INTEGRATING SITUATIONAL FACTORS AND HETEROGENEOUS DATA FOR ENHANCED TRAFFIC ACCIDENT ANALYSIS

UTHAIB Masar Abed Uthaib
Postgraduate Student
TYUTYUNNIK Vyacheslav Mikhailovich
Doctor of Sciences in Technology, Professor
Professor of the Department of Design of Radioelectronic and Microprocessor Systems
Tambov State Technical University
Tambov, Russia

Traffic accidents pose significant risks to both persons and infrastructure, necessitating the employment of advanced modelling techniques to investigate the causes, consequences, and trends of these incidents. Traffic incident modelling is difficult due to the large amount of data needed to account for factors like infrastructure problems, environmental factors, spatiotemporal changes, vehicle and driver characteristics, and more. Recent breakthroughs in machine learning, including Graph Neural Networks (GNNs), attention-based deep learning models, and hybrid tree-based classifiers, have improved predicted accuracy and interpretability. These techniques utilize multimodal data to facilitate real-time accident prediction, severity assessment, and hotspot discovery. Nonetheless, obstacles such data sparsity, class imbalance, and model interpretability remain. This paper reviews current approaches to crash modeling, illustrating the impact of diverse information on crash frequency and severity to improve road safety, optimize traffic management, and improve crash prevention policy design. Future directions include multimodal data integration, causal machine learning techniques, and real-time decision support systems to improve crash modeling.
Keywords: modelling, heterogeneous information, traffic accidents situation, machine learning.

Introduction. Traffic accidents pose a substantial threat to public safety and transportation infrastructure globally. The intricacy of modelling traffic accidents stems from the necessity to manage many data sources, spatio-temporal fluctuations, and the infrequency of accident occurrences. Innovations in data analytics and machine learning in the last few years have provided novel approaches to these problems [1; 2]. This paper examines the modelling of traffic accidents, demonstrating diverse information and situational aspects, and deriving lessons from several research studies.

Statement of the Research Problem. Traffic accidents constitute a grave danger to transportation safety, resulting in injuries, fatalities, and economic detriment globally. Although many studies on traffic accident prediction and prevention, existing models are inadequate in dealing with the complexity of accidents as a result of the variety of circumstances that can cause them. These elements encompass spatio-temporal variability, environmental conditions, human behavior, vehicle characteristics, and infrastructural disparities. Traditional statistical models often fail to accurately depict these dynamic relationships. Problems with the quality of information, interpretation of models, and generalising remain in spite of advances in ML and AI. The overarching goals of this study are to increase our knowledge of the various data types that contribute to traffic accidents, to improve our prediction models through the use of cutting-edge machine learning techniques, and to identify efficient methods for avoiding and responding to accidents. The study's overarching goal is to improve road safety by throwing light on the factors that contribute to traffic accidents and providing data to back decisions that need them.

Approaches to modelling traffic incidents.

1 – Time and Place Analyses using Graph-Neural Networks.

Because they capture the spatial and temporal relationships within road networks, graph neural networks have emerged as an efficient tool for traffic accident modelling [3].

2 – Combination of Tree-Based and Deep

Learning Models.

Hybrid models that integrate tree-based classifiers with deep learning approaches have shown capability in forecasting traffic accidents. Methods such as AdaBoost, XGBoost, and Random Forest, when integrated with deep learning models, have attained elevated performance measures, including precision, recall, and F1-scores beyond 90%. These models are very proficient in managing imbalanced datasets, characterized by the infrequency of incidents [4].

3 – Comprehensive Analysis based on multimodal large models.

To provide a thorough approach to accident and avoidance analysis, the multimodal large model called Accident GPT integrates data collected from multiple sources, including V2X (Vehicle-to-Everything) perceptions [5].

Traffic incidents involving heterogeneous data.

1. The Built Environment.

The built environment, which includes things like road types, land uses, and urban densities, has a major impact on how bad accidents are. While there is a positive correlation involving land use variety and incident incidence in rural areas, research shows that metropolitan locations with varied land use typically have higher incident rates [5; 6].

2 – Variation across space and time.

Significant spatiotemporal variation in road accidents indicates that both the frequency and intensity of these incidents differ depending on where and when they occur. Time and place determine the relative importance of factors such movement of traffic and roadway features in relation to events [6; 7].

3 – Meteorological and Environmental Variables.

In the presence of precipitation, snow, and fog, the probability of vehicular accidents is substantially increased. After accounting for variables such as road kind as well as time of day, scholars who used meteorological data to train machine learning models discovered that inclement weather still raises the probability of accidents. There has been an unusually high incidence of catastrophic events due to snow, water, and insufficient brightness. The research found that models trained using deeper neural networks were quite effective at predicting catastrophes [7]. Additionally, explainable machine learning approaches have shown that wind chill and velocity of the wind are important meteorological factors in predicting the length and severity of accidents [8]. A separate research demonstrates that the climate, in addition to location and roadway type, contributes to collisions, and machine learning approaches successfully correlate incident variables to severity degrees [9].

4 – Vehicle and The driver Characteristics.

Critical components of accident models include the kind of car, driver behaviour, and road user attributes. Motorcyclists are more likely to be involved in accidents in some urban areas, while truck-car accidents predominate on highways. Variables such as the kind of car, driver situation and crash place were identified as important factors influencing the likelihood of injury [10]. The importance of real-time driving information has been brought to light in studies using deep learning and radial basis function neural networks. These studies show that vehicle dynamics and driver behaviors greatly impact accident probability [11]. The level of severity of driver injury can be significantly impacted by factors including reckless driving, high speed, and the type of occurrence when using algorithmic machine learning [12].

Applications of Traffic Accident Models.

1 – Incident Risk Identification.

A combination of historical accident data, present road conditions, climatic variables, and traffic flow allows, machine learning techniques to pinpoint dangers regions. Finding accident-prone areas and putting preventative measures into action is achieved through the use of gradient boost and random forest algorithms [13].

2 – Analyzing Geography and Planning Cities.

Geographical information system (GIS) facilitate better infrastructure building, data consolidation on traffic flows, and visibility into accident hotspots. Better safety at crossings and on roads can be achieved with the help of these models [14].

3 – Estimating the Difficulty Level of an Accident.

Algorithms classifiers like support vector machines (SVM) and randomly generated forests can be used to classify events based on their severity. Helps first responders in determining order of importance [15].

4 – Assistive systems for Drivers in the real time.

Modern deep learning algorithms, such as convolution recurrent neural networks (CRNNs), permit the delivery of accident predictions and driver alerts in real-time. Decreases accident rates by modifying driver actions through predictive risk assessment [15].

Restrictions of modelling traffic accidents.

1 – The scarcity of data.

Because of the limitation of data used in the traffic incidents predicting as a result data imp lanced will be generated, this effecting classification accidents. Techniques like SMOTE and another ADASYN help to solve this issue [4].

2 – Temporary and Spatial differentiation.

To identify the various influence of each element in multiple positions and times. Models need to take consideration to this variation. Model like GWTR and GWTPH used to detect this issue [16].

3 – Interpreted deep learning models

In spite of the power of deep learning approaches but might be there is lack of interpretation. Models like SHAP used to solve this problem [17; 18].

Conclusion. With the use of numerous sources of information and advanced machine learning techniques, traffic accident modelling has come a long way. To better capture spatiotemporal heterogeneity and improve forecast accuracy, hybrid approaches, attention-based models, and graph neural networks have all exhibited potential. Enhancing traffic safety through the use of multimodal data, causal machine learning, and real-time decision support systems should be the focus of future research [19; 20].

REFERENCES

1. *Gao X. et al.* SMA-HYPER: Spatiotemporal Multi-View Fusion Hypergraph Learning for Traffic Accident Prediction // arXiv (Cornell University). 2024. DOI: https://doi.org/10.48550/arXiv.2407.17642.

2. *Liu H. et al.* Intelligent traffic accident detection system in complex dynamic scenarios based on the dual-stream spatiotemporal-fusion model: research paper // Preprint. DOI:http://dx.doi.org/10.2139/ssrn.5145098.

3. *Nippani A. et al.* Graph Neural networks for road safety modeling: Datasets and evaluations for accident analysis // arXiv (Cornell University). 2023. DOI:https://doi.org/10.48550/arXiv.2311.00164.

4. *Ameksa M. et al.* Predictive analysis for road accidents using a tree-based and deep learning fusion system // Journal of Intelligent & Fuzzy Systems. 2023. Vol. 46, № 1. P. 2381-2397. DOI:10.3233/jifs-232078.

5. *Wang L. et al.* AccidentGPT: Accident Analysis and Prevention from V2X Environmental Perception with Multi-modal Large Model // arXiv (Cornell University). 2023. – URL:https://arxiv.org/abs/2312.13156.

6. *Liu H. et al.* Exploring the spatiotemporal heterogeneity of freeway secondary crashes using GTWR model // Journal of Transportation Safety & Security. 2023. Vol. 16, № 3. P. 323–346. DOI: 10.1080/19439962.2023.2211028.

7. *Jiang J., Miao Y., Wu D.* Machine learning-based prediction analysis of potential factors in traffic accidents // Applied and Computational Engineering. 2024. Vol. 99, № 1. P. 112-120. DOI:https://doi.org/10.54254/2755-2721/99/20251788.

8. *Sukonna R.T., Swapnil S.I.* A Bi-level Framework for Traffic Accident Duration Prediction: Leveraging Weather and Road Condition Data within a Practical Optimum Pipeline // arXiv (Cornell University). 2023. DOI:https://doi.org/10.48550/arXiv.2311.00634.

9. *Soedirman S. et al.* It is about weather: Explainable machine learning for traffic accident understanding // 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC). 2023. P. 2689–2694. DOI:10.1109/SMC53992.2023.10393997.

10. *Zhang Z. et al.* Exploring spatial heterogeneity in factors associated with injury severity in

speeding-related crashes: An integrated machine learning and spatial modeling approach // Accident Analysis & Prevention. 2024. Vol. 206. P. 107697. DOI: 10.1016/j.aap.2024.107697.
11. *Arciniegas-Ayala C. et al.* Prediction of Accident Risk Levels in Traffic Accidents Using Deep Learning and Radial Basis Function Neural Networks Applied to a Dataset with Information on Driving Events // Applied Sciences. 2024. Vol. 14, № 14. P. 6248. DOI:https://doi.org/10.3390/app14146248.
12. *Sorum N.G., Pal D.* Identification of the best machine learning model for the prediction of driver injury severity // International Journal of Injury Control and Safety Promotion. 2024. Vol. 31, № 3. P. 360-375. DOI: https://doi.org/10.1080/17457300.2024.2335478.
13. *Mehdi M.* Machine learning and multisource data analysis approach towards traffic accident risk prediction // InterConf. 2024. № 50(221). P. 401-406. DOI:10.51582/interconf.19-20.10.2024.040.
14. *Nayak A., Goyal K.* Traffic modeling and accidental data analysis using GIS: A Review // IOP Conference Series Earth and Environmental Science. 2024. Vol. 1327, № 1. P. 012028. DOI:10.1088/1755-1315/1327/1/012028.
15. *Priyanka S. et al.* Machine Learning Applications in Traffic Safety: Assessing accident severity Automatically // 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). 2023. P. 1197-1203. DOI:10.1109/ICECA58529.2023.10394978.
16. *Liu H. et al.* Exploring the spatiotemporal heterogeneity of freeway secondary crashes using GTWR model // Journal of Transportation Safety & Security. 2023. Vol. 16, № 3. P. 323-346.DOI:10.1080/19439962.2023.2211028.
17. *Benfaress I., Bouhoute A., Zinedine A.* Enhancing traffic accident severity prediction using RES-NET and SHAP for interpretability // AI. 2024. Vol. 5, № 4. P. 2568-2585. DOI:https://doi.org/10.3390/ai5040124.
18. *Kashifi M.T.* Robust spatiotemporal crash risk prediction with gated recurrent convolution network and interpretable insights from SHapley additive explanations // Engineering Applications of Artificial Intelligence. 2023. Vol. 127. P. 107379 DOI:https://doi.org/10.1016/j.engappai.2023.107379.
19. *Uthaib M.A., Tyutyunnik V.M.* Pattern recognition of the state registration plate of a vehicle // Мир науки без границ [Электронный ресурс]: материалы XI Всерос. науч.-практ. конф. молодых учёных, Тамбов, ТГТУ, 26 апр. 2024 г. – Тамбов: издат. Центр ФГБОУ ВО «ТГТУ», 2024. – С. 285-287. – 1 электрон. опт. диск (CD ROM).
20. *Тютюнник В.М., Удаиб М.А.* Нейросетевые методы и модели прогнозирования дорожно-транспортных происшествий и снижения их рисков // Промышленные АСУ и контроллеры. – 2024. – № 12. – С. 47-56. – DOI:10.25791/asu.12.2024.15483.

ИНТЕГРАЦИЯ СИТУАЦИОННЫХ ФАКТОРОВ И РАЗНОРОДНЫХ ДАННЫХ ДЛЯ РАСШИРЕННОГО АНАЛИЗА ДОРОЖНЫХ ПРОИСШЕСТВИЙ

УДАИБ Масар Абед Удаиб
аспирант
ТЮТЮННИК Вячеслав Михайлович
доктор технических наук, профессор
профессор кафедры конструирования радиоэлектронных и микропроцессорных систем
Тамбовский государственный технический университет
г. Тамбов, Россия

Дорожно-транспортные происшествия представляют значительные риски как для людей, так и для инфраструктуры, что требует применения современных методов моделирования для исследования причин, последствий и тенденций этих инцидентов. Моделирование дорожных происшествий является сложной задачей из-за большого объема данных, необходимых для учета таких факторов, как проблемы инфраструктуры, экологические факторы, пространственно-временные изменения, характеристики транспортных средств и водителей и многое другое. Недавние прорывы в области машинного обучения, включая графовые нейронные сети (GNN), модели глубокого обучения на основе внимания и гибридные деревья классификаторов, улучшили точность и интерпретируемость прогнозов. Эти методы используют мультимодальные данные для обеспечения предсказания аварий в реальном времени, оценки их серьезности и выявления горячих точек. Тем не менее, остаются такие препятствия, как разреженность данных, дисбаланс классов и интерпретируемость моделей. В данной статье рассматриваются современные подходы к моделированию аварий, иллюстрируя влияние разнообразной информации на частоту и тяжесть аварий для повышения безопасности дорожного движения, оптимизации управления движением и улучшения разработки политики предотвращения аварий. Будущие направления включают интеграцию мультимодальных данных, методы причинного машинного обучения и системы поддержки принятия решений в реальном времени для улучшения моделирования аварий.

Ключевые слова: моделирование, гетерогенная информация, ситуация с дорожно-транспортными происшествиями, машинное обучение.
