Anomaly Detection in Simulated Vehicle Dynamics Using BeamNG.tech and the TEDA Framework

Morsinaldo Medeiros[†], Matheus Andrade[†], Thaís Medeiros[†], Marianne Silva[†‡], Ivanovitch Silva[†]
[†]Federal University of Rio Grande do Norte, Natal, Brazil
[‡]Federal University of Alagoas, Penedo, Brazil
E-mails: {morsinaldo.medeiros.075, matheus.diniz.122, thais.araujo.707}@ufrn.edu.br,
marianne.silva@penedo.ufal.br and ivanovitch.silva@ufrn.br

***Abstract*—Simulations are crucial in the automotive industry for enhancing safety and cost efficiency, especially given the growing complexity of vehicle systems. This study proposes a methodology that integrates the BeamNG.tech simulation platform with a data collection system and the Typicality and Eccentricity Data Analytics (TEDA) framework to detect anomalies in vehicle speed data, facilitating driver behavior analysis. A case study was conducted where a driver navigated a predefined route while performing maneuvers with intentional speed variations. The TEDA framework successfully identified 1,110 outliers out of 6,388 speed samples, representing approximately 17.38% of the total data. The proposed methodology enables the detection of atypical driving behaviors, demonstrating its potential to contribute to developing advanced driver assistance systems and improving safety technologies.**

***Index Terms*—vehicle simulation, anomaly detection, and driver behavior.**

I. Introduction

The increasing complexity of vehicle systems has elevated the importance of simulations in the automotive industry for enhancing safety and cost efficiency [1]. Simulations provide a controlled environment to evaluate driving behaviors and identify potential risks without endangering drivers or damaging equipment [2]. They are essential tools for testing scenarios impractical or hazardous to replicate in real life [3].

Analyzing driver behavior, notably through monitoring vehicle speed, is crucial for preventing accidents and improving safety technologies such as advanced driver assistance systems (ADAS) [4], [5]. Detecting anomalies in vehicle speed can reveal unsafe driving practices, system malfunctions, or other factors that may compromise safety [6]. Traditional anomaly detection methods often struggle with the volume and variability of data generated by modern vehicles and simulations [7].

To address this challenge, we propose an approach that integrates the BeamNG.tech simulation platform [8] with a data collection system and the Typicality and Eccentricity Data Analytics (TEDA) framework [9]. BeamNG.tech[1] offers a high-fidelity vehicle simulation environment capable of producing detailed speed data under a wide range of driving conditions. The TEDA framework provides an online, unsupervised method for anomaly detection, efficiently processing streaming data to identify outliers [10].

[1]https://www.beamng.com/

In our case study, a driver operates a simulated vehicle along a predefined route within BeamNG.tech, intentionally performing maneuvers with varying speeds to introduce anomalies. The collected speed data is then analyzed using the TEDA framework. Our results demonstrate that this integrated approach effectively identifies outliers in the speed data, capturing approximately 17.38% of the total samples as anomalies. These findings indicate instances of atypical driving behavior, highlighting the potential of our method for enhancing driver behavior analysis.

This study contributes to the development of robust methodologies for anomaly detection in vehicle simulations. By facilitating the identification of unsafe driving patterns and system irregularities, our approach can aid in the advancement of safety technologies and the design of more reliable ADAS [11].

II. Related Works

Several studies have explored vehicle simulation in virtual environments, emphasizing the analysis of real-time generated data. For instance, Birchler et al. [11] utilized BeamNG.drive to collect simulated sensor data in autonomous vehicle scenarios. By applying digital twins with BeamNG.drive, they reproduced test scenarios in a controlled environment, offering an alternative to costly real-world experiments. This approach enhanced the reproducibility and comparability of results, enabling consistent studies without the need for extensive infrastructure. However, their study did not address the integration of the simulator with real-time artificial intelligence models.

In addition, Azevedo et al. [12] investigated the Simulation of Urban MObility (SUMO) simulator as a digital twin for modeling and simulating vehicular traffic in urban environments. Their study highlighted the integration of this platform with a data collection architecture that includes a specialized time-series database (TimescaleDB) and Node-RED technology for efficient data processing. Nevertheless, the research did not explore the application of real-time artificial intelligence models or integration with simulators that provide detailed data on vehicle physics.

From another perspective, Medeiros et al. [13] proposed a methodology for analyzing driver behavior by employing TinyML and edge computing for real-time data processing.

Their approach integrated sensors, the TEDA framework, and an incremental clustering algorithm to detect and classify driving patterns. Validated in a real-world case study, the methodology demonstrated its potential for accurately classifying driving behaviors. However, this approach was limited to data collected in real scenarios and did not explore the possibility of simulated environments for additional testing.

Although these studies have advanced simulator research, gaps remain. Specifically, there is a need to integrate high-fidelity simulators with real-time artificial intelligence models to enhance anomaly detection in-vehicle data. This study aims to bridge this gap by combining the BeamNG.tech simulator with a vehicular data collection platform and the TEDA framework for outlier detection, enabling real-time vehicle speed analysis and expanding existing simulators' capabilities.

III. Proposed Approach

This section presents the proposed approach, which integrates the BeamNG.tech simulator with a vehicular data collection platform and the TEDA framework for outlier detection. The overall methodology is illustrated in Figure 1.

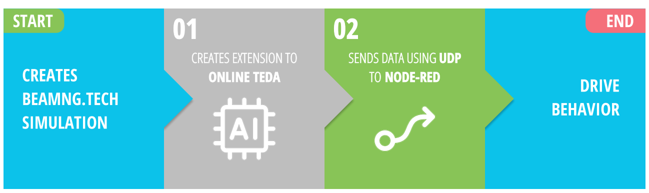

Fig. 1. Overview of proposed Approach.

A. Creating a BeamNG.tech simulation

The first step involves creating a simulation using BeamNG.tech, an expandable platform offering a wide selection of customizable scenarios, vehicles, and situations designed to meet various research and development needs [14]. Version `v0.31.3.0` of the software was utilized in this study, which allows flexibility in modifying the source code.

An extension was developed to enhance the simulator's functionality by adding an extra sensor to the vehicle's telemetry and enabling the execution of artificial intelligence models during simulation. A flat, obstacle-free environment also allowed the vehicle to accelerate and brake freely. This simplified scenario ensures that the generated data reflects the vehicle's dynamic behavior without external interference, facilitating the analysis of results.

B. Applying TEDA

In this step, the vehicle's speed sensor data are processed using the TEDA framework for anomaly detection. This algorithm leverages the concepts of typicality and eccentricity, as described in [15].

As new speed readings are acquired at discrete time instants k, the model assesses whether the current value $x_k \in \mathbb{R}$ is an outlier compared to the accumulated data distribution. The *eccentricity* $\xi_k(x_k)$ quantifies the deviation of the current sample from the rest of the dataset, while the *typicality* $\tau_k(x_k)$ evaluates the similarity of the sample to the dataset.

The eccentricity is defined as:

$$\xi_k(x_k) = \frac{1}{k} + \frac{(\mu_k - x_k)^2}{k\sigma_k^2}, \qquad k > 2, \tag{1}$$

where μ_k is the mean and σ_k^2 is the variance at time k. The typicality is given by:

$$\tau_k(x_k) = 1 - \xi_k(x_k). \tag{2}$$

To compute these metrics recursively, the mean and variance are updated as follows:

$$\mu_k = \frac{k-1}{k}\mu_{k-1} + \frac{1}{k}x_k, \qquad \mu_1 = x_1, \tag{3}$$

$$\sigma_k^2 = \frac{k-1}{k}\sigma_{k-1}^2 + \frac{1}{k}\left(x_k - \mu_{k-1}\right)^2, \qquad \sigma_1^2 = 0. \tag{4}$$

To determine whether a data point is an outlier, Chebyshev's inequality is applied:

$$\xi_k(x_k) \geq \frac{m^2+1}{k}, \tag{5}$$

where m represents the number of standard deviations from the mean μ_k, acting as a sensitivity threshold for detection. If this condition holds, the sample is classified as an outlier.

This real-time anomaly detection process efficiently assesses each new speed reading against the established data distribution, allowing immediate identification of deviations without reprocessing large datasets. This makes the solution particularly suitable for dynamic simulation environments.

C. Data Transmission

The algorithm's output adds a new column to the dataset returned by BeamNG.tech, which is then sent to the server. The transmission protocol used is `OutGauge`, which is based on data packets sent via UDP [16]. Within the simulator, the IP address and port can be configured to specify the destination for the data, facilitating integration with external systems. Leveraging this functionality, a flow was implemented in Node-RED to receive and process the data in real time, as illustrated in Figure 2.

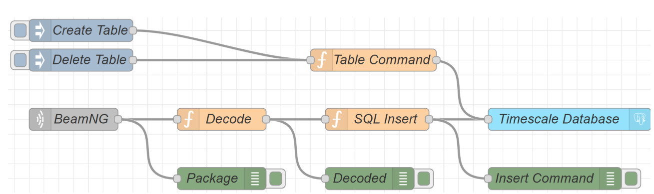

Fig. 2. Node-RED flow for real-time data processing.

The Node-RED flow consists of a UDP input node for data reception, a node for decoding packets, a node for structuring the SQL `INSERT` command, and a node responsible for connecting to the database, ensuring efficient storage in TimescaleDB.

TimescaleDB was chosen for its efficiency in handling temporal data, allowing optimized storage and querying of large volumes of real-time information [17]. Both Node-RED and TimescaleDB were deployed in Docker containers, providing greater flexibility and ease in managing the services. The Docker-based architecture allows for system scalability and ensures an isolated, reproducible environment for processing and storing data received from the simulator [18].

D. Driver Behavior Analysis

Once the data are stored in the database, they can be utilized for various analyses, such as evaluating driver behavior. For instance, if many outliers are identified in the vehicle speed time series, it can be inferred that the driver's actions involved abrupt accelerations and braking.

IV. Case Study

A case study was conducted to evaluate the proposed approach, focusing on two main research questions: (a) How does the integration of the BeamNG.tech simulator with a data collection platform and the TEDA framework enhance the detection of outliers in the vehicle's speed sensor? (b) Is it possible to perform a real-time driver behavior analysis using this integrated system?

A. Instrumentation

An extension was developed in Lua, adapting the `outgauge.lua` file to include a new column containing outlier detections. This modification enabled the real-time transmission of information to the server using the OutGauge protocol via UDP.

B. Preparation

A specific route was defined within the simulator, and the driver was instructed to operate the vehicle while intentionally incorporating speed peaks and reductions. This setup facilitated the assessment of the TEDA framework's effectiveness in identifying outliers, as illustrated in Figure 3.

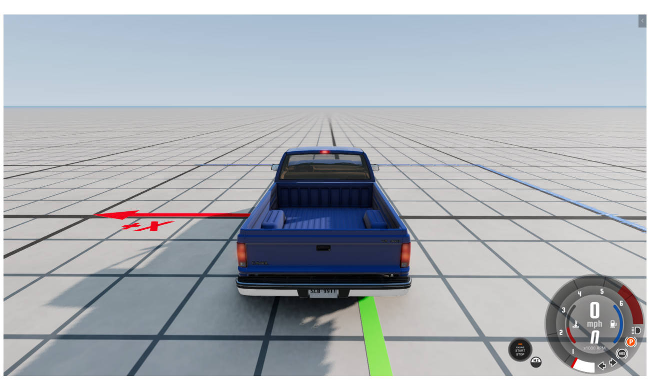

Fig. 3. Screenshot of the simulator during the route simulation.

C. Data Collection

Data were collected along the predefined route, during which various maneuvers were executed to create significant variations in speed. The data collection focused on:

1) Instantaneous speed of the vehicle.
2) Driver actions, including accelerations, decelerations, and sudden braking.
3) Patterns of driver behavior in response to different simulated scenarios.

D. Data Analysis

Analyses were conducted based on the outlier detections throughout the speed time series to identify moments of atypical driver behavior. These behaviors serve as indicators of risks or critical events during driving. The analysis confirmed that the integrated system effectively detects and analyzes anomalous driving patterns, providing a deeper understanding of the dynamics involved in driver behavior.

V. Results and Discussion

This section presents the results obtained from the case study. The analysis was based on telemetry data collected during the simulation, comprising a total of 6,388 vehicle speed samples. The TEDA algorithm identified 1,110 of these samples as outliers, representing approximately 17.38% of the total.

Figure 4 illustrates the time series of vehicle speed, with outliers highlighted in red. Excluding the outliers identified by TEDA, the average speed was 26.68 km/h, with a standard deviation of 20.63 km/h. The high variability indicated by the standard deviation reflects significant fluctuations in speed, evidenced by the peaks throughout the series. These detected outliers correspond to moments of sudden acceleration and deceleration, which may indicate atypical driver behavior or critical situations in vehicle control.

Addressing the first research question regarding the enhancement of outlier detection through the integration of the BeamNG.tech simulator, data collection platform, and the TEDA algorithm, the results demonstrate that the TEDA algorithm effectively identified speed peaks and abrupt variations. This indicates a significant improvement in outlier detection compared to traditional methods. The accurate classification of these speed anomalies as outliers showcases the algorithm's capability to detect events deviating from typical patterns.

Regarding the second research question on the feasibility of real-time driver behavior analysis, the results confirm that real-time processing, combined with the simulator-to-server communication via the OutGauge protocol, enabled effective analysis of driver behavior. The integration with Node-RED and storage in the TimescaleDB database ensured efficient organization and accessibility of the data for future analyses. These findings validate that the proposed approach successfully addresses both research questions, confirming the effectiveness of integrating the simulation environment with advanced data analysis capabilities.

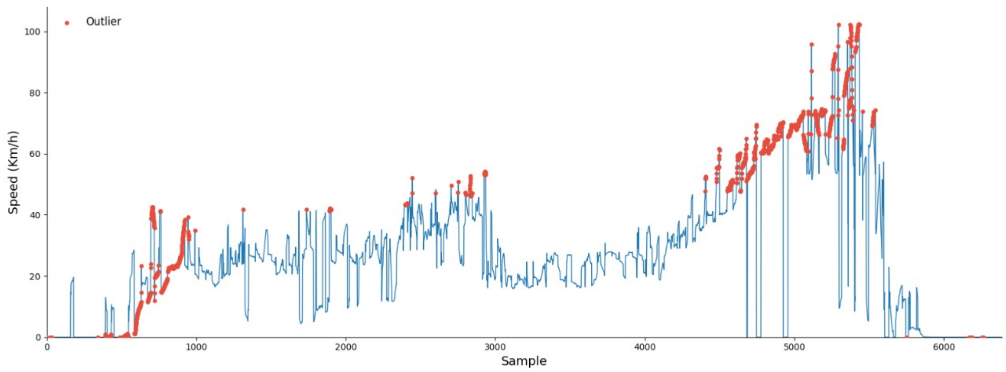

Fig. 4. Time series of speed with detected outliers highlighted in red.

VI. Conclusion

This work presented an effective approach for detecting outliers in vehicle speed time series by integrating the TEDA algorithm with the BeamNG.tech simulator and a data collection platform. The implementation enabled real-time identification of anomalous patterns during driving simulations, demonstrating the system's capability to detect atypical driving behaviors that may indicate critical events or risks, thus contributing to the development of automotive safety systems.

Future work includes developing soft sensors, such as radar chart area calculations, to classify driver behavior based on vehicle resource utilization. Additionally, integrating machine learning models—such as decision trees, deep neural networks, and Kolmogorov-Arnold Networks (KAN) — is proposed to enhance behavior classification and predict potential driving risks.

Acknowledgment

The research presented in this paper was financially supported by the Brazilian National Council for Scientific and Technological Development (CNPq) under Grant No. 405531/2022-2, the Research Development Foundation—FUNDEP/Mover/Line VI under Grant No. 29271.01.01/2023.03-00, and BeamNG through a research license.

References

[1] J. Zhou, Y. Zhang, S. Guo, and Y. Guo, "A survey on autonomous driving system simulators," in *2022 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*. IEEE, 2022, pp. 301–306.

[2] F. Khan, H. Anwar, and D. Pfahl, "Simulation-based safety testing of automated driving systems," in *International Conference on Product-Focused Software Process Improvement*. Springer, 2023, pp. 133–138.

[3] R. E. Amini, E. Michelaraki, C. Katrakazas, C. Al Haddad, B. De Vos, A. Cuenen, G. Yannis, T. Brijs, and C. Antoniou, "Risk scenario designs for driving simulator experiments," in *2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*. IEEE, 2021, pp. 1–6.

[4] E. Castellano, A. Cetinkaya, and P. Arcaini, "Analysis of road representations in search-based testing of autonomous driving systems," in *2021 IEEE 21st International Conference on Software Quality, Reliability and Security (QRS)*. IEEE, 2021, pp. 167–178.

[5] T. Zhang, H. Liu, W. Wang, and X. Wang, "Virtual tools for testing autonomous driving: A survey and benchmark of simulators, datasets, and competitions," *Electronics*, vol. 13, no. 17, p. 3486, 2024.

[6] D. Garikapati and S. S. Shetiya, "Autonomous vehicles: Evolution of artificial intelligence and the current industry landscape," *Big Data and Cognitive Computing*, vol. 8, no. 4, p. 42, 2024.

[7] N. Fouladinejad, M. K. A. Jalil, and J. M. Taib, "Reduction of computational cost in driving simulation subsystems using approximation techniques," in *2014 International Conference on Industrial Automation, Information and Communications Technology*. IEEE, 2014, pp. 111–117.

[8] C. Birchler, N. Ganz, S. Khatiri, A. Gambi, and S. Panichella, "Cost-effective simulation-based test selection in self-driving cars software with sdc-scissor," in *2022 IEEE international conference on software analysis, evolution and reengineering (SANER)*. IEEE, 2022, pp. 164–168.

[9] P. Angelov, "Anomaly detection based on eccentricity analysis," 12 2014, pp. 1–8.

[10] M. Silva, T. Flores, P. Andrade, J. Silva, I. Silva, and D. G. Costa, "An online unsupervised machine learning approach to detect driving related events," in *IECON 2022–48th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2022, pp. 1–6.

[11] C. Birchler, C. Rohrbach, T. Kehrer, and S. Panichella, "Sensodat: Simulation-based sensor dataset of self-driving cars," in *2024 IEEE/ACM 21st International Conference on Mining Software Repositories (MSR)*. IEEE, 2024, pp. 510–514.

[12] M. Azevedo, M. Andrade, M. Medeiros, T. Medeiros, M. Silva, I. Silva, E. Sisinni, and P. Ferrari, "Optimizing vehicle iot systems: Sumo-digital twin performance analysis," in *2024 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0 & IoT)*. IEEE, 2024, pp. 204–209.

[13] M. Medeiros, T. Flores, M. Silva, and I. Silva, "A multi-layered methodology for driver behavior analysis using tinyml and edge computing," in *2024 IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)*. IEEE, 2024, pp. 1–8.

[14] M. Biagiola and S. Klikovits, "Sbft tool competition 2024-cyber-physical systems track," in *Proceedings of the 17th ACM/IEEE International Workshop on Search-Based and Fuzz Testing*, 2024, pp. 33–36.

[15] P. Andrade, I. Silva, M. Silva, T. Flores, J. Cassiano, and D. G. Costa, "A tinyml soft-sensor approach for low-cost detection and monitoring of vehicular emissions," *Sensors*, vol. 22, no. 10, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/10/3838

[16] B. Team, "Outgauge support — beamng documentation," https://documentation.beamng.com/tech/integration/outgauge.html, 2023, accessed: 2024-10-11.

[17] M. Medeiros, M. Azevedo, M. Amaral, C. M. Viegas, M. Silva, I. Silva, and D. G. Costa, "A framework for efficient communication in iot applications in the automotive industry with obd-ii edge," in *2023 Symposium on Internet of Things (SIoT)*. IEEE, 2023, pp. 1–5.

[18] A. García, X. Oregui, J. Franco, and U. Arrieta, "Edge containerized architecture for manufacturing process time series data monitoring and visualization." in *IN4PL*, 2022, pp. 145–152.