Advanced Traffic Modelling for Smart Cities Using HPC Platform

Radim Cmar, Martin Šurkovský, Jan Martinovič, Kateřina Slaninová

Precise traffic model and prediction helps municipalities to understand their everyday every minute traffic conditions, which is the result of its actual infrastructure capacity and mobility demand patterns of its citizens. Only with this knowledge the municipality can take proper measures to improve its traffic conditions, such as mitigation of traffic jams. Our use case is about a precise and fast calculation of traffic model and prediction, both short-term, say for a subsequent hour, and long-term, for any given time point in future.

Our use case is represented with *traffic ecosystem* comprised of a mixture of processing and storage components, which define a complex multi-thread processing executed on the HPC backbone with an option to accommodate critical calculations on accelerators. The system deals with a big data input, namely (a) floating car data (FCD), collected from mobile devices running Sygic navigation, which define vehicle speeds on GPS positions scattered across the road network, and (b) origin destination matrix data (ODM), which defines mobility flow on a city grid. We boost the collected FCD data with our in-house *traffic simulator*, which fills the void data of insufficiently covered areas through the combined use of FCD and ODM data.

Using computationally intensive algorithms we calculate the traffic model, which is represented by (a) macroscopic parameters for each road segment: *speed*, *flow*, and *intensity*, for each 15 minutes interval over weekday and (b) coefficients of prediction model for each road segment. On top of these models, we build the services such as traffic prediction and intelligent routing. The *traffic ecosystem* regularly updates its models with a daily ingestion and processing of FCD and ODM data. Working with big data and executing complex algorithms poses challenges both on storage and processing. Data sets need to be transferred across components, efficiently stored, and readily available for processing. Computations need to be executed fast and economically so that it is cost efficient and meets a daily processing cycle. Moreover, the system needs to be permanently responsive to serve prediction and routing requests.

The HPC computation is advantageous in several processing steps. First, the collected FCD data and O/D matrix are preprocessed and used to set-up the traffic simulator. *Hidden Markov model (HMM)* is used for map matching of sparse and noisy FCD points on a road network to recover realistic driving trajectories. Second, the deterministic traffic simulator, running on Dask/distributed system, is used to boost the FCD data. The simulator performs two tasks: 1) it fills the void data, and 2) simulates various scenarios to cover unseen situations. The simulator makes use of demanding *probability time-dependent routing (PTDR)* algorithm to infer correct arrival times. Finally, having augmented data a *convolutional neural network (CNN*) model training for speed predictions for each road segment is performed.

The use case is being solved within EVEREST (dEsign enVironmEnt foR Extreme-Scale big data analyTics on heterogeneous platforms) project. The project aims at developing a holistic design environment that simplifies the programmability of High Performance Big Data analytics for heterogeneous, distributed, scalable and secure systems.

Mr. Radim Cmar (rcmar@sygic.com) is the solution architect at Sygic, Slovakia. His main expertise is in system modelling and architecture design for complex IT systems such as mobile applications, server computing systems, and client to server communication systems. He graduated at Slovak Technical University of Bratislava in 1993. From 1997 to 2001 he was the research engineer at IMEC, Leuven, with the focus on HW/SW co-design methodologies for ASIC design. From 2001 to 2005 he was the system engineer at RFMD, San Jose with the focus on the system design of the system-on-chip solutions in the field of wireless communication. In 2007 he joined Sygic and since then he has been responsible for product definitions of various navigation components, and project lead for developing products with many business partners.

Dr. Martin Šurkovský (martin.surkovsky@vsb.cz) is a researcher at IT4Innovations National Supercomputing Centre, VSB - Technical University of Ostrava. He has got a doctoral degree in computer science in 2020 at VSB - Technical University of Ostrava, Czech Republic. Martin is an active developer. His research interest includes analysis of communication in distributed systems, traffic analysis, analysis of tunnels and channels in protein structures. Hi is responsible for deterministic traffic simulator.

Dr. Jan Martinovič (jan.martinovic@vsb.cz) is currently the Head of the Advanced Data Analysis and Simulations Lab at IT4Innovations National Supercomputing Center. He has extensive experience leading substantial R&D activities. His research activities are focused on information retrieval, data processing and analysis, experimental orchestration platforms for HPC/Cloud/BigData, and traffic management. His activities also cover the development of HPC-as-a-Service solutions, which enables remote use of HPC infrastructure.

He is the coordinator of the ICT-11 project LEXIS (https://lexis-project.eu). He has previous experience with the coordination of different contracted research activities with international and national companies. Currently, he is participating as innovation manager in the ICT-51 project EVEREST (https://everest-h2020.eu), as a work package leader in the EuroHPC project LIGATE (https://www.ligateproject.eu), and as co-design manager in the EuroHPC project ACROSS (https://www.acrossproject.eu). He is also a leading researcher at IT4I for the EuroHPC project IO-SEA.

Previously, he was a leading researcher at IT4I of the two H2020-FETHPC-2014 projects ANTAREX (AutoTuning and Adaptivity appRoach for Energy efficient eXascale HPC systems, http://specs.fe.up.pt/antarex) and ExCAPE (Exascale Compound Activity Prediction). He is also responsible for the research and development team of the FLOREON+ system (http://floreon.eu) - system for disaster management support and viaRODOS system (http://viarodos.cz/) for monitoring of the Czechia highway network. He has published more than 100 papers in international journals and conferences and is HiPEAC member since January 2020.

Dr. Kateřina Slaninová (katerina.slaninova@vsb.cz) is Deputy head of Advanced Data Analysis and Simulations Lab at IT4Innovations National Supercomputing Center, Czech Republic. She has got a doctoral degree in Informatics from VSB – Technical University of Ostrava, Czech Republic. Research interests include high-performance data analysis (HPDA), machine learning, artificial intelligence, dynamic network and complex network analysis in domains like traffic management, business processes, etc. Her recent activities also cover cooperation with SMEs in areas such as traffic modelling and optimisation, image recognition, stock optimisation, bioinformatics, etc. She is author of more than 80 scientific conference and journal papers.

She is responsible for the quality management and data management plan in H2020 project LEXIS (Large-scale Execution for Industry & Society). She is responsible on IT4Innovations side for management of research and development activities in H2020 project EVEREST (dEsign enVironmEnt foR Extreme-Scale big data analytics on heterogeneous platforms). She was a work package leader of H2020 FETHPC project ANTAREX (AutoTuning and Adaptivity appRoach for Energy efficient eXascale HPC systems). She also worked within the team of the Center for the Development of Transportation Systems RODOS, and she participates in long-term cooperation with the navigation company Sygic, Slovakia. She is a member of IEEE since 2009 and acted as an observer in Research and innovation advisory group of EuroHPC JU from 2019 to 2021.