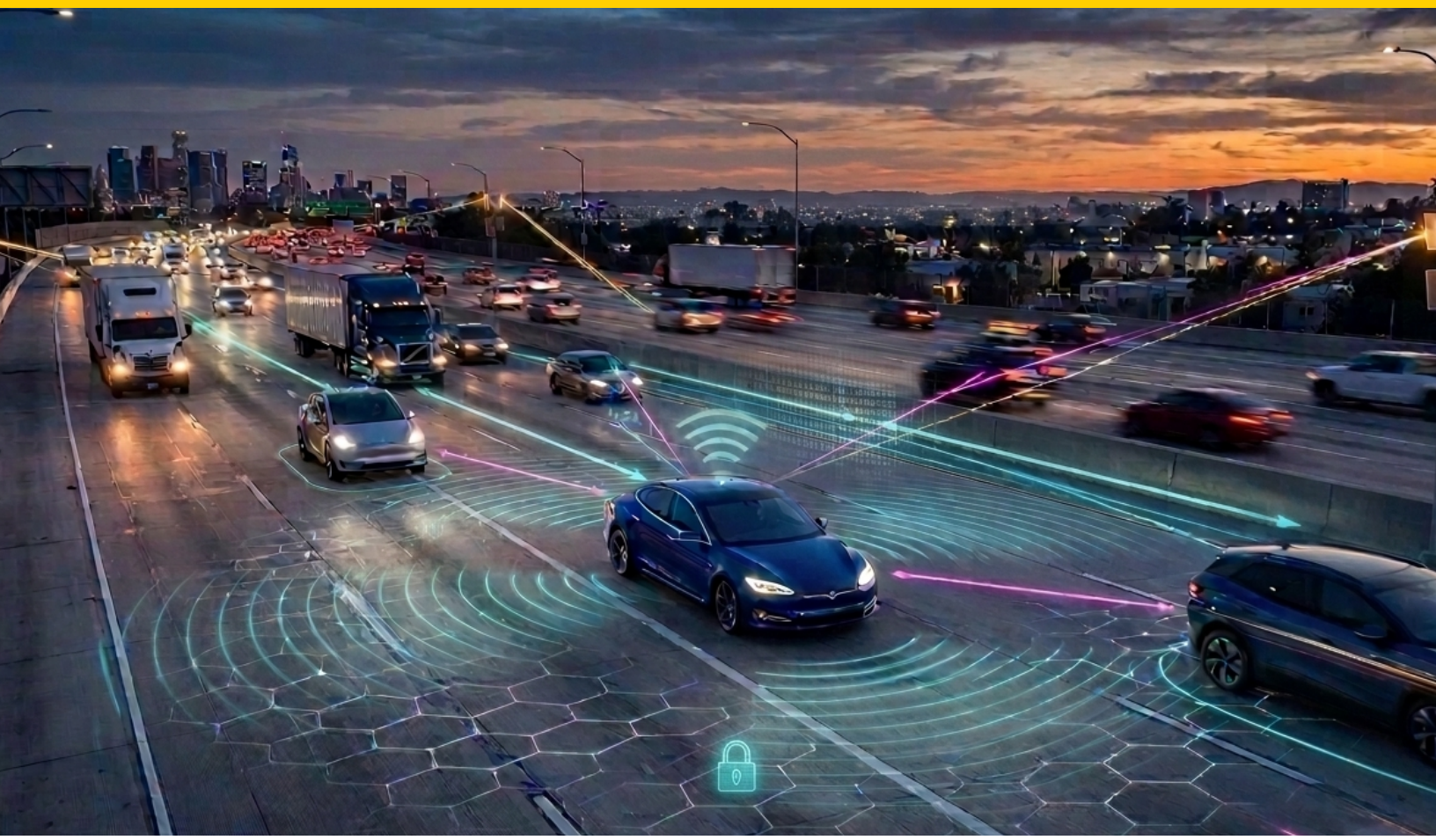
CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Final Report GR000034957
March 2026



Security Defense of Transportation Networks against Cyberattacks: A Physics-Informed AI Approach

Yongju Kim
Sikai Chen
Soyoung (Sue) Ahn
Madhav Chitturi
David A. Noyce

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Report No. GR000034957
Project Start Date: 04/01/2024
Project End Date: 03/31/2026

March 2026

Security Defense of Transportation Networks against Cyberattacks: A Physics-Informed AI Approach

Yongju Kim
Graduate Researcher

Sikai Chen
Assistant Professor

Soyoung (Sue) Ahn
Professor

Madhav Chitturi
Scientist

David A. Noyce
Professor

University of Wisconsin Madison

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

DISCLAIMER

Funding for this research was provided by the Center for Connected and Automated Transportation under Grant No. 69A3551747105 of the U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology (OST-R), University Transportation Centers Program. 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 under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Suggested APA Format Citation:

Kim, Y., Chen, S., Ahn, S., Chitturi, M., Noyce, D. (2026). Security Defense of Transportation Networks against Cyberattacks: A Physics-Informed AI Approach, CCAT Report #GR000034957, The Center for Connected and Automated Transportation, University of Wisconsin-Madison, Madison, WI.
Cover image generated using Google's Gemini Nano Banana 2.

Contacts

For more information:

Sikai Chen
1415 Engineering Drive
Madison, WI 53706
Phone: +1-213-806-0141
Email: sikai.chen@wisc.edu

CCAT
University of Michigan
Transportation Research Institute
2901 Baxter Road
Ann Arbor, MI 48152
umtri-ccat@umich.edu
(734) 763-2498

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Technical Report Documentation Page

1. Report No. GR000034957	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Security Defense of Transportation Networks against Cyberattacks: A Physics-Informed AI Approach		5. Report Date March 2026 6. Performing Organization Code N/A
7. Author(s) Yongju Kim, Sikai Chen, Sue Ahn, Madhav Chitturi, David Noyce		8. Performing Organization Report No. N/A
9. Performing Organization Name and Address Center for Connected and Automated Transportation University of Michigan Transportation Research Institute 2901 Baxter Road, Ann Arbor, MI 48109 University of Wisconsin-Madison 500 Lincoln Dr, Madison, WI 53706		10. Work Unit No. 11. Contract or Grant No. Contract No. 69A3552348305
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE Washington, DC 20590		13. Type of Report and Period Covered Final report (April 2024 – March 2026) 14. Sponsoring Agency Code OST-R
15. Supplementary Notes Conducted under the U.S. DOT Office of the Assistant Secretary for Research and Technology's (OST-R) University Transportation Centers (UTC) program.		
16. Abstract Connected and automated vehicles (CAVs) rely on digitally acquired traffic state information for real-time maneuver decisions, making them vulnerable to cyberattacks that manipulate perceived traffic states or communication channels. Such manipulation can subtly alter vehicle behavior and, in interactive traffic environments, may propagate disturbances beyond the attacked vehicle. The objective of this study is to develop frameworks for (i) detecting abnormal lane-changing under cyberattacks and (ii) enabling robust decision-making under adversarial conditions. For detection, a physics-guided neural network integrating a game-theoretic lane-changing model with LSTM learning is developed and evaluated using the I-24 MOTION dataset under simulated false data injection and denial-of-service attacks. Results show improved prediction accuracy and effective detection of falsified lane-changing behaviors. For robust control, a hierarchical adversarial reinforcement learning framework is proposed for discretionary lane-changing under bounded perturbations. Simulation results indicate improved traffic efficiency while maintaining safety under worst-case perturbations. These findings provide a modeling foundation for cybersecurity-aware monitoring and robust tactical decision-making in CAV systems.		

UMTRI · The University of Akron · Central State University · Illinois Center for Transportation · University of Minnesota · Northwestern · Purdue University · Washtenaw Community College · Wisconsin University of Wisconsin–Madison

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

17. Key Words Connected and automated vehicle, Cybersecurity, Lane-changing, Anomaly detection, Physics-guided neural network, Adversarial reinforcement learning		18. Distribution Statement No restrictions.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 63	22. Price
Form DOT F 1700.7 (8-72)	Reproduction of completed page authorized		

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AI	Artificial Intelligence
AV	Automated Vehicle
CAV	Connected and Automated Vehicle
DoS	Denial-of-Service
FDI	False Data Injection
GBM	Gradient Boosting Machine
HDV	Human Driven Vehicle
IDM	Intelligent Driving Model
LC	Lane-Changing
LSTM	Long Short-Term Memory
MDP	Markov Decision Process
PGNN	Physics-Guided Neural Network
PINN	Physics-Informed Neural Network
PPO	Proximal Policy Optimization
RL	Reinforcement Learning
RSU	RoadSide Unit
SVM	Support Vector Machine
TTC	Time-to-Collision
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything

CHAPTER 1. INTRODUCTION

1.1. Study Background

Connected and automated vehicles (CAVs) are widely regarded as a cornerstone of next-generation transportation systems, with the potential to improve safety, efficiency, and sustainability. A central source of these benefits lies in vehicle automation, which enables consistent and precise execution of driving tasks. By reducing reliance on human perception and reaction, automation can smooth vehicle operations and improve roadway capacity. Prior studies have shown that automated systems can reduce energy consumption, enhance traffic stability, and improve the allocation of limited urban resources such as parking space (Loke, 2019; Li et al., 2023).

Beyond automation, connectivity further amplifies these benefits by enabling information sharing and coordination among vehicles and infrastructure. Through real-time communication, CAVs can optimize routing decisions, coordinate maneuvers, and reduce unnecessary idling, thereby contributing to lower fuel consumption and greenhouse gas emissions (Sun et al., 2022; Yu and Long, 2022). Connectivity also allows operational decisions to extend beyond individual vehicles. For instance, shared traffic and infrastructure information can support parking guidance and demand-aware navigation, reducing circulation traffic and congestion in dense urban areas (S. Wang et al., 2021; Zhang et al., 2023).

The effectiveness of both automation and connectivity relies on advanced sensing technologies, artificial intelligence (AI)-driven decision-making systems, and communication systems. High-resolution sensors provide continuous and detailed perception of the surrounding environment, while onboard computing and learning-based algorithms enable vehicles to interpret this information and respond to evolving traffic conditions in real time. This combination allows CAVs to anticipate hazards, resolve conflicts proactively, and react with greater speed and consistency than human drivers, thereby reducing crash risk and minimizing traffic disruptions (Zheng et al., 2020; Wang et al., 2022, 2023b). However, because vehicle behavior is increasingly determined by digital information streams, this cyber-physical integration also introduces structural vulnerabilities. Disruption or manipulation of these information channels can directly influence vehicle perception and control, exposing CAVs to cyber threats that are fundamentally different from those faced by conventional human-driven vehicles (HDVs).

In this context, cyberattacks on CAVs refer to intentional and malicious interference with vehicle cyber components, including sensors, software modules, control mechanisms, or communication links. Rather than causing overt system failures, such attacks often operate by subtly manipulating the information used for decision-making. By corrupting sensor inputs, altering communicated traffic states, or perturbing control commands, an attacker can distort a vehicle's internal representation of its environment without violating basic operational constraints. Because CAV decision-making is highly sensitive to key state variables, such as speed, spacing, and acceleration, even small malicious perturbations can alter vehicle responses in systematic ways. Prior studies have shown that slight manipulations of acceleration or spacing information can destabilize traffic flow, trigger stop-and-go waves, and elevate crash risk (Y. Wang et al., 2021; Wang et al., 2023a). In CAV environments, these effects may propagate through vehicle

interactions, amplifying local disturbances into network-level inefficiencies and safety hazards. As a result, cyberattacks pose a direct threat not only to individual vehicle safety but also to the reliability and efficiency of future transportation systems (Sedjelmaci et al., 2019; Sun et al., 2023).

These characteristics fundamentally challenge conventional cybersecurity and anomaly-detection approaches. Because CAV behavior is governed by physical constraints and traffic interaction mechanisms, malicious actions must be distinguished from legitimate behavioral variability arising from normal traffic dynamics. Detection methods that rely solely on data-driven pattern recognition may struggle to make this distinction, while purely physics-based models often lack the flexibility to capture complex, context-dependent behaviors. This tension motivates the use of physics-informed AI, or more broadly physics-structured learning, which embeds physical interaction principles and interpretability into learning-based models. Such integration enables both reliable characterization of nominal behavior for anomaly detection and robust decision-making under cyberattacks.

Motivated by these considerations, this research develops modeling frameworks for cyberattack detection and robust decision-making tailored to the operational characteristics of CAVs. The framework leverages physics-informed AI to establish a behavioral prediction framework for nominal vehicle behavior and to detect deviations indicative of malicious manipulation. Building upon this foundation, the project investigates robust decision-making strategies under adversarial conditions, forming a methodological basis for enhancing the cybersecurity and resilience of CAV operations.

1.2. Problem Statement

The increasing reliance of CAVs on sensor-driven perception and automated decision-making reshapes the risk profile of vehicle operation. Most CAV control and monitoring systems implicitly assume that perceived traffic states accurately represent the surrounding environment. When this assumption is violated, cyberattacks that manipulate sensor data can directly affect vehicle decisions without causing explicit system failures. Such manipulation is particularly critical because perception underlies all downstream maneuver decisions. Even small perturbations in perceived speeds or gaps may shift decision boundaries and induce systematic changes in vehicle behavior. These deviations can propagate through vehicle interactions, potentially amplifying local disturbances into traffic flow-level impacts.

A key challenge in detecting such attacks lies in distinguishing maliciously induced deviations from normal variability in traffic systems. Vehicle trajectories naturally fluctuate due to heterogeneous driving behaviors and changing traffic conditions. Purely data-driven detection methods may struggle to differentiate these benign variations from cyber-induced anomalies and lack interpretability in safety-critical contexts. At the same time, purely physics-based models, while interpretable, often lack the flexibility to capture complex and context-dependent CAV behaviors.

Beyond detection, cyberattacks also pose challenges in existing vehicle decision-making strategies. Even when abnormal behavior is suspected, many decision and control frameworks presume reliable perception and are not designed to operate under corrupted information conditions. Therefore, ensuring robust vehicle operation under cyberattacks requires decision-making models that explicitly account for adversarial observations while maintaining physical

feasibility and safety.

Together, these challenges motivate the development of principled approaches for both cyberattack detection and attack-aware decision-making, grounded in physics-informed AI and tailored to the operational characteristics of CAVs.

Within this context, this study focuses on lane-changing as a representative safety-critical tactical maneuver. Lane-changing decisions depend critically on perceived gaps and relative speeds, making them particularly sensitive to observation perturbations. Even small perturbations can shift maneuver timing or alter lane-selection outcomes. Because such maneuvers directly affect surrounding vehicles, cyber-induced deviations may propagate beyond the subject CAV and generate traffic-level disturbances. Therefore, detecting falsified lane-changing decisions and designing robust lane-changing policies under adversarial conditions are central to ensuring both safety and operational efficiency in CAV environments.

1.3. Objectives of the Study

The overarching objective of this project is to improve the cybersecurity and operational resilience of CAVs by developing principled models for cyberattack detection and attack-aware decision-making.

Specifically, this study aims to:

- Develop a cyberattack detection model that identifies anomalous CAV behaviors resulting from malicious manipulation of sensor observations.
- Incorporate physics-based traffic interaction principles into learning-based models to predict nominal CAV behavior for anomaly detection.
- Develop a robust vehicle decision-making model that enables discretionary lane-changing under adversarial observation conditions while balancing safety, efficiency, and traffic stability.

1.4. Study Approach

To address the identified challenges, this project adopts a modeling-oriented approach that develops two complementary components: (i) a physics-guided anomaly detection model and (ii) a robust decision-making model under adversarial observation perturbations. Both components are grounded in vehicle interaction principles while leveraging learning-based techniques to capture behavioral complexity.

First, cyberattack detection is addressed through the development of a physics-guided neural network for identifying falsified lane-changing decisions. The detection model integrates a game-theoretic lane-changing decision model with a long short-term memory (LSTM) network. The physics-based component captures interaction mechanisms and gap-acceptance structure, while the neural network captures temporal dependencies in trajectory data. The model detects anomalies through cumulative discrepancies between predicted decision probabilities and observed maneuver execution. This approach enables the identification of cyberattacks without relying solely on purely data-driven classification.

Second, cyberattack defense is addressed by developing a hierarchical adversarial

reinforcement learning framework for robust discretionary lane-changing. The framework explicitly models adversarial observation perturbations and learns maneuver decisions that remain reliable under worst-case bounded sensor disturbances. The upper-level policy determines the maneuver decision under perturbed observations, while a lower-level Stackelberg execution module resolves lane-changing timing and longitudinal control by accounting for interaction with surrounding vehicles. Through multi-objective reward design, the robust decision-making model balances safety, self-efficiency, and flow-level stability under adversarial conditions.

Together, these two modeling components address complementary aspects of CAV cybersecurity: detection of cyber-induced behavioral anomalies and robust tactical control under adversarial information environments. The study emphasizes interpretable, physics-informed learning frameworks that remain consistent with traffic interaction principles while accommodating complexity.

1.5. Organization of this Report

This report is organized as follows: **Chapter** 2 presents the development of a physics-guided neural framework for detecting falsified lane-changing behavior under cyberattacks, including model formulation, training methodology, and evaluation results. **Chapter** 3 introduces a hierarchical adversarial reinforcement learning framework for robust and flow-efficient discretionary lane-changing under adversarial observation perturbations. **Chapter** 4 provides overall conclusions and discusses directions for future research. **Chapter** 5 presents a synopsis of performance indicators, and **Chapter** 6 outlines the study outcomes and outputs.

CHAPTER 2. DETECTING FALSIFIED LANE-CHANGING BEHAVIOR IN CONNECTED AUTOMATED VEHICLES USING PHYSICS-GUIDED NEURAL NETWORKS

2.1. Introduction

The adoption of connected automated vehicles (CAVs) in modern transportation systems promises substantial benefits, including improved traffic efficiency and safety (Shladover et al., 2012; Ye and Yamamoto, 2019). However, connectivity that enables real-time communication and coordination also creates new vulnerabilities that malicious actors can exploit to pose threats to traffic systems. As CAV penetration increases, so does the risk of cyberattacks, raising critical concerns about the security and resilience of CAV-based transportation networks. Cyberattacks − such as sensor spoofing, communication disruptions, data poisoning, or software infiltration − can extend beyond individual vehicles and impact the broader transportation system (Sedjelmaci et al., 2019; Wang et al., 2024). Even minor disruptions, like falsified accelerations, can propagate through traffic as stop-and-go waves, amplifying disturbances and elevating crash risks (Zheng et al., 2010; Wang et al., 2023).

In response, detecting falsified vehicle behavior has emerged as a critical research area. Existing studies have explored network-level anomalies such as route spoofing (Shoukry et al., 2018), spoofed trajectory data in vehicular communication (Amoozadeh et al., 2015; Huang et al., 2021), and manipulated vehicular sensor data (Vyas et al., 2023) at the vehicle level. Most vehicle-level studies, however, focus on car-following behavior with adaptive cruise control, but overlook falsified lane changing (LC) trajectories, despite their serious impacts on traffic flow and safety. LC maneuvers represent a fundamental yet complex aspect of traffic dynamics. Although LC can harmonize flow and speed across lanes, it often disrupts traffic by reducing bottleneck discharge rates (Cassidy and Rudjanakanoknad, 2005; Laval and Daganzo, 2006; Laval and Leclercq, 2008) and triggering stop-and-go oscillations (Mauch and J. Cassidy, 2002; Ahn and Cassidy, 2007). Moreover, improper LC maneuvers can pose significant safety risks, leading to sideswipe crashes that often result in serious injuries or deaths. In 2023, sideswipe crashes accounted for 23.0% of all crashes, including 11.5% of injury crashes and 7.8% of fatal crashes (NCS, 2025). These statistics underscore the need for the timely and reliable detection of abnormal LC behavior. Yet, detecting abnormal LC behavior remains challenging due to the undisclosed control logic of automated vehicles (AVs) and the complexity, heterogeneity, and variability of human decision-making.

Only a limited number of studies have explored abnormal LC detection, mainly using machine learning algorithms such as support vector machine (SVM), dynamic Bayesian networks, and gradient boosting machine (GBM). For instance, Ramyar et al. (2016) used one-class SVM classification to detect dangerous LC instances using segmented vehicle information rather than complete LC trajectories. Xu et al. (2023) used lightGBM to classify normal and abnormal LC behavior using relative speeds and gaps of the subject vehicle and surrounding vehicles. Furthermore, Fan et al. (2022) combined three unsupervised algorithms, including autoencoder,

one-class SVM, and t-distributed stochastic neighbor embedding, to identify driver-specific anomalous LC events. These existing studies rely on data-driven approaches to address the complexity and variability of LC behavior. However, they primarily focus on classifying LC behavior directly from data without conceptualizing the LC maneuver as a behavioral decision-making process. Physics-guided models can complement pure data-driven models by embedding vehicle dynamics and interaction principles fundamental to LC maneuvers. This enables a more comprehensive understanding of both maneuver feasibility and the underlying decision-making, essential for accurately detecting malicious and unsafe LC behavior.

Predicting LC decision is a foundational component of identifying anomalous LC events, and extensive literature has examined this subject. Previous models can be categorized into physics-based paradigms (e.g., rule-based, discrete choice-based, game-theoretic models), artificial intelligence-based approaches (e.g., fuzzy logic), and fully data-driven methods. Physics-based models structure driving behavior through interpretable causal relationships between explanatory variables and LC outcomes. Rule-based models describe LC decisions using predefined behavioral rules, such as target lane preference or feasibility constraints (Yang et al., 2022). Among the earliest examples, the Gipps model (Gipps, 1986) structured LC into three phases: (i) intention, where the driver identifies the target lane; (ii) condition, where the driver evaluates the feasibility and safety of the lane change; and (iii) execution, where the actual maneuver is initiated. The MOBIL (Kesting et al., 2007), another rule-based acceleration-driven framework, evaluated LC decisions based on the potential acceleration gain and safety risk, considering a politeness factor to represent cooperative behavior. Discrete choice models extend this perspective by formulating LC as a probabilistic decision process. Ahmed's model (Ahmed, 1999) described the three stages of LC decision: (i) decision to consider a lane change; (ii) choice of the target lane; and (iii) gap acceptance. More recently, game-theoretic models have gained attention for capturing the interactive nature of LC. In these approaches, LC decision-making is formulated as a strategic game, where drivers anticipate the behavior of surrounding vehicles in an adjacent lane and choose the best strategy (e.g., change lane or stay) that maximizes perceived payoffs while considering safety and comfort (Yu et al., 2018; Talebpour et al., 2015).

Although these physics-based models offer interpretability and behavioral insight, they often lack flexibility and accuracy due to oversimplification of driver behavior. Beyond these physics-based paradigms, fuzzy logic serves as an artificial intelligence-based approach that encodes imprecise human reasoning into interpretable decision rules. For example, Balal et al. (2016) proposed a fuzzy logic LC model with IF-THEN rules that consider the gaps between vehicles. However, fuzzy logic models often face challenges in defining and calibrating membership functions, which can limit their generalizability.

On the other hand, data-driven methods have recently been explored thanks to large-scale trajectory data to better capture complex, stochastic LC behavior. For example, methods such as Bayesian optimization-based SVM (Liu et al., 2019), Gaussian regression (Althoff and Mergel, 2011), and deep neural network (Huang et al., 2018) demonstrate superior prediction accuracy over physics-based LC decision models. However, purely data-driven models require extensive training data and significant computational resources. They also face challenges in interpretability and generalizability across diverse traffic scenarios. Interpretability is often overlooked in the vehicle cybersecurity domain, where the focus tends to be on prediction accuracy. However, it is critical to understand when, how, and why predictions fail. This enables effective diagnosis and guides model refinement, especially in safety-critical systems. Thus, understanding and predicting

LC remains a key challenge, despite their strong fitting ability and better temporal dimension capture.

To address these challenges, we propose a physics-guided neural network (PGNN) framework tailored to detect abnormal LC behavior under potential cyberattacks. As a core component, the PGNN integrates a game theory-based LC decision-making model with a long short-term memory (LSTM) network. The former estimates dynamic vehicle interactions and gap acceptance, and is formulated with lateral acceleration and a probabilistic representation to ensure compatibility with gradient-based learning. The LSTM component captures complex temporal dependencies not captured in the game-theoretic model. The hybrid PGNN architecture combines the interpretability of a physics-based model and the powerful learning capabilities and flexibility of neural networks (Raissi et al., 2019; Mo et al., 2021; Wang et al., 2021). Using the PGNN-based LC prediction output, we propose an abnormal LC behavior detection that monitors the alignment between predicted and observed LC behaviors. In this alignment, significant deviations indicate potentially falsified behaviors. The framework is validated using real vehicle trajectory data and synthetic attack scenarios.

We mainly focus on discretionary LC behavior − voluntary maneuvers intended to improve driving conditions − which entails greater behavioral uncertainty than mandatory LC (Zheng, 2014). Within LC prediction, we also distinguish decision-making from trajectory planning, and target the former, namely determining when to initiate a lane change.

The remainder of the paper is organized as follows: **Section** 2.2 describes the problem setting and key assumptions. **Section** 2.3 introduces the PGNN framework and modeling details. **Section** 2.4 describes the dataset and processing procedures. **Section** 2.5 presents calibration, prediction, and abnormal behavior detection performances, along with a comprehensive discussion. Finally, **Section** 2.6 summarizes the findings and outlines future research.

2.2. Problem Description

This paper is concerned with identifying falsified LC decisions under cyberattack scenarios. CAVs operate at the intersection of autonomy and connectivity, relying on a tightly coupled system of onboard sensors, software modules, control mechanisms, and communication infrastructure to make real-time driving decisions. Cyberattacks can exploit this dependency by corrupting sensed or communicated information that drives decision-making, thereby inducing abnormal or unsafe LC behaviors.

Common attack types such as sensor spoofing (e.g., false data injection) and jamming (e.g., denial-of-service) can poison the input signals used for decision-making, causing a CAV to perceive incorrect traffic states. For instance, falsified speed or gap information may prematurely trigger an LC maneuver or suppress an LC decision when a lane change is actually warranted. Similarly, denial-of-service attacks can interrupt updates of critical state information, forcing the vehicle (or the monitoring system) to rely on stale data and resulting in delayed or inappropriate LC decisions. Because LC maneuvers inherently involve close interactions with surrounding vehicles, such falsified decisions are particularly dangerous: they can instigate or amplify traffic disturbances and, in safety-critical situations, increase collision risk.

This paper proposes an abnormal LC detection framework designed to operate online by (i) traffic management centers and/or (ii) CAV fleet operators, as illustrated in Fig. 1. Traffic

management centers can access real-time trajectories of CAVs (including potentially corrupted data streams) via vehicle-to-infrastructure (V2I) communication (Milan´es et al., 2012). By monitoring these data streams, they can flag vehicles or communication channels that may be compromised by cyberattacks. Fleet operators can similarly monitor and collect operational data through telecommunication (e.g., LTE), enabling centralized screening of abnormal behaviors at the fleet level. In addition, a vehicle-side monitoring module may leverage onboard sensors or vehicle-to-vehicle (V2V) communication to track surrounding vehicles and support local consistency checks. These scenarios reflect a monitoring setting in which trajectory streams are increasingly available as CAV penetration and sensing infrastructure expand.

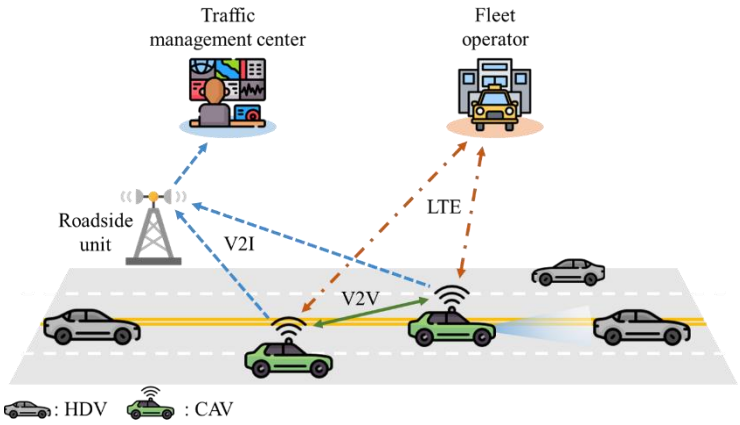


Figure 1: Operation scenarios by traffic management centers, CAVs, and fleet operator

A key operational assumption of this framework is the availability of two information channels for each subject vehicle. The first channel is the potentially corrupted trajectory stream used to compute LC probability via the PGNN-based prediction module (e.g., via onboard sensing logs or V2I/V2V messages). The second channel provides an independent and uncompromised reference of the vehicle's executed maneuver, such as LC initiation observed by infrastructure-based perception systems (e.g., roadside cameras or sensors, potentially interfaced through roadside units) or redundant sensing systems. We assume that while the first stream may be corrupted, the second stream remains attack-free and provides high-accuracy information, which we use to define the observed behavior. For clarity, we refer to this reference maneuver as the "observed behavior" in the remainder of the paper. Under this setting, detection is performed by monitoring discrepancies between (i) the LC probability predicted from the potentially corrupted input stream and (ii) the observed behavior.

Finally, we clarify the scope and assumptions used for evaluation. We assume that the majority of observed trajectories represent "normal", non-falsified driving behaviors, while falsified trajectories deviate from these norms. We also assume negligible communication delays in monitoring, so that discrepancies reflect falsification or missing updates rather than latency artifacts. For CAVs, we further assume negligible reaction delay and approximate the decision time as aligned with maneuver initiation (i.e., $t_{decision} \approx t_{start}$). These assumptions define a verification setting that enables evaluation of abnormal detection performance.

2.3. Methodology

This section presents the proposed modeling framework, comprising (i) an LC prediction module using PGNN and (ii) an abnormal behavior detection module that uses the PGNN outputs.
The modular design provides flexibility to refine each component independently and the ability to attribute errors across modules.

As shown in Fig. 2, the LC prediction module estimates LC probability over time by fusing a game-theoretic physics model with an LSTM network. The PGNN takes trajectory states as inputs and outputs LC probability. The abnormal detection module then monitors the deviations between the predicted LC probability and the observed behavior. The cumulative discrepancy serves as an anomaly score, which flags abnormal LC behaviors based on a predetermined threshold.

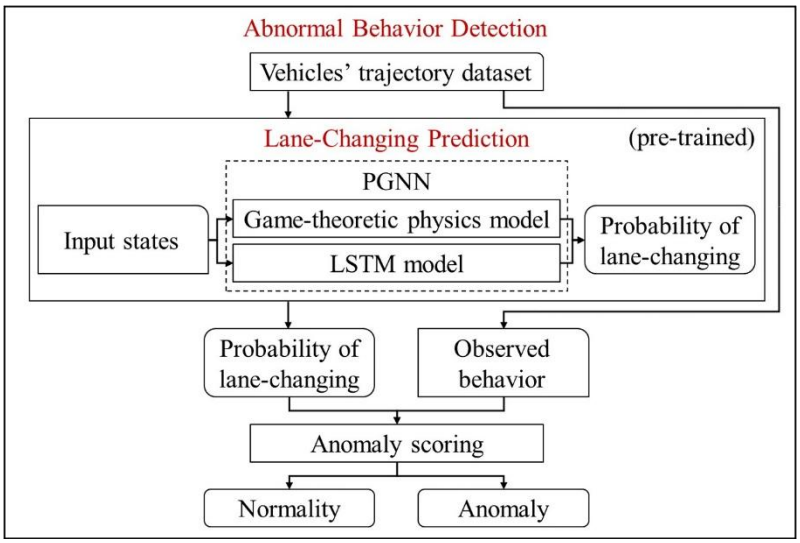


Figure 2: Modeling framework

2.3.1. Lane-Changing Prediction

A typical LC scenario involves interactions between the subject vehicle S and up to three neighboring vehicles: the current-lane leader CL, the target-lane leader TL, and the target-lane follower TF, as illustrated in Fig. 3. The subject vehicle performs a discretionary LC maneuver to improve driving conditions, such as gaining speed or accessing a larger gap.

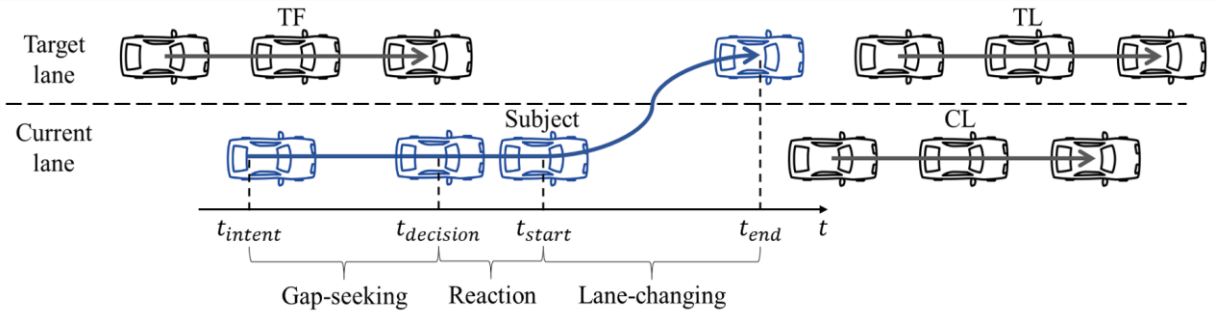


Figure 3: Vehicle in lane-changing process

Here, we focus on the decision-making interval between LC intent (t_{intent}) and LC initiation (t_{start}). Rather than modeling the formation of intent itself, we concentrate on the timing of the decision-making, during which the vehicle evaluates gap feasibility and interaction with surrounding vehicles. This formulation reflects the automated decision-making module of a CAV. While the monitoring setting assumes negligible reaction delay for CAVs, the model is developed using human-driven vehicle trajectories (further discussed in **Section** 2.4); hence we distinguish gap-seeking and reaction phases within $[t_{intent}, t_{start}]$. The trajectory data in this interval – covering both phases –are used to predict the likelihood of an LC decision.

Table 1 presents the key state variables describing the interactions among vehicles: speed, acceleration, and gap. These variables mainly represent longitudinal motion, with lateral acceleration included to account for LC dynamics. All variables are assumed to be obtainable from onboard sensing and/or V2X communication.

Table 1: Variable describing vehicle interaction in lane-changing

Notation	Definition	Unit
v	Longitudinal speed of vehicle S	m/s
v^{CL}	Longitudinal speed of leading vehicle CL in current lane	m/s
v^{TL}	Longitudinal speed of leading vehicle TL in target lane	m/s
v^{TF}	Longitudinal speed of following vehicle TF in target lane	m/s
a	Longitudinal acceleration of vehicle S	m/s^2
a^{TL}	Acceleration required for vehicle S to avoid a collision with TL	m/s^2
a^{TF}	Acceleration required for TF to avoid a collision with vehicle S	m/s^2
a^{TFL}	Acceleration required for TF to avoid a collision with vehicle TL	m/s^2
b	Lateral acceleration of vehicle S	m/s^2
g^{TL}	Longitudinal gap with leading vehicle TL in target lane	m
g^{TF}	Longitudinal gap with following vehicle TF in target lane	m

2.3.1.1. Physics-Guided Neural Network Architecture

The LC decision prediction task is carried out by a PGNN that integrates domain knowledge from a game-theoretic physics model with a data-driven LSTM. The overall structure is shown in Fig. 4. The upper part (red) represents the physics-based model, where each rectangle denotes a computational step involving model variables and parameters. The lower part (blue) represents the

neural network, where each node corresponds to a neuron with an activation function, and each edge contains trainable parameters.

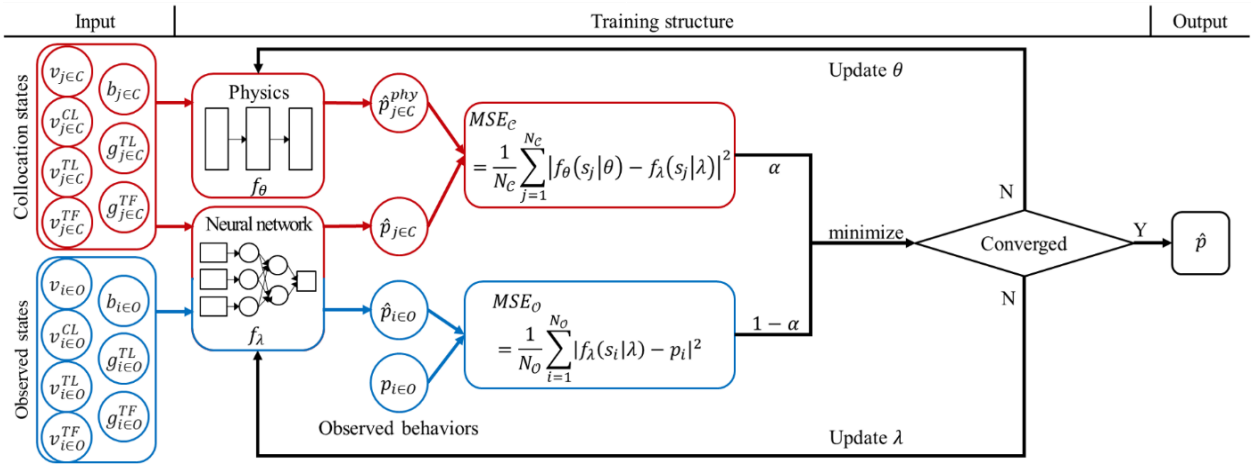


Figure 4: Architecture of the physics-guided neural network (PGNN) framework

The model is trained using two types of input states: collocation and observed states. Collocation states are a distinctive component of the PGNN framework. They are sampled from the input domain, fed into both the physics model and the neural network, and used to compute the physics discrepancy by comparing the two outputs. These states do not require observed-behavior labels and enforce that the predicted LC probabilities remain consistent with the underlying physics-based model across the state space.

Observed states are input only to the neural network, and the outputs are compared with observed behavior labels to calculate data discrepancy. Because observed behavior labels are binary and highly imbalanced over time, a vanilla LSTM trained solely on observed states may overfit to the dominant non-LC class or produce unstable probability transitions near the decision boundary.

The PGNN is trained with a total loss that incorporates both physical consistency and predictive accuracy. By leveraging both unlabeled (collocation) and labeled states, the framework exploits structural information without requiring additional observed-behavior labels. This semi-supervised mechanism improves generalization and stabilizes probability estimation, particularly under sparse or imbalanced labeling conditions.

The PGNN here mathematically maps states $\boldsymbol{s} \in S$ to the probability of an LC, p:

$$f_{\theta,\lambda} : \boldsymbol{s} \rightarrow p$$

This joint PGNN model inputs the state vector $\boldsymbol{s}(t)$, such as speed, speed differences, and gaps at time t, and outputs the LC probability $\hat{p}(t + \Delta t)$ for the next time step. The PGNN framework combines a physics-based estimator $f_\theta(\boldsymbol{s}|\theta)$, parameterized by θ, and a neural network model $f_\lambda(\boldsymbol{s}|\lambda)$, defined by a hyperparameter vector λ. Further details on the model components and training procedure are provided in the following subsections.

2.3.1.2. Physics-Based Model: Game-Theoretic Model

Our physics-based model here aims to capture the cognitive process of LC while grounding the decision in physical interaction features (e.g., speed differences and gaps). This provides a tractable formulation that supports interpretability and robustness under a limited data regime.

For the physics-based model, $f_\theta(\boldsymbol{s}|\theta)$we develop a game-theoretic formulation to represent gap acceptance through dynamic interactions between the subject vehicle (a potential lane-changer) and the target-lane follower. We model this interaction as a two-player non-cooperative game with incomplete information (Talebpour et al., 2015). At the decision time, the subject vehicle chooses whether to change lanes or stay in its current lane, while the target-lane follower chooses whether to impede the maneuver by accelerating or to yield by decelerating. We do not consider the follower's own lane change action, since it would dissolve the leader-follower interaction or have negligible impact on the subject vehicle's payoff over the decision interval. The corresponding simplified payoffs for the subject vehicle (P) and the target-lane follower (R) are summarized in Table 2. In this framework, a^{TL} denotes the acceleration required for the subject vehicle to avoid a collision with its target-lane leader (TL), and a^{TF} denotes the acceleration required for the target-lane follower (TF) to avoid a collision with the LC subject vehicle:

$$a^{TL} = \begin{cases} -\dfrac{(v^{TL}-v)^2}{2(g^{TL}-(v^{TL}-v)\Delta t)}, & \text{if } g^{TL}-(v^{TL}-v)\Delta t > 0 \\ a_0, & \text{otherwise} \end{cases}$$

$$a^{TF} = \begin{cases} -\dfrac{(v^{TF}-v)^2}{2(g^{TF}-(v^{TF}-v)\Delta t)}, & \text{if } g^{TF}-(v^{TF}-v)\Delta t > 0 \\ a_0, & \text{otherwise} \end{cases}$$

Here, a_0 is the maximum comfortable deceleration rate, set at -3.05 m/s^2, and Δt denotes the discrete time step used in the trajectory data.

Δv indicates the speed difference between the current-lane leader and the target-lane leader, while lateral acceleration b is included to represent LC intent. a_0^{TF} is defined as $\min(a^{TF}, a_0)$. Similarly, a_0^{TFL} and a_0^{TFL} are computed between the target-lane follower and the target-lane leader when the subject vehicle remains in its current lane. Accordingly, the full input state is given as $\boldsymbol{s}(t) = [v, v^{CL}, v^{TL}, v^{TF}, b, g^{TL}, g^{TF}]$. Given $\boldsymbol{s}(t)$, we compute derived interaction terms including Δv, a^{TL}, a^{TF} and a^{TFL}. To account for latent or unobserved factors, ε and δ are introduced as error terms.

Table 2: Payoffs matrix for the discretionary lane-changing game

		Subject vehicle	
	Action	Change lane	Stay in lane
Target-lane follower	Accelerate	$P_{11} = \eta_1 a^{TL} + \eta_2 \Delta v + \eta_3 b + \varepsilon_{11}$ $R_{11} = \eta_4 a^{TF} + \delta_{11}$	$P_{12} = \varepsilon_{12}$ $R_{12} = \eta_4 a^{TFL} + \delta_{12}$
	Decelerate	$P_{21} = \eta_1 a^{TL} + \eta_2 \Delta v + \eta_3 b + \varepsilon_{21}$ $R_{21} = \eta_4 a_0^{TF} + \delta_{21}$	$P_{22} = \varepsilon_{22}$ $R_{22} = \eta_5 a_0^{TFL} + \delta_{22}$

While traditional physics-based LC models typically rely on longitudinal gap-acceptance

features, our preliminary data-driven analysis identified lateral acceleration as a significant predictor. Rather than serving as a statistical correlate, lateral acceleration can be interpreted as an early-warning cue that reflects latent commitment to execute a lane change before the maneuver is initiated. This interpretation aligns with behavioral studies showing that drivers exhibit subtle lateral movements prior to the LC execution, serving as an early indication of their intent (Wissing et al., 2017; Mahajan et al., 2020). In the payoff formulation, we regard lateral acceleration as representing this transient maneuver cue.

Note that conventional game-theoretic models determine actions by maximizing the subject vehicle's payoff; our approach introduces a probabilistic formulation to enable gradient-based learning in joint PGNN training. We first solve for the mixed-strategy Nash equilibrium of the game, then compute the expected total payoff for the subject vehicle (U_S) and target-lane follower (U_{TF}). We use the difference ($U_S - U_{TF}$) as a compact conflict metric that increases when lane changing is beneficial for the subject while costly to the follower, and project it through a sigmoid function to compute the LC probability:

$$\hat{p}^{phy} = \frac{1}{1 + e^{-(U_S - U_{TF})}}$$

The model parameters, $\theta = [\eta_1, \eta_2, \eta_3, \eta_4, \eta_5]$, are calibrated using observed trajectory data. Calibration is performed by minimizing the mean squared error (MSE) between the predicted and observed behavior labels:

$$\min_{\theta} \frac{1}{N_O} \sum_{i=1}^{N_O} \left| \hat{p}_i^{phy} - p_i \right|^2 \qquad \text{s.t. } \hat{p}_i^{phy} = f_\theta(\boldsymbol{s}_i|\theta), i = 1, \cdots, N_O$$

where, $\hat{p}_i^{phy}$ is the physics-model-predicted LC probability for vehicle i; pi is binary observed behavior label; N_O is the number of observed samples (vehicles); and $\boldsymbol{s}_i$ is the input state for sample i. A genetic algorithm (Kumar et al., 2010) is employed for calibration, as it does not require gradient information. The optimal LC decision-making threshold μ is selected based on a receiver operating characteristic (ROC) curve analysis: if $\hat{p}^{phy} < \mu$, the model predicts "Stay in lane"; otherwise, it predicts "Change lane."

2.3.1.3. Neural Network Model: Long-Short Term Memory

To complement the physics-based component, we develop a data-driven model that captures temporal dependencies in LC decision-making without imposing additional hand-crafted structure. This allows us to generalize under complex and high-dimensional data.

The neural network component, denoted by $f_\lambda(s|\lambda)$, is implemented using an LSTM network (Hochreiter, 1997). The LSTM is well suited for modeling the subtle temporal evolution of LC behavior because it can selectively retain informative elements that are influential to the prediction task while filtering out less relevant variations (Wang et al., 2021). As illustrated in Fig. 5, the LSTM takes as input a sequence of state vectors over a time window of length t_w. For vehicle i, the state vector at time step t is defined as:

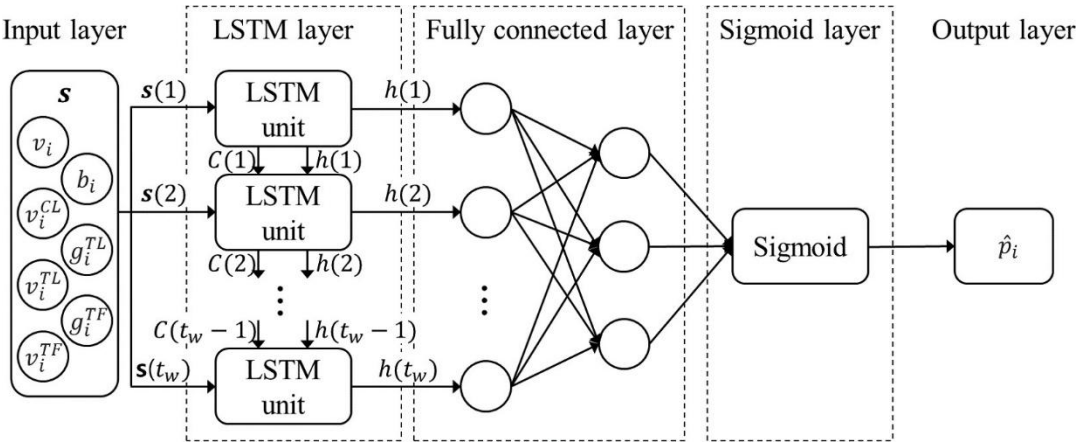


Figure 5: Structure of the long short-term memory (LSTM) model

Given the input sequence $\boldsymbol{s}_i(1), \cdots, \boldsymbol{s}_i(t_w)$, the LSTM produces a hidden representation, which is passed through a fully connected layer and a sigmoid output layer to generate the predicted LC probability $\hat{p}_i(\mathrm{t} + \Delta\mathrm{t})$. Standard LSTM update equations are provided in **Appendix** A for completeness.

2.3.1.4. Training of Physics-Guided Neural Network Model

Training LC prediction models with empirical data poses significant challenges due to the limited labeled LC trajectories and the resulting generalization issues. To mitigate this limitation, we use collocation states – trajectory data sampled from a separate dataset – as further discussed in **Section** 2.4.4. Let O denote the observed-state set and C denote the collocation-state set, where $O \cap C = \emptyset$ and $O \subset S, C \subset S$, with S representing the state space. Both consist of a set of state vectors $O = s_i, i = 1, \cdots, N_O, C = s_j, j = 1, \cdots, N_C$.

To regularize the neural network model by incorporating physics knowledge, the PGNN loss minimizes two discrepancies: (i) the physics discrepancy between the physics-model and neural-network probabilities evaluated on collocation states, and (ii) the data discrepancy between the neural-network probabilities and the observed behavior labels associated with observed states:

$$\begin{aligned} Loss_{\theta,\lambda} &= \alpha MSE_C + (1-\alpha) MSE_O \\ &= \alpha \frac{1}{N_C} \sum_{j=1}^{N_C} \left| \hat{p}_j^{phy} - \hat{p}_j \right|^2 + (1-\alpha) \frac{1}{N_O} \sum_{i=1}^{N_O} |\hat{p}_i - p_i|^2 \\ &= \alpha \frac{1}{N_C} \sum_{j=1}^{N_C} \left| f_\theta(s_j|\theta) - f_\lambda(s_j|\lambda) \right|^2 + (1-\alpha) \frac{1}{N_O} \sum_{i=1}^{N_O} \left| f_\lambda(s_j|\lambda) - p_i \right|^2 \end{aligned}$$

Here α controls the relative weight of the physics discrepancy in PGNN training.

We use the Adam optimizer (Kingma, 2014), a stochastic gradient descent algorithm, to jointly update θ and λ. The calibrated parameters from the game-theoretic model are used to initialize θ.

2.3.2. Abnormal Behavior Detection

Accurate LC prediction forms the basis for identifying falsified LC decisions. We define an anomaly score that monitors deviations between predicted LC probabilities and the observed behavior over time, thereby taking full advantage of the proposed PGNN framework. For vehicle i, the anomaly score $AC_i(t)$ is defined as:

$$AC_i(t) = \sum_{k=t-l}^{t} |\hat{p}_i(k) - p_i(k)|$$

Here, $\hat{p}_i(k)$ is the predicted LC probability at time step k, and $p_i(k)$ is the binary observed behavior label. The cumulative absolute deviation over a scoring window of length l serves as the anomaly score.

Detection is then performed using a threshold τ:

$$\text{Detection} = \begin{cases} \text{Normal behavior}, & \text{if } AC_i(t) < \tau \\ \text{Abnormal behavior}, & \text{if } AC_i(t) \geq \tau \end{cases}$$

The detection threshold τ plays a critical role in determining the sensitivity of anomaly detection. It directly impacts the true and false positive rates, where false negatives – misclassifying abnormal behavior as normal – can severely compromise traffic safety and efficiency. The value of τ may be set by the user and should be calibrated in a sensitivity analysis that considers both operational reliability and safety implications.

2.4. Data Preparation

Vehicle trajectory data were processed to calibrate the physics-based model and train the PGNN for LC prediction. In this section, we introduce the dataset and the data processing procedures.

While AV-specific trajectory data would ideally be preferred, such datasets are limited in availability and scope. Accessible datasets are often collected in urban environments, lack complete surrounding-vehicle information during LC maneuvers, or contain insufficient numbers of AV-specific LC events to support robust model training. In this study, abnormal detection is framed as identifying deviations from nominal LC behavior; thus, learning a reliable baseline from real trajectory is essential. We therefore used human-driving trajectory data to facilitate the methodological development and to ground the model in real-world driving environments where AVs interact with human-driven vehicles (HDVs) predominantly. By doing so, our PGNN model captures realistic behavioral variability and complexity, improving its robustness in detecting falsified LC behaviors in heterogeneous traffic systems. Furthermore, the proposed framework remains data-agnostic in structure.

2.4.1. Vehicle Trajectory Data

The vehicle trajectory data used in this study were derived from the I-24 MOTION dataset (Gloudemans et al., 2023). I-24 MOTION is a camera-based trajectory generation system deployed

along Interstate 24 in Nashville, Tennessee. This comprehensive dataset includes interactive driving scenarios such as freeway driving, ramp merging, and LC maneuvers. We used the publicly available "INCEPTION" dataset, which spans November 21 to December 2, 2022, with daily recordings ranging from 4 hours to 11 hours starting at 6 a.m. Trajectories are sampled at 25 Hz.

For this study, data from November 29 were used to form the observed-state set. Data from November 30 were used to form the collocation-state set and the test set. The collocation and test sets were constructed from disjoint vehicle samples to avoid overlap.

2.4.2. Data Preprocessing

The vehicle trajectory data were processed according to the following criteria:

- West-bound trajectory data from pole 20 to pole 32 were extracted. This 2-km segment is a basic freeway section away from merge and diverge areas. There is no high-occupancy vehicle lane in this segment.
- Vehicles traveling in lanes 1 (left-most) and 2 were considered to increase the chance of capturing discretionary lane changes. Vehicles in lanes 3 to 5 were excluded to reduce the influence of mandatory lane changes near the merge area.
- Only sedans and midsize passenger cars were selected as subject vehicles. Trucks were excluded due to their distinct LC behavior and a smaller sample size.
- Vehicles executing more than two LCs within a three-second window were excluded to isolate single LC maneuvers.
- LC trajectories with incomplete surrounding vehicle data were removed.
- For each identified subject vehicle S, the LC start time was defined as the first instance when the center of S exhibited lateral speed exceeding 0.2 m/s (Wang et al., 2014).
- The final LC trajectories were manually reviewed to ensure correctness, addressing any missing or mismatched data during the LC decision-making phase.

As a result, 1,028 LC trajectories from November 29 were selected as observed states. From the November 30 data, 948 LC trajectories were extracted and used to construct the collocation-state set and the test set.

2.4.3. Lane-Changing Labeling

In this section, we estimate the LC decision time prior to maneuver initiation and construct training labels that reflect the inferred decision-making process, rather than labeling only the execution moment.

A key limitation of trajectory-based LC labeling is class imbalance induced by binary, maneuver-based labels. Specifically, LC probability p is set to 0 for all frames before the start time of LC execution (t_{start}) and switches abruptly to 1 at t_{start} (Fig. 6). This maneuver-based labeling fails to reflect the decision-making process, because the driver's intention and decision moments are not directly observed from trajectory data – particularly without head-motion or eye-tracking information. In addition, reaction times vary across drivers, implying that a single fixed reaction time cannot represent individual decision timing. To address these limitations, we infer three key time points: (i) t_{intent}, the moment when the driver forms the intention to change lanes; (ii) $t_{decision}$, the decision point at which a suitable gap is identified; and (iii) t_{start}, the observed onset of the LC maneuver. While t_{start} is directly observable, t_{intent} and $t_{decision}$ are estimated as

described below.

2.4.3.1. Estimating the Decision Time via Reaction Time

We estimate $t_{decision}$ from reaction time by assuming that $t_{start} - t_{decision}$ represents the driver's reaction time. To estimate personalized reaction time from trajectories, we identify informative explanatory variables through feature importance analysis and train a GBM using the 5-second trajectory segment preceding t_{start}. The GBM outputs a probabilistic LC curve over time (labeled as the "gradient-boosted LC probability" in Fig. 6), which provides a continuous representation of the evolving decision state prior to initiation.

This probabilistic curve is estimated for each driver; however, a decision threshold is required to determine a specific decision instant $t_{decision}$. While the decision threshold is heterogeneous and endogenously determined at the individual level, we work with a representative (population-average) threshold identified from aggregate data. Specifically, we construct an aggregated curve across drivers and select a threshold corresponding to an average reaction time of 1 second (Zhang et al., 2022). The same threshold is then applied to each driver-specific curve to determine $t_{decision}$.

2.4.3.2. Estimating the Intent Time

To estimate t_{intent}, we adopt the empirical finding that the duration from LC intention to LC initiation is approximately 3 seconds (Doshi and Trivedi, 2009). Given the average reaction time of 1 second, we define the preceding 2-second interval as the gap-seeking phase. Accordingly, once $t_{decision}$ is identified, tintent is set to 2 seconds prior to $t_{decision}$. The trajectory segment from tintent to t_{start} is used as the model input.

2.4.3.3. Binary Labeling for Model Training

After estimating $t_{decision}$, we assign binary labels to represent two phases within $[t_{intent}, t_{start}]$: the 2-second gap-seeking phase (t_{intent} to $t_{decision}$) is labeled $p = 0$, and the reaction phase ($t_{decision}$ to t_{start}) is labeled $p = 1$. Although sigmoid curves are used to infer personalized decision timing, binary labels are used for model training to enforce a clear decision boundary. This relabeling aligns training supervision with the inferred decision moment and accounts for inter-driver variability in reaction time.

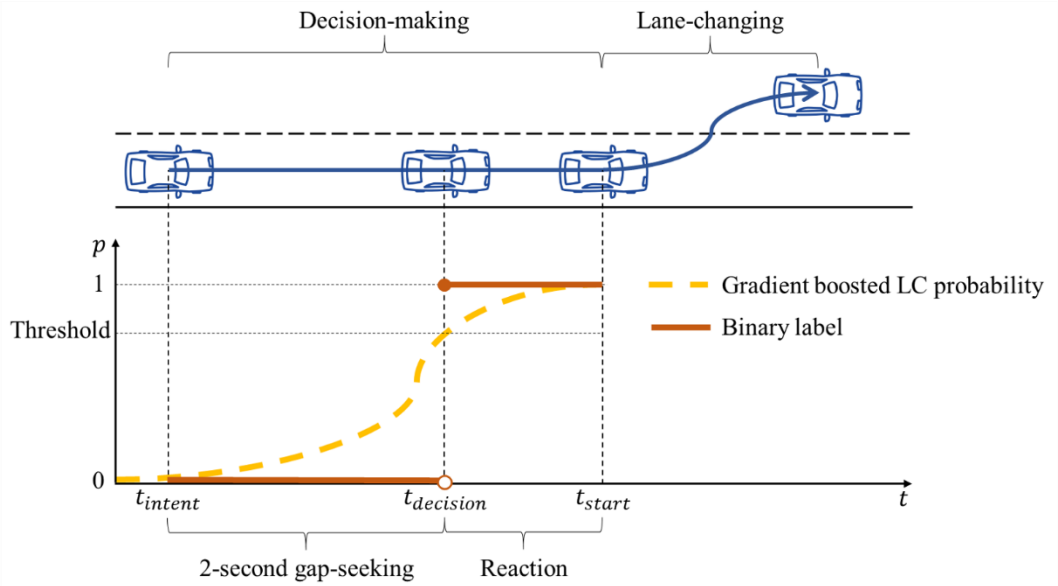


Figure 6: Estimation of gradient-boosted LC probability curve and binary labeling

2.4.4. Collocation States

Unlike observed states, collocation states do not require that the corresponding actions be observed in real trajectories. However, because the neural network component is an LSTM that models temporal dependencies, collocation states must form trajectory segments over consecutive time steps rather than isolated state samples. To expand state coverage beyond the variability of observed states, we selected additional data collected under similar traffic conditions yet individually distinct. Specifically, we used data from November 30, which exhibits similar speed profiles to the November 29 data (training data) and is incident-free (Fig. 7). Among the 948 LC trajectories available from November 30, we randomly selected 474 trajectories to serve as collocation states, and reserved the remaining 474 trajectories as the test set.

2.5. Evaluation and Discussion

In this section, we evaluate the LC prediction module and the abnormal LC detection module under two attack scenarios.

LC prediction performance is assessed using MSE between predicted LC probabilities and the corresponding observed behavior labels, together with a confusion matrix. For LC prediction, false positives indicate cases where a lane change is predicted but does not actually occur, while false negatives indicate missed lane changes. Because the abnormal detection module relies on the predicted LC probabilities, prediction errors can propagate to detection performance.

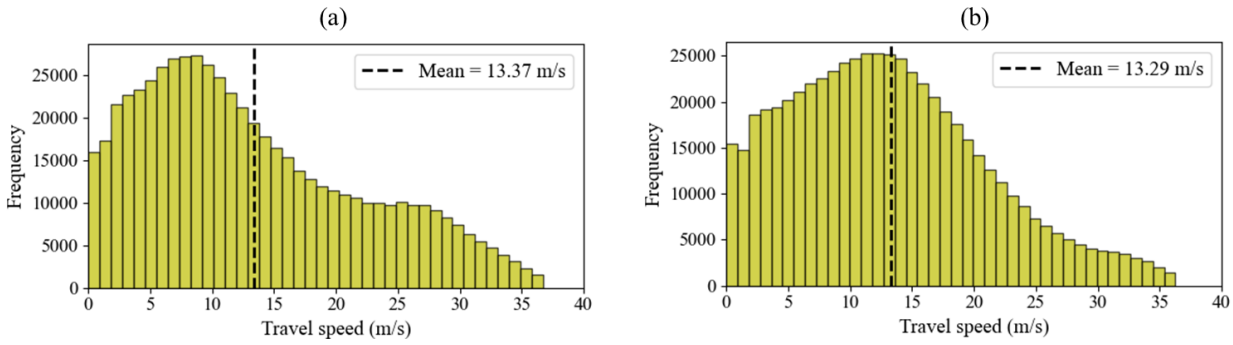


Figure 7: Travel speed distribution: (a) On November 29; and (b) On November 30

The anomaly detection module is also evaluated using a confusion matrix. For detection, false positives correspond to normal behaviors that are misclassified as abnormal. Such cases may lead to occasional false alarms and increase the operational workload of a cybersecurity system, but they do not represent missed safety-critical LC events. False negatives, where falsified LC behaviors remain undetected, are of greater concern because they directly compromise safety, especially when manipulated behaviors could trigger collisions.

To summarize overall performance, we report accuracy, precision, recall, and F1-score. Accuracy is the proportion of correct predictions among all instances, while precision measures the fraction of true positives among all predicted positives. Recall (sensitivity) measures the fraction of true positives identified among all actual positives; it is especially important in this context because missed detections can have severe safety consequences. The F1-score, defined as the harmonic mean of precision and recall ($= 2\frac{Precision \cdot Recall}{Precision + Recall}$), provides a balanced comparison under class imbalance. For LC prediction, the "stay in lane" class dominates the "change lane" class; similarly, for anomaly detection, the normal (unattacked) class is typically larger. In both cases, the F1-score is therefore useful for evaluating model performance.

2.5.1. Lane-Changing Prediction

2.5.1.1. Game-Theoretic Model Results

We evaluated the game-theoretic model on the designated test dataset. The model achieved an MSE of 0.097. To classify the LC decisions, the probability threshold μ^* was selected using validation with a 60:40 training-validation split and was determined to be 0.396. Accordingly, state vectors with $\hat{p}^{phy} \geq 0.396$ are classified as "Change lane." Table 3 reports the calibrated model parameters.

Table 3: Parameter calibration results for game-theoretic model

Parameter	η_1	η_2	η_3	η_4	η_5
Calibrated Value	-0.1710	0.0687	0.9961	0.7488	-0.9262

Table 4 summarizes the confusion matrix results. Among 35,550 evaluation vectors, the overall accuracy was 0.913 and the F1-score was 0.848, indicating acceptable performance across both classes. The true rate for $p = 0$ (stay in lane) was 0.930, while the true rate for $p = 1$ (change

lane) was 0.870. These results demonstrate that the model distinguishes LC decision labels within the decision-making interval with stable performance, compared to values ranging from 0.83 to 0.97 reported in prior literature (Balal et al., 2016).

Table 4: Confusion matrix for the game-theoretic model

		Game-theoretic model prediction			
		Stay in lane ($\hat{p}^{phy} < \mu^*$)	Change lane ($\hat{p}^{phy} \geq \mu^*$)	Total	True rate
Observation	Stay in lane ($p = 0$)	23,876	1,800	25,676	0.930
	Change lane ($p = 1$)	1,281	8,593	9,874	0.870
	Total	25,157	10,393	35,550	

Since the effectiveness of physics-neural network integration depends on the performance of the underlying physics model (Zhong et al., 2024), we compared our game-theory model with three representative LC decision models: Ahmed's gap-acceptance model (Ahmed, 1999), MOBIL model (Kesting et al., 2007), and Talebpour's game-theoretic model (Talebpour et al., 2015). Detailed formulations are provided in **Appendices** B-D. Both Ahmed's and Talebpour's models were calibrated by minimizing MSE, consistent with our approach. MOBIL, which produces binary decisions, was calibrated by maximizing the sum of accuracy and F1-score to balance predictive precision and recall.

Table 5 summarizes the comparative results. The three benchmark models achieved accuracies around 0.57, indicating limited performance in estimating LC decision timing within the decision-making interval under the class-imbalanced dataset. Among them, Talebpour's model achieved the highest accuracy and F1-score. Our model further improves performance by incorporating lateral acceleration and refining LC labeling. As shown in Table 3, the weight on lateral acceleration (η_3) is dominant, while the collision-avoidance-related terms (η_4 and η_5) also contribute. This suggests that improved performance is achieved by jointly representing latent LC intention and interactive vehicle dynamics in a unified payoff structure.

Table 5: Prediction accuracy of different physics-based models

Model	Accuracy	Precision	Recall	F1-score
Ahmed's model	0.570	0.292	0.386	0.333
MOBIL model	0.563	0.273	0.347	0.306
Talabpour's model	0.577	0.340	0.380	0.359
Our model	0.913	0.827	0.870	0.848

It is also worth mentioning that all models are evaluated on the relabeled dataset described in **Section** 2.4.3, performance values may differ from those originally reported in prior studies, which were typically calibrated and evaluated under different labeling schemes.

2.5.1.2. Physics-Guided Neural Network Model Results

We trained the PGNN model with $\alpha = 0.3$ in loss function, placing greater weight on the data discrepancy term. Fig. 8 shows the training and validation losses over epochs. The training loss

continuously decreases, and the data discrepancy remains lower than the physics discrepancy throughout training. This indicates that the mismatch between the LSTM model and observed behavior label is smaller than that between the LSTM and the game-theoretic model. The validation loss decreases initially and increases slightly after around 70 epochs, indicating overfitting. We therefore retain the parameter set (θ^*, λ^*) at the epoch with the minimum validation loss.

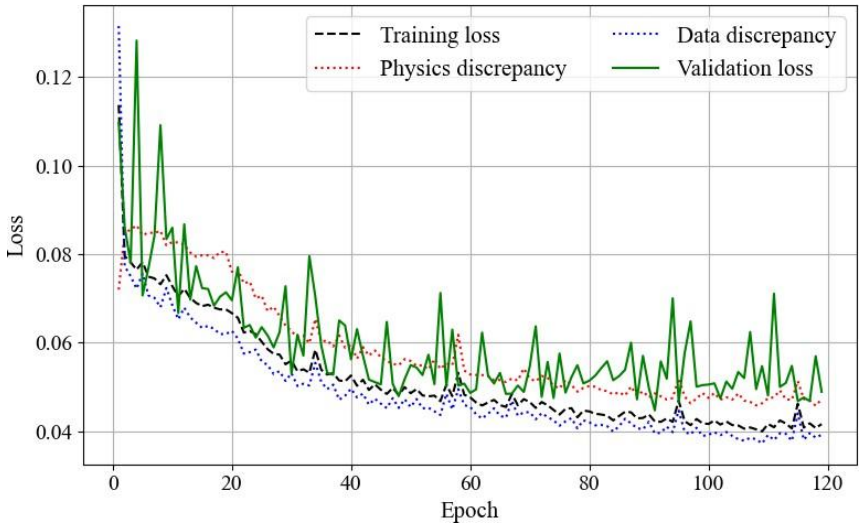


Figure 8: Training and validation loss curves of the PGNN

Table 6 summarizes the hyperparameters and evaluation metrics for the LSTM and PGNN models. Hyperparameters were selected via cross-validation using the area under the ROC curve. On the test dataset, the LSTM and PGNN achieved MSE values of 0.076 and 0.069, respectively, indicating improved predictive performance with the PGNN. While accuracy is comparable across models, the PGNN achieves lower MSE and higher F1-score, indicating improved probability estimation under class imbalance.

Table 6: Hyperparameters and performance summary for LSTM and PGNN

		LSTM model	PGNN model
Parameter	Hidden size	16	16
	Learning rate	0.05	Physics–0.01, NN–0.05
	Epoch, patience, batch size	{200, 50, 128}	{200, 50, 128}
	Window, stride	{5, 1}	{5, 1}
Metric	MSE	0.076	0.069
	Accuracy	0.901	0.912
	Precision	0.761	0.779
	Recall	0.969	0.975
	F1-score	0.852	0.866

Consistent with these metrics, the confusion matrices in Fig. 9 show fewer misclassifications for

the PGNN than for the LSTM. Relative to the LSTM, the PGNN reduces false positives from 3,021 to 2,740 and false negatives by 21% (from 303 to 239). Moreover, the PGNN's true positive rate reaches 0.975. These results indicate that incorporating physics-based structure into the LSTM via the PGNN framework improves LC probability prediction, which is essential for the subsequent abnormal detection module.

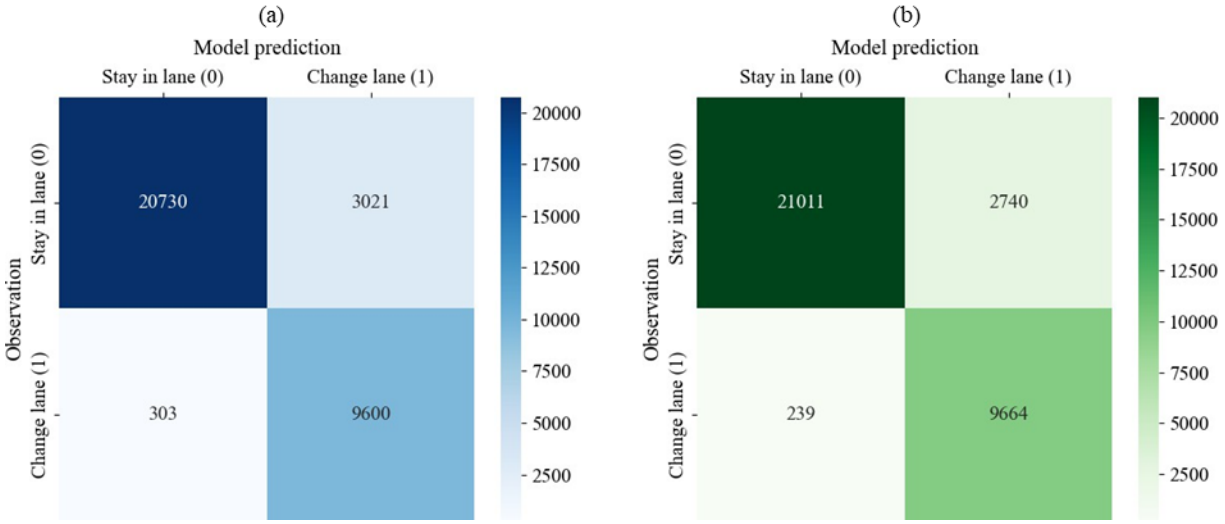


Figure 9: Confusion matrices for lane-changing prediction: (a) LSTM model; (b) PGNN model

2.5.2. Cyberattack Scenario

This section describes two cyberattack scenarios used in the evaluation of the abnormal LC detection framework. Consistent with the monitoring setting defined in **Section** 2.2, the attacks are modeled as the corruption of the input trajectory stream used for PGNN-based LC probability prediction, while the executed LC maneuver remains available through an independent, attack-free channel (i.e., the observed behavior). We consider two representative attacks: false data injection (FDI) and denial-of-service (DoS). These attacks are designed to distort either the magnitude or the freshness of the input signals, thereby altering the inferred LC decision timing and causing discrepancies between the predicted LC probability and the observed behavior.

2.5.2.1. False Data Injection Attack

We simulate FDI attacks, a type of sensor spoofing in which falsified information is introduced into sensor measurements or communicated trajectories (Vyas et al., 2023). In our scenario, the vehicle operates normally, but its recorded speed data in the input stream is intentionally manipulated. A step input of a certain magnitude is added to both the longitudinal and lateral speed profiles of the subject vehicle starting from frame 15 onward, creating discrepancies from the nominal profiles; see 10a for an example of falsified longitudinal speed. Each frame corresponds to 0.04 seconds, as the trajectory data are sampled at 25 Hz. Falsified speed profiles tend to induce prematurely high LC probability predictions.

For this experiment, FDI attacks were injected into the trajectories of 50 randomly selected vehicles, representing approximately 10% of the test dataset. The magnitude of the injected

perturbation is constrained to within 10–15% of the vehicle's original longitudinal and lateral speeds to represent a subtle yet impactful manipulation. Using these falsified speed profiles, the state vectors are recomputed at each time step, as these vectors serve as inputs to the PGNN model for LC prediction and subsequent anomaly scoring.

2.5.2.2. Denial-of-Service Attack

We also simulate a DoS attack that prevents new measurements from being updated and forces the monitoring stream to rely on stale information. As shown in Fig. 10b, the speed profile is frozen at the last value received prior to the attack, starting from frame 50 onward. In some cases, near-zero or static speeds may suppress LC probability predictions or distort their timing. For this experiment, DoS attacks are injected into the trajectories of 50 vehicles, independently from the FDI experiment.

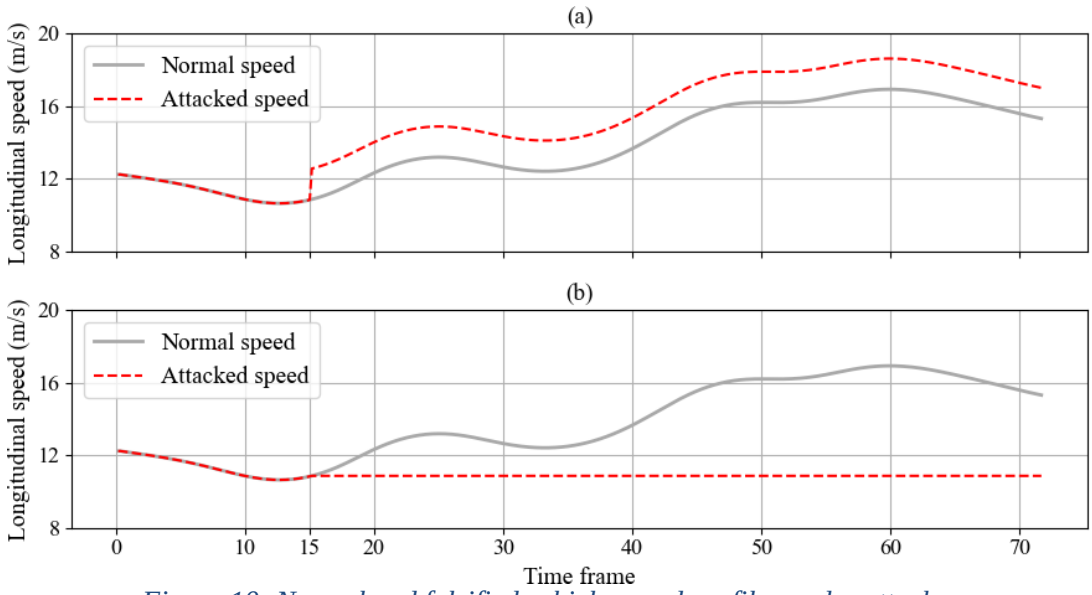


Figure 10: Normal and falsified vehicle speed profiles under attacks: (a) False data injection (FDI) attack; (b) Denial-of-Service (DoS) attack

2.5.3. Abnormal Behavior Detection

2.5.3.1. Anomaly Scoring

To assess detection performance under cyberattacks, we input the falsified trajectories into the trained PGNN model to predict LC probabilities. Figure 11 illustrates how predicted LC probabilities and anomaly scores evolve in response to a falsified input. The anomaly score was computed by accumulating the deviation between the predicted LC probability (based on the corrupted input) and the observed behavior over a fixed scoring window of 10 frames (orange-shaded). This temporal accumulation mitigates frame-level fluctuations and stabilizes anomaly assessment.

As shown in Fig. 11a, the attack begins at frame 15, resulting in a sudden rise in the predicted LC probability. This early prediction deviates from the observed behavior, which occurs after frame 40. Consequently, the anomaly score rises and surpasses the detection threshold (τ in

Eq. 9) at frame 26 (Fig. 11b), and the behavior is classified as abnormal. The corresponding detection delay is 11 frames (0.44 s at 25 Hz), demonstrating responsiveness to subtle behavioral inconsistencies.

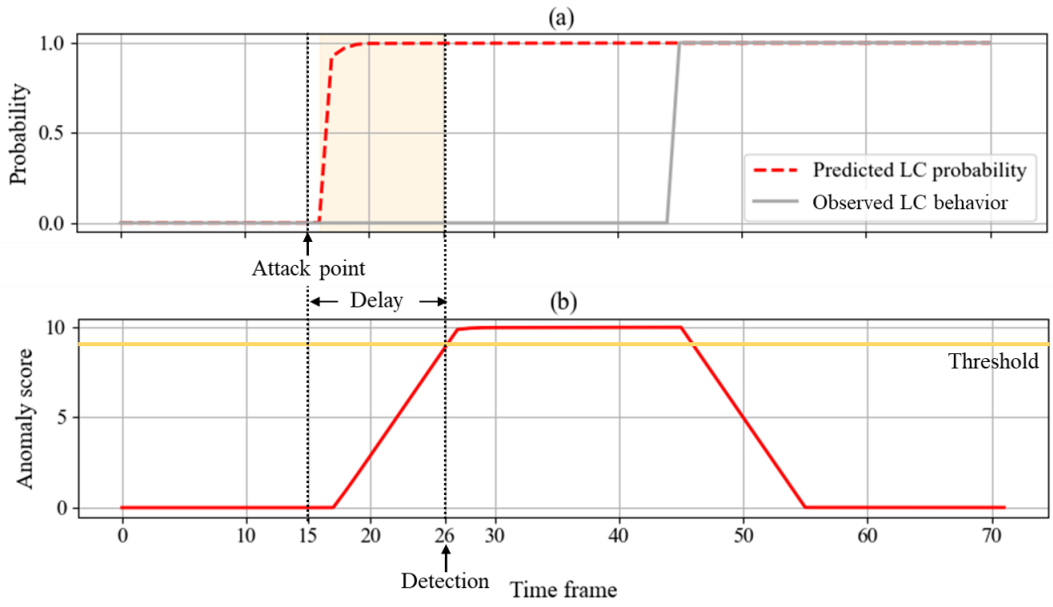


Figure 11: Anomaly detection under FDI attack:
(a) Observed vs. predicted lane-changing probabilities; (b) Example of anomaly scoring

Fig. 12 illustrates the evolution of LC probabilities and anomaly scores under a DoS attack. In this example, the predicted LC probability remains at 0 and matches the observed behavior until frame 44, when execution of the LC maneuver is observed. This creates a discrepancy between the predicted LC probability and the observed behavior. As a result, the anomaly score exceeds the detection threshold at frame 54 (Fig. 12b), and the behavior is classified as abnormal. The PGNN detects the anomaly 39 frames (1.56 s) after the onset of the attack. Because the proposed framework detects abnormality based solely on LC behavior, it cannot flag a DoS attack before the observed LC initiation occurs. Nevertheless, once LC initiation is observed, detection requires only the scoring window length (10 frames, 0.4 s), indicating prompt response after observable behavioral deviations emerge.

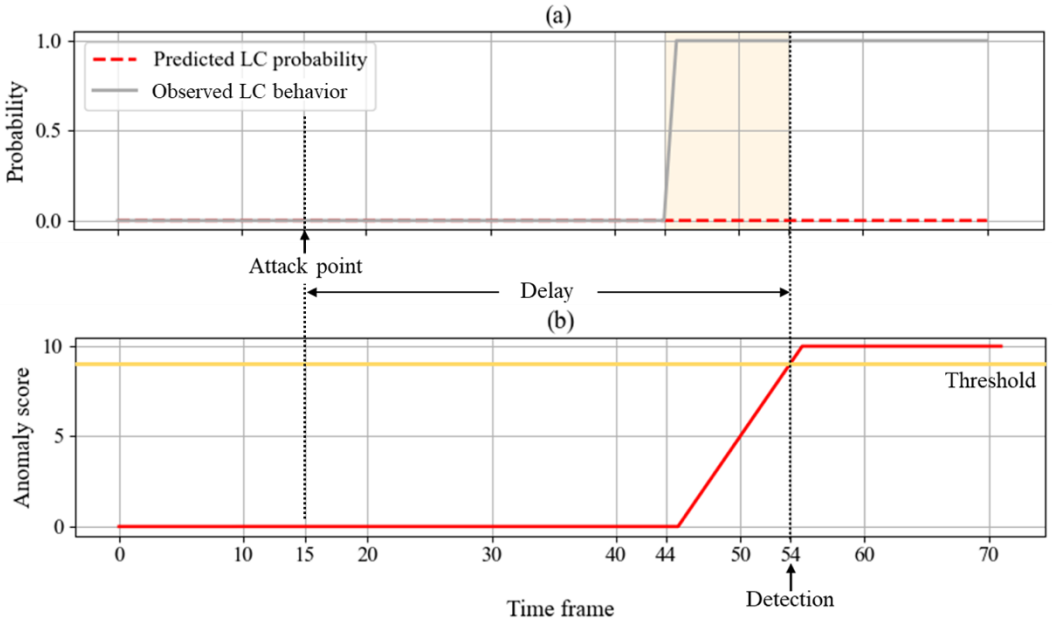


Figure 12: Anomaly detection under DoS attack:
(a) Observed vs. predicted lane-changing probabilities; (b) Example of anomaly scoring

2.5.3.2. Detection Performance for Fixed Scoring Window

We fixed the scoring window at $l = 10$ and evaluated detection performance using $\tau = 9.0$. Figure 13a shows the confusion matrix under FDI attack, averaged over 50 random selections of attacked vehicles. Among the 474 normal test trajectories and 50 falsified trajectories, the PGNN-based method detected all 50 attacks, achieving a recall of 1.0. However, 22 normal trajectories were misclassified as abnormal. This is likely attributed to aggressive or risky LC behaviors in the human-driven dataset, which were flagged as anomalies. The average detection delay was 0.39 s. This prompt detection demonstrates the PGNN model's potential for real-time anomaly identification in CAV environments.

We further investigated sensitivity to the threshold τ by varying τ and comparing detection recall, accuracy, and detection delay. As shown in Fig. 13b, increasing τ increases accuracy but also increases detection delay. This reflects a reduction in false positives at the cost of delayed alerts and, beyond a certain point, missed attacks. Accuracy exceeds 0.9 for $\tau \geq 8.1$, and the trade-off point is observed at $\tau = 9.0$, where recall remains 1.0 and accuracy reaches 0.96. Beyond this point, recall declines sharply (e.g., below 0.9 at $\tau = 9.6$).

In abnormal detection, the ability to reliably classify falsified trajectories as attacks is paramount. Because missing behaviors is more consequential than occasional false alarms, thresholds in the range $\tau = 8.1$ to 9.0 (yellow-shaded) provide desirable performance, maintaining full recall while improving accuracy. Within this range, detection delay remains below 0.4 s. In practice, τ should be selected based on operational tolerance for false positives versus false negatives by traffic management centers or fleet operators.

Figures 13c and 13d present the confusion matrix at $\tau = 8.9$ and the sensitivity analysis under DoS attack. The PGNN-based method detected all 50 attacks, while 24 normal trajectories

were misclassified as abnormal. In Fig. 13d, recall remains 1.0 over a wider range of τ . Accuracy exceeds 0.9 for $\tau \geq 8.1$, and the trade-off point is observed at $\tau = 8.9$. Beyond this value, recall decreases slightly but remains above 0.9. Unlike the FDI case, detection under DoS attack typically occurs only after LC initiation is observed; detection delay exceeds one second.

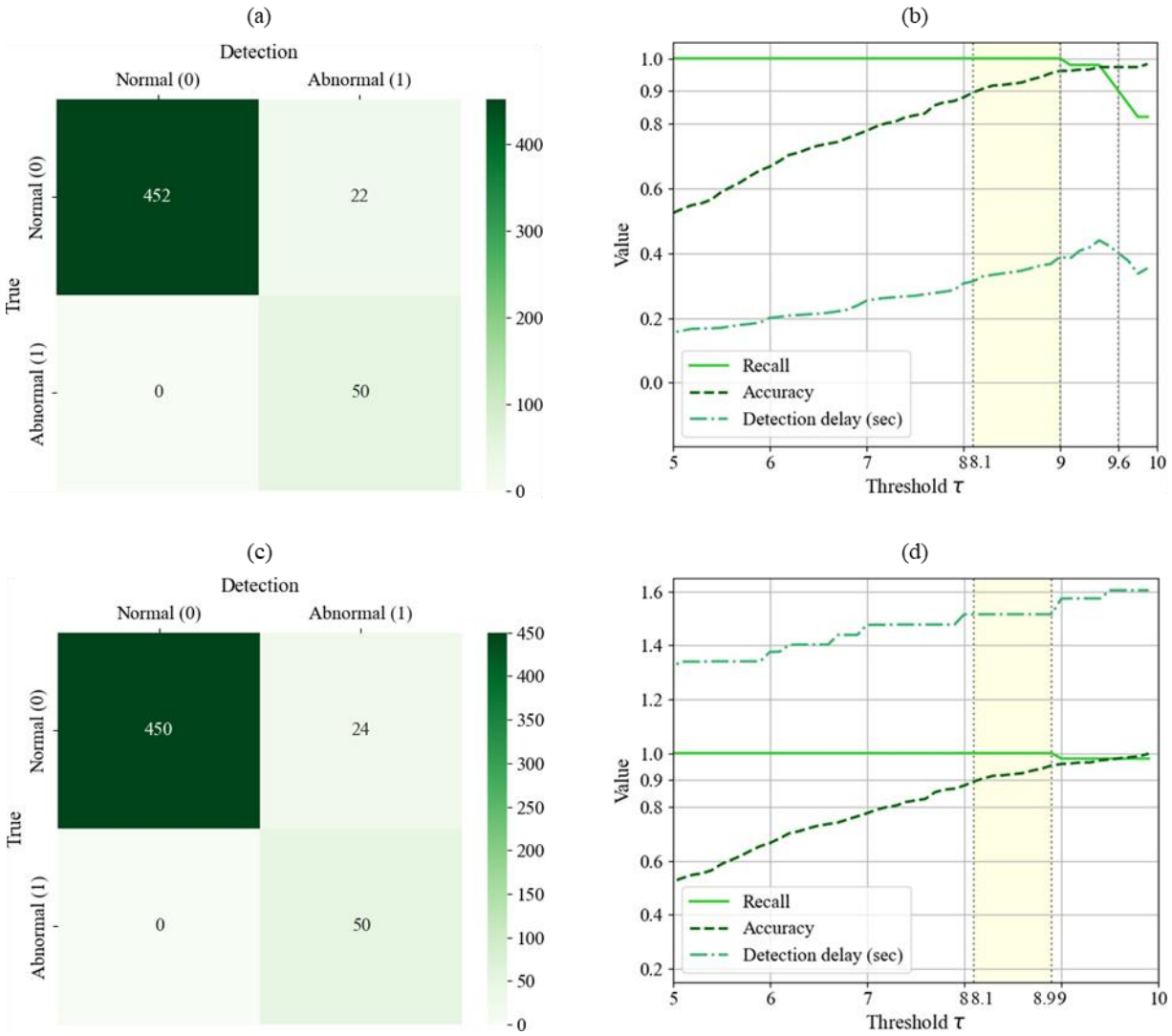


Figure 13: : Abnormal behavior detection using PGNN: (a) Confusion matrix at threshold τ = 9.0 under FDI attack; (b) Sensitivity to detection threshold under FDI attack; (c) Confusion matrix at threshold τ = 8.9 under DoS attack; (d) Sensitivity to detection threshold under DoS attack

2.5.3.3. Comparison with Physics-Only and Purely Data-Driven Detection

To evaluate the contribution of physics knowledge, we compared PGNN-based with LSTM-based and physics-only (game-theoretic) detection under FDI attack (Fig. 14). PGNN demonstrates more stable recall across detection thresholds τ . The recall curve of the PGNN model (solid green) remains relatively stable and begins to decline after $\tau = 9.0$, reaching a minimum of 0.82. In

contrast, the LSTM-based recall (solid blue) drops sharply after $\tau = 7.7$, and eventually falls to zero. This occurs because the LSTM model, without physics guidance, produces smaller deviations between falsified predictions and the observed behavior, resulting in lower anomaly scores.

The physics-only game-theoretic model's recall (solid gray) declines earlier, dropping rapidly after $\tau = 5$ and eventually reaching zero. Although the game-theoretic model achieves reason-able performance in binary LC classification, its larger MSE indicates less accurate probability estimation. Because anomaly detection relies on deviations in predicted probabilities, inaccurate probability estimation reduces detection capability. Accuracy curves (dashed lines) further show that the PGNN model is less sensitive to threshold changes. Overall, these results indicate that incorporating physics-based structure yields more robust detection across threshold settings.

Although the physics-only game-theoretic model does not achieve the lowest MSE, it provides a structural prior that sharpens decision boundaries and yields higher precision than the PGNN and LSTM in LC prediction. When embedded within the PGNN architecture, this physics knowledge serves as an inductive bias (Yu and Wang, 2024; Wang and Yu, 2025), while the LSTM component refines probability estimates and captures temporal dependence. This combination explains why the PGNN achieves more reliable anomaly detection performance than either the physics-only or purely LSTM-based model.

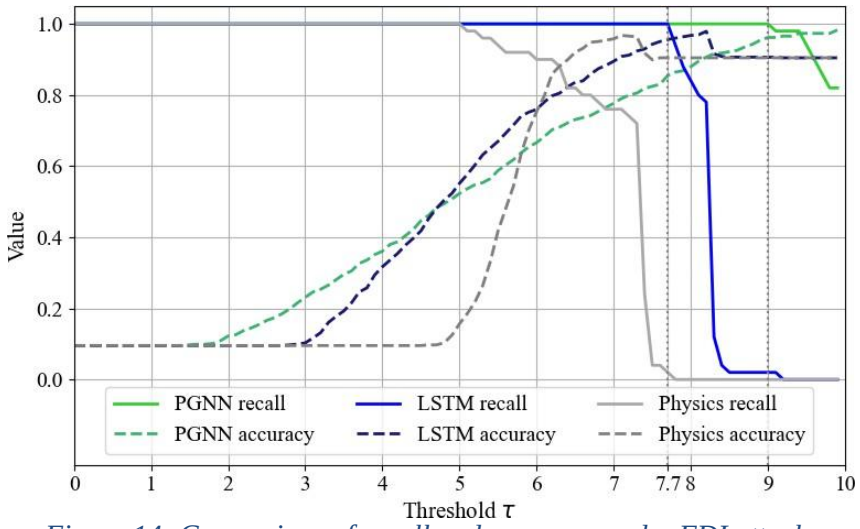


Figure 14: Comparison of recall and accuracy under FDI attack: PGNN-based, LSTM-based, and physics-only detection

2.5.3.4. Sensitivity Analysis: Scoring Window and Threshold

We extended the analysis by varying the scoring window l from 1 to 50 frames (up to 2 seconds at 25 Hz) and evaluated recall, accuracy, and detection delay across detection thresholds. Figure 15a presents recall over a grid of scoring windows and threshold-to-window ratio (τ/l) under FDI attack. Shorter scoring windows yield higher minimum recall across the threshold range. In addition, the threshold-to-window ratio required to maintain perfect recall (recall = 1.0) increases as the scoring window shortens. These results suggest that shorter windows provide more reliable recall performance under varied threshold settings.

Figure 15b illustrates, for each scoring window length, the best accuracy and the corresponding detection delay under the constraint that recall = 1.0. These results highlight a trade-off: shorter scoring windows reduce detection delays (i.e., faster response) but also reduce accuracy, increasing false positives. For instance, when $l < 4$, accuracy falls below 0.9, increasing the likelihood of false positives – normal behaviors mistakenly flagged as suspicious. From $l = 4$, accuracy stabilizes above 0.9, while detection delay increases with l. For $l > 43$, detection delay exceeds 1 second, which may hinder timely response in safety-critical settings. These findings underscore the importance of selecting l to balance detection accuracy and responsiveness.

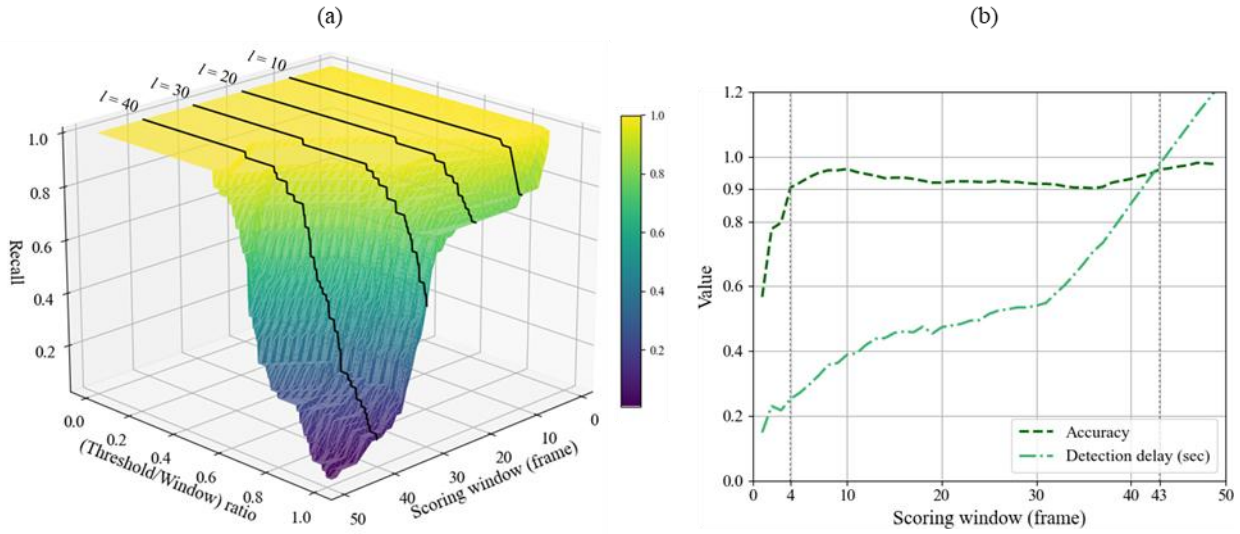


Figure 15: Detection sensitivity analysis under FDI attack:
(a) Recall by scoring window and threshold-to-window ratio τ/l;
(b) Best accuracy and corresponding detection delay under recall= 1.0

2.5.3.5. Comparison with Existing Methods

We compared the proposed PGNN-based detection against two LC-focused baselines: a one-class SVM (Ramyar et al., 2016) and a lightGBM model (Xu et al., 2023). The one-class SVM represents a classic anomaly detection approach, while the lightGBM represents a stronger machine-learning baseline that leverages complete LC trajectories and surrounding-vehicle information. Together, these models provide a meaningful benchmark to evaluate the added value of physics-guided inference beyond purely data-driven approaches. All models were trained using the same training data and evaluated on identical test and attacked datasets. For this comparison, 50 attacked datasets were randomly generated, and aggregated detection performance was reported. For the proposed PGNN framework, we used the best-performing combination of scoring window and detection threshold identified in the sensitivity analysis.

Table 7 summarizes the results. The one-class SVM achieved accuracy below 0.90 under both attack scenarios and showed poor recall (0.26) in the DoS attack. This can be attributed to its reliance on subject-vehicle speed and acceleration features and its segment-wise modeling approach (Ramyar et al., 2016). LightGBM, which uses information from surrounding vehicles (Xu et al., 2023), achieved perfect recall and an accuracy of 0.995 under FDI attack, outperforming

the PGNN-based model. However, under DoS attack, its accuracy and recall dropped to 0.936 and 0.985, indicating both false alarms and missed attacks. In contrast, the proposed PGNN-based model achieved perfect recall and maintained accuracy above 0.97 for both attack types.

All three models show higher detection performance under FDI attack than under DoS attack. This difference reflects the nature of the two attacks. FDI directly perturbs signal magnitudes and shifts the input distribution, making anomalies easier to identify using models trained on nominal data. In contrast, DoS interrupts data updates and holds recent data, which can remain within the nominal distribution and is therefore more difficult to detect. This limitation is most evident in the one-class SVM, which relies only on subject-vehicle states and often misclassified DoS attacks as normal. The robustness of the PGNN model under DoS attack stems from physics-guided learning combined with LSTM-based temporal modeling, which encodes linkage between trajectory states and LC behavior and enables detection of abnormal stagnation patterns.

Table 7: Anomaly detection of different models

Model	FDI attack		DoS attack	
	Accuracy	Recall	Accuracy	Recall
One-class SVM	0.873	0.946	0.807	0.256
LightGBM	0.995	1.000	0.936	0.985
Our PGNN-based model	0.991	1.000	0.979	1.000

2.6. Conclusions, Study Limitations and Direction for Future Research

This study proposed a PGNN framework to detect abnormal LC decisions under cyberattacks. The framework consists of (i) LC decision prediction module and (ii) an abnormal LC detection module. To model LC decision-making, we first developed a game-theoretic model that captures rational LC strategies based on interactions with surrounding vehicles. The game-theoretic outputs (LC probabilities) were then incorporated into an LSTM-based PGNN to regularize data-driven learning with physics-domain knowledge. Evaluation using the I-24 MOTION dataset demonstrated that the PGNN model achieved lower prediction error (MSE) compared to both the game-theoretic model and the standalone LSTM. The PGNN also substantially reduced false negatives in LC prediction, which is critical for maintaining safety in traffic networks.

To evaluate the framework's capability in a cybersecurity context, we simulated FDI and DoS attacks by manipulating vehicle speed profiles. The PGNN-based anomaly score detected falsified LC behaviors with high recall and a detection delay of less than 0.5 seconds under FDI attack. Sensitivity analysis of the detection threshold and scoring window suggested a trade-off among false alarms, missed detections, and detection delay. While shorter windows allowed for faster detection and greater robustness to threshold selection, they also resulted in more false positives. These findings suggest that the scoring window and threshold should be calibrated based on the operational context and tolerance for false alarms.

A limitation of this study arises from the use of human-driving trajectory data for model development and validation. While the proposed framework is adaptable to automated driving data, some behavioral cues used in this study (e.g., lateral acceleration as an intent-related signal)

may not directly correspond to variables in automated driving systems. Thus, future studies should incorporate AV-specific trajectory data to recalibrate the model and verify its applicability in fully automated driving environments. In addition, this study focuses on LC behavior, whereas cyberattacks that manipulate speed or disrupt communication can also affect car-following dynamics. Joint monitoring of both LC and car-following behaviors may provide complementary signals for anomaly detection, particularly under DoS attacks, where LC abnormalities are detected only after the maneuver initiation is observed. Finally, exploring additional cyberattack scenarios would help assess robustness and generalizability under more complex and realistic threat conditions. Despite these limitations, the proposed framework establishes an extensible foundation for trajectory-based anomaly detection that remains flexible and interpretable.

CHAPTER 3. ROBUST AND FLOW-EFFICIENT DISCRETIONARY LANE-CHANGING

3.1. Introduction

Automated vehicles (AV) rely on onboard sensing and software modules to perceive the environment, predict surrounding behaviors, and execute driving maneuvers through planning and control. Among tactical maneuvers, lane-changing (LC) is uniquely challenging due to its interactive nature and system-level consequences. An LC maneuver not only alters the subject vehicle's trajectory but also disturbs local traffic by inducing acceleration or deceleration responses from neighboring vehicles. While an optimal LC can harmonize flow and speed across lanes, empirical and theoretical studies have shown that poorly executed maneuvers can disrupt traffic, reduce bottleneck discharge rates, and trigger stop-and-go oscillations (Laval and Daganzo, 2006; Ahn and Cassidy, 2007). This duality in impacts is especially pronounced in discretionary LC, where an AV actively initiates a maneuver to optimize its own utility, such as speed gain, rather than to satisfy a mandatory routing constraint. In this setting, the LC decision creates a tension between ego-centric gain and system-level efficiency. Even when safety constraints are satisfied at the individual level, excessive or aggressive gap exploitation can introduce externalities that degrade traffic stability. Therefore, the fundamental challenge in autonomous discretionary LC is not merely safe execution, but coordinating ego-centric gain with interaction-aware behavior that mitigates unnecessary disturbances to the surrounding traffic stream.

A critical vulnerability in such gain-seeking decisions lies in their reliance on accurate state observations. In practice, autonomy stacks are highly susceptible to observation uncertainty arising from physical sensor degradation (e.g., LiDAR or camera noise), communication latency, or malicious cyberattacks such as sensor spoofing or jamming. Because discretionary LC decisions are typically formed by comparing expected utilities across lanes, they are particularly sensitive to perturbations in perceived gaps or relative speeds. Even small adversarial perturbations can shift the decision boundary and flip the selected action. Such decision flips may result in suboptimal maneuvering, overly conservative lane-keeping behavior, or poor LC initiation, each of which severely amplifies traffic-level disturbances and propagates instability upstream. Moreover, improper LC maneuvers can pose significant safety risks, potentially leading to crashes. Consequently, there is a critical need for robust LC decision-making that remains reliable under adversarial conditions.

To address tactical control challenges in LC, reinforcement learning (RL) has emerged as an attractive foundational method (Aradi, 2020; Kiran et al., 2021), as it can optimize long-horizon objectives without requiring explicit behavior models. A common trend in RL-based LC is to develop policies that optimize target-lane decisions and safety under nominal observations. For example, Mirchevska et al. (2018) proposed a deep Q-Network (DQN)-based RL approach for lane keeping and lane changing with a safety verification mechanism, reporting collision-free performance with high average speed. Hoel et al. (2018) trained a DQN agent to jointly decide longitudinal speed and LC decisions, showing that the learned policy can match or surpass a reference rule-based model in numerical simulation. Subsequent studies have incorporated environmental uncertainty (Alizadeh et al., 2019) and multi-agent formulations balancing safety,

passenger comfort, and traffic efficiency (Zhou et al., 2022).

Recognizing that RL policies are highly sensitive to adversarial perturbations and observation uncertainties (Pinto et al., 2017; Gleave et al., 2019), recent efforts have adopted adversarial learning paradigms to improve robustness against perturbed inputs. For example, He et al. (2022a) designed a constrained adversarial RL framework with reward components considering safety, self-efficiency, and vehicle dynamics. He et al. (2004) further proposed a defense-aware robust RL framework that trains a robust defender against worst-case observational perturbations to reduce safety-critical failures. Beyond single-agent paradigm, recent work by Wang et al. (2025) explored a multi-agent dueling DQN to improve safety and performance in mixed autonomy environments. However, prior studies predominately target mandatory LC scenarios such as highway merging or exiting (He et al., 2022b; Bagwe et al., 2023), with limited focus on discretionary maneuvers from a robustness perspective.

Despite these advances, existing robust learning frameworks are not designed to simultaneously capture the three dynamics central to discretionary LC: (i) the ego-centric incentive for speed gain, (ii) interaction-aware execution that anticipates a target-lane follower's response, and (iii) traffic flow-level disturbances induced by the maneuver. This holds even when considering adversarial perturbations to the state observation at these three layers of the control objective. Consequently, policies learned without these considerations often become overly self-centered and aggressive in gap exploitation, satisfying individual safety constraints at the expense of broader traffic stability.

This limitation motivates our primary research question: *Can an autonomous agent learn a discretionary LC decision-making policy that remains safe against adversarial observation perturbations while simultaneously achieving meaningful speed gains and minimizing traffic flow disturbances?* To address this, we propose a hierarchical adversarial reinforcement learning (HARL) framework for robust and flow-efficient discretionary LC. HARL explicitly models the interaction among a subject vehicle (defender) performing discretionary LC, a target-lane follower that responds to the LC maneuver, and an attacker that injects sensor-level perturbations into the defender's observations. At the upper level, the attacker selects a worst-case bounded perturbation to maximize safety-related risk, while the subject vehicle acts as the defender and learns a robust discrete maneuver decision policy (keep/left/right) under corrupted observations. At the lower level, conditioned on the defender's decision and the perturbed observation, a Stackelberg interaction resolves maneuver execution by determining the LC initiation timing along with longitudinal control actions. This lower-level interaction shapes both individual-level performance and flow-level outcomes, thereby coupling robustness with interaction-aware execution.

The main contributions of this paper are summarized as follows:

- A two-level dynamic framework for discretionary LC that couples robust decision-making with interaction-aware execution. This is achieved by integrating the upper-level adversarial policy learning with a lower-level behavioral Stackelberg game to resolve LC initiation timing and longitudinal accelerations.
- A sensor-level white-box perturbation model for worst-case robustness evaluation, where the attacker applies bounded observation perturbations directly to the defender's policy inputs.
- A multi-objective reward design that jointly optimizes safety, self-efficiency, and flow-stability. The reward encourages socially compatible discretionary LC behaviors that achieve speed gains while mitigating traffic disturbances under adversarial observation

perturbations.

The remainder of this paper is organized as follows. **Section** 3.2 presents the problem formulation and the HARL framework. **Section** 3.3 details the implementation. **Section** 3.4 reports the experimental results. **Section** 3.5 concludes the paper.

3.2. Methodology

In this section, we formulate the robust discretionary LC decision-making under adversarial observation perturbations. The subject vehicle (defender) first determines a maneuver decision under potentially corrupted observations and resolves execution-level interactions that govern LC initiation timing and longitudinal acceleration.

3.2.1. Problem Formulation

The discretionary LC decision-making problem is formulated as a Markov Decision Process (Howard, 1960), defined by the tuple (S, A, p, r, γ). At each time step t, the subject vehicle observes the current traffic state $s_t \in S$, selects an action $a_t \in A$, receives an instantaneous reward $r_t = r(s_t, a_t)$, and transitions to the next state s_{t+1} according to $p(s_{t+1}|s_t, a_t)$.

In this study, the action corresponds to the discretionary maneuver decision:

$$a_t \triangleq g_t \in G = \text{keep, left, right}$$

which indicates whether the subject vehicle keeps its lane or attempts a lane change to the left or right.

To reflect the behavioral motivation of discretionary LC, the maneuver decision is invoked by a speed-based trigger:

$$Trig_t = 1\left[v_{SV,t} \leq v_{des} - \Delta_{trig}\right]$$

where v_{des} is the desired speed and Δ_{trig} is a tolerance margin. When the trigger is not satisfied or the system is in cooldown, the maneuver decision is fixed as g_t keep.

Given g_t, the subject vehicle determines the LC initiation timing and longitudinal acceleration while accounting for the response of the target-lane follower. The realized execution outcomes govern both the reward and the state transition.

The objective is to determine a policy π^* that maximizes the expected cumulative discounted reward over a finite horizon T:

$$\pi^* \in \arg\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t r(s_t, a_t)\right]$$

where $\gamma \in [0,1)$ denotes the discount factor.

3.2.2. Framework

The proposed HARL framework consists of two interconnected levels: (i) the upper-level maneuver decision under adversarial observation, and (ii) the lower-level behavioral execution through a Stackelberg interaction.

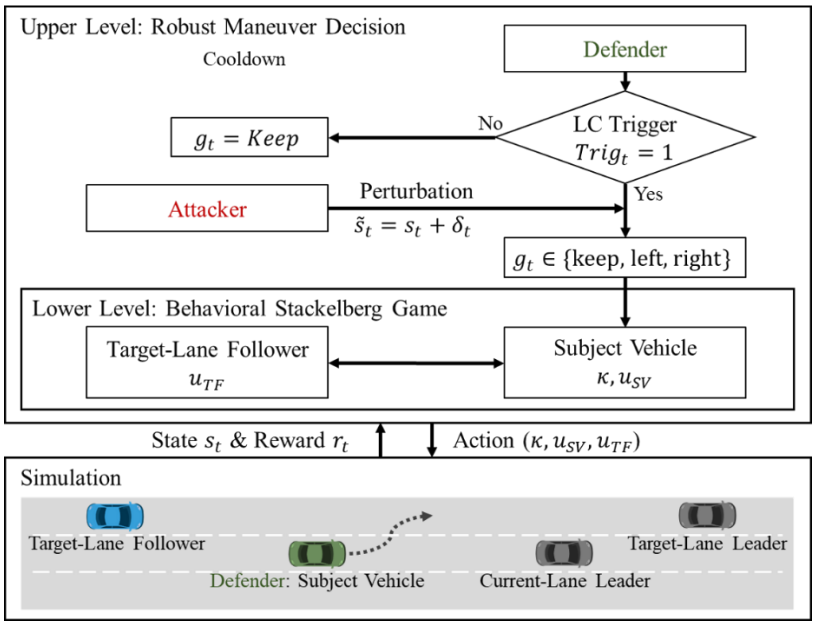


Figure 16: Structure of HARL

At each time step t, the LC trigger is evaluated. When $Trig_t = 0$ or during cooldown, the defender (subject vehicle) follows keep-lane behavior. Otherwise, the observed state may be perturbed by an attacker, and the defender selects a maneuver decision $g_t \in$ keep, left, right based on the perturbed observation $\tilde{s}_t$. If $g_t \in$ left, right, the lower-level module determines the LC initiation timing κ_t and longitudinal accelerations over a finite horizon of K steps with early termination. After execution, a cooldown is applied to prevent unrealistic maneuver frequency.

3.2.3. Upper-Level: Robust Maneuver Decision

At the upper level, the defender determines the discretionary maneuver decision under potentially perturbed observations. The trigger $Trig_t$ is assumed to be evaluated from reliable signals. Once activated, the observed state s_t may be corrupted by a perturbation δ_t bounded within an L_2-ball of radius ϵ_s:

$$\tilde{s}_t = s_t + \delta_t, \quad \| \delta_t \|_2 \leq \epsilon_s$$

The maneuver decision is therefore based on $\tilde{s}_t$ rather than the true state. To characterize robustness, we consider worst-case perturbations within the feasible set that increase safety-related risk. Safety risk is quantified using a time-to-collision (TTC) metric computed from the realized execution outcome:

$$C_{\text{safe}}(s_t, g_t) = \omega_{ttc,up} 1[TTC_{i,t} < \tau]$$

where ω is a nonnegative coefficient. The TTC for surrounding vehicle i measures the time remaining before a potential collision and is defined as:

$$TTC_i = \begin{cases} \dfrac{d_i}{v_{SV} - v_i}, & \text{if } v_{SV} > v_i \\ +\infty, & \text{otherwise} \end{cases}$$

where v_{SV} and v_i denote the longitudinal speeds of the subject vehicle and a relevant surrounding vehicle i, respectively, and di is the longitudinal gap between the subject vehicle and vehicle i. The minimum TTC among relevant surrounding vehicles is used to represent the most critical interaction.

Importantly, the perturbed observation $\tilde{s}_t$ is also used in the lower-level module, allowing adversarial perturbations to influence both maneuver selection and execution planning.

3.2.4. Lower-Level: Behavioral Stackelberg Game

At the execution level, the interaction between the subject vehicle and the target-lane follower is modeled as a Stackelberg game (Simaan and Cruz, 1973), reflecting the leader-follower structure in LC maneuvers. Conditioned on the upper-level maneuver decision g_t, the subject vehicle commits to an execution plan – specifically, when to initiate the LC and how to approach the target gap – while the target-lane follower responds by adjusting its longitudinal acceleration to preserve safety and driving comfort (Yoo and Langari, 2020).

We consider a fixed execution horizon of K steps, (i.e., $K\Delta t$ seconds). The execution terminates early when the lane change is completed or aborted. After termination, a cooldown of T_{cd} seconds is applied, during which the trigger is not evaluated to prevent unrealistically frequent lane changes.

Given $g_t \in \text{left}, \text{right}$, the subject vehicle selects (i) the LC initiation timing κ_t and (ii) the longitudinal acceleration u_{SV} applied before initiation, while anticipating the optimal response of the target-lane follower. The resulting optimization is formulated as:

$$\max_{\kappa, u_{SV}} \mathbb{R}_{SV}\left(s, g, \kappa, u_{SV}, u_{TF}^*(\kappa, u_{SV})\right)$$

subject to

$$u_{TF}^*(\kappa, uu_{SV}) = arg\max_{u_{TF}} \mathbb{R}_{TF}\ (s, g, \kappa, u_{SV}, u_{TF}))$$

where s denotes the traffic state, $g \in G$ is the maneuver decision, and u_{SV}, u_{TF} represent the longitudinal accelerations of the subject vehicle and the target-lane follower, respectively. The LC initiation timing decision is discrete:

$$\kappa \in 0,1, \ldots, K$$

which indicates the number of steps to wait before initiating the lane change.

Rewards are accumulated at each step during the execution horizon. The choice of κ affects the reward trajectory by altering the realized safety risk, achievable speed gain, and the disturbances imposed on the following vehicles.

1) Subject Vehicle Reward: The reward function of the subject vehicle (which also serves as the defender in the adversarial setting) is defined as:

$$\mathbb{R}_{SV=D} = -(\omega_{ttc} 1[TTC < \tau]) + (\omega_v \Delta v - \omega_u |u_{SV}|) - \omega_{flow} \sum_{j \in \mathcal{N}_{TF}} |\Delta v_j|$$

where all weighting parameters ω are nonnegative and regulate the relative importance of each objective.

The first term represents the safety objective, penalizing critically unsafe interactions when the minimum relevant TTC falls below a predefined threshold τ (1.5 seconds in this study).

The second term captures self-efficiency, where Δv represents the potential (or realized) speed gain associated with executing maneuver g with timing κ. The acceleration penalty $\omega_u |u_{SV}|$ discourages aggressive longitudinal control and promotes smooth maneuvering.

The last term represents the flow-stability, penalizing speed disturbances imposed on affected target-lane followers $\mathcal{N}_{TF}$. This discourages lane changes that induce excessive deceleration or instability in the target lane.

2) Target-Lane Follower Reward: The reward of the target-lane follower is formulated as:

$$\mathbb{R}_{TF} = -\omega_{ttc} 1[TTC < \tau] - \omega_{comfort} |u_{TF}|$$

where the first term ensures collision avoidance and the second term penalizes the longitudinal discomfort in response to the subject vehicle's maneuver.

3.3. Hierarchical Adversarial Reinforcement Learning Implementation

We now describe how the formulation is implemented using HARL. The training process alternates between (i) generating worst-case observation perturbations (inner maximization) and (ii) updating the defender policy using Proximal Policy Optimization (PPO) (outer minimization).

3.3.1. White-Box Adversarial Perturbation

We consider a sensor-level white-box attacker that has full access to the defender's policy network and exploits gradient information of the safety surrogate with respect to the input state to generate worst-case perturbations during training (see Algorithm 1). Given a state s_t, the attacker seeks a bounded perturbation δ_t within the budget ϵ_s (Eq. 4). The perturbation aims to maximize safety-related risk after maneuver selection:

$$\delta_t^* \approx arg \max_{||\delta_t||_2 \leq \epsilon_s} C_{\text{safe}}(s_t, g_t(\tilde{s}_t)), \quad g_t(\tilde{s}_t) \sim \pi_D(\cdot | \tilde{s}_t)$$

Because the TTC indicator in C_{safe} is non-differentiable, we use a differentiable surrogate:

$$\tilde{C}_{\text{safe}}(s_t, g_t) = \omega_{ttc,up} \cdot \text{softplus}(\tau - TTC_t)$$

The inner maximization is approximated via projected gradient ascent. After each gradient step, the perturbation is projected onto the L_2-ball to enforce feasibility.

Algorithm 1 summarizes the perturbation procedure.

Algorithm 1 White-Box Adversarial Attack via Projected Gradient Ascent

---

Input: current state s_t, defender policy $\pi_D(\cdot \mid \cdot)$, step size α, attack steps I, perturbation budget ϵ_s

Input: lower-level execution module $\Psi(\cdot)$for evaluating TTC (used in $\tilde{C}_{\text{safe}}$)

Initialize perturbation $\delta^{(0)} \leftarrow 0$ (or random s.t. $\left\|\delta^{(0)}\right\|_2 \leq \epsilon_s$)

for $i = 0$ **to** $I - 1$ **do**

$\tilde{s}_t^{(i)} \longleftarrow s_t + \delta^{(i)}$

$g_t^{(i)} \sim \pi_D\left(\cdot \middle| \tilde{s}_t^{(i)}\right)$

Compute $J\left(\delta^{(i)}\right) \leftarrow \tilde{C}_{safe}\left(s_t, g_t^{(i)}\right)$ using Ψ

$\delta^{(i+1)} \leftarrow \delta^{(i)} + \alpha \nabla_\delta J\left(\delta^{(i)}\right)$

$\delta^{(i+1)} \leftarrow \Pi_{\|\delta\|_2 \leq \epsilon_s}\left(\delta^{(i+1)}\right)$

end for

Output: $\delta_t^* \leftarrow \delta^{(I)}$, perturbed observation $\tilde{s}_t \leftarrow s_t + \delta_t^*$

3.3.2. Defender Policy Optimization

The defender aims to learn a maneuver decision policy that remains robust under worst-case observation perturbations. This corresponds to the robust objective:

$$\min_{\pi_D} \max_{\pi_A} \mathbb{E}_{s,g\sim(\pi_D,\pi_A)}\left[C_{\text{safe}}(s, g)\right]$$

In practice, the attacker generates adversarial perturbations using the procedure above, and the defender updates its policy using rollouts collected under perturbed observations. We adopt PPO due to its stability in on-policy learning.

Given rollouts collected under the current policy $\pi_{D,old}$ with perturbed observations, the defender maximizes the clipped surrogate objective:

$$L^{PPO}(\theta_D) = \mathbb{E}\left[\min\left(\rho_t(\theta_D)\hat{A}_t, \text{clip}(\rho_t(\theta_D), 1-\epsilon, 1+\epsilon)\hat{A}_t\right)\right]$$

where

$$\rho_t(\theta_D) = \frac{\pi_D(g_t|\tilde{s}_t)}{\pi_{D,old}(g_t|\tilde{s}_t)}$$

Here, θ_D denotes the defender parameters and $\hat{A}_t$ is the advantage estimate computed from perturbed rollouts. The reward rt is obtained from the lower-level Stackelberg execution outcomes in **Section** 3.2.4. Therefore, policy improvement directly reflects discretionary LC objectives – safety, self-efficiency, and flow-stability – under adversarial observations. Following standard PPO, we include value regression and an entropy bonus (Schulman et al., 2017):

$$L = L^{PPO} + c_1 L^{value} - c_2 \mathcal{H}(\pi_D)$$

where c_1 and c_2 are coefficients and $\mathcal{H}$ denotes policy entropy.

3.3.3. Algorithm Implementation

Algorithm 2 outlines the overall HARL training process. At each time step, the trigger is evaluated. If the trigger is inactive or the system is in cooldown, the defender defaults to keep-lane behavior without adversarial perturbation. When the trigger is active, the attacker computes a worst-case bounded perturbation, and the defender selects the maneuver decision g_t under the perturbed observation. If g_t = keep, the simulator advances by one step. Otherwise, the lower-level execution module runs for up to K steps with early termination. A cooldown is applied only after execution terminates.

Algorithm 2 Hierarchical Adversarial Reinforcement Learning (HARL)

Input: I, ϵ_s, PPO epochs E, horizon K, time step Δt, cooldown T_{cd}, trigger margin Δ_{trig}
Initialize defender parameters θ_D and value parameters ϕ
for iteration $n = 1$ **to** N **do**
  Reset environment; $\mathcal{B} \leftarrow \emptyset$; cooldown counter c ← 0
  for timestep $t = 0$ **to** $T - 1$ **do**
    Observe state s_t and compute trigger $\text{Trig}_t = 1[v_{SV,t} \le v_{des} - \Delta_{trig}]$
    If $c > 0$, set $\text{Trig}_t \leftarrow 0$ and c ← c − 1
    if $\text{Trig}_t = 0$ **then**
      Set $g_t \leftarrow$ keep, $\tilde{s}_t \leftarrow s_t$ {no attack}
    else
      Compute δ_t^* via Alg. 1, set $\tilde{s}_t \leftarrow s_t + \delta_t^*$ {attack}
      Sample $g_t \sim \pi_D(\cdot|\tilde{s}_t; \theta_D)$
    end if
    if g_t = keep **then**
      Step simulator; store $(\tilde{s}_t, g_t, r_t, s_{t+1}, d_t)$ in $\mathcal{B}$
    else
      Execute lower-level Stackelberg module with early termination:
        $\{(\tilde{s}_{t+j}, r_{t+j}, s_{t+j+1}, d_{t+j})\}_{j=0}^{J-1}$, term ← $\Psi(\tilde{s}_t, g_t; K)$
      Store all transitions in $\mathcal{B}$; set t ← t + J − 1
      if term=1 **then**

set c ← $\lceil T_{cd}/\Delta t \rceil$
end if
end if
end for
Update (θ_D, ϕ) using PPO with $\mathcal{B}$ (omitted for brevity)
end for

The defender state encodes local traffic conditions relevant to discretionary LC, including accelerations, speeds, and gaps of the subject vehicle and surrounding vehicles (Table 8).

Table 8: State variables of the defender

Variable (unit)	Definition
a_{SV} (m/s^2)	Acceleration of the subject vehicle (SV)
a_{TF} (m/s^2)	Acceleration of the target-lane follower (TF)
v_{SV} (m/s)	Speed of the SV
v_{TF} (m/s)	Speed of the TF
v_{CL} (m/s)	Speed of the current-lane leader (CL)
v_{TL} (m/s)	Speed of the target-lane leader (TL)
v_{FTF} (m/s)	Speed of the TF's follower (FTF)
d_{TF} (m)	Gap from SV to TF
d_{CL} (m)	Gap from SV to CL
d_{TL} (m)	Gap from SV to TL

Note: All variable are longitudinal.

The defender's action is the discrete maneuver decision $g_t \in \{\text{keep}, \text{left}, \text{right}\}$. Conditioned on g_t, the lower-level execution module determines the LC initiation timing κ_t and longitudinal accelerations (u_{SV}, u_{TF}) over a horizon of K steps. The reward at each time step is computed using the lower-level reward formulation in **Section** 3.2.4, after resolving execution outcomes such as TTC and induced speed disturbances.

Hyperparameters are summarized in Table 9.

Table 9: The main hyperparameters

Parameters	Value	Parameters	Value
Discount factor γ	0.99	Time step Δt	0.1 s
Execution horizon K	30	Cooldown T_{cd}	2 s
Trigger margin Δ_{trig}	0.1 m/s	Adversarial learning rate α	0.001
Attack steps I	5	Perturbation budget ϵ_s	3.5 m
PPO clip ϵ	0.2	PPO learning rate λ	0.0005
PPO epochs E	5	Mini-batch size	64

3.4. Experiments and Results

3.4.1. Simulation Environment

We evaluate the proposed robust discretionary LC framework using SUMO. The simulation network is configured as a three-lane ring road to provide a closed traffic environment. To ensure frequent discretionary LC opportunities, 20% of vehicles are designated as slow vehicles with reduced desired speeds. This configuration induces sustained speed loss for fast vehicles, thereby activating the LC trigger repeatedly and enabling sufficient learning of discretionary maneuvers.

All background vehicles (including slow vehicles) follow SUMO's LC2013 model, and their longitudinal motion is governed by the intelligent driver model (IDM). The subject vehicle follows a hybrid control scheme. When the trigger is inactive or during cooldown, it follows IDM and keeps its lane. When the trigger is active, the maneuver decision g_t is selected by HARL, and the subsequent execution (LC initiation timing and longitudinal accelerations) is resolved by the lower-level Stackelberg module.

The maximum speed is set to 60 km/h. Adversarial observation perturbations are applied to the subject vehicle's perceived gaps to surrounding vehicles. The perturbation budget is set to 3.5 m, approximately one vehicle length, which is sufficient to alter perceived gap without being unrealistically large.

To isolate the role of execution-level interaction modeling, we compare the HARL against a maneuver-policy-only baseline ("policy-only"). To evaluate the effect of the flow-stability objective, we further compare against HARL-noFlow, which sets $\omega_{flow} = 0$.

3.4.2. Results and Discussion

We first examine training performance in terms of episode return. Fig. 17 presents the learning curves of policy-only, HARL-noFlow, and HARL, where solid lines denote mean returns and shaded regions indicate standard deviations. HARL learns to achieve higher returns as training progresses. The fluctuations observed in the curves are mainly due to environmental stochasticity in the simulation. Although HARL-noFlow initially achieves high returns, its performance becomes unstable over training. This suggests that neglecting flow-level considerations may lead to policies that cannot sustain consistent performance. In contrast, policy-only yields stable yet substantially lower returns because its LC decisions are driven by immediate reward signals, resulting in a more myopic maneuver policy compared to the other two strategies.

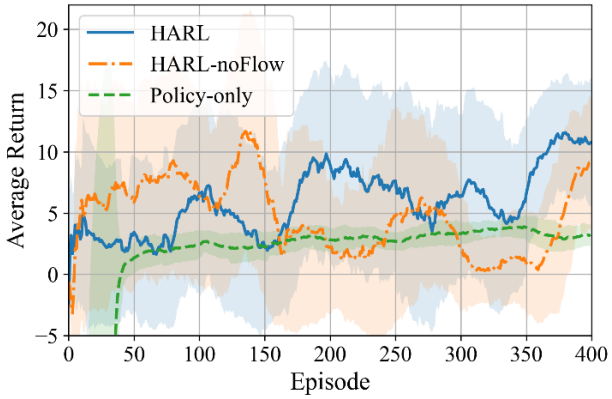


Figure 17: Training curves using Policy-only, HARL-noFlow, and HARL

Fig. 18 shows the ratio of discrete maneuver decisions during training for the three strategies. Policy-only exhibits a dominant preference for "keep" throughout training, reflecting an overly conservative maneuver policy. Compared to HARL-noFlow, HARL selects "keep" more frequently in the early stage of training. This pattern aligns with the flow-stability objective: the agent initially suppresses lane changes and gradually increases left/right decisions as it learns to obtain speed gains without inducing excessive disturbances to traffic.

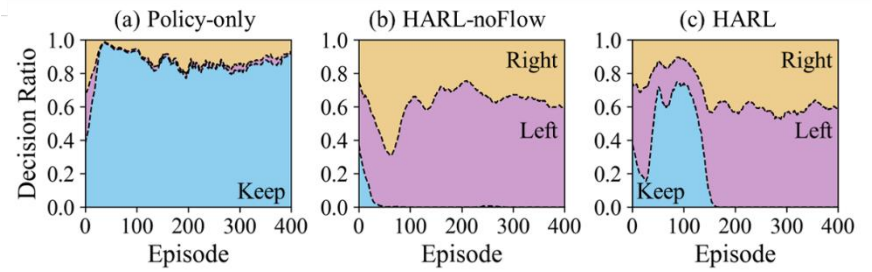


Figure 18: Ratio of lane-changing maneuver decision g_t

Table 10 summarizes evaluation results over 40 random seeds. HARL achieves the highest mean speed and flow rate while exhibiting a smaller standard deviation of acceleration than HARL-noFlow, indicating enhanced traffic stability. All three strategies result in near-zero potential collisions, demonstrating that the learned policies maintain safety even under worst-case observation perturbations.

Table 10: Evaluation of different LC strategies

Metric	Policy-only	HARL-noFlow	HARL
Return	1.76 ± 9.69	52.13 ± 20.24	46.69 ± 15.19
Mean speed (m/s)	7.42 ± 0.35	7.46 ± 0.51	7.60 ± 0.30
Flow (veh/hr)	2226.3 ± 169.6	2241.8 ± 167.5	2287.2 ± 96.9
Std. Accel (m/s^2)	0.30 ± 0.04	0.37 ± 0.05	0.33 ± 0.04
Potential collision	0.00 ± 0.00	0.65 ± 1.56	0.10 ± 0.30

Fig. 19 reports the average number of potential collisions under attacks with different perturbation budgets, where the subject vehicle is randomly selected from the set of fast vehicles. HARL consistently yields fewer than five potential collisions across perturbation levels, indicating stronger robustness to observation perturbations.

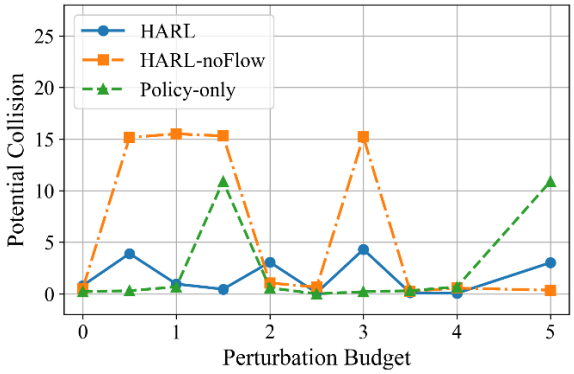


Figure 19: Number of potential collisions with perturbation budget ϵ_s

3.5. Conclusions, Study Limitations and Direction for Future Research

This paper presented a HARL framework for robust and flow-efficient discretionary LC under adversarial observation perturbations. The central objective was to address a fundamental tension in discretionary LC maneuvers: achieving ego-centric speed gains while mitigating traffic-level disturbances, and ensuring that such coordination remains reliable under worst-case sensor-level perturbations.

The proposed framework integrates two layers. At the upper-level, a defender policy learns a robust maneuver decision (keep/left/right) under bounded adversarial observation perturbations generated by a white-box attacker. At the lower-level a Stackelberg execution game resolves LC initiation timing and longitudinal accelerations by explicitly modeling interaction with the target-lane follower. This two-level architecture enables the maneuver decision to internalize both robustness and interaction-aware execution outcomes, thereby coupling adversarial resilience with flow-level responsibility. Simulation results show that HARL achieves higher mean speed and flow compared with baseline strategies, while maintaining low acceleration variability, indicating improved traffic stability. Under worst-case bounded observation perturbations, the HARL maintains potential collisions below five and exhibits stronger robustness. These findings suggest that integrating interaction-aware execution, flow-level consideration, and adversarial learning yields socially compatible and robust discretionary LC behavior.

Despite these contributions, several limitations remain. First, the flow-stability term assumes access to upstream vehicle information (e.g., the follower of the target-lane follower), but the communication pathway for obtaining such information is not explicitly modeled. In practice, this may rely on V2V communication or cooperative perception, which warrants further study under partial observability. Second, the attacker is modeled as a white-box adversary with full gradient access, representing a worst-case benchmark. More realistic grey-box or black-box attack settings should be examined to evaluate robustness under limited attacker knowledge. Third, performance should be systematically evaluated across broader traffic demand levels and heterogeneity conditions to assess scalability. Ultimately, validating the proposed hierarchical robust LC structure with empirical trajectory data or experimental platforms would further strengthen its practical relevance.

CHAPTER 4. CONCLUDING REMARKS

This study investigated cybersecurity challenges in CAVs with a focus on lane-changing behavior. Recognizing that CAV decision-making is highly sensitive to perception integrity, the study addressed both (i) falsified behavior detection and (ii) robust decision-making under cyberattacks. By combining physics-informed modeling with learning-based approaches, this work provides a structured foundation for improving the resilience of tactical vehicle maneuvers.

The contributions of this study are summarized as follows:

- Development of a PGNN framework for detecting falsified lane-changing decisions under cyberattacks. The model integrates a game-theoretic lane-changing structure with LSTM-based temporal learning to regularize data-driven prediction of lane-changing probabilities and enable anomaly detection.
- Design and evaluation of an anomaly scoring mechanism capable of detecting falsified lane-changing behaviors with high recall and accuracy under simulated FDI and DoS attacks.
- Proposal of a HARL framework for robust and flow-efficient discretionary lane-changing. The framework integrates upper-level robust maneuver decision-making with a lower-level Stackelberg execution module to account for interaction with surrounding vehicles.
- Demonstration that robust and interaction-aware lane-changing policies can achieve improved traffic efficiency and stability while maintaining safety under worst-case bounded observation perturbations.

Together, these components address complementary aspects of CAV cybersecurity: monitoring abnormal behavior through physics-informed detection and enabling resilient tactical control through adversarial learning.

There are several limitations of this study. First, the detection framework was developed and validated using human-driving trajectory data. Although the modeling structure is adaptable to automated driving systems, recalibration with AV-specific trajectory data is necessary to ensure full applicability. Second, the detection module focuses on lane-changing behavior; extending monitoring to jointly consider car-following dynamics may provide more comprehensive anomaly signals, particularly under DoS attacks. Third, the flow-stability objective in the robust lane-changing model assumes access to surrounding vehicle states when evaluating traffic-level disturbances. Although such information may be available in connected environments, the communication and perception mechanisms required to obtain these signals are not explicitly modeled. Future research should investigate the impact of communication delays partial observability, and sensing uncertainty on the robustness and efficiency of the learned policies. Fourth, the framework is evaluated under structured and bounded adversarial perturbations within a controlled simulation environment. While this provides a principled benchmark for worst-case robustness, the attack model does not fully capture the diversity of real-world cyber threats. Examining robustness under more realistic adversarial assumptions would provide a more comprehensive understanding of the model's resilience.

Finally, translating the proposed frameworks to real-world CAV systems requires validation beyond controlled datasets and simulation settings. Experimental validation using field data or controlled testbeds would further strengthen the feasibility and scalability of approaches.

CHAPTER 5. SYNOPSIS OF PERFORMANCE INDICATORS

5.1. USDOT Performance Indicators I

Three (3) transportation-related courses were offered during the study period by the PIs and the teaching assistant associated with the research project. These courses include CEE 370 (Transportation Engineering), CEE 572 (Transportation Operations), and CEE 679 (AI & Data Science in Transportation).

One (1) graduate student participated in this research project during the study period. CCAT grant funds associated with this project supported the student's doctoral training in a transportation-related advanced degree program.

5.2. USDOT Performance Indicators II

Research Performance Indicators: one (1) journal article under review in *Transportation Research Part C*, one (1) journal article in preparation, three (3) conference presentations were produced from this project, including two (2) completed presentations and one (1) under review. These presentations include contributions to the 2025 and 2026 Transportation Research Board (TRB) Annual Meetings.

Leadership Development Performance Indicators: This research project generated two (2) academic engagements and one (1) industry engagement. The PIs held positions in four (4) national and international organizations addressing issues related to the scope of this research, including Transportation Research Board Transportation Operations and Management Section, International Advisory Committee for International Symposium on Traffic and Transportation Theory, ASCE Connected & Autonomous Vehicle Impacts Committee, and IEEE Emerging Transportation Technology Testing Technical Committee. The PIs also held editorial positions in four (4) international journals, *IEEE Transactions on Intelligent Transportation Systems*, *Transportation Research Part B*, *Transportation Science*, and *Transportation Research Record*.

Education and Workforce Development Performance Indicators: The research outcomes, including the developed methodologies and processed datasets, will be incorporated into course content for CEE 572 (Transportation Operations) and CEE 679 (AI & Data Science in Transportation) at the University of Wisconsin-Madison. The students in these classes will soon be entering the workforce. Thereby, the research helped enlarge the pool of people trained to develop knowledge and utilize at least a part of the technologies developed in this research, and to put them to use when they enter the workforce.

The outputs, outcomes, and impacts are described in **Chapter** 6.

CHAPTER 6. STUDY OUTCOMES AND OUTPUTS

6.1. Outputs

6.1.1. Publications, Conference Presentations, and Presentations

(a) Journal Papers

Kim, Y., Zhong, X., Shi, L., Kontar, W., Chen, S., and Ahn, S. Detecting Falsified Lane-Changing Behavior in Connected Automated Vehicles Using Physics-Guided Neural Networks. *Under review*. Available at SSRN: http://dx.doi.org/10.2139/ssrn.6353060.

(b) Conference Presentations

Kim, Y., Kontar, W., Chen, S., and Ahn, S. (2026) Robust and Flow Efficient Discretionary Lane-Changing. IEEE International Conference on Intelligent Transportation Systems (ITSC), Naples, Italy. *Under review*.

Kim, Y., Zhong, X., Shi, L., Kontar, W., Chen, S., Ahn, S., Chitturi, M., and Noyce, D. (2026). Learning to Detect Cyberattacks on Lane-Changing via Physics-Informed Neural Inference. 105th Annual Meeting of the Transportation Research Board, Washington DC, United States.

Kim, Y., Kontar, W., Zhong, X., Zhang, Y., Chen, S., and Ahn, S. (2025). Traffic Consequences of Automated Vehicle Lane-Changing Design Philosophy. 104th Annual Meeting of the Transportation Research Board, Washington DC, United States.

(c) Presentations

"Security Defense of Transportation Networks against Cyberattacks: A Physics-Informed AI Approach." Center for Connected and Automated Transportation Working Group Meeting (virtual). September 8th, 2025.

"Security Defense of Transportation Networks against Cyberattacks: A Physics-Informed AI Approach." Center for Connected and Automated Transportation Working Group Meeting (virtual). November 11th, 2024.

6.1.2. Other Outputs

This project outlines a structured modeling approach to enhancing CAV cybersecurity through:

- A physics-guided cyberattack detection algorithm tailored to lane-changing behavior in CAVs. The proposed method enables trajectory-based monitoring and supports timely identification of falsified lane-changing decisions under cyberattacks.
- A hierarchical adversarial reinforcement learning framework for robust discretionary lane-changing under adversarial observation perturbations. The framework integrates interaction-aware execution with robust maneuver decision-making to mitigate traffic-level disturbances under corrupted sensor observation.

6.2. Outcomes

The outcomes of this project contribute to the advancement of modeling methodologies for CAV cybersecurity, particularly in the context of tactical maneuvering. By integrating physics-informed anomaly detection and robust decision-making, the study provides a structured framework for analyzing cyber-induced behavioral deviations and evaluating resilience under adversarial conditions. The proposed detection and robust control frameworks offer technical insights that may support future development of cybersecurity-aware vehicle monitoring strategies and control systems. The findings may also inform researchers and practitioners working at the intersection of traffic flow theory, machine learning, and cyber-physical system security. In addition, this project has contributed to graduate-level research training in CAV cybersecurity, adversarial learning, and physics-informed modeling, supporting workforce development in emerging transportation technologies.

6.3. Impacts

The impacts of this project are primarily methodological and knowledge-driven. The proposed frameworks advance the understanding of how cyberattacks may influence tactical vehicle maneuvers and how such impacts can be monitored and mitigated through physics-informed and interaction-aware learning models.

From a transportation systems perspective, the study highlights the importance of integrating physical traffic interaction principles into cybersecurity analysis. This perspective may encourage future research that bridges traffic flow theory and cyber-physical system security, contributing to more resilient CAV deployment strategies.

For industry and transportation-related agencies, the modeling insights developed in this study may support future exploration of anomaly monitoring systems and robust maneuver control policies within CAV ecosystems. Although the frameworks were evaluated in controlled settings, they provide a conceptual basis for incorporating cybersecurity considerations into tactical vehicle decision-making.

Finally, the project contributes to academic capability building by expanding research efforts in CAV cybersecurity and training researcher in interdisciplinary methods that combine traffic flow theory and machine learning.

REFERENCES

Ahmed, K.I., 1999. Modeling drivers' acceleration and lane changing behavior. Ph.D. thesis. Massachusetts Institute of Technology.

Ahn, S., Cassidy, M.J., 2007. Freeway traffic oscillations and vehicle lane-change maneuvers. Transportation and Traffic Theory 1, 691–710.

Alizadeh, A., Moghadam, M., Bicer, Y., Ure, N. K., Yavas, U., & Kurtulus, C. 2019, October. Automated lane change decision making using deep reinforcement learning in dynamic and uncertain highway environment. In 2019 IEEE intelligent transportation systems conference (ITSC) (pp. 1399-1404). IEEE.

Althoff, M., Mergel, A., 2011. Comparison of markov chain abstraction and monte carlo simulation for the safety assessment of autonomous cars. IEEE Transactions on Intelligent Transportation Systems 12, 1237–1247.

Amoozadeh, M., Raghuramu, A., Chuah, C.N., Ghosal, D., Zhang, H.M., Rowe, J., Levitt, K., 2015. Security vulnerabilities of connected vehicle streams and their impact on cooperative driving. IEEE Communications Magazine 53, 126–132.

Aradi, S. 2020. Survey of deep reinforcement learning for motion planning of autonomous vehicles. IEEE Transactions on Intelligent Transportation Systems, 23(2), 740-759.

Bagwe, G., Yuan, X., Chen, X., & Zhang, L. 2023, May. RAMRL: Towards robust on-ramp merging via augmented multimodal reinforcement learning. In 2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST) (pp. 23-33). IEEE.

Balal, E., Cheu, R.L., Sarkodie-Gyan, T., 2016. A binary decision model for discretionary lane changing move based on fuzzy inference system. Transportation Research Part C: Emerging Technologies 67, 47–61.

Biroon, R.A., Biron, Z.A., Pisu, P., 2022. False Data Injection Attack in a Platoon of CACC: Real-Time Detection and Isolation With a PDE Approach. IEEE Trans. Intell. Transport. Syst. 23, 8692–8703. https://doi.org/10.1109/TITS.2021.3085196.

Cassidy, M.J., Rudjanakanoknad, J., 2005. Increasing the capacity of an isolated merge by metering its on-ramp. Transportation Research Part B: Methodological 39, 896–913.

Doshi, A., Trivedi, M.M., 2009. On the roles of eye gaze and head dynamics in predicting driver's intent to change lanes. IEEE Transactions on Intelligent Transportation Systems 10, 453–462.

Fan, P., Guo, J., Wang, Y., Wijnands, J.S., 2022. A hybrid deep learning approach for driver anomalous lane changing identification. Accident Analysis & Prevention 171, 106661.

Gipps, P.G., 1986. A model for the structure of lane-changing decisions. Transportation Research Part B: Methodological 20, 403–414.

Gleave, A., Dennis, M., Wild, C., Kant, N., Levine, S., & Russell, S. 2019. Adversarial policies: Attacking deep reinforcement learning. arXiv preprint arXiv:1905.10615.

Gloudemans, D., Wang, Y., Ji, J., Zachar, G., Barbour, W., Hall, E., Cebelak, M., Smith, L., Work, D.B., 2023. I-24 motion: An instrument for freeway traffic science. Transportation Research Part C: Emerging Technologies 155, 104311.

He, X., Huang, W., & Lv, C. 2024. Trustworthy autonomous driving via defense-aware robust reinforcement learning against worst-case observational perturbations. Transportation Research Part C: Emerging Technologies, 163, 104632.

He, X., Lou, B., Yang, H., & Lv, C. 2022. Robust decision making for autonomous vehicles at highway on-ramps: A constrained adversarial reinforcement learning approach. IEEE Transactions on Intelligent Transportation Systems, 24(4), 4103-4113.

He, X., Yang, H., Hu, Z., & Lv, C. 2022. Robust lane change decision making for autonomous vehicles: An observation adversarial reinforcement learning approach. IEEE Transactions on Intelligent Vehicles, 8(1), 184-193.

Hirschberg, J., Manning, C.D., 2015. Advances in natural language processing. Science 349, 261–266. https://doi.org/10.1126/science.aaa8685.

Hochreiter, S., 1997. Long short-term memory. Neural Computation MIT-Press.

Hoel, C. J., Wolff, K., & Laine, L. 2018, November. Automated speed and lane change decision making using deep reinforcement learning. In 2018 21st international conference on intelligent transportation systems (ITSC) (pp. 2148-2155). IEEE.

Howard, R. A. 1960. Dynamic programming and markov processes.

Huang, L., Guo, H., Zhang, R., Wang, H., Wu, J., 2018. Capturing drivers' lane changing behaviors on operational level by data driven methods. IEEE access 6, 57497–57506.

Huang, S.E., Feng, Y., Liu, H.X., 2021. A data-driven method for falsified vehicle trajectory identification by anomaly detection. Transportation research part C: emerging technologies 128, 103196.

Kesting, A., Treiber, M., Helbing, D., 2007. General lane-changing model mobil for car-following models. Transportation Research Record 1999, 86–94.

Kingma, D.P., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kiran, B. R., Sobh, I., Talpaert, V., Mannion, P., Al Sallab, A. A., Yogamani, S., & Pérez, P. 2021. Deep reinforcement learning for autonomous driving: A survey. IEEE transactions on intelligent transportation systems, 23(6), 4909-4926.

Kumar, M., Husain, D.M., Upreti, N., Gupta, D., 2010. Genetic algorithm: Review and application. Available at SSRN 3529843 .

Laval, J.A., Daganzo, C.F., 2006. Lane-changing in traffic streams. Transportation Research Part B: Methodological 40, 251–264.

Laval, J.A., Leclercq, L., 2008. Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. Transportation Research Part B: Methodological 42, 511–522.

Li, J., Yu, C., Shen, Z., Su, Z., Ma, W., 2023. A survey on urban traffic control under mixed traffic

environment with connected automated vehicles. Transportation Research Part C: Emerging Technologies 154, 104258. https://doi.org/10.1016/j.trc.2023.104258.

Liu, Y., Wang, X., Li, L., Cheng, S., Chen, Z., 2019. A novel lane change decision-making model of autonomous vehicle based on support vector machine. IEEE access 7, 26543–26550.

Loke, S.W., 2019. Cooperative Automated Vehicles: A Review of Opportunities and Challenges in Socially Intelligent Vehicles Beyond Networking. IEEE Transactions on Intelligent Vehicles 4, 509–518. https://doi.org/10.1109/TIV.2019.2938107.

Mahajan, V., Katrakazas, C., Antoniou, C., 2020. Prediction of lane-changing maneuvers with automatic labeling and deep learning. Transportation research record 2674, 336–347.

Mauch, M., J. Cassidy, M., 2002. Freeway traffic oscillations: observations and predictions, in: Transportation and Traffic Theory in the 21st Century: Proceedings of the 15th International Symposium on Transportation and Traffic Theory, Adelaide, Australia, 16-18 July 2002, Emerald Group Publishing Limited. pp. 653–673.

Milan´es, V., Villagra, J., Godoy, J., Sim´o, J., P´erez, J., Onieva, E., 2012. An intelligent v2i-based traffic management system. IEEE Transactions on Intelligent Transportation Systems 13, 49–58.

Mirchevska, B., Pek, C., Werling, M., Althoff, M., & Boedecker, J. 2018, November. High-level decision making for safe and reasonable autonomous lane changing using reinforcement learning. In 2018 21st international conference on intelligent transportation systems (ITSC) (pp. 2156-2162). IEEE.

Mo, Z., Shi, R., Di, X., 2021. A physics-informed deep learning paradigm for car-following models. Transportation research part C: emerging technologies 130, 103240.

NCS, N.S.C., 2025. Motor-vehicle deaths, injuries, and number of crashes by type of crash, 2023. https://injuryfacts.nsc.org/motor-vehicle/overview/type-of-crash/. Accessed: 2025-07-24.

Otter, D.W., Medina, J.R., Kalita, J.K., 2021. A Survey of the Usages of Deep Learning for Natural Language Processing. IEEE Transactions on Neural Networks and Learning Systems 32, 604–624. https://doi.org/10.1109/TNNLS.2020.2979670.

Pinto, L., Davidson, J., Sukthankar, R., & Gupta, A. 2017, July. Robust adversarial reinforcement learning. In International conference on machine learning (pp. 2817-2826). PMLR.

Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics 378, 686–707.

Ramyar, S., Homaifar, A., Karimoddini, A., Tunstel, E., 2016. Identification of anomalies in lane change behavior using one-class svm, in: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE. pp. 004405–004410.

Roy, T., Tariq, A., Dey, S., 2022. A Socio-Technical Approach for Resilient Connected Transportation Systems in Smart Cities. IEEE Transactions on Intelligent Transportation Systems 23, 5019–5028. https://doi.org/10.1109/TITS.2020.3045854.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. 2017. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.

Sedjelmaci, H., Hadji, M., Ansari, N., 2019. Cyber Security Game for Intelligent Transportation Systems. IEEE Network 33, 216–222. https://doi.org/10.1109/MNET.2018.1800279.

Shladover, S.E., Su, D., Lu, X.Y., 2012. Impacts of cooperative adaptive cruise control on freeway traffic flow. Transportation Research Record 2324, 63–70.

Shoukry, Y., Mishra, S., Luo, Z., Diggavi, S., 2018. Sybil attack resilient traffic networks: A physics-based trust propagation approach, in: 2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS), IEEE. pp. 43–54.

Sun, R., Luo, Q., Chen, Y., 2023. Online transportation network cyber-attack detection based on stationary sensor data. Transportation Research Part C: Emerging Technologies 149, 104058. https://doi.org/10.1016/j.trc.2023.104058.

Sun, W., Wang, S., Shao, Y., Sun, Z., Levin, M.W., 2022. Energy and mobility impacts of connected autonomous vehicles with co-optimization of speed and powertrain on mixed vehicle platoons. Transportation Research Part C: Emerging Technologies 142, 103764. https://doi.org/10.1016/j.trc.2022.103764.

Talebpour, A., Mahmassani, H.S., Hamdar, S.H., 2015. Modeling lane-changing behavior in a connected environment: A game theory approach. Transportation Research Procedia 7, 420–440.

Vyas, S.D., Padisala, S.K., Dey, S., 2023. A physics-informed neural network approach towards cyber attack detection in vehicle platoons, in: 2023 American Control Conference (ACC), IEEE. pp. 4537–4542.

Wang, F., Wang, X., Ban, X.J., 2024. Data poisoning attacks in intelligent transportation systems: A survey. Transportation Research Part C: Emerging Technologies 165, 104750.

Wang, Q., Li, Z., Li, L., 2014. Investigation of discretionary lane-change characteristics using next-generation simulation data sets. Journal of Intelligent Transportation Systems 18, 246–253.

Wang, R., Yu, R., 2025. Physics-guided deep learning for dynamical systems: A survey. ACM Computing Surveys 58, 1–31.

Wang, S., Levin, M.W., Caverly, R.J., 2021. Optimal parking management of connected autonomous vehicles: A control-theoretic approach. Transportation Research Part C: Emerging Technologies 124, 102924. https://doi.org/10.1016/j.trc.2020.102924.

Wang, S., Levin, M.W., Stern, R., 2023a. Optimal feedback control law for automated vehicles in the presence of cyberattacks: A min–max approach. Transportation Research Part C: Emerging Technologies 153, 104204. https://doi.org/10.1016/j.trc.2023.104204.

Wang, S., Shang, M., Levin, M.W., Stern, R., 2023b. A general approach to smoothing nonlinear mixed traffic via control of autonomous vehicles. Transportation Research Part C: Emerging Technologies 146, 103967. https://doi.org/10.1016/j.trc.2022.103967.

Wang, S., Stern, R., Levin, M.W., 2022. Optimal Control of Autonomous Vehicles for Traffic Smoothing. IEEE Transactions on Intelligent Transportation Systems 23, 3842–3852. https://doi.org/10.1109/TITS.2021.3094552.

Wang, T., Ma, M., Liang, S., Yang, J., & Wang, Y. 2025. Robust lane change decision for

autonomous vehicles in mixed traffic: A safety-aware multi-agent adversarial reinforcement learning approach. Transportation Research Part C: Emerging Technologies, 172, 105005.

Wang, W., Qie, T., Yang, C., Liu, W., Xiang, C., Huang, K., 2021. An intelligent lane-changing behavior prediction and decision-making strategy for an autonomous vehicle. IEEE transactions on industrial electronics 69, 2927–2937.

Wang, Y., Sarkar, E., Li, W., Maniatakos, M., Jabari, S.E., 2021. Stop-and-Go: Exploring Backdoor Attacks on Deep Reinforcement Learning-Based Traffic Congestion Control Systems. IEEE Transactions on Information Forensics and Security 16, 4772–4787. https://doi.org/10.1109/TIFS.2021.3114024.

Wissing, C., Nattermann, T., Glander, K.H., Hass, C., Bertram, T., 2017. Lane change prediction by combining movement and situation based probabilities. IFAC-PapersOnLine 50, 3554–3559.

Xu, D., Liu, M., Yao, X., Lyu, N., 2023. Integrating surrounding vehicle information for vehicle trajectory representation and abnormal lane-change behavior detection. Sensors 23, 9800.

Yang, D., Lyu, M., Dai, L., Wang, X., Guo, Q., 2022. Decision-making model for lane selection of automated vehicles in connected vehicle environment. China Journal of Highway and Transport 35, 243–55.

Ye, L., Yamamoto, T., 2019. Evaluating the impact of connected and autonomous vehicles on traffic safety. Physica A: Statistical Mechanics and its Applications 526, 121009.

Yoo, J., & Langari, R. 2020. A game-theoretic model of human driving and application to discretionary lane-changes. arXiv preprint arXiv:2003.09783.

Yu, H., Tseng, H.E., Langari, R., 2018. A human-like game theory-based controller for automatic lane changing. Transportation Research Part C: Emerging Technologies 88, 140–158.

Yu, M., Long, J., 2022. An Eco-Driving Strategy for Partially Connected Automated Vehicles at a Signalized Intersection. IEEE Transactions on Intelligent Transportation Systems 23, 15780–15793. https://doi.org/10.1109/TITS.2022.3145453.

Yu, R., Wang, R., 2024. Learning dynamical systems from data: An introduction to physics-guided deep learning. Proceedings of the National Academy of Sciences 121, e2311808121.

Zhang, Y., Xu, Q., Wang, J., Wu, K., Zheng, Z., Lu, K., 2022. A learning-based discretionary lane-change decision-making model with driving style awareness. IEEE transactions on intelligent transportation systems 24, 68–78.

Zhang, Z., Liu, W., Zhang, F., 2023. On the joint network equilibrium of parking and travel choices under mixed traffic of shared and private autonomous vehicles. Transportation Research Part C: Emerging Technologies 153, 104226. https://doi.org/10.1016/j.trc.2023.104226.

Zheng, Y., Wang, J., Li, K., 2020. Smoothing Traffic Flow via Control of Autonomous Vehicles. IEEE Internet of Things Journal 7, 3882–3896. https://doi.org/10.1109/JIOT.2020.2966506.

Zheng, Z., 2014. Recent developments and research needs in modeling lane changing. Transportation research part B: methodological 60, 16–32.

Zheng, Z., Ahn, S., Monsere, C.M., 2010. Impact of traffic oscillations on freeway crash occurrences. Accident Analysis & Prevention 42, 626–636.

Zhong, X., Kontar, W., Sheng, Z., Kim, Y., Chen, S., Li, X., Ahn, S., 2024. Understanding physics and ai synergy in car-following models, in: Conference in Emerging Technologies in Transportation Systems (TRC-30), Transportation Research: Part C Symposium Committee.

Zhou, W., Chen, D., Yan, J., Li, Z., Yin, H., & Ge, W. 2022. Multi-agent reinforcement learning for cooperative lane changing of connected and autonomous vehicles in mixed traffic. Autonomous Intelligent Systems, 2(1), 5.

APPENDIX

A. LSTM Update Equations

We summarize the standard LSTM update equations used in this study. At each time step t, the LSTM cell updates its memory using the current input $\mathbf{s}(t)$, the previous hidden state $h(t-1)$, and the previous cell state $C(t-1)$. The gating mechanism consists of forget, input, and output gates. Weight matrices and bias terms are denoted by W and β, respectively, and $\sigma(\cdot)$ and $\tanh(\cdot)$ denote the sigmoid and hyperbolic tangent functions.
Forget gate

$$f(t) = \sigma\big(W_f \cdot [h(t-1), \boldsymbol{s}(t)] + \beta_f\big)$$

Input gate and candidate cell state

$$i(t) = \sigma(W_i \cdot [h(t-1), \boldsymbol{s}(t)] + \beta_i)$$
$$\tilde{C}(t) = tanh(W_C \cdot [h(t-1), \boldsymbol{s}(t)] + \beta_C)$$

Cell state update

$$C(t) = f(t) \odot C(t-1) + i(t) \odot \tilde{C}(t)$$

Output gate and hidden state

$$o(t) = \sigma(W_o \cdot [h(t-1), \boldsymbol{s}(t)] + \beta_o)$$
$$h(t) = o(t) \odot tanh\big(C(t)\big)$$

where $\odot$ denotes element-wise multiplication.

Finally, the hidden state $h(t)$ is passed through a fully connected layer and a sigmoid activation to produce the LC probability output $\hat{p}(t + \Delta t)$.

B. Ahmed's Model

Ahmed (Ahmed, 1999) categorized LC maneuvers into three types: mandatory lane-changing (MLC), discretionary lane-changing (DLC), and forced merging (FM). In the DLC model, when a driver is dissatisfied with driving conditions in the current lane, the adjacent lanes are evaluated and compared to select a target lane. According to Ahmed's LC decision model, the probability that driver i performs MLC, DLC, or FM at time t is given by:

$$P_i(LC|v_i)(t) = \frac{1}{1 + exp\big(-X_i^{LC}(t)\beta^{LC} - \alpha^{LC}v_i\big)}, \quad LC \in \{MLC, DLC, FM\}$$

where, $P_i(LC|v_i)(t)$ denotes the probability that driver i executes MLC, DLC, or FM at time t; $X_i^{LC}(t)$ is the vector of explanatory variables affecting the LC decision; β^{LC} is the corresponding

parameter vector; v_i is the driver-specific random term; and α^{LC} is the coefficient of v_i.

After the LC decision-making process, the gap acceptance model captures whether the available gaps are accepted. A lane change occurs if both the target lead and lag gaps exceed their critical values. The critical lead and lag gaps are presented as:

$$G_i^{cr,g}(t) = exp\left(X_i^g(t)\beta^g + \alpha^g v_i + \varepsilon_i^g(t)\right), g \in \{lead,lag\}$$

where, $G_i^{cr,g}(t)$ denotes the critical lead and lag gaps for driver i at time t; $X_i^g(t)$ is a vector of explanatory variables affecting the critical gap g; β^g is the corresponding parameter vector; v_i is the driver-specific random term; α^g is the coefficient of v_i; and $\varepsilon_i^g(t) \sim N(0, \sigma_{\varepsilon_g}^2)$ is a normally distributed random term.

The probability that driver i accepts the available gap at time t is given as:

$$\begin{aligned} P_i(gap\ acceptance|v_i) &= P_i(lead\ gap\ acceptance|v_i) \times P_i(lag\ gap\ acceptance|v_i) \\ &= P_i\left(G_i^{lead}(t) > G_i^{cr,lead}(t)\middle|v_i\right) \times P_i\left(G_i^{lag}(t) > G_i^{cr,lag}(t)\middle|v_i\right) \end{aligned}$$

where $G_i^{lead}(t)$ and $G_i^{lag}(t)$ are the available lead and lag gaps in the target lane, respectively.

C. MOBIL Model

The MOBIL model defines two criteria for determining LC. The first is the safety criterion, formulated as:

$$\tilde{a}_{TF} \geq -a_{safe}$$

where, $\tilde{a}_{TF}$ denotes the acceleration of the follower in the target lane; and $-a_{safe}$ is a given safe deceleration limit. The second is the incentive criterion as follows.

$$\tilde{a} - a + p(\tilde{a}_{TF} - a_{TF} + \tilde{a}_{CF} - a_{CF}) > \Delta a_{th}$$

where, a and $\tilde{a}$ refer to the acceleration of the subject vehicle before and after the lane change, respectively; p is the politeness factor; and Δa_{th} is a predefined threshold. When implementing the MOBIL LC model, $\mathrm{t_{start}}$ and $\mathrm{t_{end}}$ denote the start and end time of LC maneuver, used to determine $a, \tilde{a}, a_{TF}, \tilde{a}_{TF}, a_{CF}$, and $\tilde{a}_{CF}$.

The politeness factor p is determined based on the number of lane changes made by the vehicle during its journey:

$$\text{Politeness factor} = \begin{cases} 0.5, & number\ of\ lane-changing = 0 \\ 0.25, & number\ of\ lane-changing = 1 \\ 0.0, & number\ of\ lane-changing = 2 \end{cases}$$

D. Talebpour's Model

Talebpour (Talebpour elt a., 2015) classified LC maneuvers into two types: mandatory LC and

discretionary LC. Under inactive V2V communications, where drivers' perceptions of surrounding traffic conditions are subjective, LC behavior is modeled as a two-person, non-zero-sum, non-cooperative game under incomplete information. The subject vehicle has two pure strategies (change lane and stay in lane), while the target-lane follower has three pure strategies (accelerate, decelerate, and change lane). Table 11 summarizes the corresponding payoff matrix for the subject (P) and the target-lane follower (R).

Table 11: Payoff matrix of discretionary lane-changing game with inactive V2V communication

		Subject vehicle	
	Action	Change lane	Stay in lane
Target-lane follower	Accelerate	$P_{11} = \eta_1 a^{TL} + \eta_2 \Delta v + \varepsilon_{11}$ $R_{11} = \eta_3 a^{TF} + \delta_{11}$	$P_{12} = \varepsilon_{12}$ $R_{12} = \eta_3 a^{TFL} + \delta_{12}$
	Decelerate	$P_{21} = \eta_1 a^{TL} + \eta_2 \Delta v + \varepsilon_{21}$ $R_{21} = \eta_4 a_0^{TF} + \delta_{21}$	$P_{22} = \varepsilon_{22}$ $R_{22} = \eta_4 a_0^{TFL} + \delta_{22}$
	Change lane	$P_{31} = \eta_2 \Delta v + \varepsilon_{31}$ $R_{31} = \eta_1 a^{TF} + \eta_2 \Delta v + \delta_{31}$	$P_{32} = \varepsilon_{32}$ $R_{32} = \eta_1 a^{TF} + \eta_2 \Delta v + \delta_{32}$

When the target-lane follower chooses to change lane, a^{TF} is recalculated with respect to its new leader. In this model, the parameters are assumed to be homogeneous across all drivers.