Yijun Lin

Email: yijunlin@usc.edu

Mobile: (+1) 323-717-7259

Linkedin: yijun-lin

Personal Website: https://linyijun.github.io

Spatial Computing Lab Website: http://spatial–computing.github.io/

Education

Ph.D. Computer Science   GPA: 4.0/4.0   Aug. 2018 - PRESENT
University of Southern California

M.S. Data Informatics   GPA: 3.95 / 4.0   Aug. 2015 - May 2017
University of Southern California

B.S. Information Securities   GPA: 87.4/100.0   Aug. 2015 - May 2017
Tongji University (Shanghai, China)

Publications

Chiang, Y.-Y., **Lin, Y.**, Franklin, M., Eckel, S. P., Ambite, J. L., and Ku, W.-S. 2019. **Building Explainable Data Analytics for Location-Dependent Time-Series Data** (Vision Paper). The First IEEE International Conference on Cognitive Machine Intelligence.

Lin, Y., Mago, N., Gao, Y., Li, Y., Chiang, Y.-Y., Shahabi, C. and Ambite, J. L. 2018. **Exploiting spatiotemporal patterns for accurate air quality forecasting using deep learning**. In Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 359-368).

Li, K., Habre, R., Deng, H., Urman, R., Morrison, J., Gilliland, F. D., Ambite, J. L., Stripelis, D., Chiang, Y.-Y., **Lin, Y.** and Bui, A. A. T., King, C., Hosseini, A., Van Vliet, E., Majid, S., Eckel, S. P. 2019. **Applying Multivariate Segmentation Methods to Human Activity Recognition From Wearable Sensors' Data**. JMIR mHealth and uHealth, 7(2), p.e11201.

Nguyen, K., Yang, J., **Lin, Y.**, Lin, J., Chiang, Y.-Y. and Shahabi, C. 2018. **Los Angeles Metro Bus Data Analysis Using GPS Trajectory and Schedule Data (Demo Paper).** In Proceedings of 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 560-563).

Yu, X., Cheng, Y., **Lin, Y.**, Chiang, Y.-Y., Stripelis, D. and Ambite, J. L. 2018. **MAPINS: An intra-city PM2.5 modeling web application using a scalable data management and analysis system integrating public multi-source data.** In proceedings of the 2018 UCGIS/AutoCarto (pp. 132-134).

Lin, Y., Chiang, Y.-Y., Pan, F., Stripelis, D., Ambite, J. L., Eckel, S. P. and Habre, R. 2017. **Mining Public Datasets for Modeling Intra-City PM2.5 Concentrations at a Fine Spatial Resolution.** In Proceedings of 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (p. 25).

Work Experience

Research Intern   May. 2020 - Aug. 2020
Mentor: John Krumm
Adaptive Systems and Interaction Group, Microsoft Research

Data Scientist   Jun. 2017 - Jul. 2018
Spatial Sciences Institute (SSI), University of Southern California

Software Engineering (Intern)   Jul. 2014 - Sep. 2014
Global Business Service (GBS)/AI Department of IBM (China, Shanghai) Company Ltd.

Awards

IEEE TPS NSF Student Travel Award   Dec. 2019
ACM SIGSPATIAL Student Travel Awards   Nov. 2019
ACM SIGSPATIAL Student Travel Awards   Nov. 2018
Expediation Hackathon in Los Angeles, Runner-up Winner   Mar. 2017
Tongji University "Outstanding Graduate"   Dec. 2016
The 8th Tongji University Undergraduate Innovation Prize, Third Class   Apr. 2015
Tongji University "Outstanding Students"   Nov. 2014
Tongji University Scholarship, Second Class   Oct. 2014
Tongji University "Social Activities Scholarship" Prize, Second Class   Oct. 2014
Tongji University Scholarship, Third Class   Oct. 2013

Research Projects

An Explainable Deep Learning Method for Fine-Scale Air Quality Prediction
Spatiotemporal Prediction, Feature Selection, Semi-Supervised Learning, PyTorch, Postgres (PostGIS)
- Identify important features affecting PM2.5 concentrations from a variety of imbalanced environmental factors (e.g., weather and geographic features) for providing model explainability with a sparse layer
- Model the interactions between selected features over time and space at varying spatiotemporal scopes for fine-scale PM2.5 prediction using a convolutional, long-short-term memory network
- Add a semi-supervised loss function on both labeled and unlabeled predictions in a neighborhood for overcoming the limitation of sparse sensors

Exploiting Spatiotemporal Patterns for Accurate PM2.5 concentrations Forecasting

Spatiotemporal Forecasting, Graph-Convolutional RNN, TensorFlow, Postgres (PostGIS)

- Extract PM2.5-related geographic feature types, named "Geo-Context", for constructing the spatial correlation between sensor locations with a graph
- Jointly model spatial and temporal dependencies using the geo-context based diffusion convolutional recurrent neural network for accurately forecasting PM2.5 concentrations

Mining Important PM2.5-related Geographic Features for Fine-Grained Prediction

Spatial Prediction, Clustering, Feature Selection, Python, Postgres (PostGIS)

- Automatically select important PM2.5-related geographic feature types (built environment information from OpenStreetMap) that lead to the similar temporal patterns of PM2.5 concentrations
- Build an expert-free machine learning model with the extracted features for predicting PM2.5 concentrations at a fine spatial resolution

Metrans: Bus Delay Prediction System

GPS Trajectory Data, Map Matching, Java, Postgres (PostGIS)

- Align LA Bus GPS data to the bus scheduled trajectory data using map matching approach
- Measure the performance of the public transportation system in Los Angeles by predicting the bus delay time and estimated waiting time

Mining Periodic Patterns in People's Trajectories

GPS Trajectory, Clustering, Mining Frequent Itemsets, Recommendation, Python

- Identify potential stationary locations from users' trajectory data using StayPoint and OPTICS algorithm
- Mine non-strict (with a threshold) periodic patterns in people's trajectories with multiple periods (e.g., weekly, biweekly, and monthly) using a modified FP-Growth algorithm
- Predict people's schedule and recommend nearby locations that people might be interested

Professional Services

Services

Teaching Assistant   Jan. 2020 - May 2019
USC INF553 Foundations and Applications of Data Mining

Teaching Assistant   Jan. 2019 - May 2019
USC INF553 Foundations and Applications of Data Mining

Presentations

Oral Presentation, "Exploiting Spatiotemporal Patterns for Accurate Air Quality Forecasting using Deep Learning", 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, Washington, USA, 2018

Oral Presentation, "Mining Public Datasets for Modeling Intra-City PM2.5 Concentrations at a Fine Spatial Resolution", 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Redondo Beach, CA, USA, 2017

Poster Presentation, "Linking Historical Maps to the USC Shoah Foundation Visual History Archive", 28th International Cartographic Conference, Washington, DC, USA, 2017

Invited Lectures

Guest Lecturer, "Mining Public Online Data for Fine-Scale Air Quality Prediction"
USC SSCI680 Advanced Spatial Computing, 2019 Fall

Guest Lecturer, "Recommendation System" and "Spatial Data Visualization"
Soongsil University Visit, 2019 Summer

Guest Lecturer, "Map-Reduce" and "Introduction to Spark and Scala"
USC INF553 Foundations and Applications of Data Mining, 2018 Summer