Traffic Anomaly Detection through Generative Modeling of Multi-Agent Interactions in Traffic Flow

Extended Abstract

Zhuojun Chen
ASTRI
Hong Kong, China
georgechen@astri.org

Tacitus Hui
ASTRI
Hong Kong, China
tacitushui@astri.org

Xinghua Zhu
ASTRI
Hong Kong, China
stellaxhzhu@astri.org

Dongzhe Su
ASTRI
Hong Kong, China
dzsu@astri.org

ABSTRACT

In intelligent transportation systems, effectively modeling and interpreting the complex interactions among autonomous agents, such as vehicles and pedestrians, is crucial for traffic management and safety. This paper introduces a novel generative traffic flow model that employs a generative pretrained Transformer to capture multi-agent interactions and detect anomalies in traffic patterns. We introduce a two-stage tokenization process for set-structured traffic data, efficiently encoding variable-sized agent states into fixed-length sequences suitable for generative modeling. We demonstrate that anomalies, which often indicate potential hazards or non-compliant behaviors, can be identified as deviations from learned normal interaction patterns among agents through a zero-shot detection mechanism. Our experimental results in simulated urban settings highlight the model's capability to detect various types of traffic anomalies with high accuracy. This work significantly advances agent-based traffic modeling and underscores its potential for enhancing traffic safety and efficiency in multi-agent systems.

KEYWORDS

Anomaly Detection, Traffic Flow Modeling, Generative AI, Intelligent Transportation Systems, Multi-Agent Systems

ACM Reference Format:
Zhuojun Chen, Tacitus Hui, Xinghua Zhu, and Dongzhe Su. 2025. Traffic Anomaly Detection through Generative Modeling of Multi-Agent Interactions in Traffic Flow: Extended Abstract. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

1 INTRODUCTION

The increasing complexity of urban traffic systems necessitates advanced modeling techniques that account for the autonomous and interactive nature of individual traffic participants. Each vehicle and pedestrian operates as an autonomous agent with its own objectives, making decisions based on personal goals and environmental perceptions. In multi-agent systems, the interactions among these agents give rise to emergent traffic patterns, which are critical to understand for effective traffic management and anomaly detection. Traditional traffic models often fall short in capturing these complex agent interactions, as they typically rely on aggregated data or simplified assumptions about agent behaviors [9].

We introduce a generative traffic flow model that incorporates multi-agent interactions and allows capturing nuanced behaviors and interaction that define urban traffic dynamics. Inspired from recent advancements in generative modeling and large language models (LLMs) in natural language processing (NLP), such as GPT [1], we adapt the remarkable capabilities in capturing sequential patterns and contexts from LLMs to model agent behaviors in traffic systems. By framing traffic data as a sequential flow of agent interactions, we leverage generative pre-training to learn normal traffic patterns and subsequently identify anomalies as deviations from these learned behaviors.

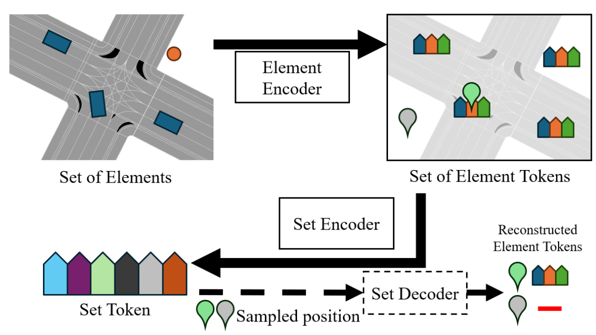

Figure 1: Two-stage tokenization process for traffic state representation. Dotted lines represent training only process.

2 GENERATIVE TRAFFIC FLOW MODEL

The core architecture of our generative traffic flow model is built upon the decoder-only Transformer model, which leverages the self-attention mechanism to capture the interactions among the tokens[10]. This architecture is well-suited for both the sequence

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

modeling task and the set tokenization task. The following works had been done to enable the implementation of the model.

Two-Stage Tokenization. The continuous and set-structured nature of traffic data necessitates a more sophisticated tokenization scheme, carried out in two stages: 1) Tokenizing the individual elements of the input set into discrete tokens; and 2) Tokenizing the entire set, where the elements have already been transformed into discrete tokens.

In Stage 2, the tokenization process involves encoding the entire set of elements, which have already been transformed into discrete finite tokens in Stage 1. This stage turns the reconstruction task into a classification problem, where the model predicts the tokens of the input set. Consequently, the reconstruction loss becomes the cross-entropy loss between the ground truth tokens and the predicted tokens.

Set Encoder. The architecture of the set encoder mirrors that of the traffic flow model but includes modifications to accommodate set-structured data. The fixed-size output representation is achieved by incorporating placeholder tokens $\mathbf{P} \in \mathbb{R}^K$ into the input set $\mathbf{I}$, forming an augmented set $\dot{\mathbf{I}}$. The augmented set $\dot{\mathbf{I}}$ is then processed by the Transformer model to produce the output tokens $\dot{\mathbf{H}}$. Only the last K tokens in $\dot{\mathbf{H}}$ (i.e. $\dot{\mathbf{H}}_{[-K:]}$) are used as the output representation of the set. The self-attention mechanism within the Transformer allows each token, including the placeholder tokens, to attend to all others in the set.

Surrogate Reconstruction. Given the variable size of the input set, the encoder $\mathcal{E}$ is commissioned to encode a fixed number (K) of latent vectors from the input set $\mathbf{I}$. To ensure the decoder $\mathcal{D}$ can accurately reconstruct or query specific elements from this fixed set of tokens, we employ a surrogate strategy that involves randomly sampling a fixed number of elements from the input set and from empty space during the training process. The encoder is then trained to minimize the discrepancy between the original input $\mathbf{I}$ and its reconstructed version, and maximize the accuracy on recognition of presence of elements at any position.

3 EXPERIMENT

We evaluate the performance of the model with anomaly detection in traffic flow, which is the identification of atypical traffic behaviors. Such anomalies may indicate potential accidents, unexpected congestion, or unlawful driving maneuvers. Previous methods have approached this challenge using image recognition techniques [7, 12, 13] and statistical approaches [2, 11].

We collect a dataset of abnormal traffic scenarios by manually introducing anomalies into the traffic environment in the CARLA simulator[4], an open-source agent-based simulator proven successful for dataset generation for autonomous driving research [3, 5, 6]. These anomalies include sudden stops, illegal lane changes, wrong-way driving, running red lights, speeding, etc. Data from these scenarios provide a labeled dataset for training a linear classifier. The dataset comprises 25 distinct abnormal traffic samples, and an equal number of normal traffic scenarios. In every traffic sample, different modalities such as agent trajectories, states of traffic signals and map data are recorded.

We evaluate the model using leave-one-out cross-validation (LOOCV) to assess its performance rigorously while having small dataset. Each scenario is sequentially excluded from the training set and used as the test set, while the remaining scenarios contribute to training the linear classifier.

Table 1: Anomaly detection results.

	Timestep level				Scenario level
Method	*AUPR-N*	*AUPR-A*	*AUROC*	*F1*	*Avg-TPR*
STGAE [11]	0.817	**0.390**	0.597	0.372	0.32
LaneGCN [8]	0.781	0.272	0.504	0.370	0.36
QCNet [14]	0.815	0.252	0.567	0.368	0.04
Ours	**0.855**	0.325	**0.643**	**0.413**	**0.92**

Embeddings are generated for each time step in the traffic scenarios by the frozen flow model pretrained with a larger dataset outside the 50 traffic samples. Only end-of-time-step token embeddings are taken. A logistic regression classifier is then fitted on these embeddings. Configured with L2 regularization with a strength of $C = 1$, the training iterates up to 2000 times to ensure convergence and accuracy in distinguishing normal from abnormal traffic states.

We evaluate the model's performance using several metrics, specifically, Scenario-level true positive rate (TPR) is used. 500 additional normal scenarios are generated as the negative class for this test. With "abnormal" as the positive class, we have a false positive rate of 0.05. The score of all the time steps in the normal/abnormal scenarios are averaged to get the average score. Anomaly is determined by whether the score of the scenario drops below a threshold, and then the TPR is determined by whether the abnormal scenario is correctly classified. The average TPR across all LOOCV iterations is reported.

Note that each validation iteration includes one abnormal and one normal scenario. Within each scenario, there are variable-length time steps. The scenario-level metric is calculated for every scenario and averaged across all scenarios. In contrast, the time step-level metrics are computed using the predictions from all time steps across all test scenarios.

We compare our model against an array of state-of-the-art anomaly detection models, including STGAE [11], LaneGCN [8], and QCNet [14]. STGAE and LaneGCN are graph-based models that leverage graph neural networks to model traffic interactions, while QCNet is a transformer-based model that takes advantage of the attention mechanism. All three models are trained on the same dataset to predict trajectories. Similarly, the final latent embeddings are used as input to a linear classifier for anomaly detection. The results are summarized in Table 1, where our model outperforms others. Remarkably, our model achieves a scenario-level true positive rate of 92%, significantly higher all other models.

REFERENCES

[1] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.
[2] Giacomo D'amicantonio, Egor Bondarau, et al. 2024. uTRAND: Unsupervised Anomaly Detection in Traffic Trajectories. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7638–7645.
[3] Jean-Emmanuel Deschaud. 2021. KITTI-CARLA: a KITTI-like dataset generated by CARLA Simulator. *arXiv preprint arXiv:2109.00892* (2021).

[4] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. 2017. CARLA: An Open Urban Driving Simulator. In *Proceedings of the 1st Annual Conference on Robot Learning*. 1–16.

[5] Yuhang Han, Zhengtao Liu, Shuo Sun, Dongen Li, Jiawei Sun, Ziye Hong, and Marcelo H Ang Jr. 2023. Carla-loc: synthetic slam dataset with full-stack sensor setup in challenging weather and dynamic environments. *arXiv preprint arXiv:2309.08909* (2023).

[6] Jaesung Jang, Hyeongyu Lee, and Jong-Chan Kim. 2022. CarFree: Hassle-Free Object Detection Dataset Generation Using Carla Autonomous Driving Simulator. *Applied Sciences* 12, 1 (2022). https://doi.org/10.3390/app12010281

[7] Shuo Li, Fang Liu, and Licheng Jiao. 2022. Self-training multi-sequence learning with transformer for weakly supervised video anomaly detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 1395–1403.

[8] Ming Liang, Bin Yang, Rui Hu, Yun Chen, Renjie Liao, Song Feng, and Raquel Urtasun. 2020. Learning lane graph representations for motion forecasting. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part II 16*. Springer, 541–556.

[9] Martin Treiber and Arne Kesting. 2013. *Traffic flow dynamics: data, models and simulation*. Springer.

[10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).

[11] Julian Wiederer, Arij Bouazizi, Marco Troina, Ulrich Kressel, and Vasileios Belagiannis. 2022. Anomaly detection in multi-agent trajectories for automated driving. In *Conference on Robot Learning*. PMLR, 1223–1233.

[12] Yan Xu, Xi Ouyang, Yu Cheng, Shining Yu, Lin Xiong, Choon-Ching Ng, Sugiri Pranata, Shengmei Shen, and Junliang Xing. 2018. Dual-mode vehicle motion pattern learning for high performance road traffic anomaly detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 145–152.

[13] Yuxiang Zhao, Wenhao Wu, Yue He, Yingying Li, Xiao Tan, and Shifeng Chen. 2021. Good practices and a strong baseline for traffic anomaly detection. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*. 3993–4001.

[14] Zikang Zhou, Jianping Wang, Yung-Hui Li, and Yu-Kai Huang. 2023. Query-Centric Trajectory Prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.