Methods for Multimodal Traffic State Estimation to Identify the Optimal Sensor Distribution in Urban Road Networks

Master‘s Thesis of Victoria Dahmen

Mentoring:
Dr.sc. ETH Allister Loder
M.Sc. Gabriel Tilg

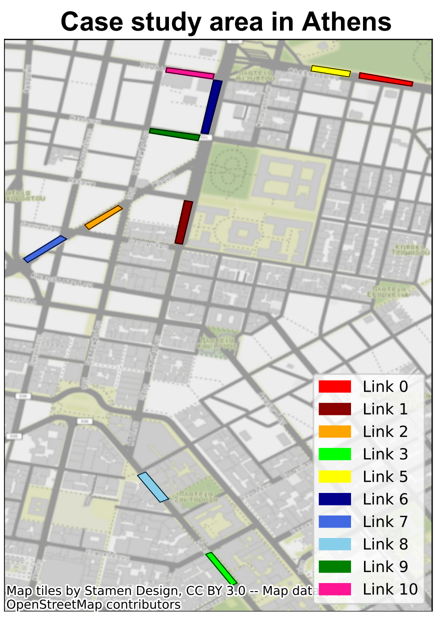

Figure 1 Selected links

Sensor Distribution Scenarios	LDD	Car 5%	Taxi	Bus
0	■			
1	■	■	■	
2		■	■	■
3	■	■		■
4	■		■	■
5		■	■	
6			■	■
7	■		■	
8	■			■
9		■		■
10	■	■		

Figure 2 Considered scenarios. Included sources are dark grey.

Traffic on urban roads is complex due to factors like the multiple modes of transport, traffic signals, and varying road characteristics. Knowledge about the current traffic state can improve traffic control measures, passenger information and fleet management. While sensor data from vehicles and loop detectors provide information that can be used to infer the overall traffic state in real-time, it is not known what distribution of such sensors is best. Additionally, few methods exist for the multimodal urban context, with a particular lack of methods that estimate the mode-specific traffic conditions. Here novel methods for urban multimodal traffic state estimation (TSE) are proposed based on selected links from the pNEUMA dataset (see **Fig. 1**), two of which are physics-informed models. These methods can also be used for the identification of the most suitable sensor combination, which may depend on the TSE method used. Four sensor sources are considered: loop detector data (LDD), and trajectory data from taxis, busses and cars (at a 5% penetration rate), and combined as shown in **Fig. 2**.

The proposed models include are a multiple linear regression (MLR) and a neural network (NN). They both consider a combination of sensor sources in the input and several link-specific features like the number of bus stops. The former entails the flow and density of each mode that is in the corresponding sensor distribution scenario. They estimate the traffic flow and density of all modes that are not already known. The MLR does this individually, one variable at a time, while the NN makes all estimations jointly. · Furthermore, two variants of the above TSE methods were proposed. They only estimate the overall traffic state at a link-level, rather than the conditions of each mode. A physics-informed loss constraint is designed that is based on the fundamental diagram (FD) of traffic flow theory. The aim is to reduce the number of unrealistic flow-density combinations and extremely high speed values among the estimations, by encouraging a state below the curve (**Fig. 5**). The loss takes the form of $L = (1-\gamma) * L_{MSE} + \gamma * L_{PHY}$. The factor γ is used to scale the physics-informed loss L_{PHY} in relation to the regular loss L_{MSE}.

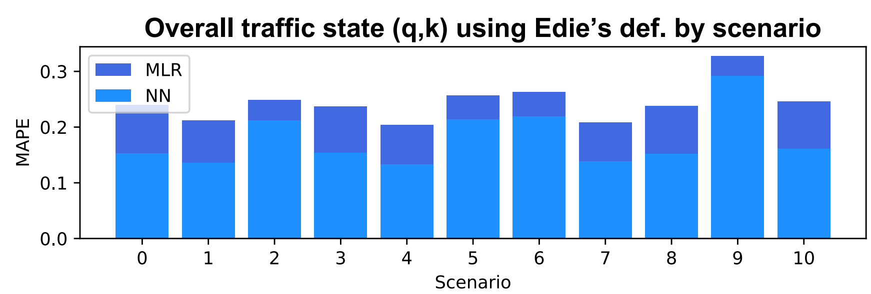

Figure 3 Average MAPE of the flow and density for each scenario

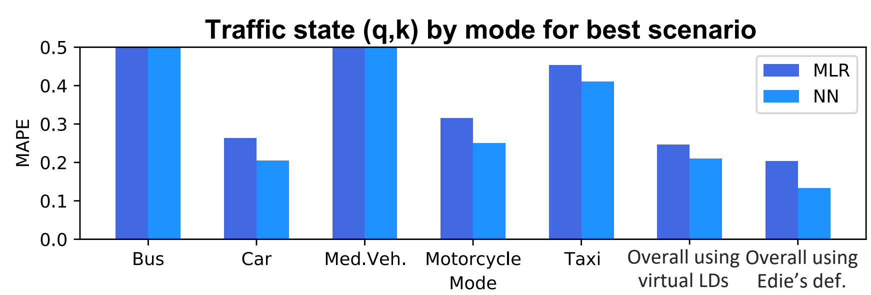

Figure 4 Best average MAPE of the flow and density for each mode

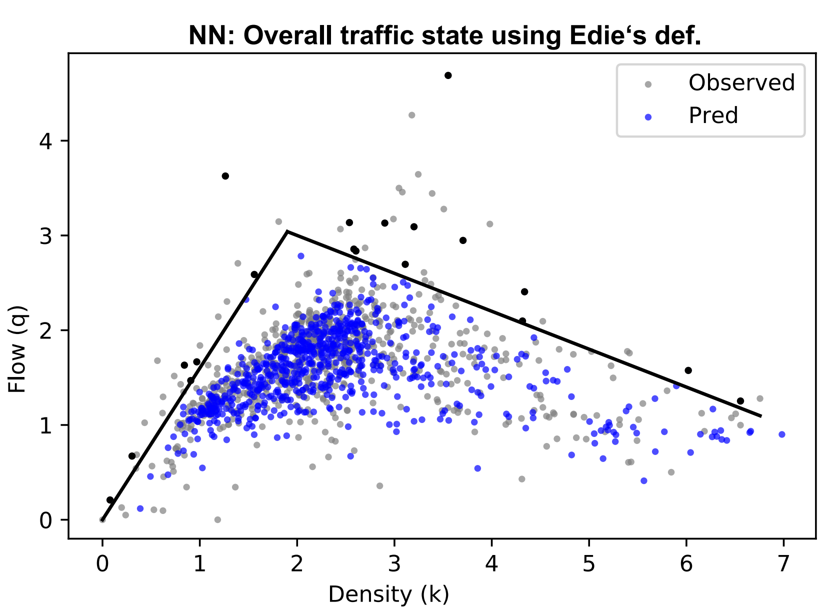

Figure 5 Estimation results of the NN and the curve used for L_{PHY}

For MLR and NN, comparisons were made for the method, output, scenario, link, bus stops, road rank and the derived speed. There are three key findings: 1) the NN performs better than the MLR in most cases. 2) The estimation results of the overall state, cars, motorcycles, and taxis are reasonable, while those of busses and medium vehicles are poor (**Fig. 4**). 3) Generally, the scenarios with LDD perform better (**Fig. 3**), yet the LDD alone (Scenario 0) is poor for mode-specific estimations. Thus, the mobile sensor data provides valuable information to the models. · For the physics-informed methods, the RMSE of the speed estimates for the overall traffic state (derived from the flow and density estimations) improved substantially, by up to 15%. The number of points far above the FD curve, associated with unrealistic behaviour, was reduced too. · The proposed methods were found to have satisfactory estimation results for most modes of transport, compared to current literature. On the whole, the sensor combinations with taxi and LD data led to the best accuracies.