

LisaNet: Learning-Integrated Space Partitioning Networks for Traffic Accident Forecasting on Heterogeneous Data

Bang An, Amin Vahedian Khezerlou, Xun Zhou, W. Nick Street, Yanhua Li

Index Terms—Supplementary Material

APPENDIX A FEATURE SUMMARY

In this section, we explain how we generate features.

A.1 Feature Generation

The details of generated features are summarized in Table 1. Features are generated on partitioned grid cells and at different time intervals.

Temporal Features F_T such calendar features are generated from the date of Vehicle Crash Records, where all grid cells share a vector of temporal features in a time interval.

Spatial Features F_S are generated based on each grid cell and remain the same over different time intervals. First, POI features are the number of POI data in each grid cell for different categories. For example, one of the POI types is shopping, we count the number of shopping instances in each grid cell. Second, basic road condition features are extracted from road network data, in which we calculate the summation or average of provided data for road segments in each grid cell. Third, we use top eigenvectors of the Laplacian matrix of road networks as spatial graph features [1], which represent the topological information for each grid cell.

Spatio-Temporal Features F_{ST} such as weather information and real-time traffic conditions are generated based on observation stations. However, Some data is only collected at sampling sites or observation stations such as weather monitoring data, thus the data is missing at the rest of the locations. To fill in the data for the entire study area, Ordinary Kriging and Universal Kriging [1] are used to estimate the weather-related features and traffic-related features respectively.

TABLE 1: Feature Table

Feature Group	Feature List
F_T	5 calendar features: day of the week, day of the year, month of the year, whether this is a holiday, whether this is weekend
F_S	13 POI features: eat-drink, going-out, sights-museums, transport, accommodation, shopping, leisure-outdoor, administrative-areas-buildings, natural-geographica, petrol-station, atm-bank-exchange, toilet-rest-area, hospital-health-care-facility 6 basic road condition features: Annual Average Daily Traffic, average speed limit, average mileage for each road, number of intersections, the total mileage of road system, and total annually traffic volume 10 SpatialGraph features
F_{ST}	9 weather features: average air temperature, highest temperature, lowest temperature, wind speed, precipitation, snowfall, snow depth, dew point temperature, and MERRA 4 real-time traffic condition features: average traffic speed, normal vehicle traffic volume, truck traffic volume, and Occupancy

REFERENCES

- [1] Zhuoning Yuan, Xun Zhou, and Tianbao Yang. Hetero-convlstm: A deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 984–992, 2018.

- B.An, X.Zhou, and W.N.Street are with the Department of Buiness Analytics, University of Iowa, Iowa City, IA 52242 USA. E-mail: bang-an, xun-zhou, nick-street@uiowa.edu.
- A.V.Khezerlou is with the Northern Illinois University, DeKalb, IL 60115 USA. E-mail: avahediankhezerlou@niu.edu.
- Y.Li is with the Worcester Polytechnic Institute, Worcester, MA 01609 USA. E-mail: yli15@wpi.edu
- Xun Zhou is the corresponding author

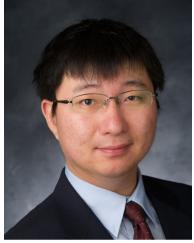
Manuscript received April 19, 2005; revised August 26, 2015.



Bang An received a Bachelor's degree in Computer Science and Business Analytics & Information Technology from Rutgers University, New Brunswick in 2019. He is currently a Ph.D student in the department of Business Analytics at Tippie College of Business, the University of Iowa. His research interests are big data analytics, spatio-temporal data mining, and urban computing.



Amin Vahedian Khezerlou received the PhD degree in business administration from the University of Iowa, in 2019. He is currently an assistant professor of Operations Management and Information Systems with College of Business, Northern Illinois University. His research interests include big data analytics and spatial and spatio-temporal data mining. He has published articles in ACM and IEEE Transactions and ACM conference proceedings.



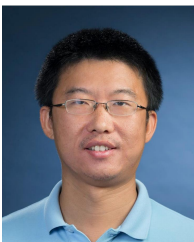
Xun Zhou (S'13-SM'21) is currently an Associate Professor in the Department of Business Analytics at the University of Iowa. He received a PhD degree in Computer Science from the University of Minnesota, Twin Cities, in 2014. His research interests include big data management and analytics, spatial and spatio-temporal data mining, and Geographic Information Systems (GIS). He has published over 80 papers in these areas and has received five best paper awards. He also served as a co-editor-in-chief

of Springer's Encyclopedia of GIS, 2nd Edition.



W. Nick Street is the Henry B. Tippie Research Professor in Business Analytics and Associate Dean for Research & PhD Programs in the Tippie College of Business, with joint appointments in the Computer Science Department, the College of Nursing, and the Interdisciplinary Graduate Program in Informatics. His research interests are in algorithmic approaches to machine learning and data mining, particularly the use of mathematical optimization in inductive learning techniques. His recent work has focused on

ensemble construction methods, federated learning, counterfactual reasoning, and personalized healthcare decision making. He has published over 130 journal, conference and workshop papers, and is the prior recipient of an NSF CAREER award. He is currently serving as a Senior Editor for the INFORMS Journal on Data Science.



Yanhua Li (S'09-M'13-SM'16) received two Ph.D. degrees in electrical engineering from Beijing University of Posts and Telecommunications, Beijing in China in 2009 and in computer science from University of Minnesota at Twin Cities in 2013, respectively. He has worked as a researcher in HUAWEI Noah's Ark LAB at Hong Kong from Aug 2013 to Dec 2014, and has interned in Bell Labs in New Jersey, Microsoft Research Asia, and HUAWEI research labs of American from 2011 to 2013. He is currently

an Associate Professor in the Department of Computer Science at Worcester Polytechnic Institute (WPI) in Worcester, MA. His research interests are urban network data analytics, smart cities, data-driven cyber-physical systems(CPS)