electronics

Article

Anomaly Detection in Connected and Autonomous Vehicle Trajectories Using LSTM Autoencoder and Gaussian Mixture Model

Boyu Wang [1], Wan Li [2,*] and Zulqarnain H. Khattak [3]

[1] Tacoma Public Utilities, Tacoma, WA 98409, USA; bwang330@ieee.org
[2] Oak Ridge National Laboratory, Oak Ridge, TN 37932, USA
[3] Systems Scientist, Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA; zkhattak@cmu.edu
* Correspondence: liw2@ornl.gov; Tel.: +1-865-341-1353

Abstract: Connected and Autonomous Vehicles (CAVs) technology has the potential to transform the transportation system. Although these new technologies have many advantages, the implementation raises significant concerns regarding safety, security, and privacy. Anomalies in sensor data caused by errors or cyberattacks can cause severe accidents. To address the issue, this study proposed an innovative anomaly detection algorithm, namely the LSTM Autoencoder with Gaussian Mixture Model (LAGMM). This model supports anomalous CAV trajectory detection in the real-time leveraging communication capabilities of CAV sensors. The LSTM Autoencoder is applied to generate low-rank representations and reconstruct errors for each input data point, while the Gaussian Mixture Model (GMM) is employed for its strength in density estimation. The proposed model was jointly optimized for the LSTM Autoencoder and GMM simultaneously. The study utilizes realistic CAV data from a platooning experiment conducted for Cooperative Automated Research Mobility Applications (CARMAs). The experiment findings indicate that the proposed LAGMM approach enhances detection accuracy by 3% and precision by 6.4% compared to the existing state-of-the-art methods, suggesting a significant improvement in the field.

Keywords: cybersecurity; anomaly detection; falsified trajectories; CAVs; LSTM; Gaussian Mixture Model

Citation: Wang, B.; Li, W.; Khattak, Z.H. Anomaly Detection in Connected and Autonomous Vehicle Trajectories Using LSTM Autoencoder and Gaussian Mixture Model. *Electronics* **2024**, *13*, 1251. https://doi.org/10.3390/electronics13071251

Academic Editors: Nikolay Hinov, Jožef Ritonja and Darius Andriukaitis

Received: 31 January 2024
Revised: 25 March 2024
Accepted: 26 March 2024
Published: 28 March 2024

1. Introduction

Connected and Antonymous Vehicles (CAVs) have the potential to transform the transportation system. CAVs can continuously share information, such as vehicle speed, location, and acceleration/deceleration to their surrounding vehicles or infrastructure by taking advantage of Vehicle-to-Vehicle (V2V) and the Vehicle-to-Infrastructure (V2I) communication. These technologies are expected to help alleviate traffic congestion and improve traffic efficiency and safety. Although these new technologies provide several advantages, safety, security, and privacy are the major concerns for real-world applications [1]. CAVs heavily rely on data from sensors. Anomalous sensor readings caused by errors or cyberattacks can have severe consequences, including crashes and traffic delays. There are different types of internal and external cyberattacks in CAV systems. Khattak et al. [2] identified three types of cyberattacks in CAVs environments: (i) attack points that are compromised in V2V communication at the application level, e.g., message falsification attack, shutdown attack, and reply attack; (ii) attack points that are compromised in V2I communication at the network level, e.g., spoofing attach, denial of service attack, and radio jamming attack; and (iii) attacks that occur in CAVs themselves by gaining physical access to the CAVs. More details and examples of different attacks can be found in [2–4].

Cybersecurity is one of the most significant concerns in the CAV applications. The threat of cyberattacks on CAVs poses a significant risk not only to the travelers within these

vehicles but also endangers pedestrians, cyclists, and other road users who interact with them. These attacks can lead to many safety issues, from loss of vehicle control to breaches of data privacy, showing the urgent need for robust anomaly detection methodologies. To address these challenges effectively, it is imperative to develop sophisticated anomaly detection systems capable of identifying and mitigating such threats in real time. These systems need to analyze vehicle behavior patterns continuously and detect deviations that may indicate a cyberattack or system failure. Most of the existing studies have relied on improving the security of the communication channels or sensors within the CAV network to avoid the type (i) attacks [5–7]. Yet, type (iii) attacks are significantly more threatening to the system. For instance, the hacker can gain access to these attack points and take control of CAVs. They can manipulate advisories of V2I lane control, making the target CAV accelerate when another CAV merges to the same lane in front of the target CAV. These attacks can disrupt key functionalities of CAVs, e.g., acceleration, vehicle speed, position, braking, lane changing, and ramp metering behaviors. It will not only put the target CAVs in danger but also disrupt the traffic flows by creating more congestion and causing accidents. Guo et al. [8] introduced a non-linear control scheme to handle the dynamic uncertainties of platoons. The methodology can also be employed to gain back the control of the attacked CAVs and maintain the stability of the network. Unfortunately, it cannot manage large disturbances and operate under certain assumptions regarding uncertainties. Some researchers have contributed by using CAV bus data from a normal human-driven vehicle to develop machine leaning-based anomaly detection algorithms while neglecting the temporal relationship between anomalous trajectories. Neglecting these temporal correlations can result in models that fail to recognize complex, time-dependent circumstances, which compromise the system's performance to address potential threats. Therefore, designing an efficient and effective anomaly detection algorithm for CAVs is significant to identify anomalous behavior in real time, thereby preventing the severe consequences of cyberattacks or sensor anomalies.

This study counters these limitations and focuses on the detection of the attacks by utilizing realistic CAV data from a platooning experiment. In this paper, we proposed an anomaly detection model, namely Long Short-term Memory Autoencoder Gaussian Mixture Model (LAGMM), to detect anomalous behavior in a platoon of CAVs. An LSTM Autoencoder was applied to generate low-rank representations and reconstruction the error for each input data point. The Gaussian Mixture Model (GMM) was leveraged to deal with density estimation tasks. The proposed model was jointly optimized for the LSTM Autoencoder and GMM simultaneously. There are generally two types of behaviors/CAV trajectories: normal operations with vehicle sensor data and falsified trajectories caused by cyberattacks (different types of cyberattacks are considered in this study). The study utilizes realistic CAV data from a platooning experiment conducted for Cooperative Automated Research Mobility Applications (CARMAs). Details of the experiment and data are provided in Section 3. This study has the following contributions:

1 This research employed data from an actual CAV platooning test to simulate anomalies and identify falsified activities in vehicle platoons containing leading and following vehicles.
2 We created a real-time anomalous CAV trajectory detection algorithm LAGMM. This model considers the temporal dynamics of CAV trajectories, which is often overlooked in previous research. It enhances detection accuracy compared to the existing techniques.
3 LAGMM applied a two-step strategy of decomposition and density estimation, which were optimized concurrently. It helps avoid local sub-optimal solutions and further reduce reconstruction errors.
4 Different types of cyberattacks in CAV systems were investigated and validated with the proposed model. The proposed model demonstrates its effectiveness in detecting the various attacks on the influences of CAV operations.

This work is organized as follows: the recent studies and literature are reviewed to send a clear picture to the audience in the second section. The experiment and data are organized and explained in Section 3. Anomalies emulation is introduced in the following Section 4. The formulation and improvements of the existing study are introduced in Section 5 as well as the model topology. Lastly, the experiments are analyzed, and the final conclusion is drawn in the Sections 6 and 7, respectively.

2. Literature Review

There are some existing studies utilizing deep learning models for anomaly detection. Most of the studies aim at detecting falsified Autonomous Vehicle (AV) trajectories in an offline manner. That is, after collecting AV data (normal and falsified trajectories), different types of supervised and unsupervised learning methods are applied to classified the trajectories. Van Wyk et al. [9] developed a Convolutional Neural Network (CNN) with Kalman filtering to detect and identify the anomalous behaviors of CAVs. Their experiment results demonstrated that the proposed model can achieve high accuracy, sensitivity, and F1 score. Javed et al. [4] proposed a multi-stage attention mechanism with a Long Short-Term Memory (LSTM)-based CNN model. The method achieved the gain of up to 3.24% in F-score. Huang et al. [3] developed a trajectory-embedding model from a natural language processing (NLP) community. The model generated vector representation of CAVs trajectories to compute the similarities between trajectories. The algorithm first trained a neural network to obtain vector representation of trajectories. Then, a hierarchical clustering algorithm was applied to estimate the distance matrix between each pair of trajectories and identify falsified trajectories. The proposed method could achieve a high detection rate (>97%). Wang et al. [10] proposed an Adaptive Extended Kalman Filter (AEDF) method to detect anomalies. This study focused on the platoon movements instead of that of a single vehicle. They used the speed and location data of the surrounding traffic of the target vehicle in anomaly detection. Kamel et al. [11] created a misbehavior detection algorithm (MDA) by utilizing data from a simulated environment in SUMO. They produced compromised data through six different attack methods and applied machine learning techniques for detecting misbehavior. Dong et al. [12] conducted simulated attacks on cooperative adaptive cruise control platoons to study their impact on traffic flow and safety. They found that an escalation in both the frequency and severity of these attacks on targeted vehicles resulted in a negative influence on traffic and increased collision risks. Yen et al. [13] analyzed the performance different back pressure-based traffic signal control algorithms such as delay-based and queue-based under cyberattacks. They observed that delay-based algorithms are vulnerable to spoofing attacks where a vehicle's arrival time is altered. Singh et al. [14] studied the impact of cyberattacks on a platoon of ten cooperative adaptive cruise control vehicles. They found a severe variation in acceleration profiles and the stability of the platoon. Nguyen et al. [15] developed a misbehavior detection algorithm by tracking host vehicle signals and verification of industry consensus Roadside Units (RSU) in a V2X (Vehicle-to-Everything) environment. So et al. [16] analyzed location spoofing using plausibility checks by a k-neural network and SVM to classify misbehavior using a vehicular reference misbehavior dataset simulated through VEINS. They observed their algorithm to improve the plausibility checks by 20%. Another study [17] analyzed a Cooperative Adaptive Cruise Control (CACC) platoon of six vehicles against time-delay attacks and observed their CACC algorithm to be stable against attacks with no cruise or jerks.

Some studies have qualitatively assessed the cyber risks. Bertini et al. [18] carried out a survey among staff members of the Oregon Department of Transportation (DOT) to assess their preparedness for the deployment of CAV systems. Their findings indicated that approximately 39–40% of the survey participants expressed concerns regarding the security risks associated with CAVs. Bhavsar et al. [19] investigated the potential failure from autonomous vehicle malfunctions within a mixed traffic flow. By employing fault tree analysis to estimate the likelihood of failure for each autonomous element, the research

found a 14% rate of failure for the components within autonomous vehicles. Hasan et al. [20] conducted a comprehensive survey of the V2X system, examining security measures, standards, and current defense strategies. Additionally, they pinpointed existing shortcomings in security solutions and highlighted unresolved problems.

Several research efforts have employed quantitative approaches to evaluate the effects of cyberattacks on CAVs. Amoozadeh et al. [21] employed OMNET++ for the analysis of ten CACC vehicles in a single-lane scenario, focusing on the consequences of jamming and message falsification attacks. During message falsification, adversaries modified the desired acceleration settings, increasing the instability across the vehicle fleet. Similarly, when subjected to radio jamming, the vehicles reverted to adaptive cruise control mode, adopting larger time gaps between each other. Islam et al. [22] explored cyberattack identification and mitigation using the CVGuard framework. CVGuard was observed to reduce the conflicts when the conflicts before and after the activation of CVGuard were compared. Another study [23] employed microsimulation to examine a fleet of ten CACC vehicles in a single-lane scenario subject to cyberattacks. Jamming emerged as a significant cyber threat, leading to fluctuations in speed and collisions. Khattak et al. [24] investigated the cybersecurity vulnerabilities within the communication channel of an Active Traffic Management (ATM) system. They created a prototype for a threat monitoring system designed to restore the compromised ATM system to normal functionality following cyberattacks. Li et al. [25] investigated the effects of cyberattacks on a single vehicle on the longitudinal safety of CAVs. The attack remained active for a brief period. The positions and speeds from leading vehicles were leveraged as elements of the attack, indicating that a minor cyberattack significantly affects the acceleration profiles more than the deceleration profiles of nine vehicles. Wardzinski [26] proposed a model for autonomous vehicle control systems, which assesses the risk of present and anticipated scenarios to create a vehicle control strategy. This approach maintains the risk level within the bounds of the minimally acceptable safety risk. It concluded that enhanced performance and safety are attainable through V2V communication and collaboration. Wang et al. [27] evaluated the impact of cyberattacks on a single platoon moving in a single lane, treating cyberattacks as similar to the dissemination of malicious information. The observation revealed that these cyberattacks notably disrupted the flow of traffic.

Recent work has explored using machine learning methods and deep neural networks like convolutional and recurrent networks for effective feature learning from time-series data. Most statistical anomaly detection methods struggle to model volatility and long-term trends in time-series data. This is because they often depend on methods that work better on simpler datasets and fail to capture the fluctuations of variables with little correlations in the system [28–30]. In recent years, deep neural networks have become very popular and are used in many advanced models. Ma et al. [31] uses a Bi-Transformer architecture to extract features from two dimensions in parallel. The model also incorporates an adaptive multi-head attention mechanism to capture relationships between different time-series variables. These contributions build on established techniques like Transformer networks and multi-head self-attention, adapting them to the multivariate time-series anomaly detection task. Wang et al. [32] proposed a framework that combines an Extended Kalman Filter (EKF) to denoise sensor readings and Support Vector Machine (SVM) to identify attacks. Notably, the EKF model incorporates heterogeneous time delays and augmented states to improve accuracy under stochastic delays and model errors. The authors also conduct a theoretical analysis relating the detection rate to "pseudo-string stability", introducing this new concept to capture platooning stability under model uncertainty. Yang et al. [33] presents an anomaly detection framework to identify attacks against CAVs by learning normal driving behaviors. This work adopts inverse reinforcement learning (IRL) to learn an optimal policy from demonstrated trajectories and then detects anomalies by comparing observed behavior to policy-based predictions. The key innovations of this work are using maximum entropy IRL to learn driving behaviors from a small number of trajectory exam-

ples and creating measures of difference between the predicted and actual driving paths to feed into a classifier that detects attacks.

The current stage of research in the field has several limitations. Firstly, due to the unavailability of real-world CAV data, most studies have used Controller Area Network (CAN) bus data from conventional human-driven vehicles to simulate cyberattacks. As a result, there is a gap in the literature regarding the assessment of cyber risks and the detection of anomalous behaviors in real CAV environments. Additionally, a significant portion of these studies have applied offline models for anomaly detection, further highlighting the need for advancements in real-time online methodologies. The existing studies suffer from heavy computational burdens, making it challenging to implement in real time. Second, existing studies did not differentiate various types of the attacks, i.e., faulty CAV sensor behaviors. For example, graduate drift sensor failure is different from bias sensor failure [9,10]. The proposed model is expected to detect different types of attacks. Third, existing studies that utilize conventional machine learning or more advanced deep learning models did not account for the temporal relations of the falsified CAV trajectory data, which can have a significant impact on anomaly prediction. Traffic flows and vehicle trajectories are continuously and dynamically updated over time. The temporal domain of speed and location should be considered in the model.

3. Data

The Federal Highway Administration (FHWA) conducted numerous field tests in partnership with the Volpe Center to collect real-world data. In [34], the researcher aimed at leveraging the data for demonstrating a platooning concept utilizing CACC and adaptive cruise control (ACC). The field tests were carried out at the Aberdeen Center in Maryland on a 4.5-mile track. The track is intended to imitate the geometry of a standard US highway. It was designed with geo-locations which are called waypoints to send signals, e.g., target speeds to the testing vehicles. A platoon of five Cadillac SRX, including leading vehicles (LVs) and following vehicles (FVs), is installed with CACC controllers and tested for various configurations. Once the waypoints send the signal to LVs, LVs utilized this information along with GPS data to fine-tune their CACC settings. The fleet will navigate from the first waypoint to the last to be counted as one complete test run. Figure 1 illustrates the layout of the test track and identifies the waypoints used during the platooning testing [34]. According to the study in [35], the speed and acceleration data are pre-processed. For example, speed profiles are smoothed with the moving average technique, while acceleration data are derived from the smoothed speed data, which effectively minimized discrepancies between acceleration values calculated from raw versus smoothed speed data.

Figure 1. Test track for field experiments.

The cleaned CAV trajectory data from CARMAs [11] were used to develop anomaly detection algorithm. Each CAV trajectory consists of four features every 0.5 s: speed of the CAV, acceleration of the CAV, average speed of non-CAV in the same vehicle platoon, and steering wheel angle of the CAV. Each CAV trajectory lasts for 10 s.

4. Cyberattack Emulation

Due to the unavailability of publicly accessible anomalous CAV sensor data, this work simulated the CAV sensor anomalies with real CAV trajectory data. The existing literature proved that these sensors are vulnerable to cyberattacks and sensor malfunctions [36–38]. We consider three types of attacks that could result in four sensor anomalies. For example, an attack involving the injection of misleading data through the CAN bus or the on-board diagnostic (OBD) system could affect in-vehicle speed and acceleration. Similarly, an adversary possessing legitimate credentials could tamper sensor readings by engaging in GPS jamming or spoofing attacks, thereby inducing anomalies. Additionally, an acoustic injection can undermine the integrity of the acceleration sensor, leading to additional irregularities. This study is focused on detecting between normal and abnormal CAV trajectories. In light of the attack scenarios described, abnormal trajectories were created, incorporating four types of anomalies across four scenarios, aligning with findings from previous research [10]:

i Short anomaly is characterized by an abrupt change in the recorded CAV trajectory data. A random Gaussian variable having a zero mean and variance of 0.001 was used to simulate the anomaly. This was scaled by $N \in (0, 0.01)$ to capture the anomaly magnitude, where N belongs to 25, 100, 1000, and 1000. The value was added to the sensor base value.

ii Noise refers to a longer-lasting shift (spanning several consecutive readings) in the variability of the tracked CAV trajectory data. The anomaly was simulated as an independent and identically distributed sequence of the random Gaussian variable with a mean of zero, length of l and variance of c.;

iii Bias is a deviation from the actual sensor measurements. This was simulated as a temporary offset from the normal readings and captured for various magnitudes using a random variable. A uniform distribution was used to sample the anomaly magnitude. The anomalous readings are generated by adding the anomaly magnitude to the true sensor readings for various durations $d \in (25, 50, 100, 1000)$.

iv Gradual drift refers to a slow and steady change in the data observed over time. This anomaly was simulated by offsetting the base values with linearly increasing values: for instance, using a linearly increasing speed of 0–5 mph denoted by $c \in (3, 5)$ using a function linspace (0, c). The anomaly was simulated for various duration.

Anomalies were intentionally introduced into the sensors of CAVs. It was assumed that the occurrence of sensor anomalies, whether due to cyberattacks or sensor malfunctions, happened independently to simplify the process. Consequently, the predictive models were designed without accounting for the possibility of correlated sensor errors. The assumption was that at any given time epoch, only a single anomaly would occur, reflecting the independent nature of attacks or sensor faults and the inherent reliability of the sensor systems. These anomalies were randomly generated and applied to various sensors in a stochastic manner. The simulated anomalies were then added on the baseline or normal readings of the affected sensor, whether in the LVs or FVs. The nature and duration of these anomalies varied. For example, some anomalies were simulated to last for 5 min, while others lasted for 20 min. Additionally, a mixed-anomaly scenario was also simulated to assess the sensitivity of the detection mechanisms.

5. Anomaly Detection Modeling Approach

5.1. Proposed Model: LAGMM

As shown in Figure 2, the proposed model is composed of two primary elements. (1) The first is a compression network designed to create a low-rank approximation, denoted as

z, of the input data using an LSTM Autoencoder. This process involves combining reduced space features z_c with features representing reconstruction errors z_r. (2) The second is a GMM model aiming at predicting likelihood or energy levels.

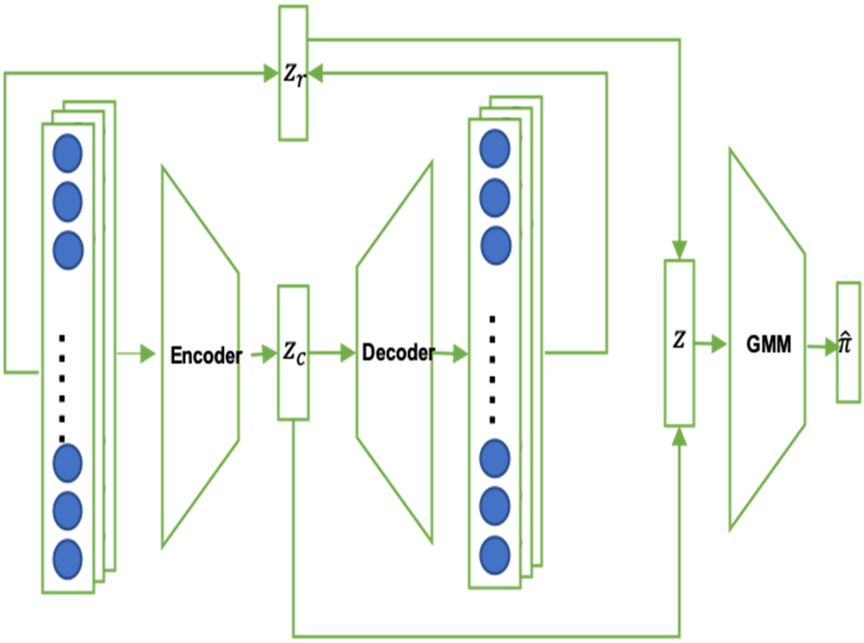

Figure 2. An overview of the LSTM Autoencoder Gaussian Mixture Model. The blue bullets in each layer indicate individual neurons that are processing and passing on information within the network.

Given an input sample x, the LSTM Autoencoder computes the low-dimensional representation z by Equations (1)–(4):

$$z_c = h(x; \theta_e) \tag{1}$$

$$x^0 = g(z_c; \theta_d) \tag{2}$$

$$z_r = f(x; x^0) \tag{3}$$

$$z = [z_c, z_r] \tag{4}$$

where z_c is the low-rank approximation learnt by the LSTM Autoencoder, z_r represents the features obtained from the reconstruction error, θ_e and θ_d signify the parameters of the LSTM Autoencoder, and x^0 is the reconstruction of x, with $h(\cdot)$ and $g(\cdot)$ indicating the encoding and decoding functions, respectively. Moreover, $f(\cdot)$ is the function used to compute the features associated with the reconstruction error.

Given the low rank approximation of the input data, the GMM-based estimation network aims at estimating the density function. GMMs are usually employed to model class distributions by learning parameters for each class during training. When classifying new data points, GMMs calculate the likelihood of belonging to each class, allowing to select the class with the highest likelihood. Additionally, GMMs can identify data points that substantially deviate from the learned normal distribution. In the GMM, the unknown parameters include the distribution of mixture components ϕ, the means of the mixtures μ and the covariance of the mixtures Σ. A multi-layer neural network (MLNN) was utilized to determine the mixture membership for each sample of data, as shown in Equations (5) and (6):

$$p = MLNN[z, \theta_m] \tag{5}$$

$$\hat{\gamma} = softmax(p) \tag{6}$$

where $\hat{\gamma}$ represents a vector indicating the predicted soft membership in mixture components, and p is the output from the MLNN. With N and the predicted memberships, the parameters of the GMM can be further estimated as follows:

$$\hat{\phi}_k = \sum_{i=1}^{N} \frac{\hat{\gamma}_{ik}}{N} \tag{7}$$

$$\hat{\mu}_k = \frac{\sum_{i=1}^{N} \hat{\gamma}_{ik} z_i}{\sum_{i=1}^{N} \hat{\gamma}_{ik}} \tag{8}$$

$$\hat{\Sigma}_k = \frac{\sum_{i=1}^{N} \hat{\gamma}_{ik} (z_k - \hat{\mu}_j)(z_k - \hat{\mu}_j)^T}{\sum_{i=1}^{N} \hat{\gamma}_{ik}} \tag{9}$$

where $\hat{\phi}_k$, $\hat{\mu}_k$, and $\hat{\Sigma}_k$ are the mixture probability, mean, and covariance for component k in the GMM. The sample energy can be estimated by Equation (10):

$$E(z) = -log(\sum_{k=1}^{K} \hat{\phi}_k \frac{exp(-\frac{1}{2}(z - \hat{\mu}_k)T)\hat{\Sigma}_k^{-1}(z - \hat{\mu}_k)}{\sqrt{|2\hat{\pi}_k|}}) \tag{10}$$

During the testing phase, energy levels are utilized to determine whether the sample data contain falsified trajectories. An increased energy level suggests a greater likelihood of anomalies.

5.2. Objective Function

The objective function of the LAGMM is shown in Equation (11). $L(x_i, x_i^{'})$ represents the loss function that quantifies the reconstruction error generated by the LSTM Autoencoder, which can be defined by the L2-norm. $E(z_i)$ denotes the probabilities that the input samples could be observed. Minimizing J aims at avoiding the singularity problem by penalizing small values of the diagonal entries.

$$J(\theta_e, \theta_d, \theta_m) = \frac{1}{N} \sum_{i=1}^{N} L(x_i, x_i^{'}) + \frac{\lambda_1}{N} \sum_{i=1}^{N} E(z_i) + \lambda_2 P(\hat{\Sigma}) \tag{11}$$

5.3. Improvements from the Existing Study

To start with, we tried an LSTM-based network as a benchmark. LSTMs were trained to recognize normal patterns of behavior. At each time step, predictions are made and the errors in these predictions signal deviations from what is considered normal. Then, a clustering method was employed to identify anomalies. It turns out that the performance was not good, especially considering multiple types of cyberattacks. Hence, we have proposed this new LAGMM for anomaly detection. Based on the study from Purohit [39], this study has two improvements: (i) accounting for the temporal relations of CAVs trajectory data by using a LSTM Autoencoder, and (ii) testing the model was against multiple types of cyberattacks. Specifically, we have implemented the LSTM Autoencoder instead of the multi-layer neural network used in the original paper [39].

6. Experiments and Results

6.1. LAGMM Configuration and Hyperparameters

For the CARMA dataset, the LSTM Autoencoder feeds a three-dimensional input into the estimation network, consisting of one reduced dimension and two dimensions derived from the reconstruction error. In particular, the LAGMM runs with an LSTM layer with dimensions ((20, 4), 128, tanh) and eight Fully Connected (FC) layers with dimensions (128, 64, tanh), FC (64, 32, tanh), FC (32, 16, tanh), FC (16, 1, none), FC (1, 16, tanh), FC (16, 32, tanh), FC (32, 64, tanh), and FC (64, 128, none). The estimation network performs with an FC (3, 10, tanh), a Dropout layer (0.5), and an FC layer (10, 4, softmax). Multiple

combinations of LSTM and Deep Autoencoder Gaussian Mixture Model (DAGMM) are also tested for our model. We have tuned the multiple hyperparameters in this study, including LSTM layers, FC layers, dropout rate, and learning rate in the Adam optimizer. The configuration is the best model we achieved so far.

6.2. Network Performance and Metrics

We consider precision, F1 score and accuracy to compare anomaly detection performance. A threshold is chosen to identify anomalous samples. For example, in the case of the LAGMM operating on the CARMA dataset, samples within the highest energy percentiles are classified as falsified trajectories. We categorize the anomaly class as positive and accordingly define metrics such as precision, F1 score, and accuracy. Three different types of attacks are mixed up to create a more realistic environment, and no specific time is set for its happening. Figure 3 illustrates the histogram of energy computed from the LAGMM model. Larger values of energy indicate a higher probability of falsified CAV trajectories. As shown in the figure, most of the samples have negative energy, while only a small portion of samples have energy higher than 5 at the horizontal axis. This graph can provide a clear picture on the dataset with a probability of cyberattacks for CAV operators.

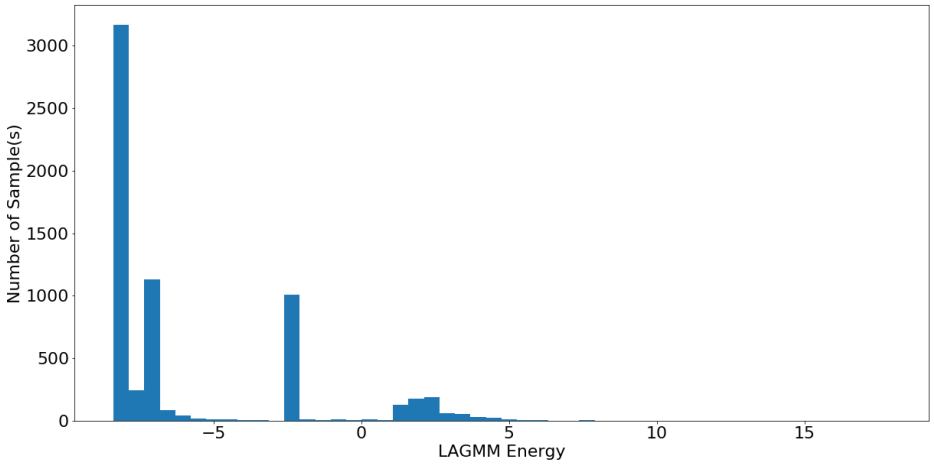

Figure 3. Histogram of LAGMM energy.

Figure 4 shows the LAGMM energy of all samples. It suggests that a small portion of samples have larger energy, e.g., larger than 10, which indicates a higher probability of CAV falsified trajectories or driving patterns.

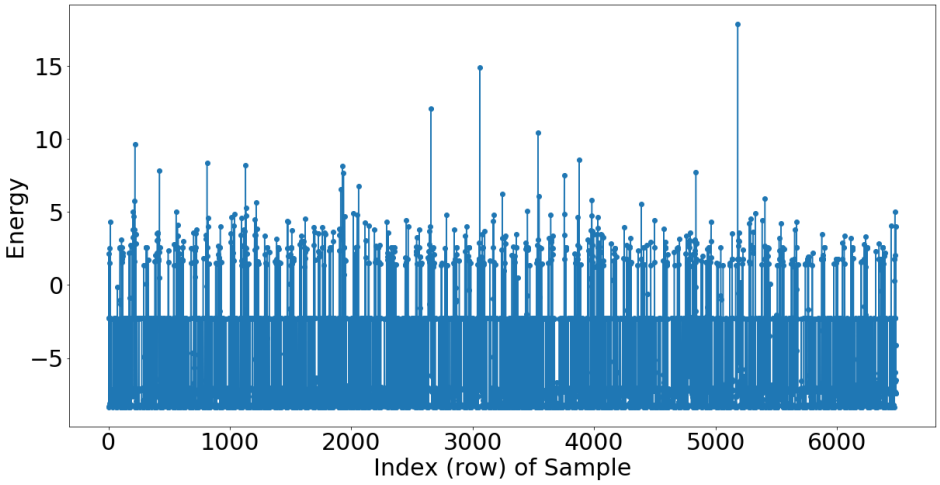

Figure 4. LAGMM energy of all samples.

Figure 5 illustrates the normal CAV trajectories (blue dots) and falsified CAV trajectories (red cross marks) in terms of each combination of two features out of a total of four features (four rows and four columns).

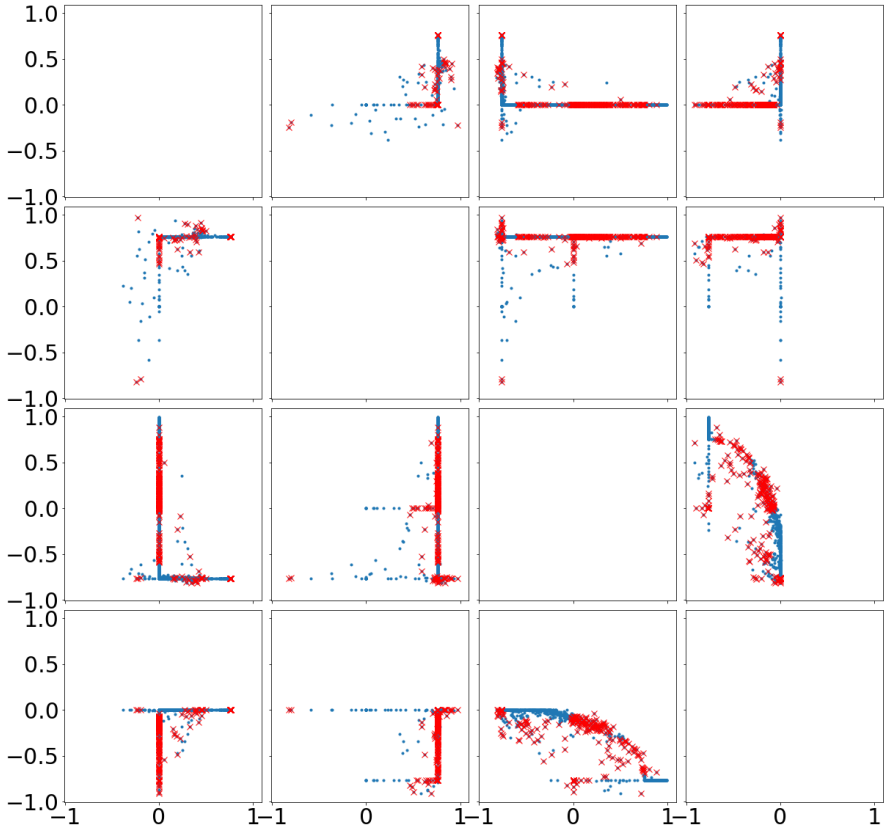

Figure 5. LAGMM energy of different features.

We compare our proposed model with benchmark studies, including an NLP model and DAGMM from [40], as shown in Table 1. We started by setting 99% of the sample as anomalies. It turns out that the model precision and accuracy is similar to random guessing with around 49.43% accuracy and 51.44% precision. As we decreased the percentile, the precision and accuracy are also observed to increase, which are all higher than the benchmark methods. The percentile drops will enable the data to have a more relax criteria; thus, it is reasonable to have more accurate results. In addition, during our testing, the proposed model can achieve a very fast prediction performance with 0.08 s on one sample. This fast response will enable the driver of CAV to make a quick decision while seeing the probability of the potential attacks in real time.

Table 1. Performance Comparisons.

Model—Percentile	Precision	Accuracy	F1 Score
NLP	60.32%	63.88%	1.27
DAGMM	64.23%	71.96%	0.092
LSTM	50.87%	51.93%	0.27
LAGMM—99%	49.43%	51.44%	0.024
LAGMM—97%	64.76%	53.55%	0.058
LAGMM—70%	70.63%	74.99%	0.53

Furthermore, the validation performance of the proposed model is promising as shown in Table 2. The LAGMM—70% still achieved a better score compared to the traditional

LSTM or NLP model with 65% accuracy. It is reasonable to believe our model has a great trade-off balance between the model complexity and the performance. Since the attacks are generated randomly, it is not possible to accurately create a precise numerical comparison between different type of attacks. However, we do notice a pattern of prediction, which means certain attacks can usually be detected more accurately than others. For example, a short anomaly can be identified more accurately than gradual drift, and noise is the least accurate attack that can be found. In this case, identification represents a higher possibility of cyberattacks, which is the energy. The reason for such behavior is likely caused by the nature of the attacks. The patterns of the first two attacks are likely to be caught by the model in the training session. On the other hand, the noise consists of more random components, which might need a more complex model structure to identify or learn. Then, the question circles back to the trade-off between model complexity and the performance again.

Table 2. Validation testing.

Model—Percentile	Precision	Accuracy	F1 Score
LAGMM—99%	50.63%	68.47%	0.034
LAGMM—97%	50.5%	55.1%	0.068
LAGMM—70%	63%	65.51%	0.54

7. Conclusions

In this paper, we proposed an anomaly detection model LAGMM to support anomalous CAV trajectories detection in real time. LAGMM consists of two major components: an LSTM Autoencoder and Decoder model which generates low-rank representations of original samples while keeping the dominant information and a GMM model that can estimate energy in a low-dimensional space. Specifically, the LSTM Autoencoder can extract long-term temporal dependencies of CAV trajectory data. The proposed anomaly detection method shows promising results in this study. The proposed LAGMM approach enhances detection accuracy by 3% and precision by 6.4% compared to the existing state-of-the-art methods, demonstrating the effectiveness of the proposed algorithm.

An important next step is validating the performance on diverse real-world autonomous vehicle data. The authors will investigate opportunities to collect such data through industry partnerships or public repositories. Expanding evaluation to various realistic scenarios would further demonstrate the generalizability of the approach. The future work also includes enhancing model capabilities to distinguish between different anomaly types such as short spikes, noise, gradual drift, etc. Being able to characterize the root cause of anomalies is important for autonomous vehicles to take appropriate actions [41]. The model could be extended with techniques such as pattern-based classification on the reconstruction errors to categorize different falsified trajectory behaviors. Incorporating Transformer-based models is another promising research direction. Self-attention mechanisms in Transformers can effectively model complex sequential dependencies in time-series data. Integrating a Transformer encoder into the proposed model architecture could potentially improve representation learning from multidimensional sensor streams. The self-attention layers can capture long-range correlations while maintaining computational efficiency. Fine-tuning on real AV data would allow the Transformer model to focus on patterns highly relevant to vehicle trajectories.

Author Contributions: Conceptualization, B.W., W.L. and Z.H.K.; Data curation, B.W. and W.L.; Formal analysis, B.W. and W.L.; Investigation, B.W., W.L. and Z.H.K.; Methodology, B.W. and W.L.; Project administration, W.L.; Resources, B.W. and W.L.; Software, B.W. and W.L.; Supervision, W.L.; Validation, B.W. and W.L.; Visualization, B.W. and W.L.; Writing—original draft, B.W., W.L. and Z.H.K.; Writing—review and editing, B.W., W.L. and Z.H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This project is funded in part by Carnegie Mellon University's Safety21 National University Transportation Center, which is sponsored by the US Department of Transportation BIL, 2023-2028 (4811).

Data Availability Statement: Data supporting the findings of this study are available from the author Wan Li at liw2@ornl.gov on request.

Conflicts of Interest: Author Boyu Wang was employed by the company Tacoma Public Utilities. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. AlSalem, T.S.; Almaiah, M.A.; Lutfi, A. Cybersecurity Risk Analysis in the IoT: A Systematic Review. *Electronics* **2023**, *12*, 3958. [CrossRef]
2. Khattak, Z.H.; Smith, B.L.; Fontaine, M.D. Impact of cyberattacks on safety and stability of connected and automated vehicle platoons under lane changes. *Accid. Anal. Prev.* **2021**, *150*, 105861. [CrossRef] [PubMed]
3. Huang, S.E.; Feng, Y.; Liu, H.X. A data-driven method for falsified vehicle trajectory identification by anomaly detection. *Transp. Res. Part C Emerg. Technol.* **2021**, *128*, 103196. [CrossRef]
4. Javed, A.R.; Usman, M.; Rehman, S.U.; Khan, M.U.; Haghighi, M.S. Anomaly detection in automated vehicles using multistage attention-based convolutional neural network. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 4291–4300. [CrossRef]
5. Zhong, H.; Cao, W.; Zhang, Q.; Zhang, J.; Cui, J. Toward trusted and secure communication among multiple internal modules in CAV. *IEEE Internet Things J.* **2021**, *8*, 17734–17746. [CrossRef]
6. Wen, X.; Chen, J.; Hu, Z.; Lu, Z. A p-opportunistic channel access scheme for interference mitigation between v2v and v2i communications. *IEEE Internet Things J.* **2020**, *7*, 3706–3718. [CrossRef]
7. Yang, T.; Lv, C. A secure sensor fusion framework for connected and automated vehicles under sensor attacks. *IEEE Internet Things J.* **2021**, *9*, 22357–22365. [CrossRef]
8. Guo, H.; Liu, J.; Dai, Q.; Chen, H.; Wang, Y.; Zhao, W. A distributed adaptive triple-step nonlinear control for a connected automated vehicle platoon with dynamic uncertainty. *IEEE Internet Things J.* **2020**, *7*, 3861–3871. [CrossRef]
9. Van Wyk, F.; Wang, Y.; Khojandi, A.; Masoud, N. Real-time sensor anomaly detection and identification in automated vehicles. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 1264–1276. [CrossRef]
10. Wang, Y.; Masoud, N.; Khojandi, A. Real-time sensor anomaly detection and recovery in connected automated vehicle sensors. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 1411–1421. [CrossRef]
11. Kamel, J.; Ansari, M.R.; Petit, J.; Kaiser, A.; Jemaa, I.B.; Urien, P. Simulation framework for misbehavior detection in vehicular networks. *IEEE Trans. Veh. Technol.* **2020**, *69*, 6631–6643. [CrossRef]
12. Dong, C.; Wang, H.; Ni, D.; Liu, Y.; Chen, Q. Impact evaluation of cyber-attacks on traffic flow of connected and automated vehicles. *IEEE Access* **2020**, *8*, 86824–86835. [CrossRef]
13. Yen, C.C.; Ghosal, D.; Zhang, M.; Chuah, C.N.; Chen, H. Falsified data attack on backpressure-based traffic signal control algorithms. In Proceedings of the 2018 IEEE Vehicular Networking Conference (VNC), Taipei, Taiwan, 5–7 December 2018; pp. 1–8.
14. Singh, P.K.; Tabjul, G.S.; Imran, M.; Nandi, S.K.; Nandi, S. Impact of security attacks on cooperative driving use case: CACC platooning. In Proceedings of the TENCON 2018-2018 IEEE Region 10 Conference, Jeju, Republic of Korea, 28–31 October 2018; pp. 138–143.
15. Nguyen, V.L.; Lin, P.C.; Hwang, R.H. Physical signal-driven fusion for V2X misbehavior detection. In Proceedings of the 2019 IEEE Vehicular Networking Conference (VNC), Los Angeles, CA, USA, 4–6 December 2019; pp. 1–4.
16. So, S.; Sharma, P.; Petit, J. Integrating plausibility checks and machine learning for misbehavior detection in VANET. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018; pp. 564–571.
17. Cui, L.; Chen, Z.; Wang, A.; Hu, J.; Park, B.B. Development of a robust cooperative adaptive cruise control with dynamic topology. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 4279–4290. [CrossRef]
18. Bertini, R.L.; Wang, H.; Knudson, T.; Carstens, K.; Rios, E. Assessing state department of transportation readiness for connected vehicle-cooperative systems deployment: Oregon case study. *Transp. Res. Rec.* **2016**, *2559*, 24–34. [CrossRef]
19. Bhavsar, P.; Das, P.; Paugh, M.; Dey, K.; Chowdhury, M. Risk analysis of autonomous vehicles in mixed traffic streams. *Transp. Res. Rec.* **2017**, *2625*, 51–61. [CrossRef]
20. Hasan, M.; Mohan, S.; Shimizu, T.; Lu, H. Securing vehicle-to-everything (V2X) communication platforms. *IEEE Trans. Intell. Veh.* **2020**, *5*, 693–713. [CrossRef]
21. Amoozadeh, M.; Raghuramu, A.; Chuah, C.N.; Ghosal, D.; Zhang, H.M.; Rowe, J.; Levitt, K. Security vulnerabilities of connected vehicle streams and their impact on cooperative driving. *IEEE Commun. Mag.* **2015**, *53*, 126–132. [CrossRef]
22. Islam, M.; Chowdhury, M.; Li, H.; Hu, H. Cybersecurity attacks in vehicle-to-infrastructure applications and their prevention. *Transp. Res. Rec.* **2018**, *2672*, 66–78. [CrossRef]

23. Cui, L.; Hu, J.; Park, B.B.; Bujanovic, P. Development of a simulation platform for safety impact analysis considering vehicle dynamics, sensor errors, and communication latencies: Assessing cooperative adaptive cruise control under cyber attack. *Transp. Res. Part C Emerg. Technol.* **2018**, *97*, 1–22. [CrossRef]
24. Khattak, Z.H.; Park, H.; Hong, S.; Boateng, R.A.; Smith, B.L. Investigating cybersecurity issues in active traffic management systems. *Transp. Res. Rec.* **2018**, *2672*, 79–90. [CrossRef]
25. Li, Y.; Tu, Y.; Fan, Q.; Dong, C.; Wang, W. Influence of cyber-attacks on longitudinal safety of connected and automated vehicles. *Accid. Anal. Prev.* **2018**, *121*, 148–156. [CrossRef]
26. Wardzinski, A. Dynamic risk assessment in autonomous vehicles motion planning. In Proceedings of the 2008 1st International Conference on Information Technology, Gdansk, Poland, 18–21 May 2008; pp. 1–4.
27. Wang, P.; Yu, G.; Wu, X.; Wang, Y.; He, X. Spreading patterns of malicious information on single-lane platooned traffic in a connected environment. *Comput. Civ. Infrastruct. Eng.* **2019**, *34*, 248–265. [CrossRef]
28. Zhou, J.; Zhang, B.; Fan, L.; Lu, Z. Aeromagnetic Anomaly Detection under Low SNR Conditions Using Multiscale Wavelet Energy Accumulation. In Proceedings of the 2020 IEEE 20th International Conference on Communication Technology (ICCT), Nanning, China, 28–31 October 2020; pp. 1641–1644.
29. Son, E.J.; Kim, W.; Kim, Y.M.; McIver, J.; Oh, J.J.; Oh, S.H. Time series anomaly detection for gravitational-wave detectors based on the Hilbert-Huang transform. *J. Korean Phys. Soc.* **2021**, *78*, 878–885. [CrossRef]
30. Jin, Y.; Qiu, C.; Sun, L.; Peng, X.; Zhou, J. Anomaly detection in time series via robust PCA. In Proceedings of the 2017 2nd IEEE International Conference on Intelligent Transportation Engineering (ICITE), Singapore, 1–3 September 2017; pp. 352–355.
31. Ma, M.; Han, L.; Zhou, C. BTAD: A binary transformer deep neural network model for anomaly detection in multivariate time series data. *Adv. Eng. Inform.* **2023**, *56*, 101949. [CrossRef]
32. Wang, Y.; Zhang, R.; Masoud, N.; Liu, H.X. Anomaly detection and string stability analysis in connected automated vehicular platoons. *Transp. Res. Part C Emerg. Technol.* **2023**, *151*, 104114. [CrossRef]
33. Yang, Z.; Ying, J.; Shen, J.; Feng, Y.; Chen, Q.A.; Mao, Z.M.; Liu, H.X. Anomaly Detection Against GPS Spoofing Attacks on Connected and Autonomous Vehicles Using Learning From Demonstration. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 9462–9475. [CrossRef]
34. Tiernan, T.; Richardson, N.; Azeredo, P.; Najm, W.G.; Lochrane, T. *Test and Evaluation of Vehicle Platooning Proof-of-Concept Based on Cooperative Adaptive Cruise Control (No. DOT-VNTSC-FHWA-17-13)*; John A. Volpe National Transportation Systems Center (US): Cambridge, MA, USA, 2017.
35. Hansun, S. A new approach of moving average method in time series analysis. In Proceedings of the 2013 Conference on New Media Studies (CoNMedia), Tangerang, Indonesia, 27–28 November 2013; pp. 1–4.
36. Trippel, T.; Weisse, O.; Xu, W.; Honeyman, P.; Fu, K. WALNUT: Waging doubt on the integrity of MEMS accelerometers with acoustic injection attacks. In Proceedings of the 2017 IEEE European Symposium on Security and Privacy (EuroS&P), Paris, France, 26–28 April 2017; pp. 3–18.
37. Currie, R. Developments in Car Hacking. 2015. Available online: https://sansorg.egnyte.com/dl/FTn9FydfUC (accessed on 29 January 2023).
38. Petit, J.; Shladover, S.E. Potential cyberattacks on automated vehicles. *IEEE Trans. Intell. Transp. Syst.* **2014**, *16*, 546–556. [CrossRef]
39. Purohit, H.; Tanabe, R.; Endo, T.; Suefusa, K.; Nikaido, Y.; Kawaguchi, Y. Deep autoencoding GMM-based unsupervised anomaly detection in acoustic signals and its hyper-parameter optimization. *arXiv* **2020**, arXiv:2009.12042.
40. Zong, B.; Song, Q.; Min, M.R.; Cheng, W.; Lumezanu, C.; Cho, D.; Chen, H. February. Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April– 3 May 2018.
41. Yun, K.; Yun, H.; Lee, S.; Oh, J.; Kim, M.; Lim, M.; Lee, J.; Kim, C.; Seo, J.; Choi, J. A Study on Machine Learning-Enhanced Roadside Unit-Based Detection of Abnormal Driving in Autonomous Vehicles. *Electronics* **2024**, *13*, 288. [CrossRef]