Safety21
INNOVATING SAFETY FOR ALL
The National University Transportation Center for Promoting Safety
Carnegie Mellon University
Penn University of Pennsylvania
The Ohio State University
CCAC
Community College of Philadelphia
The University of Texas Rio Grande Valley
Morgan State University

Cybersecurity risk assessment in connected intelligent systems for designing resilient systems

Zulqarnain Khattak, Ph.D., (PI), Carnegie Mellon University
(https://orcid.org/0000-0002-2599-4852)
Sean Qian, Ph.D., (Co-PI), Carnegie Mellon University
(https://orcid.org/0000-0001-8716-8989)
Gavin Lin, Student Assistant, Carnegie Mellon University

FINAL REPORT **:07/31/2024**

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, under [grant number 69A3552344811] from the U.S. Department of Transportation's University Transportation Centers Program. The U.S. Government assumes no liability for the contents or use thereof.

Contents

1. Introduction

Cybersecurity refers to methods and practices designed for protection of networks, computers, programs, and data from attack, damage, or unauthorized access (Rouse, 2016). Cybersecurity has emerged as a threat in every field that relies on communications. Transportation operation and management systems also utilize wired and wireless communications for managing roadways and are at significant risk of such cyberattacks. These systems were closed proprietary systems (isolated systems) in the past and had very limited cyber vulnerabilities. Those proprietary systems have now transformed into more open systems with increased accessibility due to the emergence of network computing and reliance on emerging technologies such as the internet of things (IoT), and connectivity. The National Transportation Communication for Intelligent Transportation Systems (ITS) Protocol (NTCIP) utilize center-to-center communications that rely on request-based protocols through XML messages (NTCIP9010, 2003). These protocols rely on the assumptions that most attacks are from the inside, and that hackers make up only a small portion of total intrusions, thus have no built-in security (NTCIP1105, 2011). The U.S. DOT has also taken a huge initiative to develop a security credential management system (SCMS) (Kreeb & Gay, 2014)—a message security solution for vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication. However, communication dependency opens a wide array of access points, which makes these systems vulnerable to cyberattacks and the least understood in terms of cybersecurity.

This research is based on the premise that perfect protection from cyberattacks is not realistic. Thus, the proposed research focuses on analyzing the vulnerability of cooperative driving relying on infrastructure-based communication from real-field experimental data collected at the Aberdeen center in Maryland. Multiple cyberattacks including sensor anomalies, fake BSMs, replay and denial of service were emulated. Furthermore, the driving conditions from the field experiment were emulated within a realistic simulation environment to test the consequences of different types of cyberattacks on safety effects of transportation systems and analyze crash types and severity. Long short-term memory with Gaussian mixture (LAGMM) model were utilized to design efficient and effective anomaly detection method for accounting the temporal relations of trajectories, so that anomalous behavior can be detected in real-time and the severe consequences of cyberattack or sensor anomalies can be avoided.

2. Past Literature

This section reviews past studies on cybersecurity in cooperative intelligent systems. Cyber physical systems (CPS) security has gathered a lot of attention in recent years due to the proliferation of smart devices in smart cities. A recent study (Habibzadeh et al., 2019) provided a detailed overview of theoretical and practical security challenges faced by CPS. While a few studies have conducted quantitative evaluations of cyber risks in CAVs, most prior studies have focused on qualitative evaluation of cyber risks. A study (Micro, 2017) assessed connected cars for security concerns. They identified Electronic Control Units (ECUs) as a source of vulnerability to cyberattacks. The authors asserted that the over 100 million lines of code per vehicle allow hackers to target software flaws. (Bertini et al., 2016) conducted a survey of Oregon DOT staff to analyze how prepared they are for CAVs. They found that 39-40% of respondents were concerned with security risks of CAVs. Another study (Bhavsar et al., 2017) assessed the risk of automation failure in a mixed traffic stream. The probability associated with failure of each autonomous component was estimated using fault tree analysis, resulting in a 14% failure probability of autonomous vehicular components. (Hasan et al., 2020) conducted a survey of the vehicle to everything ecosystem. The study reviewed security activities, standards, and existing defense mechanisms. They further identified existing gaps within security solutions and provided a description of open issues.

Some studies have assessed cyberattacks on CAVs using quantitative methods. (Amoozadeh et al., 2015) used OMNET++ to assess the falsified messages and radio jamming for their impact on ten CACC vehicles operating on a single lane. The desired acceleration was modified by the adversary during message falsification, which magnified instability throughout the stream. Likewise, the vehicles downgraded to ACC with larger time gaps when compromised by radio jamming. Another study (Islam et al., 2017) analyzed cyberattacks and their detection with a CVGuard architecture. They observed conflicts to increase by 10%-47% under cyberattacks while the CVGuard was observed to reduce 60% of the conflicts after the activation of CVGuard. (Cui et al., 2018) used microsimulation to analyze a ten-vehicle platoon (CACC) on a single lane under cyberattacks. Jamming was identified as a critical cyberattack and resulted in speed oscillations and crashes. They observed injury probability to increase in the range of 4.7%-40.2% under cyberattacks. (Khattak et al., 2018) analyzed an active traffic management system (ATM) for risks in its communication medium and developed a prototype threat monitoring system to revert the compromised ATM system back to normal operation under cyberattacks. The system was able to improve the speed of ATM by 13% and reduce the negative impact of cyberattacks. (Li et al., 2018) investigated the cyberattack influence on CAV safety. The attack was active for a short duration of time, and they observed rear-end risk collision index for deceleration period to be riskier with three attacked vehicles and 0.5% severity compared to nine attacked vehicles and 20% severity in acceleration period. (Wardzinski, 2008) proposed a risk based autonomous vehicle control system. Vehicle control was required to maintain the lowest acceptable safe risk. The minimum distance between vehicles was considered a risk factor for these situations based on constant speed and direction of other vehicles. The study concluded that communication and cooperation could improve performance and safety. (Y. Wang et al., 2020) developed an anomaly detection algorithm based on SVM to detect randomly modeled attacks from a car following model. They observed the algorithm to provide better prediction accuracy by 20% than the baseline algorithm. (Wyk et al., 2019) developed an anomaly detection algorithm to detect randomly generated attacks using speed and acceleration data from the USDOT Research Data Exchange (RDE) database. They used Kalman filter and convolutional neural network for anomaly detection and observed their algorithms to perform with high accuracy of 99% for attack detection.

(Javed et al., 2020) used sensor data including speed and acceleration from the RDE database to develop machine learning based anomaly detection algorithm based on normal and randomly generated attacks. They observed their ensemble algorithm to outperform classical machine learning methods 96% to 98% accuracy. (P. Wang et al., 2019) assessed cyberattack effects on a single lane platoon traveling on a single lane. They considered cyberattacks similar to the spread of malicious information and observed such cyberattacks to significantly disrupt traffic flow. (Khattak et al., 2021) analyzed the safety and stability impact of cyberattacks on CAV platoons using a lane management advisory application. They emulated multiple cyberattacks and observed a 40% increase in volatility and over 3000 crash conflicts with cyberattacks. (Kamel et al., 2020) developed a misbehavior detection algorithm (MDA) using data generated from a simulation environment in SUMO. They generated compromised data using six types of attacks and used machine learning algorithms (support vector machines and multilayer perceptron) to perform misbehavior detection. They observed their MDA to perform well with high accuracy of 93% to 95%. (Dong et al., 2020) utilized simulation-based attacks on cooperative adaptive cruise control platoons to analyze the effects on traffic flow and safety. They observed that increasing flow and severity of attacks on attacked vehicles lead to negative impacts on traffic flow and higher risks of collisions. (Khan et al., 2020) provided a synthesis of potentially critical avenues for cybersecurity to ensure reduction of the probability of cyberattack failures. The study provided a description of existing security measures in CAVs and synthesized details of cyberattacks on CAVs and mitigation strategies. (Haydari et al., 2021) studied the impact of deep reinforcement learning based attacks on traffic signal control with single and multiple intersections. They further proposed a sequential anomaly detection model and observed it to perform well for a few known attack types in their study using SUMO simulations. (Kloukiniotis et al., 2022) provides a detailed taxonomy of defense mechanisms for countering adversarial attacks in automated vehicles. A thorough investigation was carried out using adversarial noise removal technique based on deep learning based supervised approaches. (Tanksale, 2021) designed a prediction algorithm for LSTM to detect anomalies based on sensor data of real autonomous vehicles. The anomalies were emulated in the sensor data and their anomaly detection algorithm had an observed sensitivity of 97%-99%. (Haidar et al., 2021) used pseudonym certificates for misbehavior detection, where neighboring vehicles detect misbehavior and sends a report to a central entity that classifies the report as malicious or genuine. The study investigates sybil and data replay attacks. (Zheng et al., 2023) proposed learning-based algorithm for safe performance of autonomous vehicles. Their experimental results showed the ability of the proposed algorithm to enhance the efficiency of control policies while enhancing safety. The literature reveals a lack of guidance on understanding the risk of cyberattacks in cooperative driving using real experimental data and developing architectures for resilient operation of cooperative driving.

3. Objectives and Contributions

This research project has the following objectives while accounting for the limitation of past studies.

1. The study utilizes data from a real CAV platooning experiment to emulate anomalies and detect anomalous behavior within lead and following vehicles of the platoon.

2. The study develops a LAGMM to support anomalous CAV trajectories detection in real-time. The model accounts for the temporal relations of the trajectories, which have been ignored in the existing studies. The model can improve the detection rates compared to the state-of-the-art methods.

3. The two-step tasks of the LAGMM (decomposition and density estimation) would be optimized simultaneously, which helps it avoid less preferred local optima and further reduce reconstruction errors. The model aims to create a compression network for generating low-rank approximation for input data by (1) a LSTM autoencoder, which concatenates reduced space features with reconstruction error features, and (2) a GMM model to predict likelihood/energy.

4. Investigation of different types of cyberattacks in cooperative systems to assess how traffic stream stability and safety expressed through volatile behavior is affected by cyberattacks. This would indicate the impact of different types of cyberattacks on cooperative driving platoons.

3. Data Description

This study utilized data from field tests conducted by Federal Highway Administration (FHWA) (Tiernan et al., 2017) in collaboration with Volpe Center for a platooning proof of concept based on CACC and ACC. The field tests were conducted on a 4.5-mile test track at Aberdeen Center in Maryland. A fleet of five Cadillac SRX vehicles were equipped with CACC controllers, with variations of LV and FVs used for conducting the field tests. The test track geometry including grades, lane markings, width, and curvature are similar to a typical US highway. The test track was designated with geolocations called waypoints that were used to send target speeds to LVs, and the LV used these geolocations and global positioning system (GPS) to adjust its CACC parameters accordingly. Figure 1 provides a description of the test track along with the waypoints for the platooning experiments. A test run was considered complete when the test vehicle traveled from the first waypoint to the last waypoint along the track. Figure 2 provides a smoothened profile using moving average technique (Hansun, 2014). The acceleration was derived from smoothed speed data, which eliminated the difference between acceleration values estimated from raw and smoothed data. Furthermore, traffic and crash data from Virginia Department of Transportation was used to validate the observed speed profiles for the platoons and the understand crash risk with sudden speed reduction under cyberattacks. The results presented later in the report for the speed profiles account for the validation.

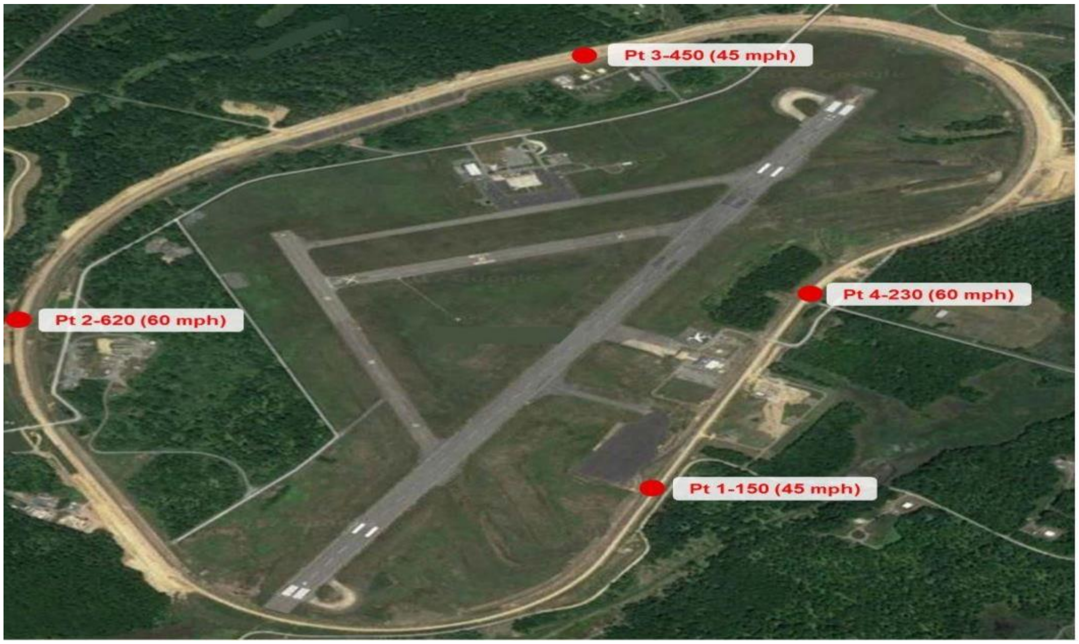


Figure 1 Test Track for Field Experiments (Tiernan et al., 2017)

4. Methodology

This section discusses the methodology for simulating cyberattacks and for the detection of cyberattack anomalies.

4.1 Methodology for Cyberattack Simulation

The cyberattack anomalies were simulated due to lack of publicly available anomalous CAV sensor data. The sensors have been shown to have vulnerability to cyberattacks and sensor failures by past research (Currie, 2015; Petit & Shladover, 2015; Trippel et al., 2017). A reasonable subset of likely attacks was subjected to a detailed investigation through a series of case studies. The cyberattacks were selected based on a higher probability for an attacker to launch these attacks, the ability to compromise CAV operation and safety, and a requirement that the attack needs a reasonable level of expertise and cost. Some examples of the types of attacks to be considered include.

1. Senor anomalies.

2.Infrastructure elements compromised by attacks at the V2I communication.

2.Communication at the vehicle level compromised by attacks at the network level (V2V).

With regards to sensor anomalies, three types of attacks were considered. For instance, a fake data injection attack through CAN bus or on-board diagnostic (OBD) can compromise the in-vehicle speed and acceleration sensors and result in several sensors anomalies. Likewise, an adversary with valid credentials can spoof the GPS through jamming or GPS spoofing attack and compromise the sensor values, thus generating anomalies. Further, an acoustic injection attack can compromise the acceleration sensor and generate anomalies. Furthermore, three types of attacks requiring communication access were selected for the case studies: fake BSM initiated by spoofing, sybil to access beacons and generate fake messages/BSMs to other surrounding vehicles, replay attack initiated when data packets stored at a previous instance of time are repeated maliciously or replayed.

These anomalies were injected into CAV sensors. The initiation of sensor anomalies due to attacks or sensor failure would be assumed as independent. A single anomaly was assumed in one time epoch due to independent nature of attacks or faults in sensors and reliability of sensors. Multiple rates alpha including 1%, 5% and 10% were used to generate several anomalous datasets. The anomalies were randomly simulated with random sensors. The simulated anomalies were then added to the base or normal sensor values of the compromised sensor (within lead vehicle or follower). The anomaly types and durations were varied for instance, anomaly to be simulated for 5mins, 20 mins etc. Further, mixed anomaly types were also considered for testing sensitivity, which includes multiple anomalies mixed. Specifically, the following anomalies and attacks were simulated.

I. Short anomaly, which is a sudden change in the observed 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 Nε(0,0.01) to capture the anomaly magnitude, where N belongs to 25, 100, 1000, 1000. The value was added to the sensor base value.

II. Noise, which is a longer-term change (multiple successive readings) in variance of the observed CAV trajectory data. The anomaly was simulated as an i.i.d sequence of random Gaussian variable with mean of zero, length l and variance c.

III. Bias, which is an offset from the true sensor readings. 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 ε(25, 50, 100, 1000).

IV. Gradual drift, which is a gradual drift in the observed data. This anomaly was simulated by offsetting the base values with linearly increasing values. For instance, using a linearly increasing speed of 0-5mph denoted by cε(3,5) using a function linspace (0,c). The anomaly was simulated for various durations.

V. Fine-grained Spoofing, where attackers can alter, or discard information exchanged through Basic Safety Messages (BSMs) between connected automated vehicles.

VI. Fine grained replay attack replaces the data at current instance of time with an older data within a replay period Δt.

4.2 Cyberattack and anomaly detection

As shown in Figure 2, the proposed LAGMM (Wang et al., 2024) consist of two major components: (1) a compression network aiming at generating low-rank approximation *z* for input data by a LSTM autoencoder, which concatenate reduced space features z_c with reconstruction error features z_r, (2) a GMM model to predict likelihood/energy.

Given an input sample *x*, LSTM autoencoder computes the load-dimensional representation z by Eq.(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 LSTM autoencoder, z_r denotes the features derived from the reconstruction error, θ_e and $theta_d$ are the parameters of the LSTM autoencoder, x^0 is the reconstructed counterpart of x, $h(\cdot)$ and $g(\cdot)$ denote the encoding and decoding function, and $f(\cdot)$ denotes the function to calculate reconstruction error features.

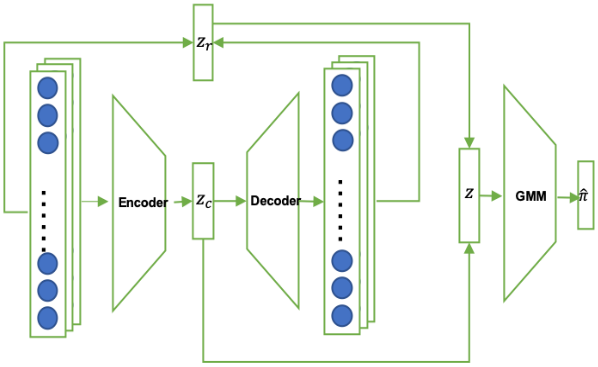


Figure 2. An overview of LSTM Autoencoder Gaussian Mixture Model.

Given the low rank approximation of the input data, the GMM-based estimation network aims at estimating the density function. The unknown parameters in GMM are mixture component distribution φ, mixture means μ, and mixture covariance Σ. A multi-layer neural network (MLNN) was leveraged to predict the mixture membership of each sample data, as shown in Eq.(5-6):

$$p = MLNN[z;\theta_m] \qquad (5)$$

$$\gamma\hat{} = softmax[p] \qquad (6)$$

where $\gamma\hat{}$ is a vector for the soft mixture-component membership prediction and p is the output of the MLNN. Given N samples and the membership prediction, we can further estimate the parameters of GMM as follows:

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

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

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

where $\varphi\hat{}_k$, $\mu\hat{}_k$, $\Sigma\hat{}_k$ are mixture probability, mean, and covariance for component k in GMM. The sample energy can be estimated with the estimated parameters:

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

where | · | denotes the determinant of a matrix.

In the testing process, the energy will be used to predict if the sample data is composed of falsified trajectories or not. Higher energy indicates a higher probability of anomalies.

4.2.1 Objective function

The objective function of the LAGMM is shown in Eq.(11). $L(x_i,x0_i)$ is the loss function that characterizes the reconstruction error caused by the LSTM autoencoder, which can be defined by L_2-norm. $E(z_i)$ denotes the probabilities that the input samples could be observed. Minimizing J aims at avoiding singularity problem by penalizing small values of the diagonal entries.

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

4.2.2 Hyperparameter Calibration

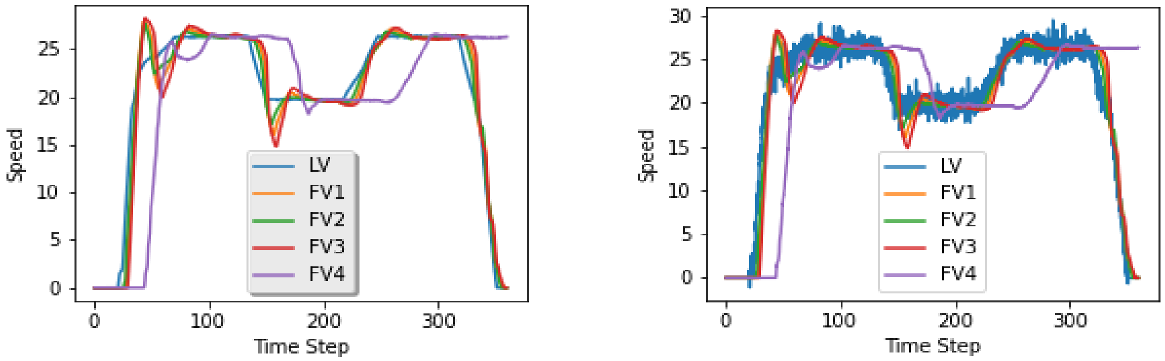
For the CARMA dataset, the LSTM Autoencoder feeds a three-dimensional input into the estimation network, consisting of one reduced dimension and two dimensions derived from the reconstruction error. In particular, the LAGMM runs with an LSTM layer with dimensions ((20, 4), 128, tanh) and eight Fully 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 have achieved so far.

5. Results and Discussion

This section discusses the results of the impact of cyberattacks on cooperative intelligent driving and the analysis for resilience and anomaly detection.

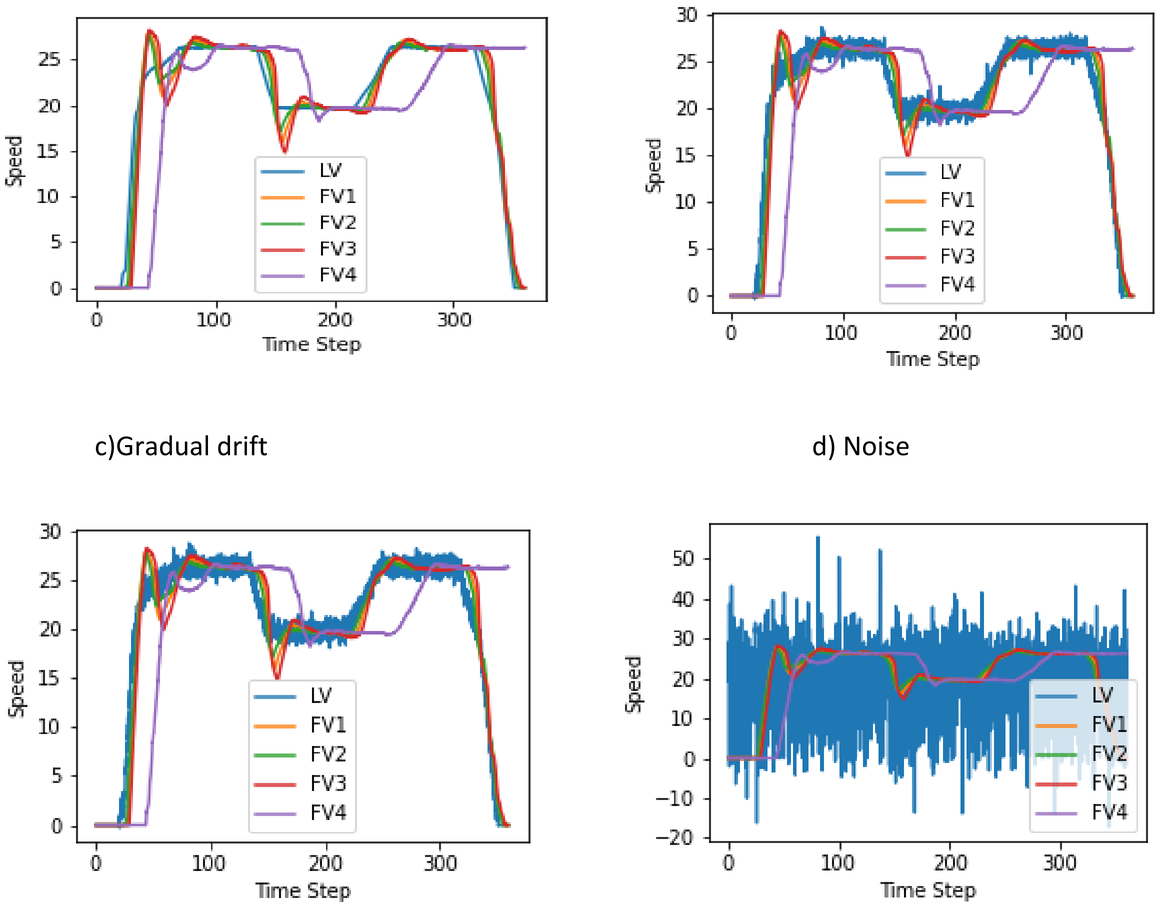
5.1 Analysis of the impact of Cyberattacks on cooperative driving

To further investigate the influence of cyberattacks on stability of cooperative driving, the influence on a 5-vehicle platoon is considered and speed profiles are used to demonstrate stability effects. The instability in speed profiles increases at varying levels as soon as the attacks begin at 40s in Figure 3 (b-f). Figure 3 (b-d) represents the cases of anomalies introduced into the vehicle sensors. The instability increases in these cases compared to the baseline (Figure 3a) with speed reduction caused by bottleneck generated over the network. This even leads to a potential collision due to pileups of multicar chain reactions. Figure 3e and Figure 3f represent spoofing and replay attacks. The replay attack case seems to be more dangerous and shows high variations with abrupt speed changes with respect to time. This leads to a highly volatile behavior of following vehicles overtaking their lead vehicles, that could lead to potential collision. The spoofing attack in Figure 3(e) represents similar but slightly lower patterns of disruption in the traffic stream. Thus, leading to uncertainty in flow patterns and potential collisions.



a) baseline- no attack

b)   Short anomaly



c)Gradual drift

d) Noise

e) Spoofing

f) Replay

FIGURE 3 Speed profiles for LV after attack. black, green, grey, white, and silver. The black vehicle serves as the Leading Vehicle (LV), while the others act as Following Vehicles (FVs).

5.2 Analysis for resilience and anomaly detection

We consider precision and accuracy to compare anomaly detection performance. We selected the threshold to identify anomalous samples. For example, when LAGMM is performed on CARMA, the top 20% samples of the highest energy will be marked as falsified trajectories. We take anomaly class as positive and define precision and accuracy accordingly.

Figure 4 illustrates the histogram of energy computed from LAGMM model (Wang et al., 2024). 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.

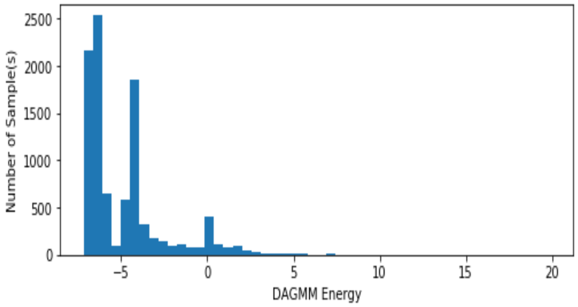


Figure 4. Histogram of LAGMM Energy.

Figure 5 shows 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.

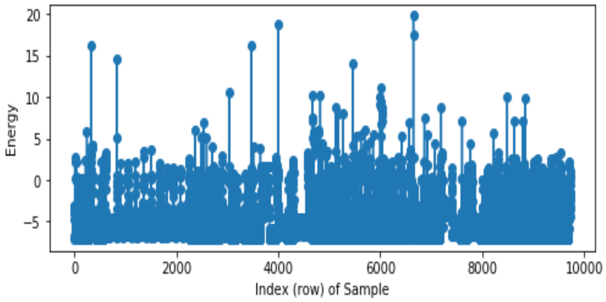


Figure 5. LAGMM energy of all samples.

Figure 6 illustrates the normal CAV trajectories (blue dots) and falsified CAV trajectories (red dots) in terms of each combination of two features out of total four features (four rows and four columns). We compare our proposed model with benchmark studies, including an NLP model and DAGMM from (Zong et al., 2018), 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 relaxed 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.

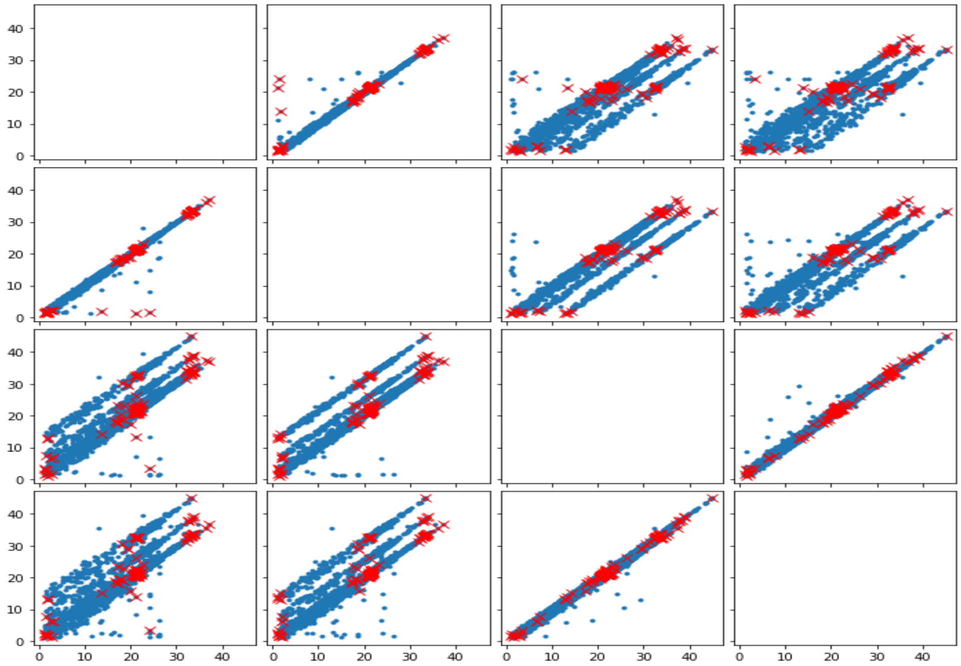


Figure 6. LAGMM energy of different features

Table 1. Performance Comparisons.

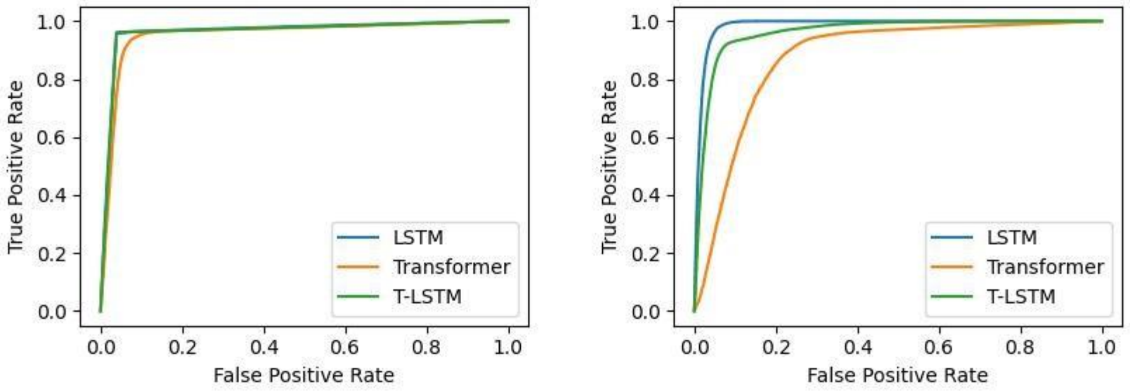
Model—Percentile	Precision	Accuracy	F1 Score
NLP	60.32%	63.88%	1.27
DAGMM	64.23%	71.96%	0.092
LSTM	50.87%	51.93%	0.27
LAGMM—99%	49.43%	51.44%	0.024
LAGMM—97%	64.76%	53.55%	0.058
LAGMM—70%	70.63%	74.99%	0.53

Furthermore, the validation performance of the proposed model is promising as shown in Table 1. 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 types 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

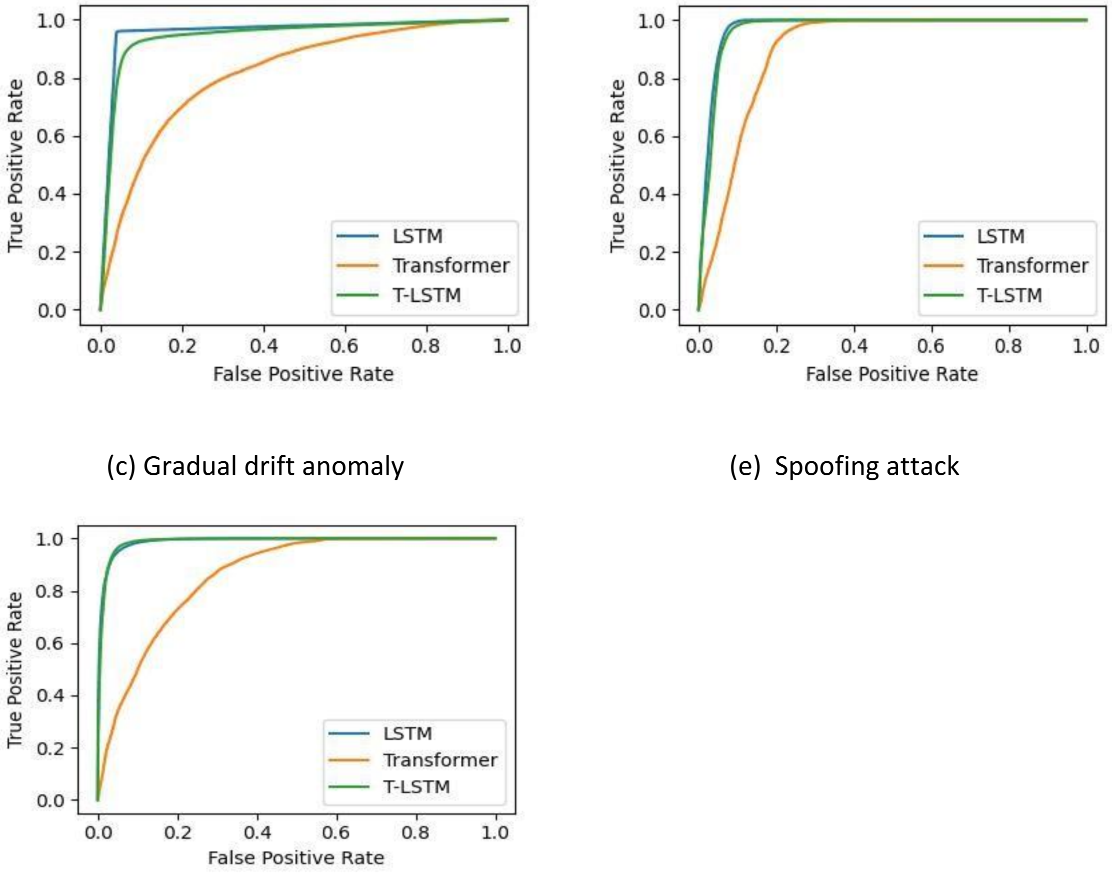
Model—Percentile	Precision	Accuracy	F1 Score
LAGMM—99%	50.63%	68.47%	0.034
LAGMM—97%	50.5%	55.1%	0.068
LAGMM—70%	63%	65.51%	0.54

We further analyzed the performance of LSTM and state of the art models for anomaly detection in terms of their accuracy (i.e. true positive and false positive rates) shown in Figure 7. The ROC curve reveals better performance when there is more Area under the ROC Curve (AUC). We observe in Figure 7 that there is a direct relationship in terms of performance with LSTM > T-LSTM > Transformer.



(a) Short Anomaly

(b) Bias Anomaly



(c) Gradual drift anomaly

(e) Spoofing attack

(f) Replay attack

Figure 7. ROC Curves showing area under the curve

6. Conclusions

This research developed a framework to study cyberattack simulation and anomaly detection within cooperative driving automation. The study utilized real-world data from field experiments of cooperative driving automation to simulate cyberattacks and perform anomaly detection. The study simulates Spoofing, Message Falsification, and Replay attacks based on three anomalies (short, bias and gradual drift) to capture complex attack patterns and employs long short-term memory neural network with Gaussian mixture model for anomaly detection, ensuring resilient operation of cooperative driving. Our aggregation strategies enhance detection performance, and real-world tests confirm the framework's effectiveness. This work advances secure, private, and efficient vehicle platooning, improving the reliability of cooperative driving automation. Future work is focusing on federated learning to enhance security of cooperative driving under adversarial attacks.

Acknowledgement

The work reported in this report is taken from published articles ‘Anomaly Detection in Connected and Autonomous Vehicle Trajectories Using LSTM Autoencoder and Gaussian Mixture Model (https://doi.org/10.3390/electronics13071251)’ and ‘Cybersecurity vulnerability and resilience of cooperative driving automation for energy efficiency and flow stability in smart cities (https://doi.org/10.1016/j.scs.2024.105368)’, which were sponsored by this project under Safety 21 BIL UTC and the work was conducted through the funding provided by the project.

Articles Published and in progress

This project has resulted in the following published papers on cybersecurity in cooperative driving and intelligent transportation systems.

- Anomaly Detection in Connected and Autonomous Vehicle Trajectories Using LSTM Autoencoder and Gaussian Mixture Model (https://doi.org/10.3390/electronics13071251)
- Cyberattack Monitoring Architectures for Resilient Operation of Connected and Automated Vehicles (10.1109/OJITS.2024.3391830)
- Cybersecurity vulnerability and resilience of cooperative driving automation for energy efficiency and flow stability in smart cities (https://doi.org/10.1016/j.scs.2024.105368)
- Guanyu Lin, Sean Qian, and Zulqarnain Khattak. Cyberattack vulnerability and resilience of cooperative driving automation using federated learning
- Guanyu Lin, Sean Qian, and Zulqarnain Khattak. Cyberattacks on leader and followers in cooperative driving automation: interpretable machine learning with federated agents

Data Statement

The raw data for the field experiments is available is available at Federal Highway Administration website https://data.transportation.gov/Automobiles/Test-Data-of-Proof-of-Concept-Vehicle-Platooning-B/ wpek-zziu/about_data. The raw cyberattack data generated is available at ZKhattak11/Cybersecurity-Safety21 (github.com) .

References

Amoozadeh, M., Raghuramu, A., Chuah, C. N., Ghosal, D., Michael Zhang, H., Rowe, J., & Levitt, K. (2015). Security vulnerabilities of connected vehicle streams and their impact on cooperative driving. *IEEE Communications Magazine*, *53*(6), 126–132. https://doi.org/10.1109/MCOM.2015.7120028

Bertini, R. L., Wang, H., Knudson, T., Carstens, K., & Rios, E. (2016). Assessing State Department of Transportation Readiness for Connected Vehicle – Cooperative Systems Deployment Oregon Case Study. *Journal of the Transportation Research Board*, *2559*(November 2015), 24–34. https://doi.org/10.3141/2559-04

Bhavsar, P., Das, P., Paugh, M., Dey, K., & Chowdhury, M. (2017). Risk Analysis of Autonomous Vehicles in Mixed Traffic Streams. *Transportation Research Record: Journal of the Transportation Research Board*, *2625*, 51–61. https://doi.org/10.3141/2625-06

Cui, L., Hu, J., Park, B. B., & Bujanovic, P. (2018). Development of a simulation platform for safety impact analysis considering vehicle dynamics, sensor errors, and communication latencies: Assessing cooperative adaptive cruise control under cyberattack. *Transportation Research Part C*, *97*, 1–22.

Currie, R. (2015). Developments in Car Hacking. *SANS Inst*.

Dong, C., Wang, H., Ni, D., Liu, Y., & Chen, Q. (2020). Impact Evaluation of Cyber-Attacks on Traffic Flow of Connected and Automated Vehicles. *IEEE Access*, *8*, 86824–86835. https://doi.org/10.1109/ACCESS.2020.2993254

Habibzadeh, H., Nussbaum, B. H., Anjomshoa, F., Kantarci, B., & Soyata, T. (2019). A survey on cybersecurity, data privacy, and policy issues in cyber-physical system deployments in smart cities. *Sustainable Cities and Society*, *50*. https://doi.org/https://doi.org/10.1016/j.scs.2019.101660

Haidar, F., Makassikis, M., Sall, M., Bakhti, H., Kaiser, A., & Lonc, B. (2021). Experimentation and Assessment of Pseudonym Certificate Management and Misbehavior Detection in C-ITS. *IEEE Open Journal of Intelligent Transportation Systems*, *2*, 128–139. https://doi.org/10.1109/OJITS.2021.3085366

Hansun, S. (2014). A new approach of moving average method in time series analysis. *IEEE Conference on New Media Studies (CoNMedia)*. https://doi.org/10.1109/CoNMedia.2013.6708545

Hasan, M., Mohan, S., Shimizu, T., & Lu, H. (2020). Securing Vehicle-to-Everything (V2X) Communication Platforms. *IEEE Transactions on Intelligent Vehicles*, *5*(4), 693–713. https://doi.org/10.1109/TIV.2020.2987430

Haydari, A., Zhang, M., & Chuah, C. N. (2021). Adversarial Attacks and Defense in Deep Reinforcement Learning (DRL)-Based Traffic Signal Controllers. *IEEE Open Journal of Intelligent Transportation Systems*, *2*, 402–416. https://doi.org/10.1109/OJITS.2021.3118972

Islam, M., Chowdhury, M., Li, H., & Hu, H. (2017). Cybersecurity Attacks in Vehicle to Infrastructure Applications and their Prevention. *Transportation Research Board 97th Annual Meeting*.

Javed, A. R., Usman, M., Rehman, S. U., Khan, M. U., & Haghighi, M. S. (2020). Anomaly Detection in Automated Vehicles Using Multistage Attention-Based Convolutional Neural Network. *IEEE Transactions on Intelligent Transportation Systems*, 1–10. https://doi.org/10.1109/tits.2020.3025875

Kamel, J., Ansari, M. R., Petit, J., Kaiser, A., Jemaa, I. Ben, & Urien, P. (2020). Simulation Framework for Misbehavior Detection in Vehicular Networks. *IEEE Transactions on Vehicular Technology*, *69*(6), 6631–6643. https://doi.org/10.1109/TVT.2020.2984878

Khan, S. K., Shiwakoti, N., Stasinopoulos, P., & Chen, Y. (2020). Cyber-attacks in the next-generation cars, mitigation techniques, anticipated readiness and future directions. *Accident Analysis and Prevention*, *148*. https://doi.org/https://doi.org/10.1016/j.aap.2020.105837

Khattak, Z. H., Park, H., Hong, S., Boateng, R., & Smith, B. L. (2018). Investigating Cybersecurity Issues in Active Traffic Management Systems. *Transportation Research Record, Journal of Transportation Research Board*. https://doi.org/https://doi.org/10.1177/0361198118787636

Khattak, Z. H., Smith, B. L., & Fontaine, M. D. (2021). Impact of cyberattacks on safety and stability of connected and automated vehicle platoons under lane changes. *Accident Analysis & Prevention*, *150*(105861). https://doi.org/10.1016/j.aap.2020.105861

Kloukiniotis, A., Papandreou, A., Lalos, A., Kapsalas, P., Nguyen, D. V., & Moustakas, K. (2022). Countering Adversarial Attacks on Autonomous Vehicles Using Denoising Techniques: A Review. In *IEEE Open Journal of Intelligent Transportation Systems* (Vol. 3, pp. 61–80). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/OJITS.2022.3142612

Kreeb, B., & Gay, K. (2014). Security Credential Management System (SCMS) Proof of Concept (POC). *US Department of Transportation*.

Li, Y., Tu, Y., Fan, Q., Dong, C., & Wang, W. (2018). Influence of cyberattacks on longitudinal safety of connected and automated vehicles. *Accident Analysis and Prevention*, *121*, 148–156.

Micro, T. (2017). *Cybersecurity Solutions for Connected Vehicles Contents*.

NTCIP1105. (2011). National Transportation Communications for ITS Protocols-CORBA Security Service Specification. *AASHTO, ITE, NEMA*.

NTCIP9010. (2003). National Transportation Communications for ITS Protocol- XML in ITS Center-to-Center Communications. *AASHTO, ITE, NEMA*.

Petit, J., & Shladover, S. E. (2015). Potential Cyberattacks on Automated Vehicles. *IEEE Transctions on Intelligent Transportation Systems*, *16*(2), 546–556.

Rouse, M. (2016). *Cybersecurity*. http://whatis.techtarget.com/definition/cybersecurity

Tanksale, V. (2021). Design of Anomaly Detection Functions for Controller Area Networks. *IEEE Open Journal of Intelligent Transportation Systems*, *2*, 312–321. https://doi.org/10.1109/OJITS.2021.3104495

Tiernan, T. A., Richardson, N., Azeredo, P., Najm, W. G., & Lochrane, T. (2017). Test and Evaluation of Vehicle Platooning Proof-of-Concept Based on Cooperative Adaptive Cruise Control. *Federal Highway Administration*.

Trippel, T., Weisse, O., Xu, W., Honeyman, P., & Fu, K. (2017). WALNUT:Waging doubt on the integrity of MEMS accelerometers with acoustic injection attacks. *Proc. IEEE Eur. Symp. Secur. Privacy*, 2–18.

Wang, B., Li, W., & Khattak, Z. H. (2024). Anomaly Detection in Connected and Autonomous Vehicle Trajectories Using LSTM Autoencoder and Gaussian Mixture Model. *Electronics (Switzerland)*, *13*(7). https://doi.org/10.3390/electronics13071251

Wang, P., Yu, G., Wu, X., Wang, Y., & He, X. (2019). Spreading Patterns of Malicious Information on Single-Lane Platooned Traffic in a Connected Environment. *Computer-Aided Civil and Infrastructure Engineering*, *34*(3), 248–265. https://doi.org/10.1111/mice.12416

Wang, Y., Masoud, N., & Khojandi, A. (2020). Real-Time Sensor Anomaly Detection and Recovery in Connected Automated Vehicle Sensors. *IEEE Transactions on Intelligent Transportation Systems*, *3*, 1–11.

Wardzinski, A. (2008). Dynamic risk assessment in autonomous vehicles motion planning. *Proc. International Conference on Information Technology*, *May*, 1–4. https://doi.org/10.1109/INFTECH.2008.4621607

Wyk, F. Van, Wang, Y., Khojandi, A., & Masoud, N. (2019). Real-Time Sensor Anomaly Detection and Identification in Automated Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *3*, 1–13.

Zheng, H., Chen, C., Li, S., Zheng, S., Li, S. E., Xu, Q., & Wang, J. (2023). Learning-Based Safe Control for Robot and Autonomous Vehicle Using Efficient Safety Certificate. *IEEE Open Journal of Intelligent Transportation Systems*, *4*, 419–430. https://doi.org/10.1109/OJITS.2023.3280573

Zong, B., Song, Q., Min, M. R., Cheng, W., Lumezanu, C., Cho, D., & Chen, H. (2018). Deep autoencoding gaussian mixture model for unsupervised anomaly detection. *In Proceedings of the International Conference on Learning Representations*.

1. Report No. 436	**2. Government Accession No.**		**3. Recipient's Catalog No.**
4. Title and Subtitle Cybersecurity risk assessment in connected intelligent systems for designing resilient systems			**5. Report Date** 31st July 2024
			6. Performing Organization Code Enter any/all unique numbers assigned to the performing organization, if applicable.
7. Author(s) Zulqarnain Khattak, Ph.D., (PI), (https://orcid.org/0000-0002-2599-4852) Sean Qian, Ph.D., (Co-PI), (https://orcid.org/0000-0001-8716-8989) Gavin Lin, Student Assistant			**8. Performing Organization Report No.** Enter any/all unique alphanumeric report numbers assigned by the performing organization, if applicable.
9. Performing Organization Name and Address Carnegie Mellon University, Department of Civil and Environmental Engineering			**10. Work Unit No.**
			11. Contract or Grant No. Federal Grant No. 69A3552344811
12. Sponsoring Agency Name and Address Safety21 University Transportation Center Carnegie Mellon University 5000 Forbes Avenue Pittsburgh, PA 15213			**13. Type of Report and Period Covered** Final Report (July 1, 2023-June 30, 2024)
			14. Sponsoring Agency Code USDOT
15. Supplementary Notes Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.			
16. Abstract Transportation operation and management systems utilize wired and wireless communications for managing roadways and are at significant risk of cyberattacks. Furthermore, perfect protection from cyberattacks is not realistic. Thus, this research proposes to focus on analyzing the vulnerability of cooperative driving relying on infrastructure-based communication using real-field experimental data collected at the Aberdeen center in Maryland. Multiple cyberattacks and sensor anomalies were emulated using conditions from field experiments to test the consequences of different types of cyberattacks. Long short-term memory with Gaussian mixture (LAGMM) model were utilized to design efficient and effective anomaly detection method for accounting the temporal relations of trajectories, so that anomalous behavior can be detected in real-time and the severe consequences of cyberattack or sensor anomalies can be avoided. The research would help agencies in monitoring the cooperative driving environment for anomalous behavior.			
17. Key Words Cybersecurity, cyberattack, anomaly detection, long short-term memory, vehicle platooning, anomaly detection, field experiments		**18. Distribution Statement** No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161. Enter any other agency mandated distribution statements. Remove NTIS statement if it does not apply.	
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 21	**22. Price** -

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized