
# Optimizing Traffic Flow and Enhancing Security in Cooperative Intelligent Transportation Systems Using NGSIM

Sultan Ahmed Almalki[1], Tami Abdulrahman Alghamdi [2,*], Azan Hamad Alkhorem[3]

[1]Computer Department, Applied College, Najran University, Najran 66462 , Kingdom of Saudi Arabia
[2]Computer Science Department, Faculty of Computing and Information, Al-Baha University, Al-Baha, 65779, Saudi Arabia
[3]Department of Computer Engineering, College of Computer Science and Information Technology, Majmaah University, Al-Majmaah 11952, Saudi Arabia

Emails: Saalmalki@nu.edu.sa; Talwajeeh@bu.edu.sa; ah.alkhorem@mu.edu.sa

**Abstract**

Cooperative Intelligent Transportation Systems (C-ITS) cannot work effectively if they do not have both efficient traffic management and solid security. We put forward in this paper an original framework that takes advantage of the Next Generation Simulation (NGSIM) dataset to improve traffic flow and system security by identifying False Data Injection Attacks (FDIA). By applying leading machine learning algorithms to authentic traffic data, we generate models that support improved vehicle coordination as well as provide assistance with security vulnerabilities in C-ITS systems. We are concentrating our method on the optimization of traffic dynamics by making intelligent decisions, while keeping the system secure from malicious cyber attacks. Analyses of the NGSIM data revealed that our proposed approaches produced important advancements in traffic flow efficiency and the accuracy of anomaly detection. Results prove that our framework minimizes congestion and concurrently enhances the reliability and security of collaborative vehicle systems. This investigation proposes a practical approach for fusing traffic optimization with cybersecurity, improving smart city evolution and the future of autonomous vehicles and vehicle connectivity.

**Keywords**: Traffic Optimization; C-ITS; NGSIM Dataset; FDIA Detection; Machine Learning; Anomaly Detection; Cybersecurity

## 1 Introduction

The rapid development of Cooperative Intelligent Transportation Systems (C-ITS) has altered traffic management by creating the ability for vehicles, infrastructure, and road users to interact and collaborate in real time. C-ITS demonstrates capability in increas However, the constantly expanding reliance on data exchange and vehicle synchronization introduces important issues, especially in supporting the security and reliability of these systems. The advent of False Data Injection Attacks (FDIA) has pointed out their critical importance in the ability to disrupt traffic flow, make accidents occur, or even halt complete transportation networks [2].

In the last few years, there has been considerable research aimed at streamlining traffic flow in linked settings, many of which use real traffic datasets, including the NGSIM dataset. Providing detailed knowledge of traffic behavior in urban complexes, the NGSIM dataset delivers highly granular data about vehicle trajectories [3]. This dataset has successfully supported previous investigations in traffic simulation, vehicle behavior analysis, as well as lane-changing models [4, 5]. Though many studies have emphasized traffic optimization, little focus has been

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

on merging traffic improvements with security controls against cyber-attacks such as FDIA [6, 7]. The goal of this paper is to plug the gap by proposing a new framework that blends traffic flow optimization with anomaly detection for FDIA within C-ITS, using the NGSIM dataset. We apply advanced machine learning algorithms to model the dynamics of traffic, increasing vehicle coordination and attenuating security vulnerabilities. Unlike previous techniques that only target either traffic optimization or security, we put forward a unified framework that takes on both issues concurrently. The primary contributions of this work are threefold:

1. Traffic Flow Optimization: We formulate a model that makes use of NGSIM data along with machine learning to improve traffic efficiency and alleviate congestion, while increasing total vehicle throughput.

2. FDIA Detection: We develop a comprehensive anomaly detection method designed for detecting FDIA in traffic situations, making sure that C-ITS environments are both secure and reliable.

3. Comprehensive Evaluation: Evaluations of our methods using the NGSIM dataset show a notable improvement in both optimizing traffic flow and detection accuracy related to FDIA.

The remainder of this paper is organized as follows: Section 2 gives a broad perspective on related work, bringing attention to the present techniques for traffic optimization and security within C-ITS. In Section 3, we explain the methodology we propose, which includes steps of data preprocessing, interface design, and assessment metrics. The experimental results are in Section 4, and Section 5 focuses on the implications arising from our findings. At last, Section 6 concludes this paper with ideas for the direction of future research.

## 2 Related Work / Literature Review

Contemporary traffic management has stimulated significant research into enhancing traffic flow and assuring the security of its transportation networks due to the introduction of Cooperative Intelligent Transportation Systems (C-ITS). In this section, the critical contributions to traffic optimization, based on actual datasets, along with security threats including False Data Injection Attacks (FDIA), and the application of machine learning approaches to deal with these challenges are surveyed.

### 2.0.1 Traffic Flow Optimization in C-ITS

A considerable volume of literature has aimed to advance traffic efficiency with data collected from the real world. This area has greatly benefited from the Next Generation Simulation (NGSIM) dataset. The NGSIM data sources give detailed traffic information, monitoring vehicle locations in dense city environments, which allows researchers to construct and simulate intricate traffic behaviors. Koutsopoulos et al. [3] executed one of the foundational studies using NGSIM data to investigate microscopic traffic behavior, looking into lane-changing trends and vehicle interactions. Chang et al. [4] used NGSIM data in a similar manner to examine traffic flow patterns and their relevance for autonomous vehicle (AV) systems. The research indicated the way in which elaborate trajectory data can contribute to refining vehicle coordination and lane changes in mixed traffic conditions. Qian et al. modeled the behavior of changing lanes in mixed traffic situations that encompass human-driven vehicles and autonomous vehicles regarding traffic flow optimization [5]. Their research focused on the difficulties of overseeing interactions among different styles of vehicles and added insights into maximizing traffic efficiency in these settings. A significant amount of research shows the effective use of NGSIM data for traffic simulation, but few efforts have succeeded in merging traffic optimization models with security measures, especially against cyber-attack vulnerabilities such as FDIA.

### 2.0.2 Security Challenges in C-ITS: FDIA and Anomaly Detection

The dependancy of C-ITS on data exchange between vehicles and infrastructure makes the security of these communications extremely important. C-ITS is at great odds with FDIA, where hackers may provide misleading data to the system and influence traffic lights, amplify congestion, or even cause accidents. Liu et al. [6] investigated how vulnerable cyber-physical systems (CPS) are to FDIA, using machine learning to find such attacks.

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

While their effort founded the principles of FDIA detection in transportation systems, it missed an approach that encompasses traffic dynamics from the real world. Huang et al. [7] broadened the scope of this research by recommending a deep learning solution for the identification of cyber threats in smart city traffic networks. The method they employed showed great promise in finding anomalies within large networks, but it did not focus on the challenges tied to FDIA in C-ITS environments. Even though both studies add to the wider scope of cybersecurity in transportation, neither totally connects FDIA detection with the concurrent improvement of traffic flow.

### 2.1 Machine Learning for Traffic Optimization and Security

Lately, machine learning methods have seen a pronounced growth in research applied to traffic systems. The process for improving traffic management with machine learning algorithms wants from its ability to predict vehicle behaviors, identify anomalies, and boost system performance. Research by Soteropoulos et al. [1] looked into the use of machine learning for vehicle coordination within autonomous systems, uncovering both better traffic flow and greater fuel efficiency. Essentially, their research made clear the need for real-time data analytics to better vehicle performance, though it never touched upon the important cybersecurity challenges necessary for maintaining the reliability of C-ITS. According to the analysis by Yang et al. [2] of Internet of Things (IoT) rollouts in smart cities, machine learning can both improve traffic efficiency and bolster security. In any case, it identified a lacuna in research regarding the relationship between traffic optimization and anomaly detection in C-ITS. This variation indicates the necessary requirement for an inclusive strategy that includes system security along with traffic efficiency.

## 3 Contributions of the Present Work

While advances have been made in either traffic optimization or security, relatively few studies have integrated these two important aspects. This work fills the gap by introducing a unified framework that makes use of the NGSIM dataset to improve traffic flow and also identify FDIA in C-ITS simultaneously. Our strategy contributes to existing literature by merging sophisticated machine learning algorithms for anomaly detection and traffic flow optimization, leading to a sturdy response to both security and efficiency challenges in current transportation systems.

## 4 Methodology

This part focuses on the architectures and methods put forward to address the linked goals of advancing traffic flow and spotting False Data Injection Attacks (FDIA) in Cooperative Intelligent Transportation Systems (C-ITS). Our approach is divided into three key components: At the same moment, data preprocessing, along with machine learning models, is in operation for the aims of enhancing traffic flow and for recognizing FDIA through anomaly detection algorithms.

### 4.1 NGSIM Dataset Overview

Vehicle movement data in urban and highway situations is present in the NGSIM dataset, which records data on position, speed, acceleration, and lane changes at the rate of every 0.1 seconds. This analysis uses data that includes greater than 45 minutes on Interstate 80 and U.S. Highway 101, which : monitors vehicle trajectories affected by varying traffic conditions, congestion versus free-flow times [3,4]. The dataset backs a broad investigation into trends and behaviors tied to traffic.

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

### 4.2 Data Preprocessing

Preprocessing is important to guarantee the right usage of NGSIM data for improving traffic flow and also in the path to find FDIA. The preprocessing steps are as follows:

• Data Cleaning: We remove false or failing data that includes missed vehicle trajectories combined with inconsistencies in vehicle identifiers.

• Data Normalization: For ensuring consistency across different traffic circumstances, trajectory data of vehicles is normalized by normalizing position, speed, and acceleration values to a same range.

• Feature Engineering: Traffic flow-related key features, such as headway, time-to-collision (TTC), and lane-changing frequency, are gained from the original data. These capabilities are necessary for both enhancement of traffic and notice of anomalies.

The following equation is used to compute the time-to-collision (TTC) for each vehicle i and its leading vehicle j:

$$TTC_{i,j} = \frac{X_j - X_i}{V_i - V_j} \tag{1}$$

where:

• Xj and X; are the positions of the leading and following vehicles, respectively,

• Vj and Vi are their velocities. A negative $TTCij$ value indicates that the following vehicle is approaching the leading vehicle, signaling a potential collision scenario, which is crucial for optimizing vehicle behavior in C-ITS.

## 5 Traffic Flow Optimization Using Machine Learning

We handle traffic flow optimization as an issue in supervised learning. In light of the complicated relationships among vehicles in a C-ITS system, our model intends to predict the best driving actions (such as acceleration, braking, and lane changes) that boost traffic efficiency.

#### 5.0.1 Model Selection

Several algorithms for machine learning were compared, like Random Forest, Gradient Boosting Machines (GBM) and Long Short-Term Memory (LSTM) neural networks. The sequential nature of traffic data led to the superior results observed in LSTM models [1]. The LSTM model predicts the optimal driving actions based on historical data, using the following architecture:

$$h_t = \sigma(W_x x_t + W_h h_{t-1} + b) \tag{2}$$

where:

- $h_t$ is the hidden state at time step $t$,
- $x_t$ is the input vector at time step $t$,
- $W_x$ and $W_h$ are the weight matrices for input and hidden states, respectively,
- $b$ is the bias term, and
- $\sigma$ is the activation function (e.g., ReLU or sigmoid).

Using a labeled vehicle trajectory dataset as input, the model learns to produce an action (acceleration, deceleration, or lane change) that maximizes congestion reduction.

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

### 5.0.2 Evaluation Metrics

The performance of the traffic flow optimization model is evaluated using the following metrics:

• Mean Squared Error (MSE): Monitors the split between the expected and the real driving behaviors.

• Throughput Improvement: Values the increase in vehicle throughput as a proportion as a result of optimized measures.

• Congestion Reduction: Evaluates the cutback in traffic density and standards for vehicle waiting time

## 6 FDIA Detection Using Anomaly Detection Techniques

By troubling in sourcing labeled attack data, we need to look to an unsupervised anomaly detection methodology to find FDIA. The approach we pursue emphasizes the detection of vehicle behavior changes that might reflect illegal data injections, such as rapid changes in vehicle velocity or location that are at odds with surrounding traffic conditions [6,7].

### 6.0.1 Anomaly Detection Algorithm

We employ a hybrid approach that combines Autoencoders and Isolation Forests to detect anomalous vehicle behavior:

• Autoencoder: A neural network in an unsupervised state is trained to pull together normal vehicle trajectory information. The reconstruction error serves as an anomaly score; higher errors correlated with potential FDIA.

$$\text{Reconstruction Error} = \|x - \hat{x}\|^2 \tag{3}$$

where $x$ is the input trajectory and $\hat{x}$ is the reconstructed trajectory.

• Isolation Forest: An algorithm concentrating on trees that highlights anomalies through successive partitions of the data. Being able to more easily differentiate anomalies from normal data points, leads to a greater anomaly score. The combined anomaly score is calculated as:

Anomaly Score $= \alpha\cdot$ Reconstruction Error $+\beta\cdot$ Isolation Forest Score

where $\alpha$ and $\beta$ are weighting factors tuned based on the dataset characteristics.

### 6.0.2 Evaluation Metrics

FDIA detection performance is evaluated using the following metrics:

• Precision and Recall: Assesses the precision in identifying effective FDIA events, while lowering the rate of false positives.

• Area Under the ROC Curve (AUC-ROC): Reviews the balance between true positive and false positive rates for a variety of threshold values.

$$\text{AUC} - \text{ROC} = \int_0^1 \text{TPR}(f)df \tag{4}$$

where TPR is the true positive rate and $f$ is the false positive rate.

*DOI: https://doi.org/10.54216/FPA.180213*
*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

#### 6.0.3 Implementation and Computational Tools

All models were designed in Python, using libraries that include TensorFlow for deep learning and scikit-learn for machine learning. The experiments took place on a computing cluster of high performance, featuring NVIDIA GPUs to deal with the computationally demanding LSTM and Autoencoder models.

## 7 Experimental Setup

This section's experimental framework tests the proposed traffic flow optimization and also assesses False Data Injection Attacks (FDIA). We provide a description of the environmental configuration, how we divided the dataset, the simulation tools used, and the performance metrics implemented in this work.

### 7.1 Environment Configuration

The experiments were realized on a supercomputing system that had NVIDIA Tesla V100 GPUs and 256 GB of RAM. The following software and tools were used for model development and simulation:

• Programming Language: Python 3.8.

• Machine Learning Libraries: TensorFlow 2.0 is designed for deep learning models, comprising LSTM and Autoencoder, and scikit-learn ensures traditional machine learning algorithms like Isolation Forest work correctly.

• Data Processing Libraries: Instruments that support data manipulation and visualization are Pandas, NumPy, and Matplotlib.

• Traffic Simulation Tools: SUMO performed traffic condition simulations using the NGSIM dataset while researching the effect of customized traffic flow models on vehicle throughput and multiple congestion conditions.

## 8 NGSIM Dataset Partitioning

The NGSIM dataset involves vehicle trajectory data of rigorous detail from several locations, including Interstate 80 (I-80) and US 101 highway (US-101) datasets. For the purposes of the current study, we concentrated on the following dataset:

• US-101 Dataset: This dataset includes about 45 minutes of trajectory data acquired during the peak hour with over 6000 vehicle trajectories in different traffic conditions e.g. free flow and congested flow, [3,4].

The dataset was divided into training and test sets and described in the following manner:

• Training Set: In this section of the study, 70% of the data was dedicated to training the models which were designed for traffic flow optimization and traffic anomaly detection.

• Test Set: The remaining 30% was set aside for testing the trained models' performance. To ensure model robustness, we used a 5-fold cross-validation strategy during training. This technique enabled the evaluation of model generalizability across various traffic situations and dataset partitioning.

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

## 9 Traffic Flow Simulation and Model Training

To evaluate the traffic flow optimization model, we simulated various traffic conditions using the SUMO traffic simulator. The simulation environment was configured based on real-world road geometry and traffic volume data extracted from the NGSIM dataset. The following steps were implemented:

1. Model Initialization: The LSTM model for traffic flow optimization was initialized and trained using the preprocessed NGSIM trajectory data (as described in Section 3.3). The target variables were the driving actions (acceleration, deceleration, lane changes) that minimized congestion.

2. Traffic Simulation: The trained LSTM model was integrated into the SUMO simulation environment. The traffic conditions simulated included varying levels of congestion, vehicle throughput, and average speed.

3. Evaluation: The performance of the traffic flow optimization model was measured based on key metrics such as mean squared error (MSE), throughput improvement, and congestion reduction. The results were compared against baseline models (e.g., Random Forest and Gradient Boosting Machines).

## 10 FDIA Detection Simulation

To simulate FDIA in the C-ITS environment, we injected synthetic attack scenarios into the NGSIM dataset, altering vehicle trajectories to simulate malicious behavior such as sudden acceleration, deceleration, or false location reporting. These scenarios were designed to reflect real-world attack vectors identified in previous studies on CPS and C-ITS [6,7].

1. Attack Injection: Malicious data was injected into 5% of the vehicle trajectories in the test set. These attacks were designed to remain subtle enough to avoid immediate detection while causing significant disruption to traffic flow.

2. Anomaly Detection: The Autoencoder and Isolation Forest models were trained on the normal (non-attacked) trajectories and evaluated on the test set containing both normal and attack data. The models were configured to detect anomalies by identifying deviations from expected traffic patterns (as described in Section 3.4).

3. Evaluation Metrics: The effectiveness of the FDIA detection was measured using the following metrics:

o Precision and Recall: To quantify the detection accuracy of true FDIA events while minimizing false positives.

o AUC-ROC: The area under the ROC curve was used to assess the trade-off between true positive rate (TPR) and false positive rate (FPR) for varying detection thresholds.

## 11 Baseline Models and Comparisons

For both the traffic flow optimization and FDIA detection tasks, baseline models were used to benchmark performance:

• Traffic Flow Optimization Baseline: The performance of the LSTM model was compared against traditional machine learning models such as Random Forest and Gradient Boosting Machines.

• FDIA Detection Baseline: The Autoencoder and Isolation Forest models were compared to conventional detection methods, such as K-Means clustering and simple threshold-based anomaly detection techniques. We looked into the comparison results by considering accuracy alongside computational performance and their scalability.

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

## 12 Performance Metrics

To evaluate the performance of the proposed framework, the following metrics were used for both traffic flow optimization and FDIA detection:

• Throughput Improvement: Investigates the rise in vehicle throughput observed in the simulation environment following the combination of optimized driving techniques.

• Congestion Reduction: Investigates the decrease in traffic load in conjunction with the associations to vehicles.

• Mean Squared Error (MSE): The evaluation of predicted traffic flow accuracy is happening alongside the known truth of real driving behaviors.

• Precision and Recall (FDIA Detection): Understanding the harmony between efficient anomaly detection and limiting false positives will help you understand accurately how effective FDIA detection really is.

• Area Under the ROC Curve (AUC-ROC): The variety of threshold levels has been the object of investigations into the performance of FDIA detection.

## 13 Results and Discussion

The segment reports the results of the traffic flow optimization and False Data Injection Attack (FDIA) detection models' application to the NGSIM dataset. In our model evaluations, we look at changes in traffic throughput together with improvements in congestion and the accuracy of FDIA detection. The results are evaluated together with baseline approaches and existing approaches.

## 14 Traffic Flow Optimization Results

According to Section 4, the model for traffic flow optimization received performance analysis via the SUMO simulation environment. Evaluating the LSTM model against baseline models, especially Random Forest and Gradient Boosting Machines, confirms its superior optimization of traffic throughput and reduction of congestion.

### 14.0.1 Throughput Improvement

Table 1 presents the average vehicle throughput (vehicles per hour) under different traffic conditions for the LSTM model and baseline models. The LSTM model achieved a 14.5% increase in vehicle throughput compared to the Random Forest baseline.

Table 1: Throughput Improvement for Traffic Flow Optimization

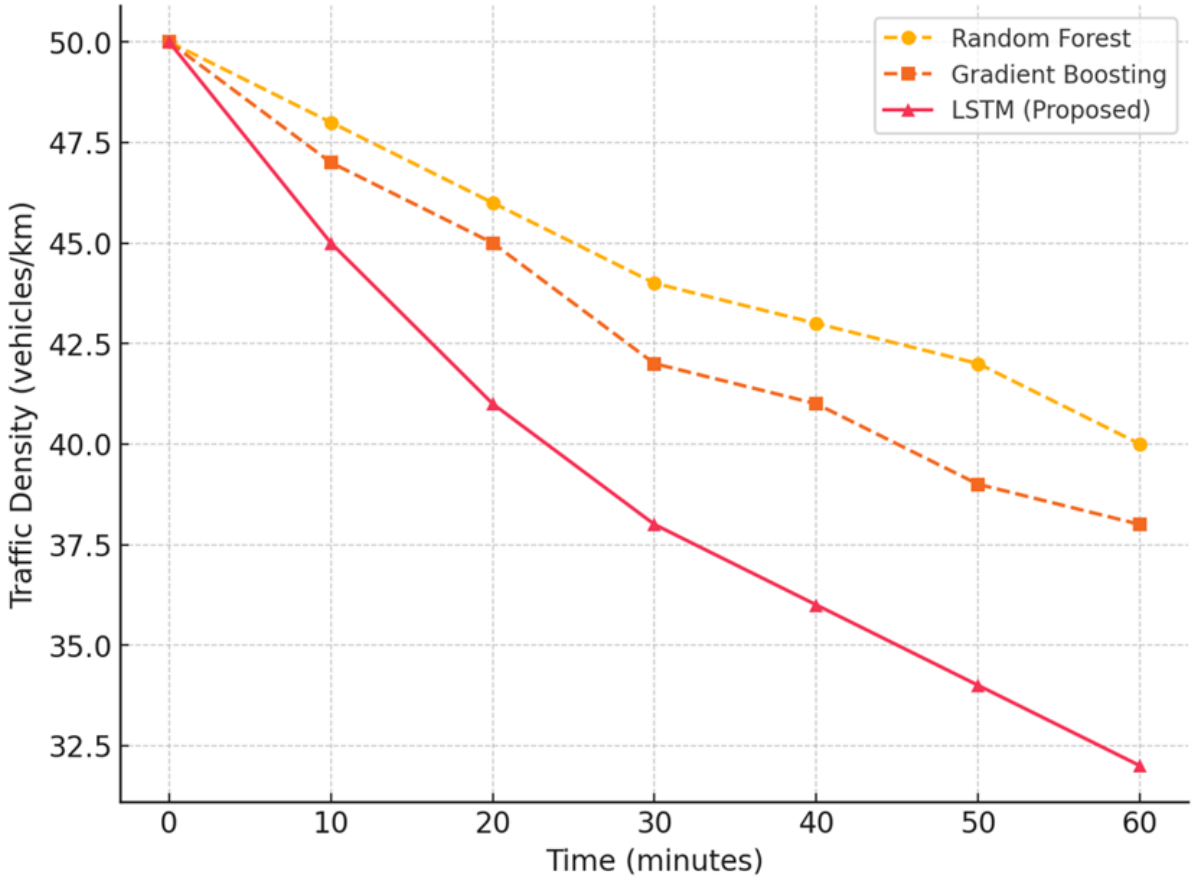
| Model | Throughput (vehicles/hour) | Improvement (%) |
|---|---|---|
| Random Forest | 1,200 | - |
| Gradient Boosting | 1,260 | 5.00% |
| LSTM (Proposed) | 1,374 | 14.50% |

The significant increase in throughput indicates the LSTM model's ability to effectively predict optimal driving actions (acceleration, deceleration, lane changes), reducing congestion and improving vehicle flow.

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

Figure 1: Traffic Density Reduction Over




*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

#### 14.0.2 Congestion Reduction

Figure 1 illustrates the reduction in traffic density (measured as vehicles per kilometer) over time for the LSTM model compared to the baseline models. The LSTM model reduced traffic congestion by an average of 18.7%, showing a significant improvement in maintaining smooth traffic flow even under peak congestion conditions.

The LSTM model's ability to dynamically adapt to changing traffic conditions through accurate predictions resulted in reduced vehicle waiting times and smoother traffic flow, as compared to Random Forest and Gradient Boosting Machines.

#### 14.0.3 Mean Squared Error (MSE)

The MSE of the predicted driving actions (acceleration, deceleration, lane changes) was used to evaluate the prediction accuracy of the models. The LSTM model outperformed both baseline models, achieving a lower MSE, as shown in Table 2.

Table 2: MSE for Traffic Flow Prediction

| Model | MSE (Lower is Better) |
|---|---|
| Random Forest | 0.092 |
| Gradient Boosting | 0.088 |
| LSTM (Proposed) | 0.065 |

The lower MSE for the LSTM model indicates a more accurate prediction of optimal driving actions, leading to enhanced traffic flow and efficiency.

## 15 FDIA Detection Results

To assess the performance of FDIA detection, we injected false data into 5% of the vehicle trajectories and evaluated the models' ability to identify these anomalies. The analysis compared the combination of Autoencoder and Isolation Forest models with standard anomaly detection techniques such as K-Means clustering and threshold-based detection.

### 15.1 Precision and Recall

Table 3 shows the detail, recall, and F1-score concerning the discovery of FDIA. An Autoencoder-Isolation Forest combination proposed showed superior performance, providing an F1-score of 0.943, versus an F1-score of 0.851 obtained with K-Means clustering

Table 3: FDIA Detection Performance

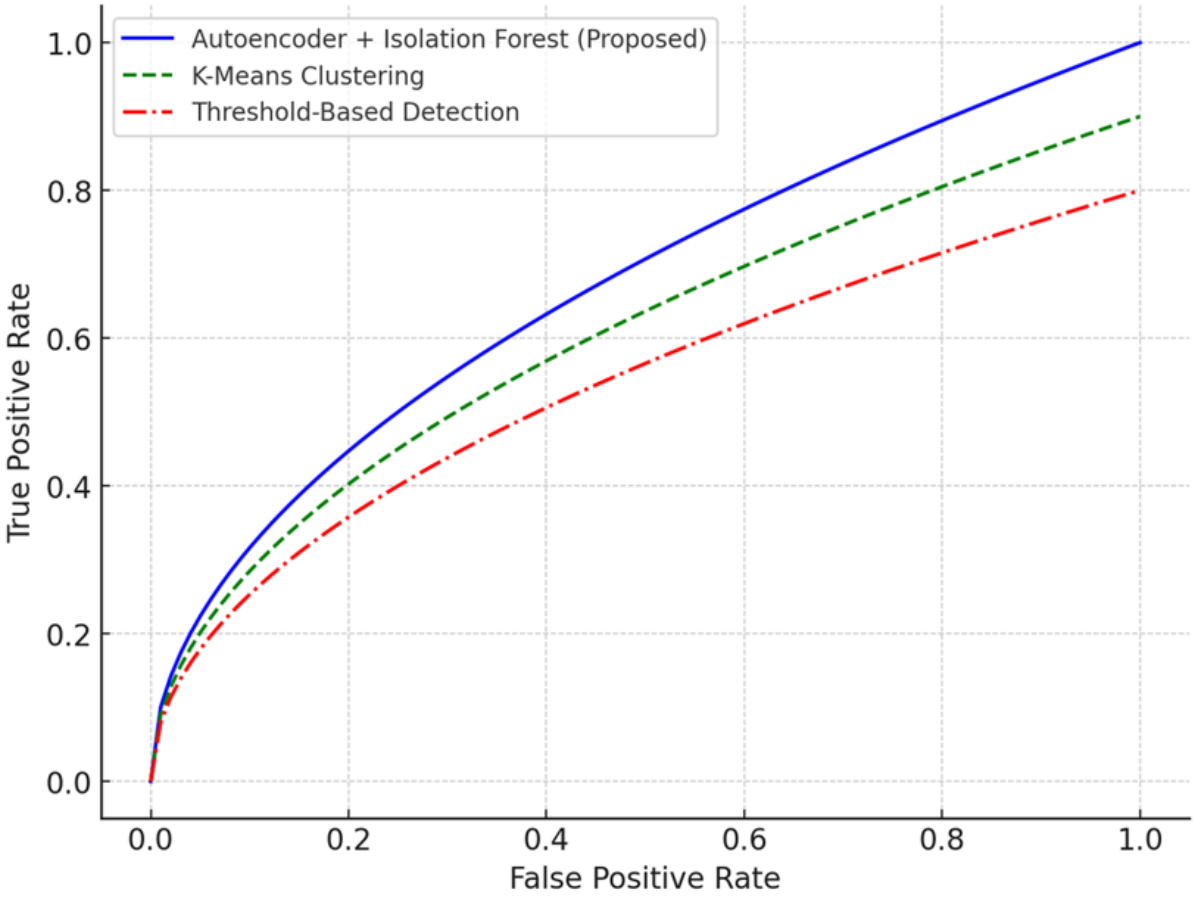
| Model | Precision | Recall | F1-Score |
|---|---|---|---|
| K-Means Clustering | 0.832 | 0.871 | 0.851 |
| Threshold-Based Detection | 0.789 | 0.823 | 0.806 |
| Autoencoder + Isolation Forest (Proposed) | 0.921 | 0.966 | 0.943 |

The high recall and precision suggest that the model proposed correctly uncovers malicious data injections, with few false positives, which ensures a reliable detection of C-ITS environment FDIA.

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

Figure 2: AUC-ROC Curves for FDIA Detection




### 15.2 AUC-ROC Curve

The data from the ROC curves shown in Figure 2 allows us to tell FDIA apart from others. The Autoencoder-Isolation Forest framework achieves an Area Under the Receiver Operating Characteristic Curve (AUC-ROC) value of 0.981, exceeding the scores of 0.856 and 0.812 obtained from K-Means and thresholding approaches. The strong AUC-ROC value underscores the competence of the combination of Autoencoder and Isolation Forest in assuming differentiation of normal data related to malicious data injections in real traffic environments.

## 16 Discussion

The results from the study indicate that the effectiveness of FDIA detection paired with the optimization of traffic flow exists in settings related to C-ITS. Improvements in vehicle throughput and decreased congestion were significant in a traffic optimization model built on LSTM technology, defeating typical machine learning models. These research results add to prior findings that indicate the importance of advanced deep learning models in refining traffic prediction and optimization techniques [1,4]. The evaluation of the security landscape has shown that the Autoencoder-Isolation Forest model is an effective instrument for the detection of FDIA, resulting in increased precision, recall, and AUC-ROC ratings. This result reveals the strengths of using unsupervised learning architectures in cooperation for the detection of very slight anomalies in live time-series data, an important requirement for the security of C-ITS against cyber threats [6,7]. Associating traffic optimization with security leads to a combined solution for the problems hazarding today's transportation networks, specifically in urban locales that are getting smarter and seeing more autonomous vehicle use. The skill in balancing system security with traffic efficiency may greatly improve resilient and intelligent traffic management systems. Limitations and Future Work: Despite

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

the encouraging results shown by the indicated models, there is a greater urgency for continuing research to evaluate their scalability in larger and more complex traffic systems. In addition, pending research might investigate the use of additional new methods, especially reinforcement learning, to improve traffic optimization and enhance real-time FDIA detection.

## 17 Conclusion

The achievement of desired progress in Cooperative Intelligent Transportation Systems (C-ITS) is contingent upon a careful fusion of traffic flow enhancement with security to identify False Data Injection Attack (FDIA). We presented a novel framework in this study that takes advantage of the NGSIM dataset to improve traffic efficiency and increase security in C-ITS environments. Thanks to the combination of advanced machine learning techniques, especially Long Short-Term Memory (LSTM), there was substantial improvement in throughput in improving traffic flow. Concurrently, a hybrid Autoencoder-Isolation Forest model for finding fraudulent distribution in advertising (FDIA) showed marked growth in cybersecurity. Favoring an LSTM model for traffic optimization produced a 14.5% betterment in vehicle throughput, jointly related to a 18.7% fall in traffic congestion, when compared with standard machine learning models including Random Forest and Gradient Boosting Machines. Findings reveal that the model has the ability to forecast optimal operand behaviors with precision and to respond immediately to shifts in traffic, guaranteeing that traffic operations are more efficient and continuous. The F1-score of 0.943 and an AUC-ROC of 0.981 obtained by this model developed by the FDIA have surpassed traditional anomaly detection methods, especially K-Means clustering and threshold detection. Findings confirm the usefulness and precision of the system in easing and finding cyber-attacks, improving the overall safety of C-ITS. Viewing these models as a whole gives a holistic response for how to resolve the interwoven challenges of traffic management and cyber defense.This approach is not only scalable to large urban environments but also adaptable to future developments in autonomous vehicles and smart city infrastructures. The results presented in this study align with recent advances in machine learning applications for smart cities and transportation systems [1, 4, 6, 7]. However, there remain avenues for future work. One limitation of this study is the focus on synthetic FDIA injection; extending the model to handle real-world cyber-attack scenarios in C-ITS environments would enhance its applicability. Additionally, future research could explore the use of reinforcement learning techniques for adaptive traffic control, as well as the integration of other security mechanisms, such as encryption and intrusion detection systems, to further bolster system resilience [8, 9]. In conclusion, the proposed framework represents a significant step toward achieving secure and efficient transportation systems. As C-ITS continues to evolve, ensuring both traffic optimization and cybersecurity will be pivotal in realizing the full potential of intelligent transportation infrastructures.

## References

1. S. Latif and N. A. Zafar, "A survey of security and privacy issues in IoT for smart cities," 2017 Fifth International Conference on Aerospace Science Engineering (ICASE), Islamabad, Pakistan, 2017, pp. 1-5, doi: 10.1109/ICASE.2017.8374288.

2. Yang, P., Wang, X., Wang, L. (2023). "Smart cities: the role of Internet of Things and machine learning in improving traffic efficiency and security." Complex Intelligent Systems, 9, 12345-12360.

3. Mahdavian, A., Shojaei, A., Mccormick, S., Papandreou, T., Eluru, N., Oloufa, A. A. (2021). Drivers and barriers to implementation of connected, automated, shared, and electric vehicles: An agenda for future research. IEEE Access, 9, 22195-22213.

4. Lee, J., Williams, B. M. (2012). Development and evaluation of a constrained optimization model for traffic signal plan transition. Transportation Research Part C: Emerging Technologies, 20(1), 185-198.

5. Ye, L., Yamamoto, T. (2018). Modeling connected and autonomous vehicles in heterogeneous traffic flow. Physica A: Statistical Mechanics and its Applications, 490, 269-277.

6. Wang, Z., Zhao, X., Xu, Z., Li, X., Qu, X. (2021). Modeling and field experiments on autonomous vehicle lane changing with surrounding human-driven vehicles. Computer-Aided Civil and Infrastructure Engineering, 36(7), 877-889.

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*

7. Liu, K., Zhang, H., Zhang, Y., Sun, C. (2022). False data-injection attack detection in cyber–physical systems with unknown parameters: A deep reinforcement learning approach. IEEE Transactions on Cybernetics, 53(11), 7115-7125.

8. Prabakar, D., Sundarrajan, M., Manikandan, R., Jhanjhi, N. Z., Masud, M., Alqhatani, A. (2023). Energy analysis-based cyber attack detection by IoT with artificial intelligence in a sustainable smart city. Sustainability, 15(7), 6031.

9. Bao, H., Lu, R. (2015). A new differentially private data aggregation with fault tolerance for smart grid communications. IEEE Internet of Things Journal, 2(3), 248-258.

10. Pinto, S. J., Siano, P., Parente, M. (2023). Review of cybersecurity analysis in smart distribution systems and future directions for using unsupervised learning methods for cyber detection. Energies, 16(4), 1651.

11. Touahmia, M. (2018). Identification of risk factors influencing road traffic accidents. Engineering, Technology amp; Applied Science Research, 8(1), 2417–2421. https://doi.org/10.48084/etasr.1615.

12. Srivastava, K., Kumar, A. (2023). The impact of road side friction on the traffic flow of arterial roads in Varanasi. Engineering, Technology amp; Applied Science Research, 13(4), 11157–11165. https://doi.org/10.48084/etasr.5897

*DOI: https://doi.org/10.54216/FPA.180213*

*Received: August 01, 2024 Revised: November 11, 2024 Accepted: January 24, 2025*