Towards Data-Efficient Mobility Analytics in Spatial Networks
on Tuesday 9th of November 2021, 17:00.

Abstract:
Vast amounts of vehicle trajectory data are being collected which has solutions to traditional transportation tasks at a finer granularity than ever before, but also enabled solutions to entirely new tasks. However, although vehicle trajectory data presents great opportunity, utilizing such data presents significant challenges.

First, trajectory data is sparse in contextual information. One way to address this issue is to map-match trajectories s.t. they are enriched with map data that provides road segment attribute information and structural information about the road network. Unfortunately, map data contains little attribute information, and the structure of the road network is difficult to leverage. Second, vehicle trajectory data is skewed s.t. it concentrates on some road segments at some times of day. This leads to the curious situation where data abundance and data sparsity co-occur in the same road network. Techniques that utilize vehicle trajectory data should therefore be sufficiently flexible to handle both situations. Finally, large sets of vehicle trajectories present challenges in terms of storage and transmission costs. It is therefore beneficial to be capable of processing data in compressed or easy-to-compress formats.

Efficient data utilization is the key to solve transportation tasks using large sets of vehicle trajectory data. In particular, data efficient techniques should utilize the available data as much as possible and perform well under both conditions of data sparsity and data abundance. Finally, they should also be capable of operating on compressed or easy-to-compress trajectory data formats, and, as a result, the thesis focuses on the use of map-matched vehicle trajectory data, i.e., trajectories represented as road segments in a road network, which can be compressed efficiently. This thesis focuses on data efficient techniques for solving transportation tasks using large sets of vehicle trajectory data.

First, the thesis investigates the application of network representation learning techniques to road networks. These techniques extract information from the structure of the road network, without relying on any attribute information, and can therefore improve the utilization of map data for contextual information when using map-matched vehicle trajectories. However, existing network representation learning techniques make assumptions that are inappropriate for road networks. This thesis therefore presents a representation learning technique designed for road networks that outperforms existing techniques on two transportation tasks.

Second, the thesis investigates how two categories of approaches to the important task travel time and speed estimation can be combined to leverage the strengths of both. The first approach relies on function-fitting and performs well under conditions of data sparsity. The other approach relies on aggregation of historical data and performs well under conditions of data abundance. However, since data abundance and data sparsity co-occur in road networks, no technique is universally applicable to all areas of the road network. This thesis therefore proposes a travel time and speed estimation framework that uses Bayesian probability theory to determine when to apply each approach. Rather than a hard decision
rule, the framework smoothly transitions from a function-fitting approach to an aggregation-based approach as the available data increases.

Third, the thesis proposes techniques for analysing driver behavior based on vehicle trajectories that have been converted into an easy-to-compress format. Specifically, the techniques can discover intermediate destinations within a vehicle trajectory and recover the preferences a driver exhibits in a trajectory w.r.t., e.g., travel time and fuel consumption. The techniques are not only able to operate on easily compressed vehicle trajectory data, but also has low processing costs with trivial parallelizability s.t. even very large trajectory sets can be processed efficiently.

Members of the assessment committee are Associate Professor Alvaro Torralba (Chairman), Aalborg University, Denmark, Professor Cyrus Shahabi, University of Southern California, USA, and Associate Professor Eleni I. Vlahogianni, National Technical University of Athens, Greece. Professor Christian S. Jensen and Professor MSO Thomas Dyhre Nielsen are Tobias’ supervisors. Moderator Associate Professor Manfred Jaeger.

All interested parties are welcome.