trucks over-congest the network
  - Reinforce or more frequently maintain roads that are more likely to be damaged by trucks
- Air quality
  - Effect of trucks on air pollution in areas they frequently pass by or drive to

Long Beach to Open New $1.5 Billion Gerald Desmond Bridge

By Howard Fine

Monday, October 5, 2020

Like 21 Share Share Tweet

With three lanes in each direction, the bridge will be able to accommodate more truck traffic. The old bridge carried up to 16,000 trucks per day. Photo by Ringo Chiu.

TRANSPORTATION AND AIR QUALITY

Heavy Trucks Cause Much Of Our Air Pollution. A New State Rule Aims To Change That

Updated June 26, 2020 1:10 PM | Published June 26, 2020 1:10 PM

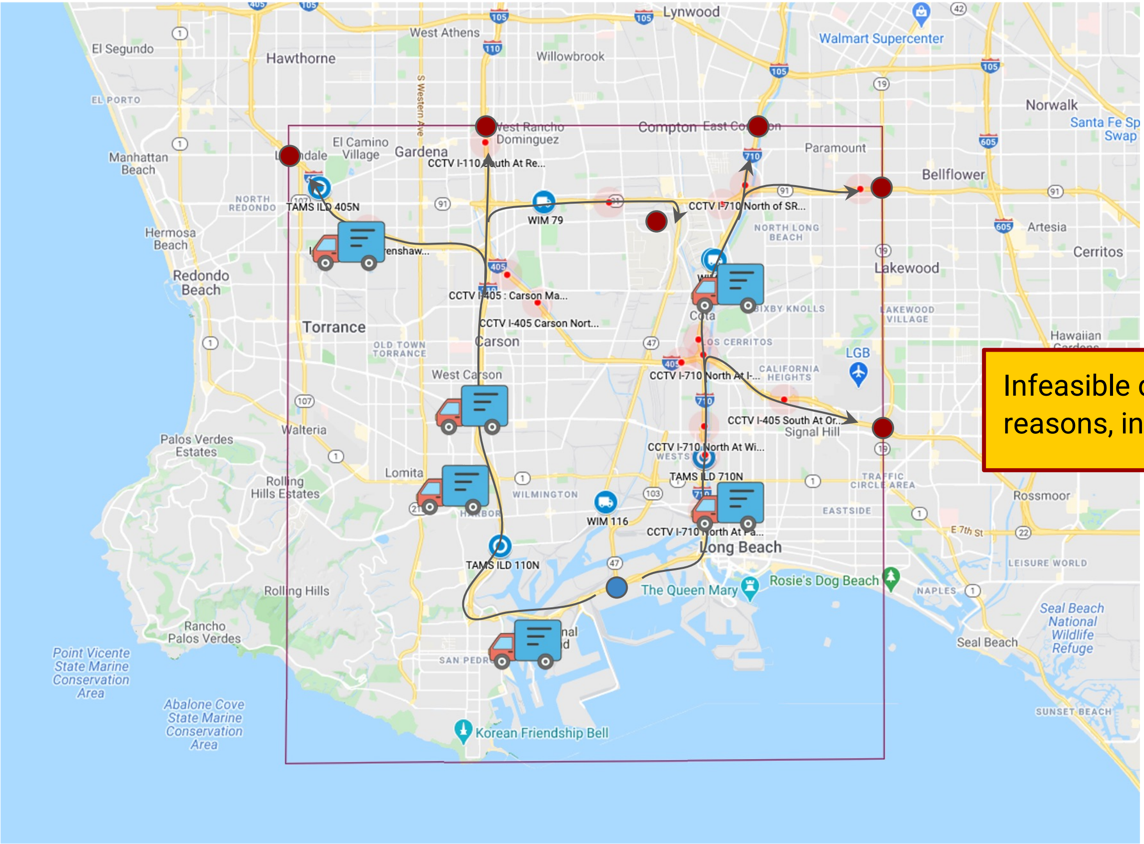
Trucks stand prepared to haul shipping containers at the Port of Los Angeles on Sept. 18, 2018. (Mario Tama/Getty Images)

COVID-19 Reveals That the Real Cure For Freight Truck Congestion is Fewer Cars

And no, adding highway lanes *still* doesn't cut down on traffic.

By Kea Wilson | Mar 25, 2020 | 25 COMMENTS

Ideal Scenario -- GPS Tracking



Legend

- Origin
- Destination

Infeasible due to many reasons, including privacy...

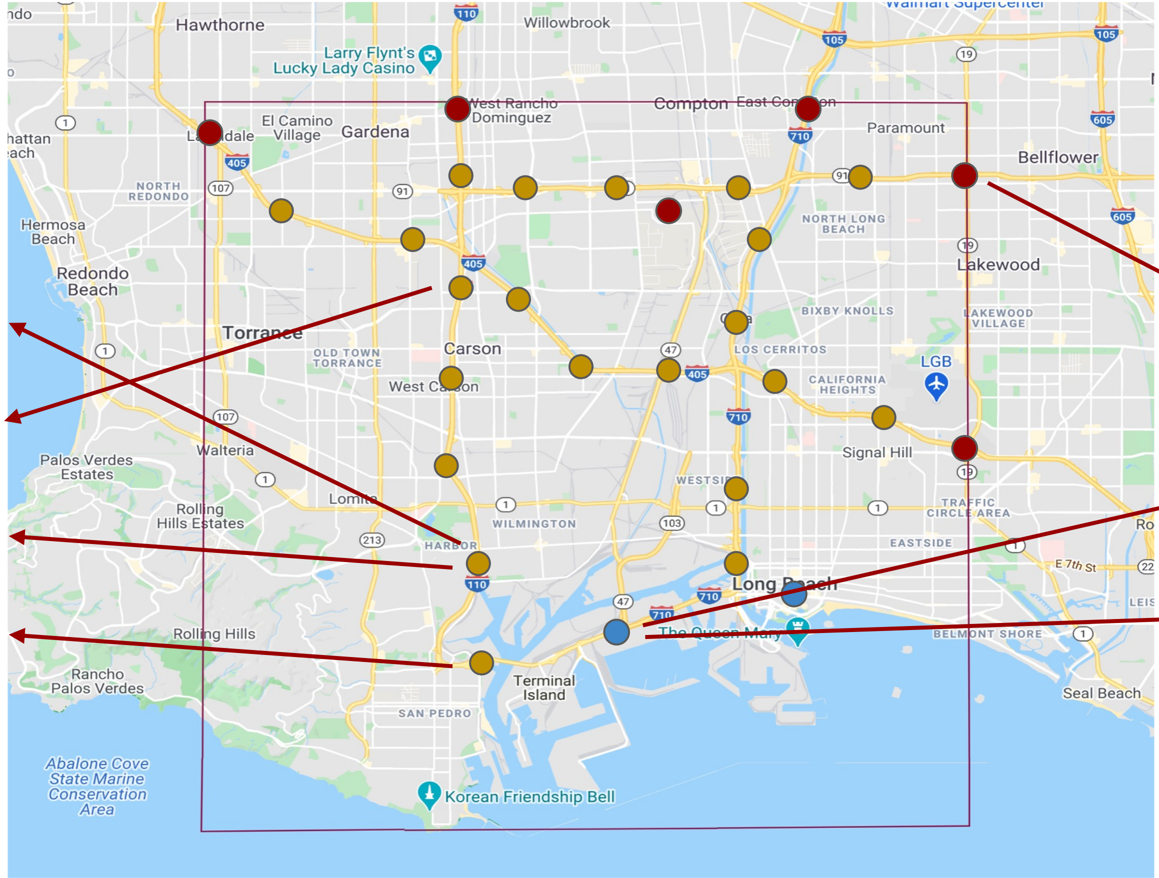
Reality -- Discrete Sensor Observations

Maybe *large* truck at time t_2.

Maybe *small* truck at time t_2.

Maybe *medium* truck at time t_3.

Maybe *small* truck at time t_1.



Legend

- Origin
- Destination
- Sensor

No idea what is going on at time t_{any}.

One *large* truck at time t_2.

One *small* truck at time t_0.

Outline

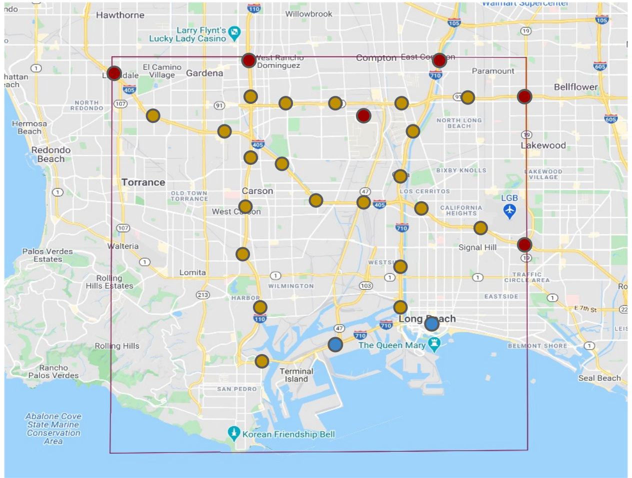
- Motivation
- Problem Statement
- Data Sources
- Algorithms
  - Baseline
  - Naive / FlowPath
  - Reachability-based
- Experiments

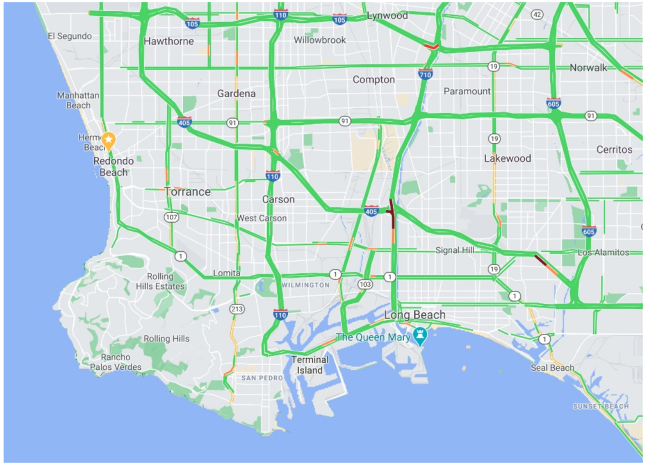
Problem Statement

Given a region of interest **R**, its road network **G**, and a sensor-based dataset **ϴ**, estimate the volume of truck movements per time unit (e.g., 1 hour).

Research goal:

To *accurately* estimate the *time-dependent* flow of trucks in a road network.





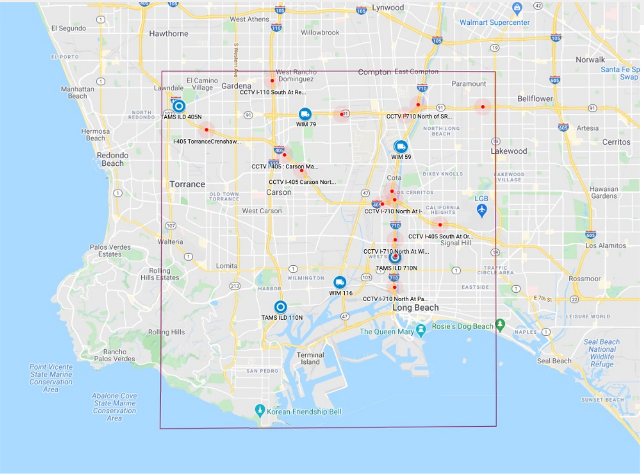
Sensor Data

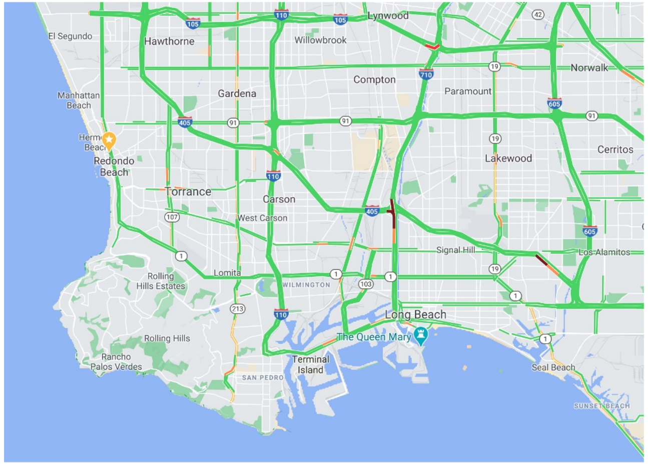
- RFID sensors **(very accurate)**
  - Typically at port exits; truck “check-out”
  - <location, timestamp, truck type> + *<truck id>*
  - Refers to a specific truck id
- Weigh-in-motion (WIM) sensors **(very accurate)**
  - Sparse but provide checkpoints
  - <location, timestamp, truck type> + *<truck id>*
- TAMS [1] sensors **(accurate)**
  - Sparse and probabilistic
  - <location, timestamp, truck type prob.>
- CCTV cameras **(variable accuracy)**
  - Sparse and probabilistic
  - <location, timestamp, truck type prob.>

Truck detection; “checkpointing”

- Inductive Loop Detectors
  - ADMS [2]

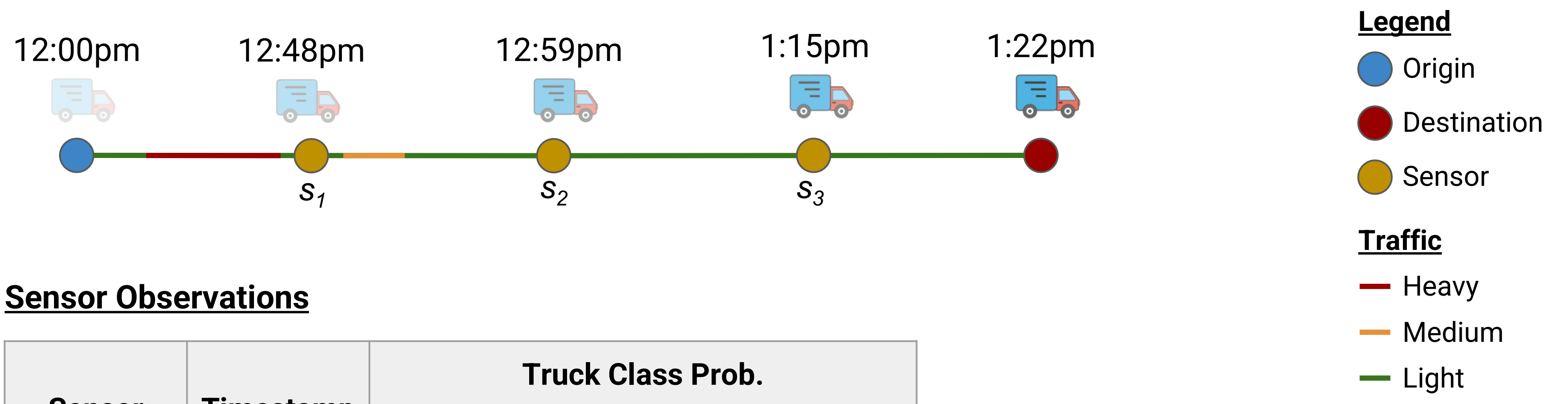
Traffic





[1] Tok, Yeow Chern & Hyun, Kate & Hernandez, Sarah & Jeong, Kyungsoo & Sun, Yue & Rindt, Craig & Ritchie, Stephen. (2017). Truck Activity Monitoring System (TAMS) for Freight Transportation Analysis. Transportation Research Record Journal of the Transportation Research Board. 2610. 10.3141/2610-11.
[2] Anastasiou, Chrysovalantis, Jianfa Lin, Chaoyang He, Yao-Yi Chiang, and Cyrus Shahabi. "Admsv2: A modern architecture for transportation data management and analysis." In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Advances on Resilient and Intelligent Cities*, pp. 25-28. 2019.

Sensor Observation Examples



Sensor Observations

Sensor	Timestamp	Truck Class Prob.		
		Small	Medium	Large
s_1	12:48pm	0.04	0.96	0.00
s_2	12:59pm	0.07	0.92	0.01
s_3	1:15pm	0.05	0.93	0.02

Truck Flow Estimation

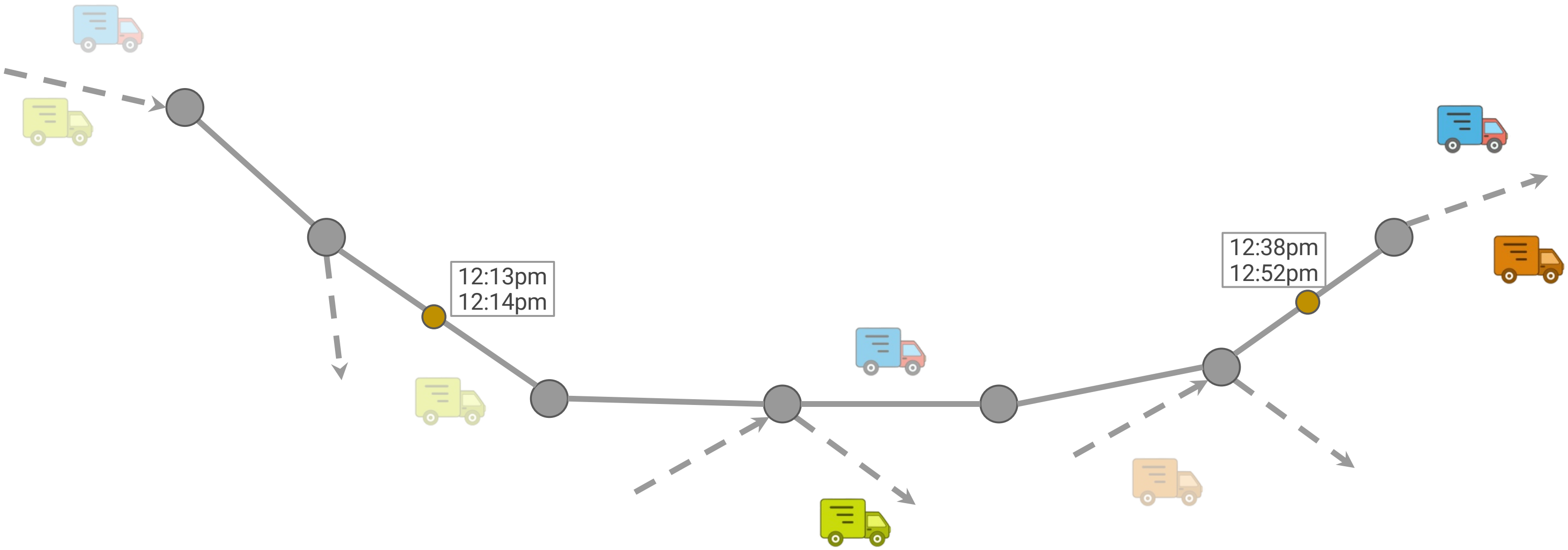
- Input
  - **G** = (V, E) the road network
  - **S** the set of sensors
  - **θ** the set of sensor observations
- Output
  - **T** the set of truck flow time-series; one per edge/road segment

Legend

Intersection

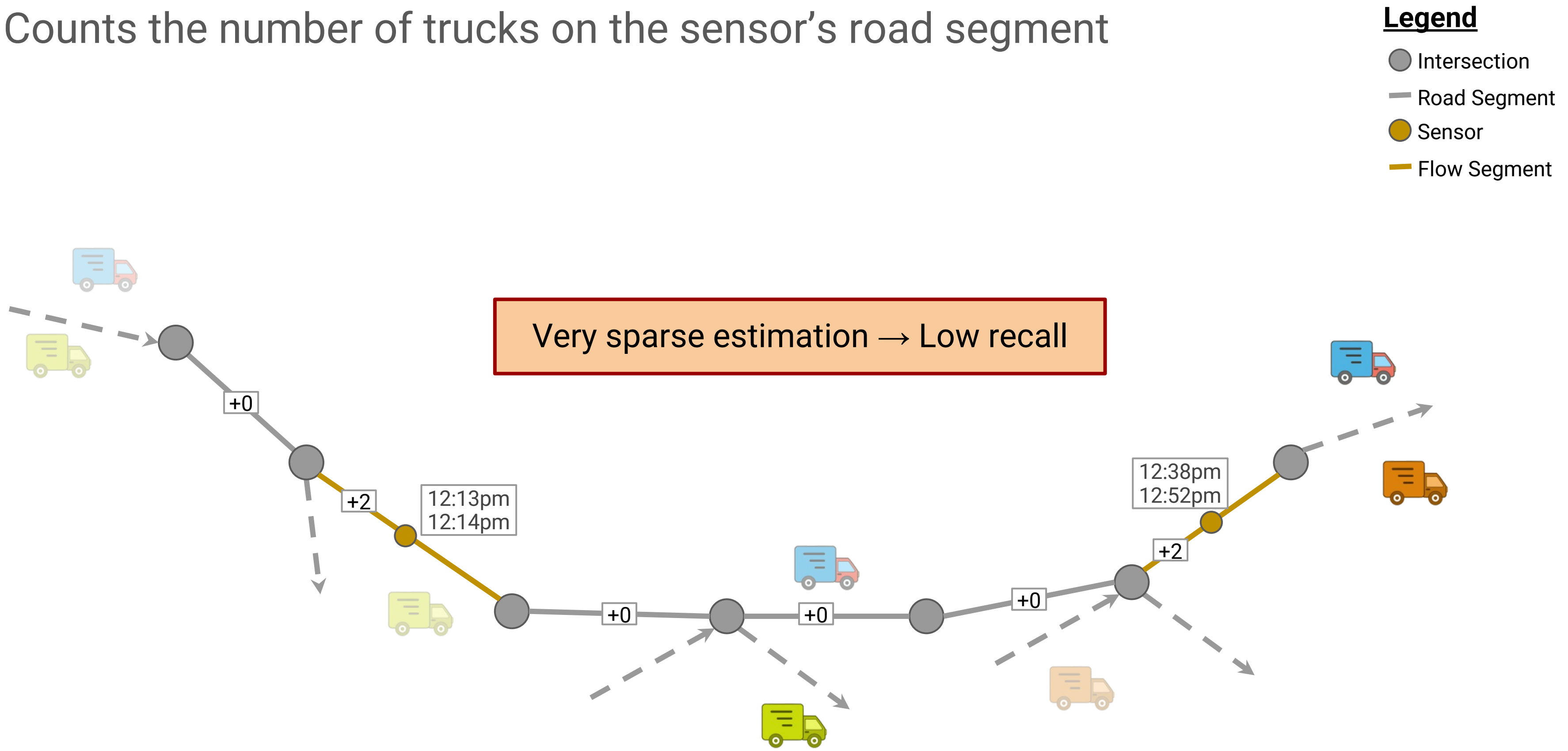
Road Segment

Sensor



Baseline Algorithm

- Counts the number of trucks on the sensor's road segment

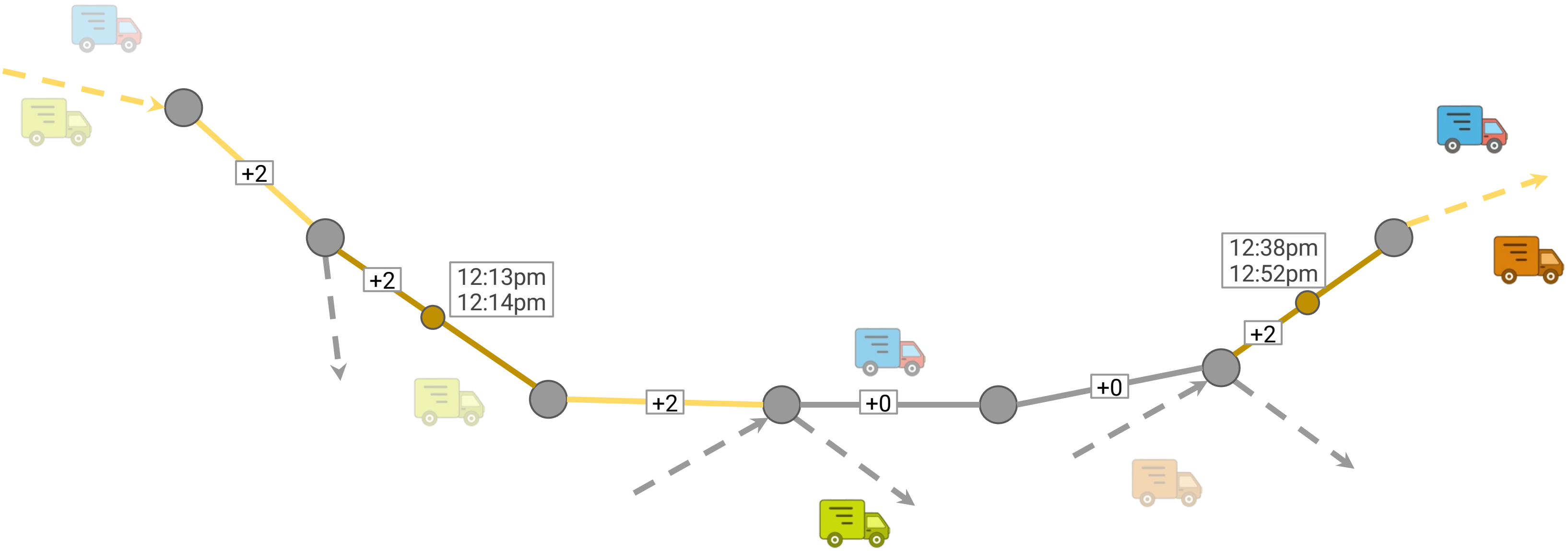


Naive Flow Path Expansion

- Counts the number of trucks on the sensor's road segment.
- Expands backwards and forwards as long as intersection does not affect flow count.

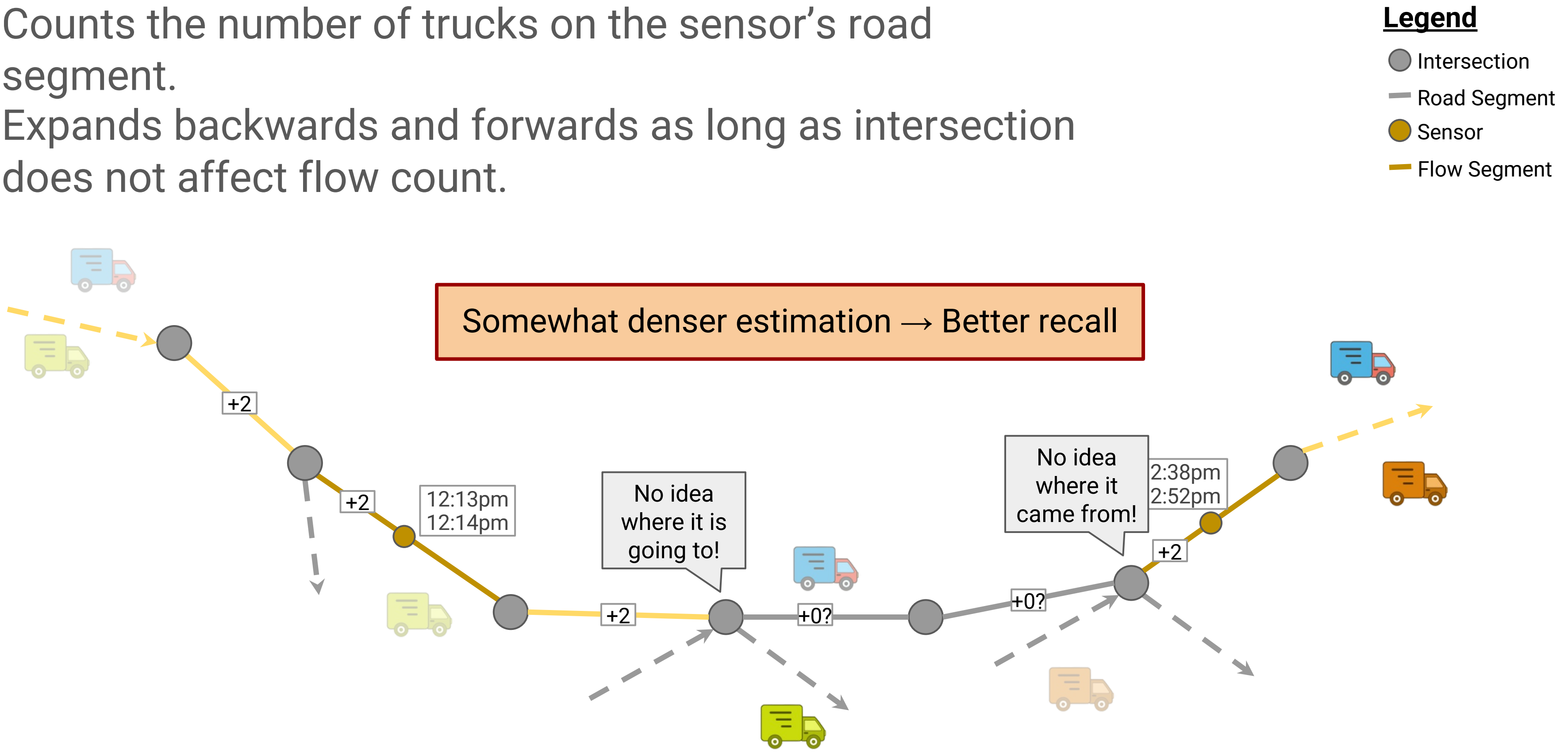
Legend

Intersection
Road Segment
Sensor
Flow Segment



Naive Flow Path Expansion

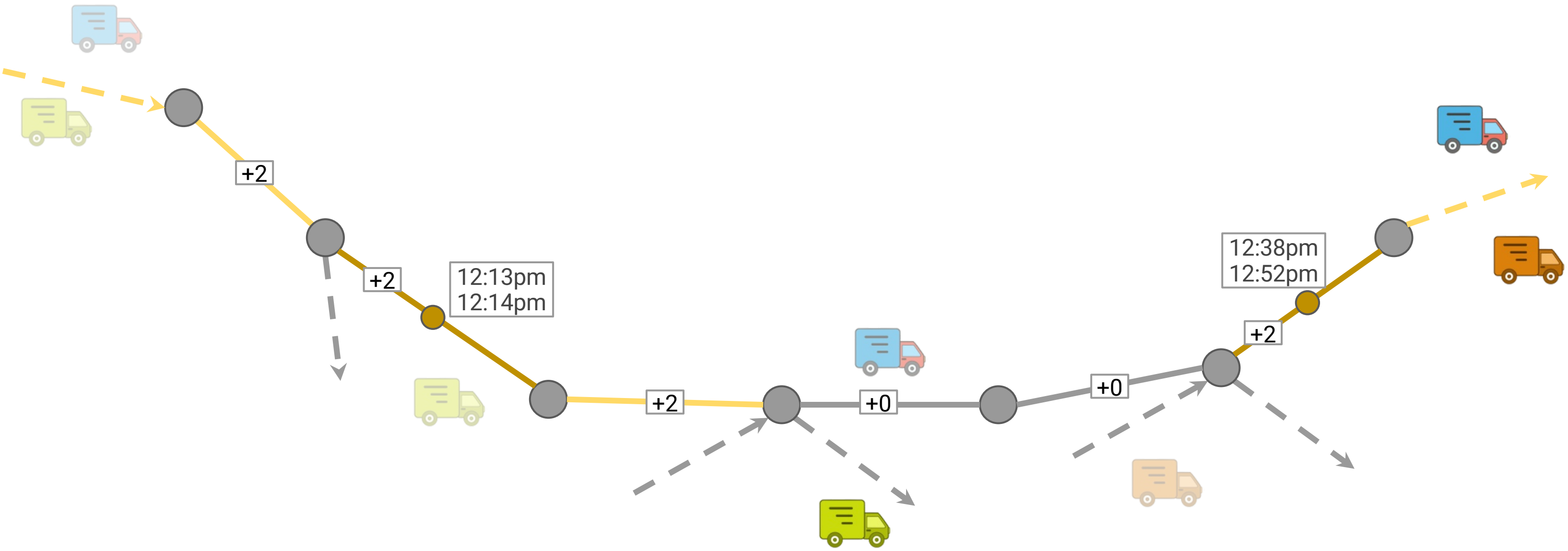
- Counts the number of trucks on the sensor's road segment.
- Expands backwards and forwards as long as intersection does not affect flow count.



Reachability-based Estimation

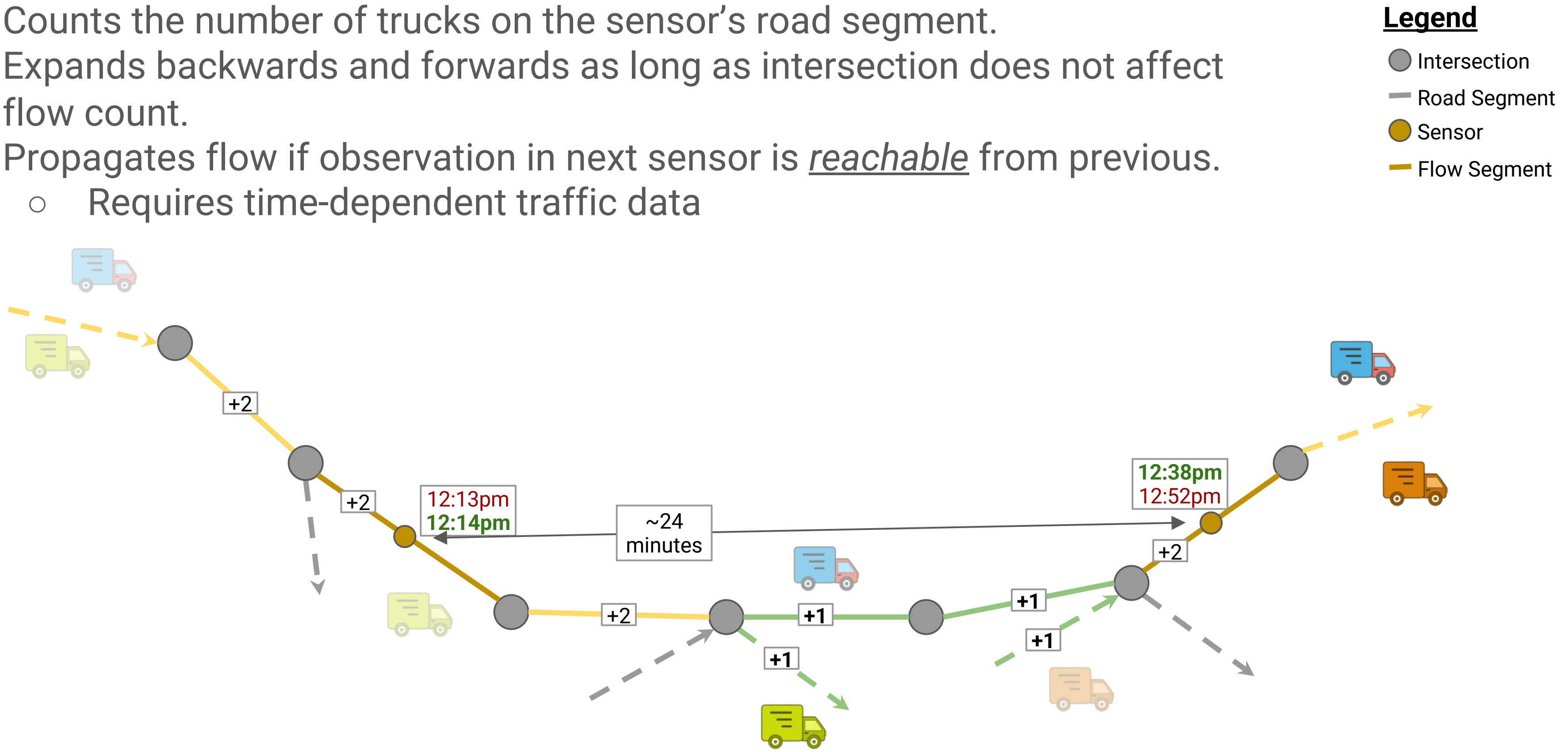
- Counts the number of trucks on the sensor's road segment.
- Expands backwards and forwards as long as intersection does not affect flow count.
- Propagates flow if observation in next sensor is *reachable* from previous.
    - Requires time-dependent traffic data

Legend
- Intersection
- Road Segment
- Sensor
- Flow Segment



Reachability-based Estimation

- Counts the number of trucks on the sensor's road segment.
- Expands backwards and forwards as long as intersection does not affect flow count.
- Propagates flow if observation in next sensor is *reachable* from previous.
    - Requires time-dependent traffic data

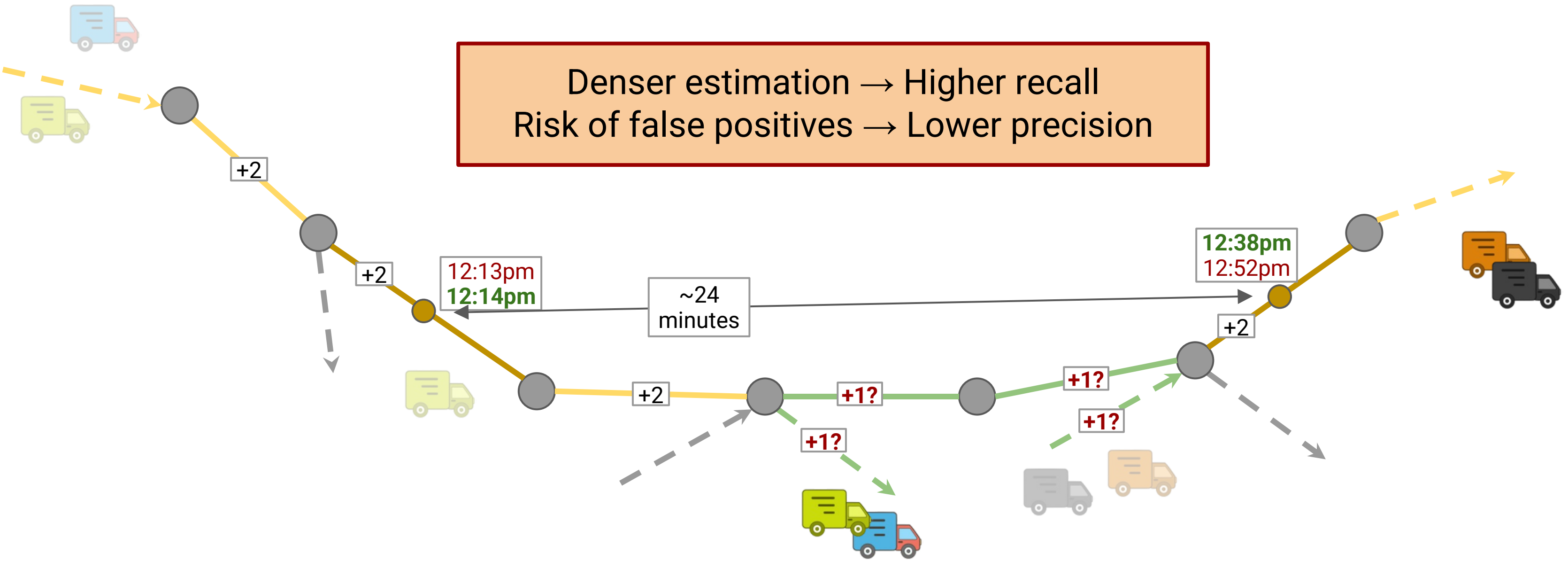


Reachability-based Estimation

- Counts the number of trucks on the sensor's road segment.
- Expands backwards and forwards as long as intersection does not affect flow count.
- Propagates flow if observation in next sensor is ***reachable*** from previous.
  - Requires time-dependent traffic data

Legend
- Intersection
- Road Segment
- Sensor
- Flow Segment

Denser estimation → Higher recall
Risk of false positives → Lower precision



Experimental Setup

- Datasets
  - SYNTH(S, T): Synthetic datasets with S sensors and T trucks
    - S = {100, 150, 200, 250, 300}
    - T = {250, 500, 750, 1000, 5000}
    - *“Simulates” truck trajectories and generate sensor observations*

- Algorithms
  - Baseline: Only estimates at edges where data is sensed.
  - FlowPath: Extrapolates flow based on logic.
  - Reachability-based

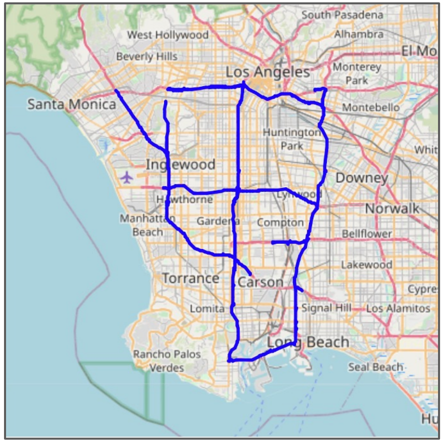
- Metrics
  - Precision: *Percentage of graph edges in estimation that exist in ground truth*
  - Recall: *Percentage of graph edges in ground truth that are in estimation*
  - MAE: Mean Absolute Error of flow estimation
  - MAPE: Mean Absolute Percentage Error

$$Precision = \frac{|GroundTruth \cap Estimation|}{|Estimation|} \qquad Recall = \frac{|GroundTruth \cap Estimation|}{|GroundTruth|}$$

Experimental Results

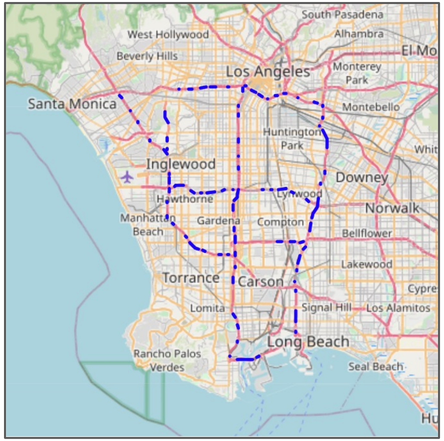
- **300 sensors, 1000 trucks**

Ground Truth



Edges: 888

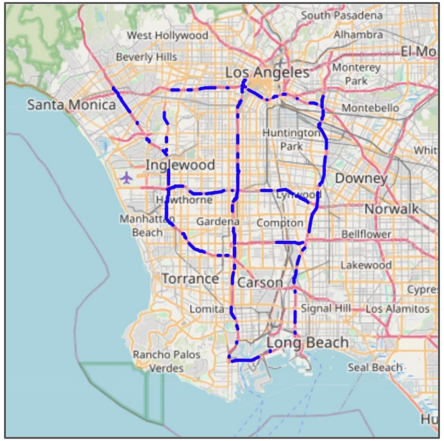
Baseline



Edges: 94 (TP/94, FP/0)
Precision: 100%
Recall: 10%

MAE: 7.464
MSE: 170.651
RMSE: 13.063
MAPE: 89.22%

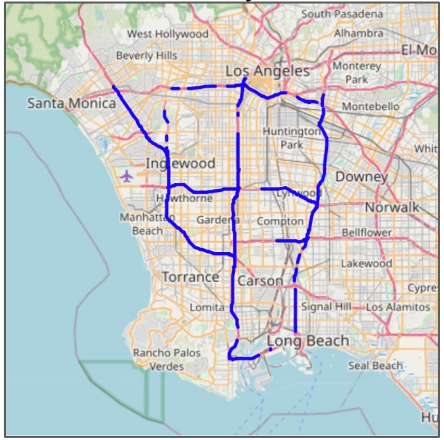
Naive Flow Path



Edges: 334 (TP/327, FP/7)
Precision: 98%
Recall: 37%

MAE: 4.858
MSE: 104.845
RMSE: 10.239
MAPE: 63.27%

Reachability-based



Edges: 702 (TP/666, FP/36)
Precision: 95%
Recall: 75%

7x

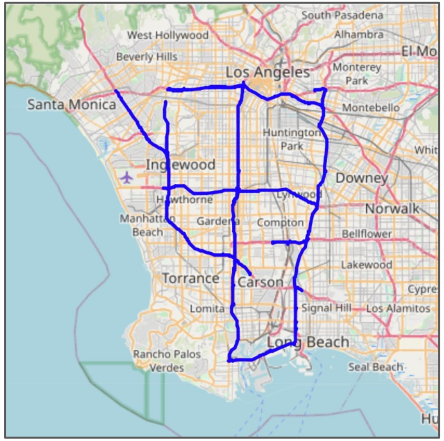
MAE: 3.054
MSE: 72.999
RMSE: 8.544
MAPE: 40.08%

2x

Experimental Results

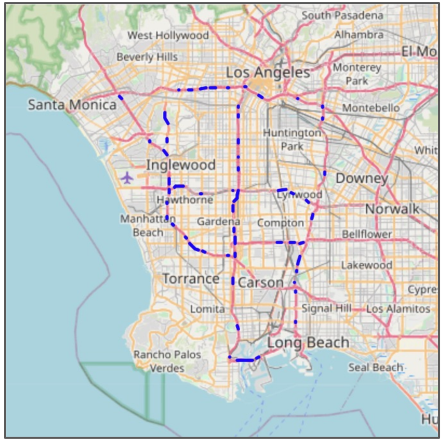
- **200 sensors, 1000 trucks**

Ground Truth



Edges: 888

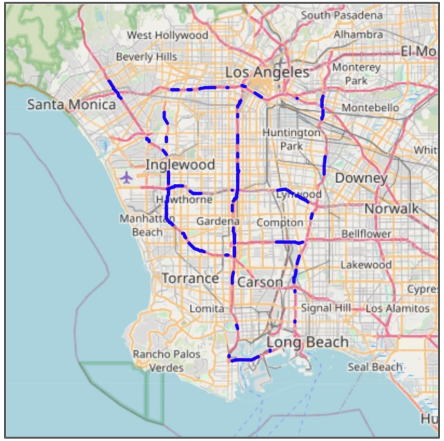
Baseline



Edges: 63 (TP/63, FP/0)
Precision: 100%
Recall: 7%

MAE: 7.807
MSE: 177.910
RMSE: 13.338
MAPE: 92.79%

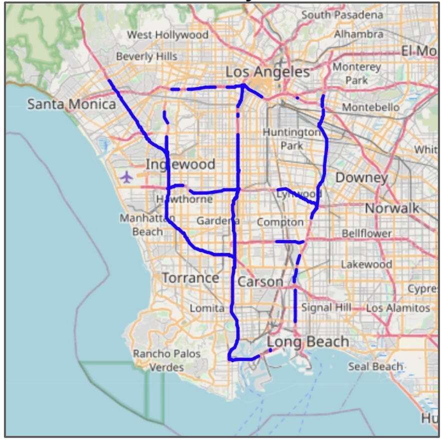
Naive Flow Path



Edges: 226 (TP/219, FP/7)
Precision: 97%
Recall: 25%

MAE: 6.065
MSE: 130.300
RMSE: 11.415
MAPE: 75.28%

Reachability-based



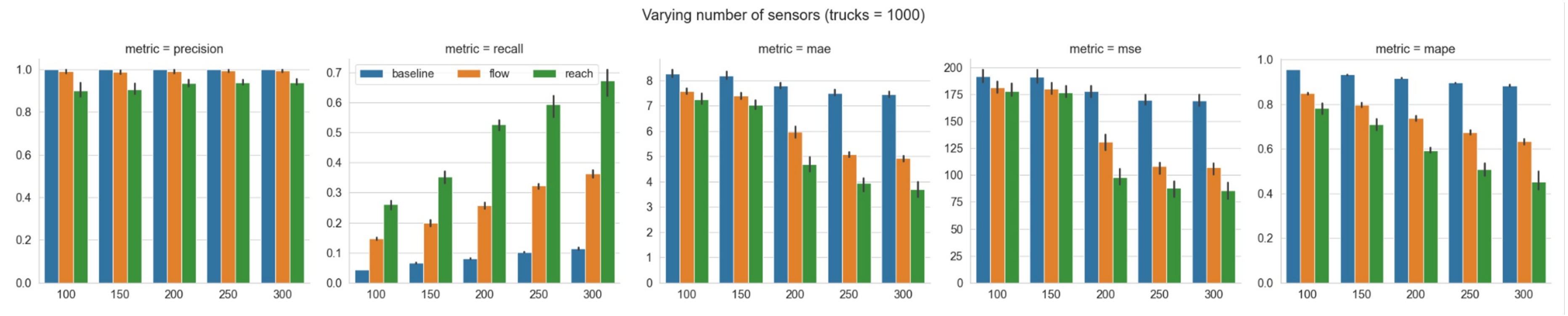
Edges: 601 (TP/564, FP/37)
Precision: 94%
Recall: 63%

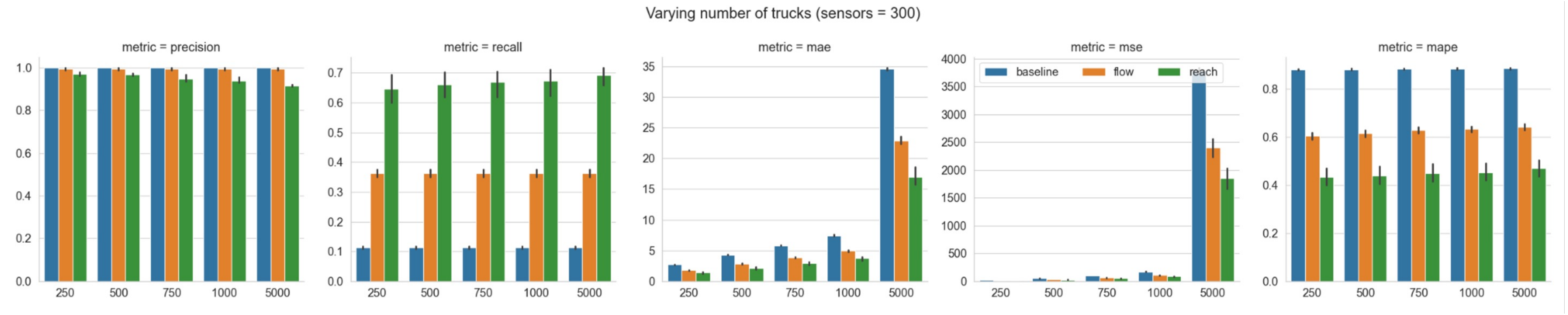
9x

MAE: 4.000
MSE: 90.450
RMSE: 9.510
MAPE: 51.20%

1.5x

Experimental Results





Summary & Future Work

- Critical for planners and decision makers to understand freight flow
- Estimating the volume of trucks from sensor data is feasible
- Reachability-based approach yields more accurate results
    - **9x** higher precision compared to the baseline
    - **2x** improvement in MAE

Future work

- Improve computational efficiency and accuracy of algorithm
- Validate approach with real-world data
- Infrastructure optimization
    - where should the next sensor be installed in order to improve accuracy?