Center for Teaching Old Models New Tricks (TOMNET)

A USDOT Tier 1 University Transportation Center

PROJECT PROPOSAL

Title: Teaching the travel demand flow estimation models: a new deep-learning approach using multi-source data

Principal Investigator: Xuesong Zhou, Associate Professor, School of Sustainable Engineering and the Built Environment, Arizona State University

Co-Principal Investigator: Vladimir Livshits, Director of Transportation Technologies and Services, Maricopa Association of Governments

---

1. Introduction

The amount of data that transportation systems collect and store today has reached a new level compared with the previously traditional collection methods. For example, as one of pioneering Metropolitan Planning Organizations (MPO) in the United States in transportation planning, Maricopa Association of Governments (MAG) can provide the regional household travel survey data of 60GB, 1-year TMC-based speed data of 26GB, and 1-year link-based speed data of 3.1TB. Facing those unstructured and structured data streaming in from heterogeneous sensor sources at an unprecedented rate, it is critical to quickly manage and mine useful information under control. Also, it should be always aware that the value of those big data is reflected not just by its high volume but also by what specific goals/problems the data are used for. With the development of new computing technologies, machine learning has currently evolved as a powerful tool to learn from data with independent adaptions to generate reliable, repeatable decisions and results in a variety of application areas. However, compared with the areas in automotive, financial services, healthcare and etc., the open-source learning materials, studies and applications of machine learning on transportation system planning and operations are still relatively weak and need to be enhanced and taught to students, researchers and practitioners in time.

Focusing on the Traffic Demand Flow Estimation (TDFE) problem, which infers the number of persons/vehicles travelling between a particular origin and destination via a particular route/link, this project aims to support the education and training of relative students, researchers and practitioners to understand and learn the knowledge of deep learning and its application procedure by developing a step-by-step tutorial and open-source software packages and providing a number of well-organized workshops.

Specifically, deep learning technologies have been applied in a number of studies mainly focusing on traffic flow prediction (Dougherty, 1995; Park et al., 1998; Dia, 2001; Yin et al., 2002; Vlahogianni et al., 2005; Zhong et al., 2005; Zheng et al., 2006; Chan et al., 2012; Kumar et al. 2013). Recently, Lv et al. (2015) demonstrated one of the early applications of deep learning

networks to uncover and identify hidden patterns from the observed traffic measurements as a time series from multiple days. It has been well recognized that the simple application of the artificial neural network (ANN) software package is insufficient in explaining the behavioral relationship between complex traffic states and underlying travelling choice parameters. In this project, our major focus is to develop a theoretically explainable and interpretable deep learning approach to estimate different layers of demand variables and behavior coefficients, and then teach and educate interested users for the learning and real-world applications.

2. Project Objectives

This project will develop a step-by-step tutorial and a self-study software package and organize several workshops for students, researchers and practitioners to learn deep-learning approaches and how to apply it to estimate and calibrate the travel demand flow and behavioral estimation models with available multi-source sensor data.

The project has the following specific goals:

- To provide interested users an opportunity to understand the complex modeling process using the deep-learning approach by constructing computation graphs and calculating the derivatives among different layers.
- To cultivate in users an interest of machine learning knowledge for other transportation problems rather than just travel demand flow estimation modeling problems.
- To make critical issues visible by visualizing traffic flow dynamics, traffic bottlenecks, traveler route choice behavior.
- To develop a learning platform and assess users' beliefs and attitudes towards machine learning materials and technology and further gather their feedbacks and advice
- To broadly disseminate machine learning education packages, findings, and products to diverse other application fields for further communication and improvements.

3. Proposed Methodology and Data

Through collecting all relative learning and study materials, a comprehensive learning guide will be prepared for users to understand the basic concepts and general procedure of the machine-learning technique. By developing an online learning platform, users are encouraged to finish a number of tasks and write teaching notes, because it is believed that learning by teaching is usually better learning by doing. Meanwhile, an open-source package will be developed for users to further learn and understand the coding process for their future applications. A number of workshops will be prepared to broadly spread the learning materials for specific users in transportation communities.

Taking the travel demand flow estimation model as our learning and study problem, the data that will be collected for our developed deep-learning approach include household travel surveys for trip generation, mobile phone sample data for trip destination choice (OD split ratio), floating car data and vehicle location/identification data for candidate routes generation for each OD pair and route choice, sensor data at fixed locations for traffic counts, and the physical traffic network with nodes, links, zones and centroids.

Focusing on the details in modeling and solutions, we will propose a modeling framework using a multi-layer Hierarchical Flow Networks (HFN) representation. This flow-oriented estimation formulation can structurally formulate the problem and simultaneously estimate different levels of unobservable or partially observed traffic demand variables and behavior parameters. By recognizing the multiple sources of information in emerging big data applications, we map different levels of traffic demand variables to various data sources in traffic demand estimation applications. The systematic linkage between each representation layer in HFN and individual sources enable planners to better conduct cross validation and data fusion. To build a theoretically sound modeling framework, we will trace back to the fundamental or low level representation for deep learning networks and construct a transportation-focused computational graph as a structured modeling language. To enable computational efficient solution methods, we use a dynamic programming principle to efficiently compute numerical gradients. By using a new set of tools (Tensor Flow), we can quickly represent modular structures and compute many compounded partial derivatives at different representation layers of the proposed HFN.

4. Work Plan

Task 1: Prepare learning materials on machine learning

We will collect a number of learning materials for the beginners to understand basic concepts and required mathematical knowledge. Then the studies on machine learning in transportation systems will be reviewed and summarized by different categories. Finally, we will focus on the travel demand flow estimation problem for users to apply the machine learning skill in practice.

Task 2: Collect data

We will collaborate with the staffs from MAG to prepare the required multi-source data, including household travel survey data, mobile phone sample data, floating car data, loop detector data, and network data. The data format and required data in each data set will be carefully designed based on the proposed deep-learning approach for the travel demand flow estimation problem. Finally, users can be familiar with the data sets and know how they are used in the proposed models.

Task 3: Develop the model and an open-source software package

Focusing on the travel demand flow estimation problem, a number of subtasks will be performed as follows.

- Construct a multi-layer Hierarchical Flow Networks (HFN) based on the physical network from MAG.
- Connect the linkage between each representation layer in HFN and each available data source from MAG.
- Apply the Tensor Flow library as a deep learning toolkit to find the computationally efficient solution under a dynamic programming principle
- Develop an open-source software package for users to learn how to solve the problem by deep-learning approach and further understand the logic of the proposed algorithm.
- Use an open-source visualization tool, NeXTA, to display the final traffic condition and travel demand.

Task 4: Prepare an online learning platform and host several workshops
Prepare an online material-sharing platform (e.g., Github, Researchgate) to share our training materials and collect the feedbacks from the public. Meanwhile, we will host several workshops for students, researchers and practitioners to teach them understand the concept and basic framework of machine learning methods so that they can finally apply those approaches to different real-world transportation problems in practice.

5. Project Schedule

The work plan consists of the following 4 tasks and will be completed in one year from Aug 2018 to Jul 2019.

Table 1 Proposed timeline

Project Week	***8-9***	***10-11***	***12-1***	***2-3***	***4-5***	***6-7***
Task 1: Prepare learning materials on machine learning						
Task 2: Collect data						
Task 3: Develop the model and an open-source software package						
Task 4: Prepare an online learning platform and host several workshops						

6. Relevance to the Center Theme/Mission

This project aims to teach the traditional four-step travel demand flow estimation model a new trick, which is a deep-learning approach capable of using the multi-source sensor data. This approach is also useful to incorporate heterogeneous data sources for better estimating and calibrating activity-travel behavior and mobility choice models.

7. Anticipated Outcomes and Deliverables

This project will provide comprehensive training materials on machine learning and its applications in transportation systems. A new travel demand flow estimation model will be developed, estimated and calibrated by this approach using multi-source sensor data. Also, the corresponding algorithm will be open-source for the learners. Several workshops will be hosted for students, researchers and practitioners in transportation communities.

Therefore, this project can have the following benefits.

- The learners in the transportation area would gain brief exposure to how machine learning can be used to solve transportation problems.
- The traditional travel demand flow estimation model will be reconstructed and solved by a new way considering the result consistency of using multi-source sensor information.
- The feedbacks from learners in online learning platform or workshops can be useful to further teach and apply the machine learning knowledge in our transportation domain.

8. Research Team and Management Plan

This research project will be performed by the collaboration of ASU team and MAG team. The ASU research team is leaded by Dr. Xuesong Zhou, who has a great reputation in dynamic traffic assignment and simulation and recently focuses on the deep-learning approaches to make full use of big data for transportation planning and operations. Also, a post-doctoral research associate, Dr. Jiangtao Liu, will help to develop the models and prepare the workshops. At the MAG team, Dr. Vladimir Livshits manages the transportation models of MAG for more than 10 years as the director of transportation technologies and services. In addition, Shuyao Hong from MAG will also contribute to the data preparation and model development as a Transportation Engineer II.

Our whole team will be managed by Dr. Zhou. The two teams will report the progress with each other every week. The ASU team will also visit MAG to present the modeling details for its calibration and comparison with the results of existing models.

9. Technology Transfer Plan

A training tutorial of deep learning applications in transportation systems will be prepared and shared to practitioners through an online learning website. The PIs will develop a webinar to introduce the basic knowledge and model framework. The coding for model development on the travel demand flow estimation problem will be open source and distributed on the project website. The PIs will host several workshops for students, researchers and practitioner to learn step by step. The research findings will also be presented at professional conferences.

FHWA Planning Method Team Lead Mr. Brian Gartner has also expressed his interests to further provide three TMIP presentation opportunities to reach more planners interested in learning the new integrated machine learning and demand flow estimation approach. Dr. Cynthia Chen from the University of Washington also expressed her interests to work with the team, FHWA team and planning organizations using other available technology transfer channels, such as the existing FHWA-NSF joint workshop on the use of big data and machine learning in transportation.

10. Workforce Development and Outreach Plan

This deep learning approach proposed in this approach will be used to improve transportation engineering education for both undergraduate and graduate programs at ASU. The step-by-step training materials and open-source software package also can be downloaded by other teachers for their classes and research. In addition, the basic part of our learning materials can be offered to K-12 students to teach them fundamentals of both machine learning and transportation engineering and how to use big data to solve transportation problems. With the help of MAG, The research team plans to reach out to professionals and decision makers for technology transfer of transportation planning models to the Federal and State transportation agencies.

11. References

- Chan, K.Y., Dillon, T.S., Singh, J., Chang, E., 2012. Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg–Marquardt algorithm. IEEE Transactions on Intelligent Transportation Systems 13(2), 644-654. doi: 10.1109/TITS.2011.2174051

- Dia, H., 2001. An object-oriented neural network approach to short-term traffic forecasting. European Journal of Operational Research 131(2), 253-261. Doi: 10.1016/S0377-2217(00)00125-9
- Dougherty, M., 1995. A review of neural networks applied to transport. Transportation Research Part C Emerging Technologies 3(4), 247-260. Doi: 10.1016/0968-090X(95)00009-8
- Kumar, K., Parida, M., Katiyar, V.K., 2013. Short term traffic flow prediction for a non urban highway usin,g artificial neural network. Procedia - Social and Behavioral Sciences 104, 755-764. Doi: 10.1016/j.sbspro.2013.11.170
- Kumar, K., Parida, M., Katiyar, V.K., 2013. Short term traffic flow prediction for a non urban highway usin,g artificial neural network. Procedia - Social and Behavioral Sciences 104, 755-764. Doi: 10.1016/j.sbspro.2013.11.170
- Park, B., Messer, C., Urbanik, T., 1998. Short-term freeway traffic volume forecasting using radial basis function neural network. Transportation Research Record: Journal of the Transportation Research Board 1651(1), 39-47. doi:10 10.3141/1651-06
- Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2005. Optimized and metaoptimized neural networks for short-term traffic flow prediction: A genetic approach. Transportation. Research. Part C 13(3), 211–234. doi:44 10.1016/j.trc.2005.04.007
- Yin, H., Wong, S.C., Xu, J., Wong, C.K., 2002. Urban traffic flow prediction using a fuzzy-neural, approach. Transportation Research Part C Emerging Technologies 10(2), 85-98. Doi:10.1016/S0968-090X(01)00004-3
- Zheng, W., Lee, D.H., Shi, Q., 2006. Short-term freeway traffic flow prediction: Bayesian combined neural network approach. Journal of Transportation Engineering 132(2), 114-121. Doi:10.1061/(ASCE)0733-23 947X(2006)132:2(114)
- Zhong, M., Sharma, S., Lingras, P., 2005. Short-term traffic prediction on different types of roads with genetically designed regression and time delay neural network models. Journal of Computing in Civil Engineering 19(1), 94-103.26 Doi: 10.1061/(ASCE)0887-3801(2005)19:1(94)

12.Qualifications of Investigators (One-page CV per Investigator)

13.Budget Including Non-Federal Matching Funds

Institution: Arizona State University

Project Title: Teaching the travel demand flow estimation models: a new deep-learning approach using multi-source data

Principal Investigator: Xuesong Zhou

Budget Period: 8/1/2018 - 07/31/2019

CATEGORY	Budgeted Amount from Federal Share	Budgeted Amount from Matching Funds	Explanatory Notes; Identify Source of Matching Funds
Faculty Salaries	2,340	7,500	Faculty regular salary as matching
Other Staff Salaries	24,000	10,000	MAG staff salary as matching
Student Salaries	739		
Fringe Benefits	5,280		
Total Salaries & Benefits	32,359		
Student Tuition Remission			
Operating Services and Supplies		7,500	MAG in-kind matching for sharing their multiple data sources
Domestic Travel			
Other Direct Costs (specify)			
Other Direct Costs (specify)			
Total Direct Costs			
F&A (Indirect) Costs	$17,636		
TOTAL COSTS	$49,995	25,000	

Grant Deliverables and Reporting Requirements for UTC Grants (November 2016)

Exhibit F

UTC Project Information	
Project Title	Teaching the travel demand flow estimation models: a new deep-learning approach using multi-source data
University	Arizona State University
Principal Investigator	Xuesong Zhou
PI Contact Information	Associate Professor, School of Sustainable Engineering and the built Environment, Arizona State University; xzhou74@asu.edu
Funding Source(s) and Amounts Provided (by each agency or organization)	$50,000 from TOMNET $25,000 from MAG
Total Project Cost	$75,000
Agency ID or Contract Number	
Start and End Dates	08/01/2018 and 07/31/2019
Brief Description of Research Project	Focusing on the Traffic Demand Flow Estimation (TDFE) problem, this project aims to support the education and training of students, researchers and practitioners to understand and learn deep learning and its application procedure.
Describe Implementation of Research Outcomes (or why not implemented) Place Any Photos Here	Develop a step-by-step tutorial and open-source software packages; Organize several workshops for training and education of machine learning.
Impacts/Benefits of Implementation (actual, not anticipated)	The learners in the transportation area would gain brief exposure to how machine learning can be used to solve transportation problems. The traditional travel demand flow estimation model will be reconstructed and solved by a new way considering the result consistency of using multi-source sensor information.
Web Links • Reports • Project Website	

XUESONG ZHOU
School of Sustainable Engineering and the Built Environment
Arizona State University, Tempe, AZ

A. APPOINTMENTS

2013 – Present: **Arizona State University, Tempe, Arizona**
Associate Professor, School of Sustainable Engineering and the Built Environment

B. PRODUCTS

1. Zhou*, X. and Taylor, J. (2014) DTALite: A queue-based mesoscopic traffic simulator for fast model evaluation and calibration. Cogent Engineering, 1(1)

2. Xing. T. Zhou*, X. (2011) Finding the Most Reliable Path With and Without Link Travel Time Correlation: A Lagrangian Substitution Based Approach, Transportation Research Part B. 45 (10) , 1660-1679.

3. Xing, T. Zhou*, X. Taylor J. (2013) Designing Heterogeneous Sensor Networks for Estimating and Predicting Path Travel Time Dynamics: An Information-Theoretic Modeling Approach. Transportation Research Part B. 57, 66-90

Software Development:

4. Principle Developer of DTALite, light-weight open-source traffic assignment/simulation engine; more than 3000 downloads from https://code.google.com/p/nexta/
5. Principle Developer of NEXTA, free visualization platform for DYNASMART, Dynus-T and TRANSIMS (4,000 downloads, used by about 100 planning and research organizations)

C. SYNERGISTIC ACTIVITIES

1. **Service for professional organizations**:
   Editor, *Transportation Research Part C: Emerging Technologies*;
   Editorial Board Member, *Transportation Research Part B: Methodological*;
   Co-Chair, IEEE Intelligent Transportation Systems Society, Technical Committee on Traffic and Travel Management;
   Subcommittee Chair and member, Transportation Research Board (TRB) Committee on Transportation Network Modeling (ADB30);
   Chair, Institute for Operations Research and the Management Sciences (INFORMS), Railway Applications Section (RAS)
2. **Promoting open transportation data hub and schema:** Help FHWA to establish open data set and schema through multi-year project titled "Effective Integration of Analysis, Modeling, and Simulation Tools"
3. **Serving as problem co-owners and student/researcher competition organizing committee member** for INFORMS Railway Application Section Problem Solving Competition, 2012, 2013, involving more than 60 participating teams each year, supported by CSX, BNSF Railway and Norfolk Southern
4. **Hosting workshops for FHWA and DOT/MPO practitioners:** Hosted three one-day FHWA-sponsored training workshops for more than 50 practitioners at Salt Lake City, Utah, Detroit, Michigan and Portland, Oregon, on the topic of dynamic traffic assignment system