Predicting spatio-temporal traffic patterns with deep learning

Author: Felix Kunde, M.Sc.

Affiliation: Beuth University of Applied Sciences, Luxemburger Straße 10, 13553 Berlin

Contact: fkunde@beuth-hochschule.de, 030 45043824

Intelligent transportation systems (ITS) all around the world are collecting and processing huge amounts of data from numerous sensors to generate a ground truth of urban traffic. Detecting macroscopic traffic parameters such as travel time (time needed per trajectory) or traffic density (number of cars per trajectory) is crucial for managing and monitoring the system, as must adapt to different traffic scenarios and provide guidance to drivers to reduce traffic congestion and road collisions. Plus, it produces input data for traffic simulation programs which are used for traffic planning.

Apart from static sensor data, probe data from driving cars (Floating Car Data - FCD) is a very valuable resource because it produces trajectory-based data with a greater yet more realistic accuracy. Some cities have contracts with companies that own a great fleet of vehicles to deliver probe data, e.g. public transport or taxi cab companies. However, such datasets are mostly biased as the driving behavior can be bound to certain tasks e.g. busses have a fixed route and schedule what might cause waiting times or intentional delaying while driving.

From the very precise traffic information of the routing engines from Google Maps[1] or Here Maps[2] we can see the benefit of private transport data which is produced from the GNSS units of cars or smartphones. Usually, such data is not available to an ITS of the public sector. A city might fear an investment in such a (potentially huge) data set because of the hardware and software requirements it takes to process and store it. Depending on the level of detail of the recorded tracks, guarantees would have to be made that user privacy is treated carefully.

Availability of configurable and scalable traffic prediction algorithms

Many ITSs only take the detected data to monitor the current state of the traffic to react to traffic congestion e.g. by switching signs or blocking roads. The time series analysis methods used on the historic data sets are mostly very simple, e.g. moving average or exponential smoothing. Traffic predictions for future temporal horizons longer than 15 minutes are neither applied nor trusted although there is plenty of sample data available to run against modern algorithms.

On the contrary short-term traffic forecasting based on sensor data has already seen many different approaches in research within the last decades, be it for freeways or arterial road networks, with univariate or multivariate inputs and for different temporal lags. Unfortunately, the developed methods are rarely adopted in reality. On the one hand, it is difficult to decide which algorithm are most suitable for the ITS because of the heterogeneous setup presented in papers. Usually, one method is engineered exhaustively and compared to only simple variants of other algorithm paradigms (base lines). On the other hand, the proposed solutions are often tested on oversimplified examples, e.g. by only taking a small subset of sensors in a certain area of a city or even only a freeway. The question of scalability is often neglected.

Analyzing sensor data with machine learning algorithms

[1] https://www.google.de/maps

[2] https://wego.here.com/

A current trend – some might call it a hype – in IT including the field of GIS is artificial intelligence (AI) and machine learning (ML). It is often used as the answer to Big Data which is true in terms of that ML algorithms require a great amount of sample data to generalize on a given learning task. For example, car companies are investing into sensors that scan the environment of a car in order to extract the driving behavior of surrounding cars, to read street signs or to detect the surface condition of the road. Self-driving cars can be trained with such input data, but how well would their AI work when driving in a street network with totally different conditions than in the training phase.

As said in the beginning, many ITS have generated lots of data over years – usually in the form of time series on different aspects – from simple measurements at single sensors to trajectory-based aggregations. These time series represent a very good resource for machine learning to filter and predict spatio-temporal traffic patterns. One of the possible methods to apply could be artificial neural networks (ANN), which got a lot attention recently under the buzz word "Deep Learning". The basic idea of ANNs is not new and many researchers have already adopted them for traffic flow forecasting. But, with grown CPU and GPU power plus newly available deep learning frameworks like Google's Tensor Flow, Facebook's Torch or SkyMind's DeepLearning4J there is a great potential for powerful additions in terms of usability and scalability. It is easier than ever to train an ANN with numerous input and target data setups and to optimize its properties – so-called hyperparameters. On the other hand, the possible number of different combinations can make it hard to find a solution that provides a solid prediction for most scenarios.

Goals of the thesis

For the PhD thesis it is planned to apply and compare different traffic prediction algorithms in depth. The current work is focused on one parametric approach with the spatial-temporal autoregression integrated moving average (STARIMA) and two nonparametric methods with Support Vector Machine (SVM) and recurrent neural network (RNN). One main research question is, if there is one method which outperforms the others in prediction accuracy, computability and generalization or if an adaptive combination of different methods is also feasible. The thesis would have to answer as well how to install a continuous learning in case of that constraints within the street network are changing, e.g. roads block because of construction. Another possible side-effect could occur when a routing engine makes recommendations based on the predictions. If enough drivers are following the instructions this might produce patterns that had not exist before which would require a new training for the prediction. For traffic planners is would be very interesting to extract the findings of a ML algorithm, to gain new insights about spatio-temporal correlations within the sensor data. The thesis will investigate into ways to classify and visualize different patterns.

A first implementation with feed forward neural networks (FFNN) has already been implemented using Tensor Flow. Here it was the task for the network to predict a future occupancy or speed value at a certain sensor location only by training it with temporal offsets of 5, 10, 15, 30 to 45 minutes including data from neighboring sensors filtered by a spatial weight matrix. This first (rather simple) test had already promising results which will be presented at conferences in march 2017.