TAP: A Comprehensive Data Repository for Traffic Accident Prediction in Road Networks

Baixiang Huang
bhuang15@hawk.iit.edu
Illinois Institute of Technology
Chicago, Illinois, USA

Bryan Hooi
bhooi@comp.nus.edu.sg
National University of Singapore
Singapore

Kai Shu
kshu@iit.edu
Illinois Institute of Technology
Chicago, Illinois, USA

ABSTRACT

Road safety is a major global public health concern, and effective prediction of traffic accidents at a fine-grained spatial scale plays a critical role in reducing roadway deaths and serious injuries. However, previous studies have either overlooked implicit spatial correlations or inadequately simulated road structures due to the lack of graph-structured datasets. To bridge this gap, we introduce a graph-based Traffic Accident Prediction (TAP) data repository, along with two representative tasks: accident occurrence and severity prediction. With its real-world graph structures, comprehensive geographical coverage, and rich geospatial features, this repository has considerable potential to facilitate various traffic-related tasks. We extensively evaluate eleven Graph Neural Network (GNN) baselines using the constructed datasets. We also develop a novel GNN-based model, which can capture additional angular and directional information from road networks. We demonstrate that the proposed model consistently outperforms the baselines. The data and code are available at https://github.com/baixianghuang/travel.

CCS CONCEPTS

• **Information systems → Spatial-temporal systems**; **Location based services**.

KEYWORDS

Traffic accident prediction, Graph neural networks, Road safety

ACM Reference Format:
Baixiang Huang, Bryan Hooi, and Kai Shu. 2023. TAP: A Comprehensive Data Repository for Traffic Accident Prediction in Road Networks. In *The 31st ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL '23), November 13–16, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3589132.3625655

1 INTRODUCTION

Road traffic accidents are the leading cause of death for young people globally [24]. The U.S. traffic fatality rate has experienced an alarming 19 percent surge from 2019 to 2021, marking the highest number of road deaths in the U.S. Fatality Analysis Reporting System's history since 2005 [1, 22]. Therefore, understanding and mitigating traffic crashes is an imperative task. Traffic crash prediction at a fine-grained spatial scale (e.g., a traffic intersection) can help governments mitigate traffic risks, such as informing the design of future road networks and planning accident-response facilities with the awareness of risk-prone accident hotspots [23].

SIGSPATIAL '23, November 13–16, 2023, Hamburg, Germany

ACM ISBN 979-8-4007-0168-9/23/11.
https://doi.org/10.1145/3589132.3625655

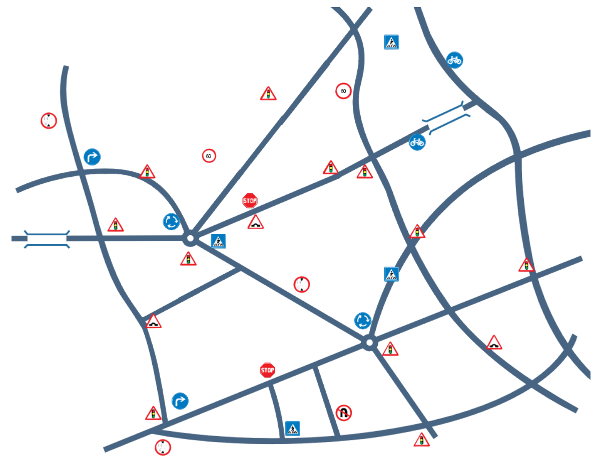

Figure 1: A road network with various environmental features (e.g., traffic signs, speed limits, and road types).

This work focuses on the environmental risk factors of traffic accidents: in particular, *can we predict how risk-prone a road intersection is only based on geospatial map data.* Specifically, we concentrate on accidents occurring around intersections (also known as junctions or crossroads) due to the following reasons: (1) Representing intersections as nodes and roads as edges aligns with real-world road networks; (2) The majority of traffic crashes tend to occur near intersections [9]; (3) Intersections typically contain transportation infrastructure such as traffic lights, stop signs, and pedestrian crosswalks, which provide valuable environmental features.

Figure 1 shows a typical urban road network that has various road types, each exhibiting distinctive attributes such as lane counts and road lengths. Consequently, various road types and features may pose different crash risks. For instance, a high-volume freeway is likely to have more roadway crashes than a less traveled two-lane residential road because the freeway has more traffic and a higher speed limit [21]. Moreover, road characteristics such as turning radius and direction also affect road safety. In general, sharper road curves are more dangerous [17].

Previous studies commonly employ hand-crafted features and discretize a region into sub-regions using a grid and overlook the underlying graph structure and implicit spatial correlations [3, 26]. In comparison, Graph Neural Networks (GNNs) incorporate surrounding spatial information by aggregating features from neighbors multiple hops away. However, the scarcity of graph-based datasets

Table 1: Comparison with existing traffic prediction datasets.

Dataset	Type	Geospatial	Graph-based	Coverage	Time
UK Traffic Accidents [14]	Accident			Nation-wide (1.6M rows)	2016
NY State Motor Vehicle Crashes [13]	Accident			State-wide (3.5M rows)	2022
Maryland Vehicle Crashes [19]	Accident			State-wide (0.9M rows)	2023
NY City Motor Vehicle Collisions [4]	Accident			City-wide (2.0M rows)	2023
Chicago Traffic Crashes [12]	Accident			City-wide (0.7M rows)	2023
TAP	Accident	✓	✓	Nation-wide (17.0M nodes and 43.0M edges)	2023

is a major obstacle of applying GNNs for this task. Previous studies generate artificial graphs in a manner that does not adequately capture real-world road structures [7, 27]. Moreover, the majority of existing datasets [4, 12, 13] have limited coverage and lack geospatial features, while also lacking graph structure information, thereby making the application of GNNs infeasible.

To construct the Traffic Accident Prediction (TAP) data repository, we first collect raw accident records, street geospatial data, and graph structure information. The crash data are integrated with graph structure and geospatial features. We then compute geometric data based on graph structures. The TAP is organized into city-level and state-level, covering 1,000 U.S. cities and 49 states, which allows users to study various traffic-related problems.

Significantly facilitated by the proposed data repository, we develop a novel framework called Traffic Accident Vulnerability Estimation via Linkage (TRAVEL). This framework aggregates features from its neighbors in a way that captures both the angles and directions of roads adjacent to a node.

Our contributions are the following:

- We construct and release a new data repository that significantly simplifies the application of graph-based machine learning methods for traffic crash prediction and analysis.
- We formulate traffic accident prediction as a node prediction problem, with the objective of forecasting accident occurrences or severity levels across a given road network.
- We propose a new GNN architecture, TRAVEL, which can capture angular and directional information from road networks. We show that our proposed model consistently outperforms 13 state-of-the-art machine learning baselines.

2 DATA REPOSITORY CONSTRUCTION

The Traffic Accident Prediction (TAP) data repository incorporates real-world graph structure and geospatial information. Our data repository is readily accessible and designed to be user-friendly. Its data format can seamlessly integrate with existing GNNs. We compare existing datasets in Table 1. The TAP incorporates real-world graph structure information, contains rich geospatial features, and has comprehensive geographical coverage. Data statistics and numerical measures can be found in our GitHub repository. We are committed to regularly updating the dataset with the latest traffic crash records and auxiliary features.

2.1 Data Collection

The raw accident events come from Bing Map Traffic [2, 20]. It contains about 2.8 million traffic accident data between January 2016 and December 2021. Traffic accident data are typically collected from police reports [4, 8, 19]. However, traffic flow and speed data are less available as traffic monitoring devices are not widely available or completely prohibited by laws [5, 28]. We also use OpenStreetMap (OSM) [16], a collaborative initiative that provides a freely available geospatial database, as our data source. OSM data contain rich environmental features such as road type, road length, and the number of lanes. OSM includes walkable, drivable, and bikeable urban road data. We use drivable public road data (private-access or service roads not included).

Table 2: Edge features (top) and node features (bottom) included in our datasets.

Graph features	Description
highway	The type of a road (tertiary, motorway, etc.).
length	The length of a road.
bridge	Indicates whether a road represents a bridge.
lanes	The number of lanes of a road.
oneway	Indicates whether a road is a one-way street.
maxspeed	The maximum legal speed limit of a road.
access	Describes restrictions on the use of a road.
tunnel	Indicates whether a road runs in a tunnel.
junction	Describes the junction type of a road.
angle	Angular information of a road.
direction	Directional information about a road.
highway	The road type of a node.
street_count	The number of roads connected to a node.

2.2 Data Preprocessing

We first build road networks using structural and geospatial features from OSM. The features are listed in Table 2. The angular and directional features will be discussed in section 3.1. We run reverse geocoding to find the corresponding addresses of accident coordinates using Nominatim [11]. Next, the geocoded data are split according to settlement hierarchy: there are 49 states and 1,585 cities. Since the spatial distribution of traffic crashes is sparse and imbalanced, there are limited positive samples for small cities. Therefore, We only select the top 1,000 cities.

Next, missing values are replaced with a new category, and feature data are encoded using one-hot encoding. Then the coordinate data of accident locations are used to find the nearest corresponding graph nodes based on the haversine distance. For the accident occurrence prediction task, binary labels are added to each node indicating whether it contains at least one accident. For the severity prediction task, average accident severities are bucketized into 8 classes (using an interval size of 0.5) to be used as labels. The severity feature uses an integer from 0 to 7 to specify the magnitude of the accident's impact from low to high.

3 METHODOLOGY

We model a road network as a weighted directed graph, where vertices represent endpoints (intersections or dead-end nodes) and edges represent roads. In the accident occurrence prediction task, we aim to output a binary prediction indicating whether a node has accidents or not. In the accident severity prediction task, We aim to forecast the severity of each node from a set of 8 classes. Therefore, both tasks are formulated as node classification problems.

3.1 TRAVEL Framework

The Traffic Accident Vulnerability Estimation via Linkage (TRAVEL) framework is a GNN-based approach that can capture geospatial and additional angular and directional information. The angular component captures angle-related information about an intersection (e.g., whether it has a right or left turn, sharp turns, etc.). The directional component contains the direction feature of a road: for instance, whether it is heading north-to-south versus east-to-west.

3.1.1 Angular Component. A graph has node features $\mathbf{x}_v$ for node v and edge attributes $\mathbf{e}_{uv}$ for edge (u, v). The angular component augments the message passing process from node u to node v with angle information between the road (u, v) and all the other roads intersecting at node v. Given points u, v, w, let $\angle(\overrightarrow{uv}, \overrightarrow{wv})$ denote the directed angle from $\overrightarrow{uv}$ to $\overrightarrow{wv}$. The angular component is designed to augment the messages passed from u to v with information about the angles between road (u, v) and other roads intersecting at v. Formally, the set of directed angles between road (u, v) and each of the other roads at v is: $\Phi_{uv} := \{\angle(\overrightarrow{uv}, \overrightarrow{wv}) : w \in \mathcal{N}_v \setminus \{u\}\}$

Next, we design a aggregation function to emphasize the presence of informative features from Φ_{uv}: 1) sharp left turns, 2) sharp right turns, and 3) nearly straight roads. We define Φ^{π}_{uv} as the set $\{|\pi - \phi| : \phi \in \Phi_{uv}\}$, then aggregating as (where $\|$ denotes concatenation): $\mathbf{a}_{uv} := \min(\Phi_{uv}) \parallel \max(\Phi_{uv}) \parallel \min(\Phi^{\pi}_{uv})$. The first two parts correspond to the sharpest angles of left and right turns to edge (u, v). The third part corresponds to the angle of (u, v) to the road which is closest to a straight road along with (u, v). Such aggregated angular information $\mathbf{a}_{uv}$ provides a concise summary of the angles information between (u, v) and other roads at v.

The angular component incorporates angular information $\mathbf{a}_{uv}$ when passing a message along (u, v). This component takes in node representations $\mathbf{h}_v$ (suppressing the layer number since we only describe a single TRAVEL layer), and outputs the angular node representations $\mathbf{h}^{\text{Angle}}_v$:

$$\mathbf{h}^{\text{Angle}}_v = \mathsf{ReLU}(\mathbf{W}\mathbf{h}_v + \mathbf{m}^{\text{Angle}}_{\mathcal{N}(v)}) \tag{1}$$

$$\mathbf{m}^{\text{Angle}}_{\mathcal{N}(v)} = \sum_{u \in \mathcal{N}(v)} \mathsf{MLP}(\mathbf{h}_u \parallel \mathbf{e}_{uv} \parallel \mathbf{a}_{uv}) \tag{2}$$

where $\mathbf{W}$ represents a trainable weight matrix. Initial 0-th layer embeddings are equal to node features $\mathbf{x}_v$. $\mathbf{m}_{\mathcal{N}(v)}$ denotes the aggregated message from node v's neighborhood $\mathcal{N}(v)$.

3.1.2 Directional Component. The directional information are useful because in some cities, north-south roads may have different characteristics from east-west roads. This component captures the direction that each road is heading in. Let LAT_u and LON_u denote the latitude and longitude of node u, respectively. We compute the edge **direction** (u, v) as: $\mathbf{d}_{uv} = (\text{LAT}_v - \text{LAT}_u, \text{LON}_v - \text{LON}_u)$

$$\mathbf{h}^{\text{Dir}}_v = \mathsf{ReLU}(\mathbf{W}\mathbf{h}_v + \mathbf{m}^{\text{Dir}}_{\mathcal{N}(v)}) \tag{3}$$

$$\mathbf{m}^{\text{Dir}}_{\mathcal{N}(v)} = \sum_{u \in \mathcal{N}(v)} \mathsf{MLP}(\mathbf{h}_u \parallel \mathbf{e}_{uv} \parallel \mathbf{d}_{uv}) \tag{4}$$

3.1.3 Combined TRAVEL Layer. The complete TRAVEL layer is the concatenation of the angular and directional components, i.e., $\mathbf{h}^{\text{Angle}}_v \parallel \mathbf{h}^{\text{Dir}}_v$. This TRAVEL layer can be straightforwardly trained using standard loss functions or plugged into any existing GNN.

4 EXPERIMENTS

In the accident occurrence prediction task. Table 3 shows the prediction results on the six sample cities. We run every experiment three times and report the average score along with the standard deviation in the format of "average score ± standard deviation". We generally observe that: (1) The proposed TRAVEL consistently achieves the best performance on all the metrics due to its ability to capture angular and directional features on top of other environmental features. (2) GNN-based approaches generally outperform XGBoost and MLP. This is because nodes in GNNs can aggregate feature information from their neighbors, while the MLP and XGBoost can only learn from local feature data. (3) GNN variants that support multi-dimensional edge features generally outperform models that do not support them. We include additional experimental results, evaluation details, and settings in our GitHub repository or the extended version of this paper [6].

5 RELATED WORK

Early work applied statistical regression models to predict traffic accidents [15]. Najjar et al. [10] trained Convolutional Neural Networks on satellite images to produce a traffic risk map. Another line of research also included temporal features. Prior research applied k-means clustering, and logistic regression [18]. Chen et al. [3] used stacked denoising autoencoders to infer traffic risk. Yu et al. [25] combined Long Short-Term Memory and stacked autoencoders for post-accident condition prediction. Some recent research applies GNNs to this task. However, their graph constructions fail to adequately capture the structures present in actual road systems. For example, Zhou et al. [26] used Graph Convolutional Networks for accident prediction over a rectangular grid. Zhou et al. [27] formulated graphs by dividing an area into grids and adding edges between grids with strong correlations. In contrast, our approach not only takes advantage of graph structures but also incorporates real-world environmental features.

6 CONCLUSION

In this paper, we affirm the benefits of GNNs for the task of roadway crash prediction. We first formulate the accident occurrence and severity prediction tasks as node classification problems. To

Table 3: Accident occurrence prediction results in terms of F1 score(%) and AUC(%).

	Miami		Los Angeles		Orlando		Dallas		Houston		New York	
Classifier	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC
XGBoost	11.8±1.6	53.1±0.4	16.5±0.4	54.5±0.1	39.4±1.1	61.4±0.3	31.0±2.2	58.5±0.8	16.1±0.6	53.8±0.2	23.8±0.8	56.8±0.3
MLP	13.0±0.8	61.3±1.9	16.0±0.5	66.3±0.1	38.8±1.9	65.6±2.0	32.5±1.1	67.8±0.4	15.9±0.7	64.0±0.4	23.7±0.9	65.6±1.2
GCN	20.0±3.3	68.5±3.3	40.2±1.1	80.4±0.3	51.6±0.8	73.1±1.2	39.8±1.9	73.1±0.4	16.4±1.3	66.7±0.2	39.2±3.7	75.5±0.4
ChebNet	20.7±2.9	71.3±3.6	39.8±1.8	81.0±0.3	53.1±0.6	76.7±1.6	42.0±0.5	75.8±0.4	23.8±0.5	69.6±0.5	40.9±4.3	78.3±1.1
ARMANet	19.2±3.3	69.5±3.5	40.8±1.0	80.9±0.4	51.5±1.3	75.7±1.4	41.2±0.5	75.6±0.2	23.1±0.4	69.2±0.7	42.4±1.1	77.7±0.6
GraphSAGE	20.7±2.4	67.6±2.8	41.6±0.5	80.5±0.3	52.6±1.3	74.1±1.2	44.2±0.5	74.4±0.3	23.7±0.4	68.5±0.4	42.5±1.1	76.3±0.1
TAGCN	25.2±1.1	73.5±2.4	49.5±0.7	84.7±0.2	53.3±2.5	77.2±1.2	45.4±0.4	77.0±0.5	23.7±0.6	70.5±0.3	42.0±1.1	81.5±0.2
GIN	22.8±1.2	72.7±2.6	41.6±0.7	81.8±0.2	54.7±1.4	76.6±1.1	41.3±2.0	75.2±0.3	20.9±1.0	68.0±0.3	41.7±2.1	79.1±0.5
GAT	22.6±1.5	68.3±3.0	41.6±0.4	80.9±0.2	55.3±1.3	74.1±1.0	42.1±1.5	73.6±0.3	17.8±0.8	67.3±0.3	42.2±0.5	76.6±0.4
MPNN	38.8±2.1	82.4±1.0	46.0±1.6	83.9±0.2	61.4±2.5	81.8±0.7	48.5±1.9	79.4±0.4	28.2±1.7	73.5±0.5	44.9±0.8	86.9±0.4
CGC	34.4±2.7	79.5±1.5	45.0±1.2	81.5±0.2	59.0±2.1	81.1±0.8	48.5±0.5	79.2±0.7	27.3±1.9	72.3±0.1	40.6±1.2	85.4±0.8
Transformer	37.7±3.3	81.0±1.9	48.9±0.3	83.8±0.3	62.9±1.6	82.0±0.7	49.8±0.7	80.0±0.7	28.4±0.7	73.9±0.4	43.1±0.7	87.2±0.4
GEN	44.9±3.1	81.0±2.4	48.6±6.2	82.7±0.9	63.0±1.1	81.2±0.9	56.5±1.7	79.5±0.1	34.1±6.0	73.7±0.4	47.3±1.4	87.7±0.9
TRAVEL	51.9±1.0	84.9±0.9	55.3±0.9	85.9±0.5	65.0±0.4	82.3±0.4	58.0±0.9	80.8±0.7	46.4±0.7	74.5±0.3	51.1±0.9	88.2±0.2

stimulate future research in this field, we construct and release a comprehensive graph-based data repository. We also evaluate state-of-the-art machine learning approaches using the proposed data. The proposed datasets feature real-world graph structures and geospatial data, making them a reliable, comprehensive, and user-friendly resource for traffic crash analysis. Furthermore, we propose our TRAVEL framework, designed to capture angular and directional attributes. The experiments show that TRAVEL consistently outperforms the baselines. For future work, we plan to investigate the explainability of geospatial and geometric features, as well as their correlations with traffic accident risk.

ACKNOWLEDGMENTS

This work is supported in part by NSF SaTC-2241068. We would like to thank the anonymous reviewers for their helpful comments.

REFERENCES

[1] National Highway Traffic Safety Administration. 2022. *Early Estimate of Motor Vehicle Traffic Fatalities in 2021.* Technical Report. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813283
[2] Microsoft Bing. 2022. *Bing Map Traffic API.* https://www.bingmapsportal.com
[3] Quanjun Chen, Xuan Song, Harutoshi Yamada, and Ryosuke Shibasaki. 2016. Learning deep representation from big and heterogeneous data for traffic accident inference. In *Thirtieth AAAI conference on artificial intelligence.*
[4] New York City Police Department. 2023. *New York City Motor Vehicle Collisions.* Retrieved June 22, 2023 from https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95
[5] Insurance Institute for Highway Safety (IIHS). 2023. *Automated enforcement laws.* https://www.iihs.org/topics/red-light-running/automated-enforcement-laws
[6] Baixiang Huang, Bryan Hooi, and Kai Shu. 2023. TAP: A Comprehensive Data Repository for Traffic Accident Prediction in Road Networks (Long Version). *arXiv preprint arXiv:2304.08640* (2023).
[7] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. 2017. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926* (2017).
[8] Daryl Lloyd. 2016. Reported road casualties in Great Britain: Main results 2015. (2016).
[9] Sobhan Moosavi, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, Radu Teodorescu, and Rajiv Ramnath. 2019. Accident risk prediction based on heterogeneous sparse data: New dataset and insights. In *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.* 33–42.
[10] Alameen Najjar, Shun'ichi Kaneko, and Yoshikazu Miyanaga. 2017. Combining satellite imagery and open data to map road safety. In *Thirty-First AAAI Conference on Artificial Intelligence.*
[11] Nominatim. 2023. *Nominatim.* https://nominatim.org
[12] City of Chicago. 2023. *Chicago Traffic Crashes.* https://catalog.data.gov/dataset/traffic-crashes-crashes
[13] New York State Department of Motor Vehicles. 2022. *New York State Motor Vehicle Crashes.* Retrieved June 22, 2023 from https://data.ny.gov/Transportation/Motor-Vehicle-Crashes-Vehicle-Information-Three-Ye/xe9x-a24f
[14] UK Department of Transport. 2016. *UK Accidents.* https://www.kaggle.com/datasets/daveianhickey/2000-16-traffic-flow-england-scotland-wales
[15] Jutaek Oh, Simon P Washington, and Doohee Nam. 2006. Accident prediction model for railway-highway interfaces. *Accident Analysis & Prevention* 38, 2 (2006), 346–356.
[16] OpenStreetMap. 2023. *OpenStreetMap.* https://www.openstreetmap.org
[17] Sarbaz Othman, Robert Thomson, and Gunnar Lannér. 2009. Identifying critical road geometry parameters affecting crash rate and crash type. In *Annals of Advances in Automotive Medicine/Annual Scientific Conference*, Vol. 53. Association for the Advancement of Automotive Medicine, 155.
[18] Seong-hun Park, Sung-min Kim, and Young-guk Ha. 2016. Highway traffic accident prediction using VDS big data analysis. *The Journal of Supercomputing* 72, 7 (2016), 2815–2831.
[19] Maryland State Police. 2023. *Maryland Statewide Vehicle Crashes.* Retrieved June 22, 2023 from https://opendata.maryland.gov/Public-Safety/Maryland-Statewide-Vehicle-Crashes/65du-s3qu
[20] Moosavi Sobhan. 2023. *US Accidents Dataset.* https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents
[21] SWOV. 2012. Fact Sheet: The relation between speed and crashes. https://shorturl.at/uJLO2
[22] TRIP. 2022. *Addressing America's Traffic Safety Crisis: Examining The Causes of Increasing U.S. Traffic Fatalities and Identifying Solutions to Improve Road User Safety.* Technical Report. https://tripnet.org/wp-content/uploads/2022/06/TRIP_Addressing_Americas_Traffic_Safety_Crisis_Report_June_2022.pdf
[23] Federal Highway Administration US Department of Transportation. 2023. *Safety Countermeasures initiative.* https://highways.dot.gov/safety/proven-safety-countermeasures
[24] WHO et al. 2018. *Global status report on road safety 2018: summary.* Technical Report. World Health Organization.
[25] Rose Yu, Yaguang Li, Cyrus Shahabi, Ugur Demiryurek, and Yan Liu. 2017. Deep learning: A generic approach for extreme condition traffic forecasting. In *Proceedings of the 2017 SIAM international Conference on Data Mining.* SIAM, 777–785.
[26] Zhengyang Zhou, Yang Wang, Xike Xie, Lianliang Chen, and Hengchang Liu. 2020. RiskOracle: A Minute-Level Citywide Traffic Accident Forecasting Framework. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 1258–1265.
[27] Zhengyang Zhou, Yang Wang, Xike Xie, Lianliang Chen, and Chaochao Zhu. 2020. Foresee urban sparse traffic accidents: A spatiotemporal multi-granularity perspective. *IEEE Transactions on Knowledge and Data Engineering* (2020).
[28] Xiaohua Zhu and Mark Antony Freeman. 2019. An evaluation of US municipal open data portals: A user interaction framework. *Journal of the Association for Information Science and Technology* 70, 1 (2019), 27–37.