

# Zepu Wang

Mobile: 1-2157307532 Email: [zepu@seas.upenn.edu](mailto:zepu@seas.upenn.edu)

Google Scholar: [https://scholar.google.com/citations?user=o40\\_p5sAAAAJ&hl=zh-CN&oi=ao](https://scholar.google.com/citations?user=o40_p5sAAAAJ&hl=zh-CN&oi=ao)

Website: <https://zepuwang.github.io/> (Keep Updating)

## Research Interest

- Urban Computing
  - Time Series Prediction
  - Spatial Temporal Graph Neural Networks
  - Uncertainty Quantification
- Intelligent Transportation Systems
  - Interested in Learning-based Control and Autonomy

## Education

**University of Pennsylvania**

Philadelphia, the US

*Master of Science in Engineering in Data Science*

08/2022 – 08/2024

- Current GPA: 3.9/4.0
- Relevant courses: *Statistics for Data Science, Machine Learning, Deep Learning, Natural Language Processing, Graph Neural Networks*

**Duke University/Duke Kunshan University**

Durham, the US/Kunshan, China

*Bachelor of Science in Interdisciplinary Studies (Subplan: Data Science) (by Duke)*

08/2018 – 05/2022

*Bachelor of Science in Data Science (by Duke Kunshan)*

- GPA: 3.94/4.0
- *Summa Cum Laude* (Undergraduate highest honor); Dean's List (2018, 2019, 2020, 2021, 2022); Merit-based Undergraduate Official Scholarship
- Relevant courses: *Statistical Learning, Data Analysis, Algorithms, Stochastic Process, Linear Algebra*

## Academic Achievement

- **Zepu Wang**, Yuqi Nie, Peng Sun, Nam H. Nguyen, John Mulvey, H. Vincent Poor. ST-MLP: A Cascaded Spatio-Temporal Linear Framework with Channel-Independence Strategy for Traffic Forecasting, under review by *AAAI Conference on Artificial Intelligence 2024*
- Gezhi Wang, **Zepu Wang**, Peng Sun, Azzedine Boukerche. SK-SVR-CNN: A Hybrid Approach for Traffic Flow Prediction with Signature PDE Kernel and Convolutional Neural Networks, under review by *IEEE ICC International Conference on Communications 2024*
- **Zepu Wang**, Dingyi Zhuang, Yankai Li, Shenhao Wang, Jinhua Zhao, Peng Sun. ST-GIN: An Uncertainty Quantification Approach in Traffic Data Imputation with Spatio-temporal Graph Attention and Bidirectional Recurrent United Neural Networks, accepted by *IEEE ITSC International Conference on Intelligent Transportation Systems 2023*
- **Zepu Wang**, Yifei Sun, Zhiyu Lei, Xincheng Zhu, Peng Sun. SST: A Simplified Swin Transformer-based Model for Taxi Destination Prediction based on Existing Trajectory, accepted by *IEEE ITSC International Conference on Intelligent Transportation Systems 2023*
- **Zepu Wang**, Peng Sun, Yulin Hu, Azzedine Boukerche. A novel hybrid method for achieving accurate and timeliness vehicular traffic flow prediction in road networks, accepted by *Computer Communications 2023*
- **Zepu Wang**, Peng Sun, Yulin Hu, Azzedine Boukerche. A Novel Mixed Method of Machine Learning Based Model in Traffic Flow Prediction, accepted by *ACM MSWIM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems 2022*
- **Zepu Wang**, Peng Sun, Yulin Hu, Azzedine Boukerche. SFL: A High Precise Traffic Flow Predictor for Supporting Intelligent Transportation Systems, accepted by *IEEE Globecom Global Communications Conference 2022*
- **Zepu Wang**, Peng Sun, Azzedine Boukerche. A Novel Time Efficient Machine Learning-based Traffic Flow Prediction Method for Large Scale Road Network, accepted by *IEEE ICC International Conference on Communications 2022*

## Campus Research Experience

Undergraduate Research Mentor (02/2024 ~ present)

Duke Kunshan University, Kunshan, China

Supervisor: Prof. Peng Sun, Ph.D.

- Mentor capstone projects for current undergraduate students in traffic flow prediction; Schedule weekly meeting to discuss progress; Assist them in designing deep learning models; Edit final academic paper.
- Applied convolutional neural networks (CNN) to extract trend of traffic flow; Adopted signature PDE kernel to predict the residual of traffic flow; Combined CNN with signature PDE kernel to design a hybrid traffic prediction algorithm.

Online Research Assistant (02/2023 ~ 08/2023)

Massachusetts Institute of Technology, the US

Supervisor: Prof. Jinhua Zhao, Ph.D., Prof. Shenhao Wang, Ph.D.

Project 1: Uncertainty Quantification in Traffic Data Imputation

- Utilized graph attention layers and bidirectional recurrent units to capture spatio-temporal traffic data patterns and predicted uncertainty of traffic data.

Project 2: Mixture Models for Uncertainty Quantification in Sparse Travel Demand Prediction

- Used mixture models (Laplace, Poisson, Gaussian) to approximate complex travel demand distribution, addressing high zero occurrences.

Research Assistant (02/2021 ~ 08/2022)

Duke Kunshan University, Kunshan, China

Supervisor: Prof. Peng Sun, Ph.D., Prof. Azzedine Boukerche, Ph.D.

Project 1: A Velocity-based Model in Traffic Flow Prediction

- Conducted evaluations and comparisons among popular traffic flow prediction models (Linear Regression, SVR, Decision Tree, Random Forest, LSTMs, and GRUs) in a single intersection based on accuracy ( $RMSE$ ,  $MAE$ ,  $R^2$ ).
- Increased 2% accuracy ( $RMSE$ ) on average by considering the vehicles' speed in surrounding intersections to adjust the original results.

Project 2: A Hybrid Model in Traffic Flow Prediction

- Combined Long Short-Term Memory (LSTM) neural networks, Wavelet Analysis and Spectral Analysis to design an accurate traffic flow forecasting algorithm.

Project 3: Traffic Flow Prediction using Auto-encoder

- Applied Auto-encoder as a dimension reduction technique for large road networks and increased the time efficiency by 27.4% with sacrifice of only 5% accuracy ( $MSE$ ).
- Provided a comprehensive analysis of trade-off between original data information loss and noises' filtration from the original data while applying dimension reduction technique.

Research Assistant (07/2021 ~ 08/2021)

Westlake University, China

Supervisor: Prof. Stan Z. Li, Ph.D.

Project: Protein-Protein Interactions (PPIs)

- Predicted PPIs based on primary protein structures, using traditional natural language processing (NLP) methods (RNNs) and advanced NLP methods (Transformers).
- Increased accuracy by 2% ( $MSE$ ) in RNNs methods by using the trick of pad sequences.
- Proposed potential negative influence on the model accuracy by the imbalanced dataset (most samples show no interaction exist) and provided Synthetic Minority Oversampling Technique (SMOTE) as a solution.

## Internship Research Experience

Dell EMC

Beijing, China

Research Intern

06/2021 – 09/2021

- Applied open sources of algorithms and packages to implement Kinect 2, a Windows designed camera in the Ubuntu operating system.
- Extracted depth data from the camera to make original gesture measuring model (OpenPose) more accurate.

Tencent

Beijing, China

Data Analysis Intern

05/2020 – 08/2020

- Applied machine learning methods (Linear Regression, SVR, and Tree-related methods) to analyze the public dataset (>100,000 people in Shanghai with more than 20 characteristics), focusing on the relationship between variable *age* and *frequency* of mobile payments users.
- Evaluated the prevalence of mobile payment in Shanghai based on data mining results and provided potential suggestions for mobile payment companies to play appropriate advertising strategies for specific age groups.

## Campus Work Experience

Developing Business Strategies for Master Kong

University of Pennsylvania, Philadelphia, the US

Wharton Analytics Fellow

10/2022 – 12/2022

- Served as a technical lead, led 5 undergraduate students, and cooperated with MBA students to find out potential salable flavors, and potential salable combinations of products.
- Figured out the most appropriate discount rate for Master Kong products based on real sales data from Jingdong.

## Additional information

- Programming/Statistical Skills:** Python, C++, Java, MATLAB, STATA, SQL
- Languages:** Mandarin Chinese (Native), English (Fluent, TOFEL iBT *MyBest* Score 111)
- Miscellaneous Interests:** Debate; Bridge (Card Game); Texas Hold'em; Cooking; Fitness; Soccer