Gaussian Mixture Model and Extended Kalman Filter for traffic data imputation: impact on perimeter control

NIRVANA PECORARI

Department of Transport and Planning
Faculty of Civil Engineering and Geosciences
Technical University of Delft, Delft, The Netherlands

ABSTRACT

In the domain of traffic engineering and transportation planning, accurate data play a pivotal role in understanding, analysing, and optimizing the performance of transportation systems. Traffic data, encompassing information such as vehicle counts, speeds, and travel times, serve as the foundation for informed decision-making aimed at improving mobility, safety, and efficiency.

However, the collection of real-world traffic data is often subject to various challenges, such as sensor malfunctions. As a result, incomplete or missing data entries are commonplace, leading to gaps in the temporal and spatial coverage of the datasets. These gaps can limit the accuracy of the insights derived from such data. Furthermore, missing or corrupted data might have a considerably negative impact on the performance of algorithms relying on them for the sake of traffic operations, e.g. in the instance of information provision, traffic management, …

Traffic data imputation emerges as a valuable technique to address the deficiencies caused by missing or incomplete data. *Imputation* refers to the process of estimating or predicting the missing values within a dataset based on the available information and patterns present in the observed data.

Most of the traffic data imputation methods can be categorised into three types: prediction methods, interpolation methods and statistical learning methods. To assess their performance, the literature compares these methods from different aspects, including reconstruction errors, statistical behaviours and running times.

Perimeter control strategies based on the Macroscopic Fundamental Diagram (MFD) aim to manage traffic congestion and improve overall traffic flow by regulating the flow of vehicles entering or leaving a specific area, such as a city centre. During the last twenty years, the MFD has gained significant attention and recognition. It builds upon the Fundamental Diagram (FD), which is a well-established concept in traffic flow theory. The FD describes the relationship between traffic flow, traffic density, and speed for a *single* road segment. The extension of this concept to a network-wide level, where interactions between *multiple* road segments are considered, led to the development of the *Macroscopic* Fundamental Diagram, which typically shows that the network production can be maximized within a certain range of traffic accumulation.

In case of missing data, imputation methods can attempt to indirectly reconstruct the network accumulation and production. To the best of our knowledge, no study in the literature investigates the impact of different imputation methods on the MFD reconstruction, when used for perimeter control purposes.

To address the research gap, we perform simulation-based experimental analysis through a two-step methodology. In the first step, we evaluate microscopic traffic scenarios in SUMO bearing different patterns and share of corrupted inductive loop detectors (missing flow and speed data). We then deploy Gaussian Mixture Model (GMM) and Extended Kalman Filter (EKF), respectively, as imputation methods to reconstruct the flow and speed data in real time. Considering simulated measurements as ground truth reference, we assess the goodness of the imputation through well-known indexes, like Root Mean Square Error (RMSE). In the second step, we base the perimeter control action on imputed data, through the reconstructed regional accumulation and assess the performance of perimeter control under differently reconstructed data through traffic performance indexes.

Key words: traffic data imputation, Gaussian Mixture Model, Extended Kalman Filter, perimeter control, Macroscopic Fundamental Diagram.

N. Pecorari (1997) is currently a second year PhD candidate at the Department of Transport & Planning of Delft University of Technology. She works under the supervision of Dr. Ir. M. Rinaldi, Prof.dr. L. Leclercq and Prof.dr.ir. S.P. Hoogendoorn. She is affiliated with AIM Lab. The main research project aims to develop, study and evaluate road network supply management approaches based on a real-time contouring of geographical areas which should be 'protected' from traffic, e.g. in response to heightened pollutant emissions, traffic congestion or incidents, vulnerable infrastructure. N. Pecorari holds a MSc in Physics from the University of Ferrara (Italy).