Kick-off Presentation

Improving spatio-temporal traffic prediction through transfer learning

João Rodrigo Olenscki

Chair of Transportation Systems Engineering
Department of Mobility Systems Engineering
Technical University of Munich

September 21st, 2023

Outline

Traffic prediction

- Traffic prediction refers to the task of predicting future traffic conditions (such as congestion levels, travel times, vehicle flows) in a specific region/road grid over a certain period of time.It's a fundamental step for the development of smart cities (Oyewola, Dada, and Jibrin 2022).
- Typical approaches fail to support starting-up models for new cities as they require a considerable investment in a city-wide sensor network and an idle wait for the data to be gathered. This is also referenced in the literature as a "cold start" problem.
- To solve this problems, one can use transfer learning models.

Transfer Learning

- The main idea of such methods is to explore the patterns that cities may share between one another.
- Practically, this means that one would use a data-rich source city to pre-train the weights of the model and then initializing training for the data-poor target city with these pre-trained weights.
- With this approach, significantly accurate models can be generated with as few as two days worth of data.
- Similarly, one can also use this approach to perform intra-city learning, i.e. learning about the entirety of the city based on a sub region of it.

Challenges

- Overfitting: to generate a set of domain invariant features, one is required to avoid overfitting of the model to the source city, as the features need to be shared patterns between cities.
- Domain shift: the source and target domains need to be independently identically distributed (i.i.d.) in order for the transferability of the knowledge to make sense.
- Source selection: in order to find shared patterns between the domains (cities), one must be assured that they are relatable and maintain similarities. One can't use, for instance, the Manhattan island to model a small town in the Midwest.

Outline

Methodology

- For this bibliographic revision 26 papers were read and analysed.
- After the reading, for each paper, a summary report of the methods and architectures implemented was written and an information table was constructed.

Model name	
Year	
Strenghts	+ +
Weaknesses	- -
Based on	
Transfer Learning?	Yes/No

Figure 1 Information table model

Highlights

Model name	Year	Spatial structure (region or segment)	Intra- or inter-city?	Takes advantage of additional data?	Capable of learning from multiple sources?
RegionTrans	2019	Region	Inter	Yes	No
MetaST	2019	Region	Inter	No	Yes
ST-MGCN	2019	Region	Inter	Yes	No
ST-DAAN	2022	Region	Both	Yes	No
CrossTReS	2022	Region	Both	Yes	No
GGCN + MGCN	2022	Region	Inter	Yes	No
DASTNET	2022	Segment	Inter	No	Yes
CityTrans	2023	Segment	Inter	No	No

Table 1 Highlights from the revision

Findings

- During this revision, several noteworthy findings could be made, for instance:
  - We could observe the evolution of some lines of research through time. In 2019 (Geng et al. 2019) introduced the idea of using multi-modal graph convolutional network (ST-MGCN). Two years later, (L. Zhang et al. 2022) enhanced this approach by combining all individual modes into a group GCN.
  - It is also perceived that there are two ways of pre-processing the spatial information: create a grid and process the data like a grid (L. Wang, Geng, et al. 2019; S. Wang et al. 2022), or generate a graph network from the regions, assigning a node for each region and connecting the region to the adjancies by an adjancy matrix (Ouyang et al. 2023; Tang et al. 2022).
  - Finally, from the revision of state-of-the-art articles, we could find that there are models that exploit the existence of external data. For instance (L. Wang, Geng, et al. 2019) makes use of Google Map's check-in information to model vehicle flow in data scarce regions.

Opportunities

- It can be seen on Table 1, most models don't use multiple source cities to enhance generalization. One could combine existent models with multiple sources (or a source optimizer) in order to enhance it.
- Most models take advantage of external data (e.g. weather data, POI, holidays' calendar), but there are some openings to enhancing this field (Google Maps' data for instance).
- There are many ways of processing the spatial structure. The most common one is using a GNN, but one could also apply computer vision algorithms to pair knowledge transfer regions.
- For the data augmentation, one could try to shift and tilt the region grid of a city in order to generate new regions.

Objectives

- Build a model architecture with the following requisites:
  - based on a region spatial structure;
  - capable of transferring learning intra- and inter-cities;
  - capable of using additional external data; and
  - capable of learning from multiple sources.
- Additionally, a variation of the model should be able to receive a single source city and perform data augmentation on it in order to generate “multiple” sources.
- The designed model is to be tested against at least for different baselines: two classical statistical (HA, ARIMA), one deep learning model, and one state-of-the-art transfer learning model.

Outline

Methodology

- Figure 2 shows a preliminary flowchart for the model.

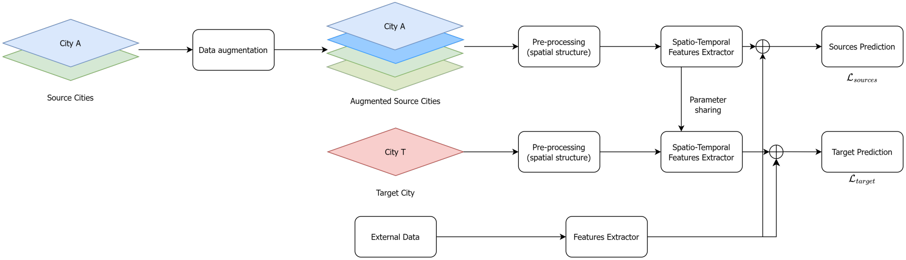


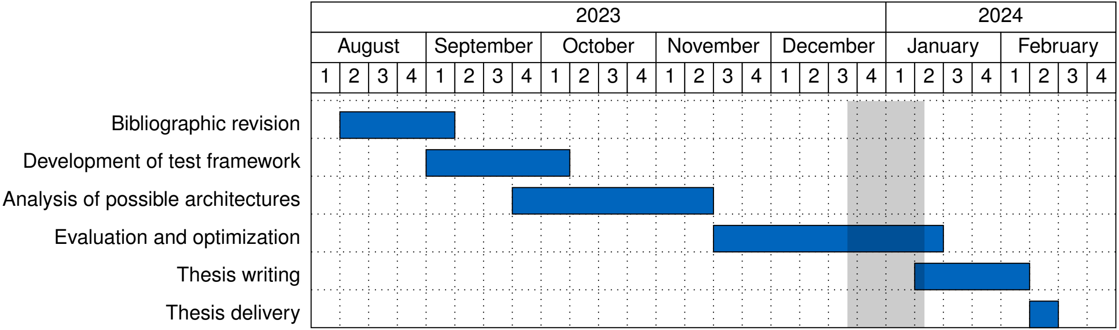
Figure 2 Methodology flowchart proposed

Outline

General information

- **Title**: Improving spatio-temporal traffic prediction through Transfer Learning
- **Description**: This research delves into state-of-art transfer learning models for the Spatio-Temporal Traffic Prediction (STTP) problem. These models employ a vast diversity of techniques and architectures to make predictions about the complex field of traffic forecasting. By studying all these models, we expect to press this field further and develop new state-of-the-art models ourselves.
- **Advisors**:
  - Cheng Lyu, M.Sc.
  - Prof. Constantinos Antoniou, Ph.D.
- Date of start: 07.08.2023
- Deadline: 07.02.2024

Thesis timeline



Thank you!

Thank you!

References I

- Oyewola, David Opeoluwa, Emmanuel Gbenga Dada, and Muhammed Besiru Jibrin (2022). "Smart City Traffic Patterns Prediction Using Machine Learning". In: *Machine Learning Techniques for Smart City Applications: Trends and Solutions*. Springer International Publishing, pp. 123–133.
- Geng, Xu et al. (July 2019). "Spatiotemporal Multi-Graph Convolution Network for Ride-Hailing Demand Forecasting". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 33.01, pp. 3656–3663.
- Zhang, Lingyu et al. (Sept. 2022). "Multi-Modal Graph Interaction for Multi-Graph Convolution Network in Urban Spatiotemporal Forecasting". In: *Sustainability* 14.19, p. 12397.
- Wang, Leye, Xu Geng, et al. (Aug. 2019). "Cross-City Transfer Learning for Deep Spatio-Temporal Prediction". In: *Proceedings of the Twenty-Eighth International Joint*

References II

Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization.

Wang, Senzhang et al. (2022). “Spatio-Temporal Knowledge Transfer for Urban Crowd Flow Prediction via Deep Attentive Adaptation Networks”. In: *IEEE Transactions on Intelligent Transportation Systems* 23.5, pp. 4695–4705.

Ouyang, Xiaocao et al. (2023). “CityTrans: Domain-Adversarial Training with Knowledge Transfer for Spatio-Temporal Prediction across Cities”. In: *IEEE Transactions on Knowledge and Data Engineering*, pp. 1–14.

Tang, Yihong et al. (2022). “Domain Adversarial Spatial-Temporal Network: A Transferable Framework for Short-term Traffic Forecasting across Cities”. In.

Zhang, Junbo, Yu Zheng, Dekang Qi, et al. (Oct. 2016). “DNN-based prediction model for spatio-temporal data”. In: *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM.

References III

- Yao, Huaxiu et al. (May 2019). “Learning from Multiple Cities: A Meta-Learning Approach for Spatial-Temporal Prediction”. In: *The World Wide Web Conference*. ACM.
- Huang, Yunjie et al. (2023). “Traffic Prediction With Transfer Learning: A Mutual Information-Based Approach”. In: *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–17.
- Nguyen, Manh-Phu, Minh-Son Dao, and Thanh-Binh Nguyen (2022). “3D-STGPCN: 3D Spatio-Temporal Graph Point-wise Convolutional Network for Traffic Forecasting”. In: pp. 5998–6007.
- Zhang, Yuan et al. (2023). “Full-scale spatio-temporal traffic flow estimation for city-wide networks: a transfer learning based approach”. In: *Transportmetrica B* 11.1, pp. 869–895.
- Jin, Yilun, Kai Chen, and Qiang Yang (2022). “Selective Cross-City Transfer Learning for Traffic Prediction via Source City Region Re-Weighting”. In: pp. 731–741.

References IV

- Xing, Jiping et al. (2023). “Transfer learning for robust urban network-wide traffic volume estimation with uncertain detector deployment scheme”. In: *Electronic Research Archive* 31.1, pp. 207–228.
- Wang, Yanlong et al. (2022). “Application-Based Cooperative Content Caching via Tensor Domain Adaptation Networks”. In: pp. 22–27.
- Li, Junyi et al. (2022). “Network-scale traffic prediction via knowledge transfer and regional MFD analysis”. In: *Transportation Research Part C: Emerging Technologies* 141.
- Lin, Jiansheng et al. (2021). “A Data-driven Base Station Sleeping Strategy Based on Traffic Prediction”. In: *IEEE Transactions on Network Science and Engineering*.
- Boudabous, Safa et al. (2021). “Dynamic Graph Convolutional LSTM application for traffic flow estimation from error-prone measurements: Results and transferability analysis”. In.

References V

- Elmi, Sayda and Kian-Lee Tan (2020). "Travel Time Prediction in Missing Data Areas: Feature-based Transfer Learning Approach". In: pp. 1088–1095.
- Xie, Peng et al. (2020). "Urban flow prediction from spatiotemporal data using machine learning: A survey". In: *Information Fusion* 59, pp. 1–12.
- Wang, Leye, Di Chai, et al. (Apr. 2023). "Exploring the Generalizability of Spatio-Temporal Traffic Prediction: Meta-Modeling and an Analytic Framework". In: *IEEE Transactions on Knowledge and Data Engineering* 35.4, pp. 3870–3884.
- Tian, Chujie et al. (Sept. 2021). "A transfer approach with attention reptile method and long-term generation mechanism for few-shot traffic prediction". In: *Neurocomputing* 452, pp. 15–27.
- Manibardo, Eric L., Ibai Lana, and Javier Del Ser (Sept. 2020). "Transfer Learning and Online Learning for Traffic Forecasting under Different Data Availability Conditions:

References VI

Alternatives and Pitfalls”. In: *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE.

- Wang, Leye, Bin Guo, and Qiang Yang (2018). “Smart City Development with Urban Transfer Learning”. In.
- Lu, Bin et al. (Aug. 2022). “Spatio-Temporal Graph Few-Shot Learning with Cross-City Knowledge Transfer”. In: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM.
- Wei, Ying, Yu Zheng, and Qiang Yang (Aug. 2016). “Transfer Knowledge between Cities”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM.
- Lin, Ziqian et al. (July 2019). “DeepSTN: Context-Aware Spatial-Temporal Neural Network for Crowd Flow Prediction in Metropolis”. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 33.01, pp. 1020–1027.

References VII

- Zhang, Junbo, Yu Zheng, and Dekang Qi (Feb. 2017). “Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction”. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 31.1.
- Shaygan, Maryam et al. (Dec. 2022). “Traffic prediction using artificial intelligence: Review of recent advances and emerging opportunities”. In: *Transportation Research Part C: Emerging Technologies* 145, p. 103921.