Vulnerability Assessment and Detection of Stealthy Sequential Cyberattacks in Hybrid Tracked Vehicles

Grace Muriithi, *Graduate Student Member, IEEE*, Behnaz Papari, *Senior Member, IEEE*, Ali Moghassemi, *Graduate Student Member, IEEE*, Anirudh Sundar, Ali Arsalan, *Graduate Student Member, IEEE*, Elutunji Buraimoh, Laxman Timilsina, *Member, IEEE*, Gokhan Ozkan, *Senior Member, IEEE*, and Christopher Edrington, *Senior Member, IEEE*

***Abstract*— This article presents an innovative approach to improve hybrid powertrains' cyber–physical security for high-speed, tracked, off-road vehicles against subtle and sophisticated cyberattacks. While cyber threats against hybrid tracked vehicles (HTVs) have been acknowledged, exploring advanced cyberattacks targeting battery lifetime and energy efficiency within the powertrain remains underexplored. This article proposes a dual-pronged intelligent methodology to construct novel FDIAs targeting vehicle control decisions from the energy management system (EMS). The optimal cyberattack strategy by a minimally informed adversary is formulated as a partial observable Markov decision process (POMDP), employing deep reinforcement learning (DRL) for online learning and attacking. Simultaneously, a comprehensive reward system is devised that incorporates sniffing features to elevate the stealthiness and efficacy of the cyberattacks. This augmentation ensures that the orchestrated attack vectors remain inconspicuous to human drivers. Evaluation metrics are formulated to assess the impact and stealth characteristics of the cyberattack. Furthermore, a sliding window-based controller data monitoring scheme combining iForest and dynamic time warping (DTW) algorithms is proposed to automatically detect the stealthy FDIAs in real-time, ensuring secure control of energy-efficient powertrain systems. Through the proposed metrics and detection module, the article thoroughly examines the impact of cyberattacks on the energy consumption of the HTV. It also provides vital insights for defending vehicles that operate in austere environments against sophisticated controller attacks.**

***Index Terms*— Cyberattack detection, deep reinforcement learning (DRL), hybrid tracked vehicles (HTVs), stealthy sequential cyberattacks, vehicle cyber–physical security, vulnerability assessment.**

Received 25 April 2024; revised 25 June 2024, 3 September 2024, and 29 October 2024; accepted 21 November 2024. Date of publication 2 December 2024; date of current version 26 March 2025. This work was supported in part by Clemson University's Virtual Prototyping of Autonomy Enabled Ground Systems (VIPR-GS) through U.S. Army DEVCOM Ground Vehicle Systems Center (GVSC) under Cooperative Agreement W56HZV-21-2-0001 and in part by Warren H. Owen Distinguished Professorship Endowment. *(Corresponding author: Grace Muriithi.)*

Grace Muriithi, Anirudh Sundar, and Ali Arsalan are with the Automotive Engineering, Clemson University, Greenville, SC 29607 USA (e-mail: gmuriit@clemson.edu; sundar@g.clemson.edu; aarsala@clemson.edu).

Behnaz Papari is with the Automotive Engineering and Electrical and Computer Engineering, Clemson University, Greenville, SC 29607 USA (e-mail: bpapari@clemson.edu).

Ali Moghassemi, Elutunji Buraimoh, and Christopher Edrington are with the Electrical and Computer Engineering, Clemson University, Clemson, SC 29631 USA (e-mail: amoghas@clemson.edu; eburaim@clemson.edu; cedring@clemson.edu).

Laxman Timilsina is with the Electrical and Computer Engineering, Clemson University, Clemson, SC 29631 USA (e-mail: ltimils@clemson.edu).

Gokhan Ozkan is with the Electrical and Computer Engineering, Clemson University, Charleston, SC 029405 USA (e-mail: gokhano@clemson.edu).

Digital Object Identifier 10.1109/TTE.2024.3510458

NOMENCLATURE

APTs	Advanced persistent threats.
CAN	Connected area network.
CPS	Cyber–physical system.
DoS	Denial of service attack.
DQN	Deep Q network.
DRL	Deep reinforcement learning.
DTW	Dynamic time warping.
EMS	Energy management system.
EVCSs	Electric vehicle charging stations.
FDIA	False data injection attack.
HTV	Hybrid tracked vehicle.
HEV	Hybrid electric vehicle.
HVB	High voltage battery.
ICE	Internal combustion engine.
ICTs	Information and communication technologies.
iForest	Isolation forest.
LSTM	Long short-term memory.
POMDP	Partially observable Markov decision process.
SoC	State of charge.

I. INTRODUCTION

IN THE recent past, the U.S. Army has made a significant commitment to transitioning its ground vehicle fleet toward electric and hybrid electric drive technologies [1]. These modern HEVs represent a fusion of cutting-edge technologies, combining autonomy, electrification, and sophisticated cyber–physical interactions. These systems seamlessly integrate ICT, intelligent control policies, and intricate physical components of the vehicle [2]. While this move promises substantial economic and environmental benefits through military vehicle modernization, it also ushers in a new era of cyber–physical security challenges [3]. Understanding the vulnerabilities within the electrified powertrain, beyond routine operations and disturbances, has now become an imperative task in the realm of modern tracked vehicles.

Security research has examined potential cyberattack strategies and countermeasures, primarily in civilian passenger vehicles and other CPSs [4]. These cyberattacks exploit various facets of hybrid powertrain vulnerabilities. FDIAs [5] involve injecting undetectable malicious data into sensor measurements and misleading vehicle operations. Conversely, replay attacks entail capturing legitimate system data and then

retransmitting it later to deceive the system into accepting the retransmission as genuine. DoS attacks, which disrupt component availability, involve overwhelming network nodes or resources with data, preventing legitimate requests from being processed, which can severely degrade the vehicle's performance. Cascaded cyberattacks [6] assess the risk of cascading failures triggered by targeting a small set of components. Time synchronization attacks [7] focus on critical temporal information within measurements, actuators, and devices. These studies of cyberattack schemes can shed light on security concerns that typically lie beyond the scope of normal operational dynamics, random faults, or minor/major disturbances [8].

When considering cyber-security for modern HEVs, it's essential to recognize that cyberattacks are not limited to a single target. It should be assumed that attackers can compromise multiple measurements or manipulate control commands from the cyber realm [9]. Therefore, forensic and security studies must comprehend the vulnerability of HEVs under the potential threat of intelligent and well-informed cyber-actors. To this day, it remains a formidable challenge to comprehensively assess the vulnerabilities of modern vehicles in the face of multiple coordinated cyberattacks. Cybersecurity research has received increasing attention to address these challenges, and much effort has been devoted to vulnerability assessment, cyberattack detection, and threat-resilient control.

Cyber–physical attacks have been extensively investigated in other CPSs like microgrids, smart grids, and EVCSs [8], [10], [11], [12], [13]. Cyber–physical security monitoring can serve as a second line of protection because an abnormal system measurement indicates potential cyberattacks. Interestingly, exploring this area in the context of off-road tracked vehicles has been notably limited. In the realm of passenger electric vehicles (EVs), a few researchers [3], [9], [14], [15] have attempted to tackle cyber–physical security issues. Guo et al. [9] developed a systematic framework for evaluating the cyber–physical security of connected EVs. They proposed metrics to evaluate the impact of cyberattacks on vehicle performance and comfort. Their analysis of various cyberattack scenarios on electronic control units provides guidelines for security measures. Guo et al. [3] expanded on this by analyzing vulnerabilities in EVs with four-wheel drive, where cyber threats targeting the motors can compromise stability control and handling. They proposed detection methods combining state estimation and performance metrics. Ye et al. [14] provided an overview of cyber–physical security challenges for connected EVs. They discussed threats to firmware, charging, and powertrain control and simulated cyberattacks impacting individual components and whole vehicles. They proposed an architecture for next-generation power electronics to address these issues. Yang et al. [5] proposed a "systematic and generalized methodology" to assess how vulnerable electric drive systems are to cyberattacks. They found that cyberattacks can increase torque ripple by up to 300% and total harmonic distortion by over 70%.

While the field of vehicular cybersecurity has been steadily gaining the attention of researchers, it is strikingly apparent that only a limited number of studies have explored the intricate realm of stealthy and sequential FDIAs as applied to (H)EVs. Stealthy sequential FDIAs represent a sophisticated and insidious threat to CPSs, including HEVs and smart grid infrastructures. The essence of these attacks lies in their ability to gradually introduce carefully crafted false data into the system's states over an extended period, manipulating specific system parameters while evading detection. Rather than executing straightforward, concurrent cyberattacks, adversaries increasingly resort to sequential strategies that enable them to target vital signals, thus amplifying the potential for damage. Furthermore, there remains a conspicuous dearth of literature addressing scenarios wherein adversaries harness the power of artificial intelligence (AI) to enhance the sophistication of the cyberattack when targeting HTVs, as seen in Table I. This signifies a glaring gap in our understanding of the evolving threat landscape in the context of HEV/HTV cybersecurity [16]. While attacker knowledge is inherently limited, APT hackers possess capabilities that make training RL models for stealthy attacks feasible. APT attackers conduct extensive reconnaissance using techniques like network scanning and phishing to gather data on target networks, controls, and normal data patterns. This data facilitates training the initial attack models.

Furthermore, these attackers can also create high-fidelity simulated environments replicating the target vehicle setup using compromised inside knowledge to generate realistic training data without risking detection. Once deployed, the RL malware provides feedback, allowing the attack model to adapt and refine its approach based on observed responses continuously. Consequently, developing untraceable attacking strategies that can hide in normal system traffic and potentially produce a high long-term impact [24]. Some studies proposed using Q-learning in the smart grid domain to identify critical cyberattack sequences that could cause cascading failures. Yan et al. [8] pioneered this approach, using Q-learning and a power flow model to find three-step attack sequences that could damage IEEE test systems. Chen et al. [10] modeled FDIAs as a POMDP and used Q-learning to enable "online learning and attacking," evaluating their method on an IEEE 39-bus system for voltage violation. Abianeh et al. [20] proposed using multiagent DRL to identify vulnerabilities in existing FDI detection methods and generate stealthy FDIAs, then used another DRL approach to address these vulnerabilities.

Multiple detection algorithms have been proposed to address destabilizing cyberattacks, broadly categorized into 1) anomaly-based and 2) signature-based. These two categories are complementary, as the latter is very effective in detecting known attacks, while anomaly-based approaches can detect "zero-day" attacks, i.e., malicious events that are neither observed nor reported yet. These methods are further categorized into model-dependent and model-independent methods. Model-based detection involves analyzing residuals between predicted system outputs and measures, often derived from physical relationships or autoregressive models. The main work of physics-based detection [25] is to obtain the residual between the predicted system outputs and the measures, which is considered a proxy for the presence of attacks [26], [27]. The prediction model can be developed from physical relationships,

TABLE I
COMPARISON BETWEEN PROPOSED WORK AND EXISTING STUDIES

Study	Investigated System	Vulnerability Assessment Approach	Attack Design (AI)	Impact Analysis / Detection Method
[3]	Electric vehicles	Predefined attack signals (constant values)	×	Physics-guided ML
[4]	Electric vehicles	Predefined attack signals (constant values)	×	Machine Learning (ML)
[5]	Electric drive system	Predefined attack signals	×	Impact analysis - No detection
[8]	Smart grid	Q learning	✓	Impact analysis - No detection
[9]	Electric vehicles	Predefined attack signals	×	Impact analysis - No detection
[10]	Smart power grid	Q learning	✓	Kernel density estimation
[11]	Smart grid	DRL-based attacks	✓	Impact analysis - No detection
[14]	Electric vehicles	Predefined attack signals	×	Impact analysis - No detection
[17]	Power grid	Optimization-based attack	✓	Impact analysis - No detection
[18]	Power grid	Principle component analysis	✓	Impact analysis - No detection
[19]	Power grid	Heuristic graph method	✓	Impact analysis - No detection
[20]	Islanded DC microgrids	DRL-based attacks	✓	DQN
[21]	Hybrid electric vehicles	Predefined attack signals	×	Probability distribution
[22]	Islanded DC microgrids	DRL-based attacks	✓	Neural network
[23]	EVCS	Predefined attack signals	×	Impact analysis - No detection
This Paper	**HTVs**	**Learning-based attacks based on DRL**	✓	**Impact analysis with defense framework**

such as Newton's laws, fluid dynamics, electromagnetic laws, or an autoregressive model [28]. Researchers have incorporated adaptive control concepts [29] to counter FDIAs and DoS attacks; however, the complexity and precision of such algorithms are highly dependent on the order size of the underlying model. These model-based methods are also prone to instability under system parameter uncertainties or the presence of multiple coordinated cyberattacks.

Data-driven approaches, gaining attention due to computing advancements, do not require system parameters or models [4], [30], [31], [32]. Detection algorithms based on supervised learning (signature-based) have been investigated [33]. Still, their performance is significantly impacted by the quality of the collected labeled dataset, and they are also prone to the overfitting phenomenon. Researchers have recently employed deep learning schemes based on the system's physical observations to mitigate these challenges [3]. Guo et al. [3] and Kandasamy et al. [34] both developed machine learning classifiers that leverage physics-guided data features to detect cyberattacks in EVs, achieving high accuracy across various driving scenarios. Zhang et al. [35] proposed a federated graph neural network (GNN) for fast anomaly detection in the CAN bus, capable of identifying attacks in as little as three milliseconds by leveraging directed attributed graphs from CAN message streams. While their approach is highly effective for real-time detection of immediate and discrete CAN bus attacks, our work targets more complex, time-dependent stealth attacks on vehicle controllers. This requires a different methodology to account for the subtle and evolving nature of these threats.

Yang et al. [36] and Darby and Gottumukkala [37] both proposed innovative detection methods for cyber threats in vehicles, with [36] presenting a fast, model-free approach using trustworthy sensor signals to detect cyberattacks in EV traction motor drives and [37] proposing a decentralized security exchange which efficiently supports cyber–physical security for autonomous vehicles. In the broader CPS domain, authors in [10], [20], and [22] introduced an RL framework to emulate an intelligent attacker to explore the susceptibilities within microgrids, then used to neural network detector, kernel density, and DQN as the proposed detectors, respectively. Although these methods have shown effectiveness in their respective domains, they may not be directly applicable or optimal for detecting and mitigating stealthy FDIAs in HTVs with high levels of uncertainty and wide variations in operating conditions. The approaches proposed in these references are designed for a single variable and detect already-known threatening situations. Once a new tampering attack occurs, the detection accuracy is reduced, and good adaptability is not guaranteed. Furthermore, most of these methods are based on supervised learning techniques, where several labeled positive and negative samples need to be learned for training, which may require large storage space and is unsuitable for vehicular lightweight detection modules [38].

A. Work and Contributions

The increasing prevalence of cyber threats in CPSs, such as EVs and HTVs, necessitates a comprehensive investigation into the potential impacts of stealthy and sequential cyberattacks on these systems. To address this critical issue, this article presents a novel FDIA strategy specifically targeting the EMS in HTVs. The core of our approach lies in developing a covert attacker capable of intelligently infusing FDIAs during specific vehicle states over an extended period with the malicious intent of degrading the vehicle's battery capacity and increasing its energy consumption. This attacker is assumed to be malicious software or a virus that can only perform attacks based on partial information and measurements, and it should be smart enough to learn its optimal strategy from its actions and gains. To counteract these threats and ensure secure power sharing, we propose a complementary detection method based on iForest and DTW, which is particularly effective in identifying stealthy sequential attacks that are localized in specific vehicle conditions within HTVs. The proposed detection framework tackles the challenge of multisignal monitoring and automatic detection by employing a sliding window approach combined with an unsupervised algorithm and a distance metric method. This approach considers the real-time nature of vehicle operation and the high variability in driving conditions. The main contributions of this work are threefold as follows.

1) A novel FDIA strategy targeting the EMS of HTVs is developed. The attacker's strategy is formulated as a POMDP, which accounts for the attacker's limited knowledge of the vehicle's state. An LSTM-DQN-learning algorithm is used to train the attacker to derive optimal strategies. The cyberattack is designed to corrupt actuation signals within the power split module during specific vehicle conditions. A unique reward structure is introduced to incentivize the attacker to be effective, stealthy, and evade detection.
2) The impact of the designed cyberattacks on vehicle energy consumption is assessed. Metrics have been developed to evaluate the degradation of energy efficiency and overall energy consumption. The effects of different cyberattack intensities and road conditions on vehicle performance are investigated, providing insights into physical system behavior under various threat scenarios.
3) A complementary monitoring and detection method based on iForest and DTW is proposed. The detection framework is designed to identify stealthy sequential attacks localized in specific vehicle conditions. The method tackles the challenge of multisignal monitoring and automatic detection by employing a sliding window approach combined with an unsupervised algorithm and a distance metric method [tracks windows with consecutive discordant indices (DIs)]. This innovative approach enhances detection efficiency and robustness, ensuring secure and reliable energy control in HTVs.

It's important to note that cybersecurity is often characterized as an arms race between attackers and defenders. By developing sophisticated attack models, we contribute to a better understanding of potential threats and gain crucial insights into their characteristics and impacts. This knowledge, in turn, informs the development of more robust defense mechanisms. This cycle of innovation is crucial for advancing the overall security of CPSs like HTVs. The rest of this article is organized as follows. Section II explains the cyber–physical model of the HTV. Section III focuses on the mathematical modeling of the proposed FDIA. Section IV presents the training process of the DRL agent. Evaluation metrics are discussed in Section V. The proposed monitoring and detection mechanism is presented in Section VI. In Section VII, the experiment setup is explained. Section VIII provides vulnerability assessments for different cyberattack intensities and driving scenarios, while Section IX presents the performances of the detection method. Finally, Section X concludes this article.

II. Cyber–Physical Modeling of HTV With Steering Loads

Fig. 1 shows the cyber–physical layout of the HTV under investigation. The vehicle's mobility is achieved physically by integrating the high-voltage battery (HVB) pack, ICE, and power electronics via a high-voltage dc bus. The power electronics consist of inverters, which convert dc power to ac power for the traction motors and generators, high voltage

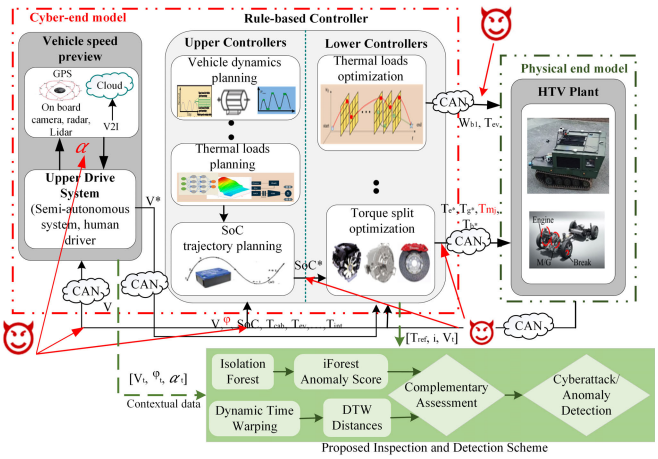


Fig. 1. Typical diagram of the cyber–physical HTV.

motors to operate cooling fans and their associated inverters, and an HV-LV dc–dc converter to service the auxiliary power requirements of low voltage systems such as the autonomous systems, vehicle controllers, communication and safety systems, and coolant pumps. In the HTV system, the upper controllers operate on a time scale of seconds to minutes for high-level planning. The lower controllers and EMS function in milliseconds to seconds, executing real-time strategies, while the power converters respond in nanoseconds to microseconds, ensuring rapid hardware response [39].

The vehicle speed reference is provided by the upper drive control system (autonomous system) that provides the torque demand, and the torque split module (EMS) focuses on tracking the speed by optimizing the four motors' torque requirements considering energy sources (HVB and ICE-GEN). Each motor aims to track its torque reference given by EMS, considering the steering load constraints. The power required for turning a skid-steered tracked vehicle is detailed in [2].

A. Vehicle Dynamics

Newton's second law of motion predominantly governs the longitudinal behavior of the vehicle. For this design, the prediction horizon is segmented into N_p discrete intervals along an axis with a constant time step Δt. Within this framework, the dynamics of the vehicle along the longitudinal axis for the duration t within the range from $t \in [1, N_p]$ are modeled as [9]

$$d(t+1) = d(t) + v(t)\Delta t \tag{1a}$$

$$v(t+1) = v(t) + \frac{\Delta t[T_{\text{total}}(t)/r_w - F(t)]}{M} \tag{1b}$$

$$F(t) = f_w(t) + f_g(t) + f_r(t) \tag{1c}$$

$$T_{\text{total}}(t) = \sum_{i=1}^{n} T_i(t) \tag{2}$$

where t represents the tth time step, d and v indicate the cumulative distance traveled and the speed of the vehicle, respectively, and $d(1) = d_0$, $v(1) = v_0$; T_{total} refers to

the aggregate torque requirement; $T_i(t)$ denotes the reference torque for motor i; r_w signifies the radius of the tire under dynamic conditions. The forces acting on the vehicle, namely, f_w for aerodynamic drag, f_g for the resistance due to gradients, and f_r for the resistance caused by tire rolling, are essential components of the vehicle's dynamics. These forces are calculated below, where α denotes the road grade, C_D denotes the drag coefficient, and f represents the rolling resistance coefficient

$$f_w(t) = \frac{1}{2}\rho_a A_f C_D v^2(t) \tag{3a}$$

$$f_g(t) = mg\sin(\alpha(t)) \tag{3b}$$

$$f_r(t) = mgf\cos(\alpha(t)). \tag{3c}$$

B. Powertrain Model

A powertrain model is developed to simulate the vehicle over a driving scenario to identify the total tractive power required at each instance of a driving scenario. The engine generator (ICE-GEN) and the HVB, as dc electrical power, meet this tractive power, as well as auxiliary power requirements and cooling power. The rate at which fuel flows through the engine is determined by a static mapping that correlates with the engine's torque, T_{ice} and its rotational speed ω_{ice}. This relationship is mathematically represented as $\dot{m}_f = f(\omega_{\text{ice}}, T_{\text{ice}})$, which effectively characterizes the engine's fuel consumption rate

$$P_{\text{fuel}} = \text{LHV} \cdot \dot{m}_f. \tag{4}$$

The LHV is the lower heating value of the fuel. Generally, the dynamics of the battery in the majority EMS of (H)EVs is simplified to an equivalent resistance model [40]. At the tth time instance, these dynamics are expressed as

$$I_{\text{bat}}(t) = \frac{V_{\text{oc}}(t) - \sqrt{V_{\text{oc}}^2(t) - 4P_{\text{bat}}(t)R_{\text{b}}(t)}}{2R_{\text{b}}(t)} \tag{5}$$

where P_{bat} signifies the power either discharged or charged by the battery; I_{bat} denotes the electric current flowing through the battery; V_{oc} represents the open-circuit voltage of the battery, and R_{b} stands for the battery's internal resistance. Each of these variables is subject to change based on the battery's SoC, computed as follows:

$$\text{SoC}(t+1) = \text{SoC}(t) - \frac{\Delta t \cdot I_{\text{bat}}(t)}{C_{\text{bat}}} \tag{6}$$

where C_{bat} denotes the battery's nominal capacity. The power contributions of the battery and engine are typically derived from the mechanical power and efficiency of the motors, the auxiliary and cooling loads calculated as

$$P_{\text{batt}} + P_{\text{ICE-GEN}} = P_{\text{trac-L}} + P_{\text{trac-R}} + P_{\text{cool}} + P_{\text{aux}} \tag{7a}$$

$$P_{\text{trac-L}} = \sum_{i=1}^{n} \omega_i(t) T_i(t)_{\text{trac-L}} \eta_i \tag{7b}$$

$$P_{\text{trac-R}} = \sum_{i=1}^{n} \omega_i(t) T_i(t)_{\text{trac-R}} \eta_i. \tag{7c}$$

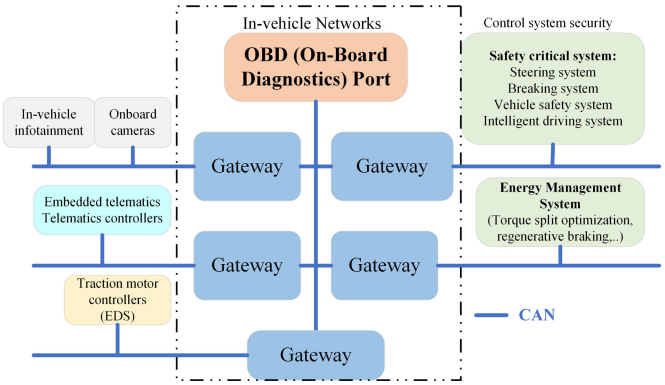


Fig. 2. In-vehicle network architectures.

The tractive power is the sum of the power drawn by the four traction motors to overcome inertia, turning power, aerodynamic drag, and rolling resistance

$$P_{\text{trac}-L} + P_{\text{trac}-R} = \frac{P_{\text{inertia}} + P_{\text{roll}} + P_{\text{aero}} + P_{\text{turn}}}{\eta_{\text{driveline}}}. \tag{8}$$

The vehicle is simulated using a "driver" PID controller following the velocity and turning profiles of the driving scenario. In (8), the relationship between the dc electrical power required by the traction motors and external forces on the vehicle, namely, rolling and aerodynamic resistances, inertial force, and power required to turn the skid-steered tracked vehicle, is described. Here, $P_{\text{trac}-L}$ and $P_{\text{trac}-R}$ are the power drawn by the left and right traction motors. $\eta_{\text{driveline}}$ is the driveline efficiency, which includes mechanical and electrical losses. The general specifications of the vehicle are given in Table A.1. The control of the vehicle requires the confluence of cyber and physical systems, where communication between multiple devices and controllers ensures effective operation. The cyber system spans the communication system between multiple controllers, sensing systems, and their signal processing, as well as control of low, mid, and high-level actuators, to mention some salient systems. The complex energy and thermal control architecture of this vehicle's powertrain is further described in [41]. As seen in Fig. 2, all the signals are transmitted by the high-speed control area network (CAN) buses. Generally, the EMS is connected to the in-vehicle CAN and is considered part of the cyber realm. Given the desired torque demand, the EMS optimizes each converter's torque references (positive or negative values) to maximize energy efficiency. Evaluating the cyber–physical security of the EMS is crucial for ensuring safe and autonomous driving capabilities. The inputs to an EMS are crucial as they influence the energy efficiency of critical components such as motors, batteries, and power electronics. For instance, a compromised EMS can lead to a fake torque request, causing the controller to deliver unintended power that could lead to erratic vehicle behavior and considerable power loss. Fig. 1 illustrates the dominating signals in the designed HTV, which utilizes CAN protocol to communicate with various subsystems. Malicious cyber actors may target these data and controller channels at different locations to compromise the vehicle functionality. In this work, the controller channel under attack is the information transmission between the torque split controller and the

traction motors. A DRL agent that targets the controller outputs ($T_i(t)$), transmitted using the CAN protocol, by injecting stealthy sequential FDIAs during specific vehicle conditions to increase the vehicle's energy consumption is considered. The FDIA design is developed and explained in Section III.

III. Proposed Cyberattack Modeling

An intelligent attacker is developed to launch cyberattacks on the control commands from the power split module, which governs the power inverters responsible for controlling the physical motors. From the attacker's point of view, assessing the states of the target powertrain system relies solely on limited local observations obtained from accessible vehicle sensors. As a result, the attacker cannot comprehensively understand all the consequences of their FDI activities. They are forced to make inferences and learn in conditions of uncertainty. The proposed FDI attacking scheme is formalized using a POMDP to tackle these challenges.

In a POMDP, the interaction between an agent and its environment is modeled. The environment is in a state denoted as s at each time step, and the agent selects an action represented as a, leading to the environment transitioning to a new state, s', with a certain probability described by $T(s'|s,a)$. At the same time, the agent receives an observation denoted as o, which is dependent on the new state with observation probability $\Omega(o|s',a)$. In addition, the agent receives a reward equal to $r(s,a)$. The primary goal of the learning agent is to maximize its expected future discounted reward, expressed $\mathbb{E}_\pi\left[\sum_{t=0}^{\infty}\gamma^t r(s_t,a_t)\right]$ where the discount factor $\gamma \in [0,1)$ determines the balance between immediate and future rewards. A successful solution to POMDP results in the agent's optimal policy, denoted as π, for interacting with the environment [10]. The states, actions, and reward functions of the experience-based FDIA are described below.

A. States and Observations

Without loss of generality, the HTV's energy demand and supply state s_t at time step t, $\forall t \in \{1,\ldots,T\}$ is denoted as

$$s_t = [\alpha(t), \phi(t), \mathrm{SoC}(t), \mathrm{Gen}(t), B_{\mathrm{regen}}(t), P_{\mathrm{cool}}(t) \times P_{\mathrm{aux}}(t), \{T_i(t)\}_{i=1}^{n}] \quad (9)$$

where $\alpha(t)$ is the road grade, $\phi(t)$ is the vehicle acceleration, $\mathrm{SoC}(t)$ is the HVB SoC, $\mathrm{Gen}(t)$ is the energy generated by the generator coupled to the ICE, $B_{\mathrm{regen}}(t)$ is the regenerative braking energy, $P_{\mathrm{cool}}(t)$ and $P_{\mathrm{aux}}(t)$ are the power consumed by the cooling system and auxiliary components, respectively, and $T_i(t)_{i=1}^{n}$ represents the set of four reference torque commands, all at time step t. As the attacker is deprived of the complete topology of the vehicle energy system, only a limited number of measurements can be observed and manipulated during each time step t. Such set of measurements is denoted as $m(t)$, where its elements are $m_j(t)$ for $j = 1,2,\ldots,m_k$. Here, the attacker observes six elements, as seen in Fig. 1, but selectively compromises only four. The observation vector at time t is represented as

$$o_t = [\alpha(t), \phi(t), T_{m_1}(t)), T_{m_2}(t), T_{m_3}(t)), T_{m_4}(t)]. \quad (10)$$

B. Action Space

The attacker aims to tamper with specific signals in an infiltrated vehicle, particularly the control commands from the power-split modules. The attacker monitors $\phi(t)$ and $\alpha(t)$ to determine the optimal moments to inject attacks. To ensure the attacks are both impactful and stealthy, the attacker injects the malicious values only when the vehicle accelerates, the road grade is positive, or both conditions are met.

Let $m_{\mathrm{atk}}(t) \subseteq m(t)$ be a subset of observed data points that the attacker tampers with. The attacker's action vector a_t at time t is defined as

$$a_t = [a_t^{\min}, a_t, a_t^{\max}] \quad (11)$$

where $a_t^{\min}$ and $a_t^{\max}$ represent the minimum and maximum values of the action, respectively. The attacker generates error messages $\Omega_t^{(j)} = 1 + a_t \cdot 0.05$, where $j = 1,2,\ldots,m$ corresponds to the number of tampered measurements. These error messages form the attack vector $av_t = \Omega_t^{(1)}, \Omega_t^{(2)}, \ldots, \Omega_t^{(k)}$. The subtle attack vector is then injected to manipulate the control signals as $T_t^{je} = T_{m_j}(t)\Omega_t^j$, where T_t^{je} represents the forged torque commands from the EMS. To prevent the corrupted commands from deviating significantly from the attacker's intended objective, the forged torque commands are clipped:

$$T_t^{je,\mathrm{CLIP}} = \mathrm{clip}(T_t^{je}, T_t^{je,\min}, T_t^{je,\max}). \quad (12)$$

By carefully monitoring $\phi(t)$ and $\alpha(t)$, the attacker ensures that the attacks are injected at the most effective moments, maximizing their impact while maintaining stealthiness.

C. Reward Function

The reward function, $r(s_t,a_t)$, reflects the attacker's objective, i.e., increasing vehicle energy consumption while aiming to maintain stealth and minimize detection risk. The reward function comprises four subrewards, each addressing specific aspects of the attack scheme. The subrewards are denoted as $R_{E(t)} = \{r_E^0,\ldots,r_E^t,\ldots,r_E^{T-1}\}$, $R_{S(t)} = \{r_S^0,\ldots,r_S^t,\ldots,r_S^{T-1}\}$, $R_{D(t)} = \{r_D^0,\ldots,r_D^t,\ldots,r_D^{T-1}\}$, and $R_{V(t)} = \{r_V^0,\ldots,r_V^t,\ldots,r_V^{T-1}\}$. The first subreward reflects the agent's primary objective, maximizing the attack's impact by deviating the corrupted torque measurements from the true values originating from the power split module. $R_{S(t)}$ measures the agent's ability to mimic expected controller commands, $R_{D(t)}$ captures the likelihood of the attack being detected, and $R_{V(t)}$ encourages the agent to explore different attacking strategies. The effectiveness subreward, $R_{E(t)}$, is designed to maximize the impact of the attack by assessing the deviation between actual torque measurements ($T_i(t)$) and modified torque measurements (T_t^{je}) across multiple dimensions (i = 1 to m). This assessment first computes the raw effectiveness score $E(t) = \sum_{i=j}^{m} |T_t^{je} - T_i(t)| \cdot \delta$. Here, δ influences the agent's decision-making under specific conditions, such as when the vehicle is simultaneously accelerating and experiencing a positive grade. The resulting value is then subject to a min-max normalizing operation, generally constraining $R_{E(t)}$ to a range between 0 and 1. Formally, it is calculated as

$$R_{E(t)} = \frac{E(t) - E_{\min}}{E_{\max} - E_{\min}} \quad (13)$$

where $E_{\min}$ and $E_{\max}$ are the minimum and maximum possible values of $E(t)$, respectively. The stealthiness subreward, $R_{S(t)}$, estimates the agent's ability to make the corrupted control signals closely mimic the expected behavior of the system under normal conditions. This subreward is approximated using the following heuristic:

$$R_{S(t)} = \frac{1}{1 + R_{E(t)}}. \tag{14}$$

The detection subreward, $R_{D_{(t)}}$, models the likelihood of the attack being detected by the system. It uses a sigmoid function of $E(t)$ to incentivize the agent to adopt strategies that reduce the risk of detection. It's mathematically modeled as

$$R_{D(t)} = \frac{1}{1 + e^{-E(t)}}. \tag{15}$$

Finally, the variety score, R_V, assesses the extent to which the agent modifies its attacking strategy. The standard deviation of the modifying factor $\Omega_t^{(j)}$ over a defined temporal window is computed to evaluate the dynamics of the modifying factor. A window size of $p = 50$ is employed, signifying consideration of the 50 previous time steps. This is emulated as

$$R_{V(t)} = \sigma\left(\Omega_0^{(j)}, \Omega_1^{(j)}, \Omega_2^{(j)}, \ldots, \Omega_p^{(j)}\right). \tag{16}$$

These subrewards are aggregated using a weighting system to form the comprehensive reward function

$$r(s_t, a_t) = \rho \cdot R_{E(t)} + \beta \cdot R_{S(t)} - \varkappa \cdot R_{D(t)} + \xi \cdot R_{V(t)} \tag{17}$$

where $\rho, \beta, \varkappa$, and ξ are weights that reflect the importance of each subreward, determined based on domain knowledge and empirical experimentation. The primary goal was to design a reward function that captures the smart attacker's main objectives: 1) maximizing the impact on vehicle energy consumption ($\rho = 0.45$) and 2) maintaining stealthiness ($\beta = 0.25$ and $\varkappa = 0.2$), while assigning a lower weight to the attacker's strategy modification ($\xi = 0.1$). These weight configurations were fine-tuned through iterative experiments to achieve the desired agent behavior that maximizes the attacker's impact and maintains a low detection profile. However, these weights may be subject to future refinement as more data becomes available or vehicle system characteristics evolve.

IV. DRL for Optimal FDI Attack in HTVs

The primary goal of policy optimization is to discover the best possible policy π^* that maximizes the expected cumulative reward, defined as $R(\pi) = \mathbb{E}_\pi\left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t)\right]$. DRL can effectively tackle this optimization challenge. In this context, the agent's objective is to learn the policy $\pi(a_t|s_t)$ that maximizes the expected cumulative reward $R(\pi)$. For a specific policy π, the action-value function $Q^\pi(s_t, a_t)$ plays a crucial role as it characterizes the expected cumulative reward when taking action a_t in state s_t and following policy π thereafter [42]. Its estimation is based on the Bellman equation, as illustrated in the following equation:

$$Q^\pi(s_t, a_t) \leftarrow Q^\pi(s_t, a_t) + l_i\Big[r_{t+1} + \gamma \max_a Q^\pi(s_{t+1}, a) - Q^\pi(s_t, a_t)\Big]. \tag{18}$$

Here, $Q^\pi(s_t, a_t)$ is the Q value estimate for state s_t and action a_t under policy π. The learning rate l_i determines the step size for the update. The term r_{t+1} represents the immediate reward after taking action a_t in state s_t, while γ is the discount factor ($0 \le \gamma \le 1$) balancing immediate and future rewards. $\max_a Q^\pi(s_{t+1}, a)$ denotes the maximum Q value for the next state over all possible actions. The expression $[r_{t+1} + \gamma \max_a Q^\pi(s_{t+1}, a) - Q^\pi(s_t, a_t)]$ is the temporal difference error, representing the difference between the estimated Q value and the target value. This update rule adjusts the Q value estimate based on the immediate reward and expected future value, gradually improving the Q-function approximation. In this study, instead of a vanilla Artificial Neural Network (ANN), an LSTM network is employed as the core component of the deep Q-network [43]. LSTM networks offer improved sequential learning capabilities, making them particularly suitable for enhancing the agent's ability to learn and adapt to complex vehicular dynamics that generate high-dimensional data. Additional costs and benefits of leveraging this approach are detailed in Table A.2. To ensure the stability of the training procedure, the agent saves every experience tuple $e = (s_t, a_t, r_t, s_{t+1})$ at each time step in a memory bank called $D = \{e_1, \ldots, e_R\}$. A randomly selected minibatch of these tuples is taken from D during each training step. The DQN agent employs a Q-network that approximates the action-value function $Q_i^\pi(s_t, a_t|\theta_Q)$, where θ_Q represents the network weights. Furthermore, to facilitate the convergence of the Q-network, an auxiliary target network denoted as $\hat{Q}_i^\pi(s_t, a_t|\theta_{Q'})$ is utilized. The optimization of the Q-network weights takes place within the training process and is guided by the loss function detailed in the following equation:

$$L(\theta_Q^i) = \mathbb{E}\Big[\left(Q_i^\pi(s_t, a_t|\theta_Q^i) - y_t\right)^2\Big] \tag{19}$$

where $y_t = r_t + \gamma \hat{Q}_i^\pi(s_{t+1}, a_{t+1}|\theta_{Q'})$. The target network parameters are updated as described in the following equation:

$$\theta_{Q'}^i \leftarrow \vartheta\theta_Q^i + (1 - \vartheta)\theta_{Q'}^i \tag{20}$$

where ϑ is the smoothing factor ($0 < \vartheta \le 1$).

Our implementation of the LSTM-based DQN algorithm includes two key mechanisms for handling larger datasets: experience replay buffer and batch processing. We utilize an experience replay buffer that allows the agent to store and reuse past experiences, improving learning efficiency and stabilizing training. This buffer can be dynamically adjusted as more data is generated during training. Furthermore, data is processed in mini-batches to handle large datasets without overwhelming computational resources, making the algorithm scalable in terms of time and memory. Future work will explore additional scalability optimizations such as prioritized experience replay, which selects the most informative samples to enhance learning efficiency, and distributed computing, enabling parallel data processing to further scale the algorithm. The complete training process is outlined in Algorithm 1.

V. Evaluation Metrics

This section investigates the impact of stealthy attacks on energy consumption and powertrain efficiency, which can

Algorithm 1 Experience-Driven FDIA Training

Initialize Q-network weights θ_Q, target network weights $\theta_{Q'}$, and experience memory D.
Initialize weights $\{l, \beta, \gamma, \xi\}$ for the comprehensive reward function.
Initialize exploration parameters and learning hyperparameters.
for episode $= 1$ to N_{episodes} **do**
  Initialize the initial state o_1.
  for timestep $= 1$ to T **do**
    Select action a_t based on the policy $\pi(a_t|o_t)$.
    Execute action a_t and observe reward r_t and next state o_{t+1}.
    Store the experience tuple (o_t, a_t, r_t, o_{t+1}) in D.
    Randomly sample a minibatch of experiences from D.
    Calculate target values for the Q-network update.
    Update Q-network weights θ_Q using the loss function $L(\theta_Q)$ with (19).
    Update target network weights $\theta_{Q'}$ with (20).
    Update exploration parameters.
  end for
end for

have long-term consequences, such as reducing the vehicle's traveling range. Two key metrics are considered to assess the impact: the percentage increase in average energy drawn from the battery and the extent of engine fuel consumption throughout the attack horizon. To assess the total energy consumption of the vehicle's engine relative to its normal consumption rate, we express the engine energy consumption at time t as

$$E_{\text{eng}} = \int_0^t T_f(\tau) \cdot \omega_f(\tau) d\tau. \tag{21}$$

In this context, E_{eng} represents the total energy consumed by the engine during the driving cycle, T_f is the engine torque (in Nm), and ω_f is the engine angular speed (in rad/s). The integral accumulates the instantaneous power $(T_f \cdot \omega_f)$ over time, capturing the total energy consumption throughout the driving cycle. By comparing E_{eng} under normal and attacked conditions, we can quantify the impact on the eng-gen side, assuming the SoC remains within defined limits. Now, let's describe the evaluated metric that focuses on change in fuel consumption as

$$I_{\text{eng},k,l} = \frac{E_{\text{atk,eng},k,l} - E_{\text{nom,eng},k}}{E_{\text{nom,eng},k}} \times 100 \tag{22}$$

where $E_{\text{atk},k,l}$ denotes the fuel consumption during the lth attack intensity within the kth drive cycle.

To quantify the battery capacity reduction, which might result in other long-term consequences like battery lifetime degradation and traveling range reduction. We examine the percentage increase in battery energy consumption between the attacked and nonattacked scenarios

$$E_{\text{bat}} = \int_{t_0}^{T_0} V_{\text{oc}}(t) \cdot I_{\text{bat}}(t) dt. \tag{23}$$

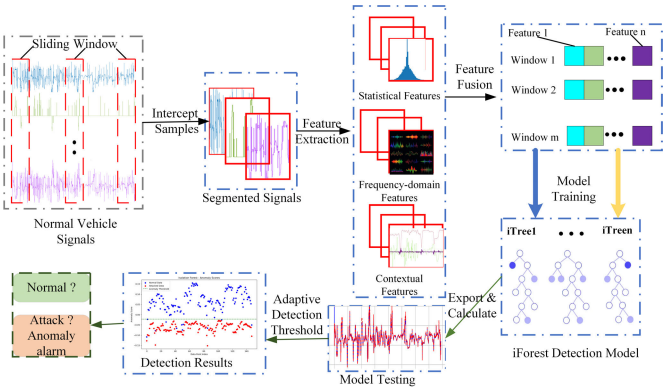


Fig. 3. Cyberattack detection procedure.

Let $E_{\text{bat},k,l}$ represent the total energy consumed by the battery in the lth attack intensity during the kth driving cycle, and $E_{\text{nom},k}$ denote the energy consumption by the battery under nonattack conditions for the kth driving cycle.

We then calculate the battery capacity reduction metric, denoted as $I_{\text{bat},k,l}$, as follows:

$$I_{\text{bat},k,l} = \frac{E_{\text{atk,bat},k,l} - E_{\text{nom,bat},k}}{E_{\text{nom,bat},k}} \times 100. \tag{24}$$

This metric $I_{\text{bat},k,l}$ reflects the increase in battery energy consumption as a percentage relative to the nonattacked scenario for the kth driving cycle and lth attack scenario. It provides insight into the impact of cyberattacks on the battery's energy state.

VI. Complementary Cyberattack Detection With iForest and DTW

This section proposes a sliding window-based controller commands monitoring scheme to tackle the challenges of detecting subtle, efficiency-degrading FDIAs. This scheme focuses on smaller data segments, making identifying localized attacks in specific conditions easier. The proposed method uses an iForest algorithm [44]. This unsupervised learning method isolates anomalies by recursively partitioning the data based on randomly selected features and split points. Furthermore, DTW distance [45] is employed alongside the iForest to screen for windows with consecutive anomalies in the controller output signals. The iForest scheme consists of the following steps.

A. Sliding Window Segmentation

The continuous time-series data of the vehicle used for anomaly detection are divided into fixed-sized, overlapping windows, as shown in Fig. 3. Let W be the window size and S be the step size. The ith window can be represented as

$$W_i = [x_i, x_{i+1}, \ldots, x_{i+W-1}] \tag{25}$$

where x_i represents a vector containing the EMS output values and other relevant vehicle contextual signals at time step i. Here, a window size of 100 and a sliding step of 10 is adopted to segment the signals for feature extraction. This window size balances capturing sufficient temporal information to

detect anomalies while maintaining computational efficiency. A larger window size could capture more complex patterns but may also introduce more noise, while a smaller window size might miss subtle anomalies. The sliding step of 10 allows for overlapping windows, ensuring continuity in detection and reducing the risk of missing critical features.

B. Feature Extraction

For each window, a set of features is extracted to capture the characteristics of the controller and vehicle-level signals. Feature extraction helps suppress noise interference and reduce information redundancy. The extracted features include the following.

Statistical Features:

1) Mean: $\mu_i = (1/W)\sum_{j=i}^{i+W-1} x_j$.
2) Standard deviation: $\sigma_i = ((1/W-1)\sum_{j=i}^{i+W-1}(x_j - \mu_i)^2)^{1/2}$.
3) Kurtosis: $K_i = ((1/W)\sum_{j=i}^{i+W-1}(x_j-\mu_i)^4/\sigma_i^4)$.

Frequency Domain Features: The frequency spectrum of the time-domain overlapping signal is obtained through the fast Fourier transform (FFT). Let $X_i(k)$ be the FFT of the ith window

$$X_i(k) = \sum_{j=i}^{i+W-1} x_j e^{-j2\pi kj/W}. \tag{26}$$

The following features are extracted from the frequency domain.

1) Mean of absolute FFT coefficients: $\mu_{FFT,i} = (1/W)\sum_{k=0}^{W-1}|X_i(k)|$.
2) Standard deviation of absolute FFT coefficients: $\sigma_{FFT,i} = (\frac{1}{W-1}\sum_{k=0}^{W-1}(|X_i(k)| - \mu_{FFT,i})^2)^{1/2}$.
3) Maximum absolute FFT coefficient: $\max_{FFT,i} = \max_{k=0}^{W-1}|X_i(k)|$.

Contextual Features: Additional vehicle operating conditions-related features are incorporated, including vehicle speed v, acceleration ϕ, and road grade α mean.

1) Feature Fusion: The extracted features from all windows are combined to form a feature tensor X. Each row of the tensor represents the features extracted from a single window.

2) Anomaly Detection Using iForest: The iForest algorithm trains the target model using the normal data (i.e., data without attacks). The algorithm assigns an anomaly score to each data point. The anomaly metric for a data point x is calculated as

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}} \tag{27}$$

where $E(h(x))$ is the average path length of the point x across all isolation trees, $c(n)$ is the average path length of unsuccessful searches in a binary search tree, and n is the number of data points. An adaptive anomaly threshold τ is established based on each window's anomaly scores calculated by the algorithm. Data points with anomaly scores surpassing this threshold are classified as anomalies, indicating potential attacks. The test data, collected when the vehicle is under attack, is fed into the trained iForest model to detect cyberattacks. Any windows with anomaly scores exceeding the threshold are marked as potential attacks. Furthermore, a DTW layer is incorporated alongside the iForest model to enhance attack detection by identifying and tracking DIs. The DIs compute the dissimilarity between time series data of normal torque signals and corrupted ones. To accomplish this, DTW distance is leveraged. DTW is a technique commonly used in data mining and clustering to determine the minimum distance between two-time sequences while accounting for variations in sequence length and time synchronization. A lower DTW distance indicates that the observed behavior closely aligns with the normal expected behavior, while a higher DTW distance signifies a deviation. To enhance the detection of these deviations, an adaptive threshold is implemented, which dynamically adjusts based on recent DTW distances.

C. Discordant Index Calculation Using DTW Distance

Consider two time sequences, $X = \{x_1, \ldots, x_i, \ldots, x_I\}$ and $Y = \{y_1, \ldots, y_j, \ldots, y_J\}$. We define a distance matrix $M = \{m_{a,b}\}$ where $m_{a,b} = |x_a - y_b|$ represents the distance between x_a and y_b.

The path $n = \{n_1, \ldots, n_T\}$ from (x_1, y_1) to (x_a, y_b) is known as the wrapping path of the two sequences [45]. This path must satisfy the following constraints:

1) $\max\{X, Y\} \le T \le A + B - 1$;
2) $n_1 = m_{11}$ and $n_T = m_{AB}$;
3) for $n_t = m_{xy}$, the next element n_{t+1} is only one of the set $\{m_{a,b+1}, m_{a+1,b}, m_{a+1,b+1}\}$.

The DTW path represents the optimal path among all possible paths, where each element signifies the local shortest distance

$$d(X,Y) = \min\sum_{a=1}^{T} n_a. \tag{28}$$

To enumerate the DTW distance, the aggregate distance matrix D_m is defined, with elements calculated recursively as follows:

$$D_m(a,b) = \begin{cases} m_{11}, a = 1 \text{ and } b = 1 \\ m_{1b} + D(1, b-1), a = 1 \text{ and } b \neq 1 \\ m_{a1} + D(a-1, 1), a \neq 1 \text{ and } b = 1 \\ m_{ab} + \min\{D(a-1,b), D(a-1,b-1), \\ D_m(a, b-1)\}, a \neq 1 \text{ and } b \neq 1. \end{cases} \tag{29}$$

The DTW distance between time sequences X and Y is obtained from the cumulative distance matrix

$$d_m(X,Y) = D_m(A,B). \tag{30}$$

Therefore, DI is calculated as follows:

$$\text{DI} = \frac{d(X,Y)}{\max(A,B)} \tag{31}$$

where $d_m(X,Y)$ is the DTW distance between time sequences X and Y, A is the length of time sequence X, and B is the length of time sequence Y. The DI quantifies the dissimilarity between normal and attacked torque signals, reflecting the degree of discordance. A higher DI value suggests a greater divergence between the two sequences, indicating a potentially

Algorithm 2 Anomaly Detection Using iForest Algorithm

Inputs: Normal data: X_{normal}, test data: X_{test}, window size: W, step size: S and anomaly score threshold: τ
Train the Isolation Forest model:
Divide the normal data X_{normal} into overlapping windows of size W with step size S.
Extract features from each window as described in the feature extraction step.
Train the iForest model using the extracted features.
Detection Process:
Divide the test data X_{test} into overlapping windows of size W with step size S.
Extract features from each window as described in the feature extraction step.
Pass the extracted features through the trained iForest model.
Obtain the anomaly scores for each window.
Output:
Flag windows with anomaly scores above the threshold τ as potential attacks.

more detectable attack. An adaptive threshold is adopted to flag off anomalous indices by computing the moving average of the DTW distances by an average window length of 10 and threshold factor 1.1 enabling quicker adaptation to local data variations. This threshold achieves the desired sensitivity for capturing windows with consecutive outliers as potential cyberattacks. Combining iForest and DTW distance proves beneficial in ensuring secure power distribution from the controller, particularly in the context of real-time driving of tracked vehicles. These vehicles exhibit high levels of uncertainty and wide variations in operating conditions, making it challenging to differentiate between abnormal and varying driving scenarios. By leveraging the iForest model to detect point anomalies and the DTW layer to identify discordant patterns, the proposed approach can more effectively identify anomalous behavior and potential cyberattacks amidst the inherent variability in real-time tracked vehicle operation.

VII. Experimental Setup

This study develops a co-simulation platform integrating MATLAB and Python to increase understanding of cybersecurity aspects of a previously constructed HTV detailed in our prior research [2]. The vehicle model was validated in real-time simulation on Speedgoat. To thoroughly understand the impacts of the stealthy sequential cyberattacks, our investigations considered three distinct driving terrains, as shown in Fig. 4. The smart attacker module, built-in Python and deployed on a Raspberry Pi model 4B (RPI), incorporates supplementary scripts simulating various attacker behaviors. Fig. 5(a) and (b) depicts the comprehensive design and sequence of the experimental setups. The Speedgoat simulator runs the real-time vehicle model by solving intricate vehicle controls, ensuring the intended driver power requirements are met, and power distribution between the battery and engine is managed effectively. Simultaneously, the trained DRL agent exploits the vulnerabilities of the torque command signals

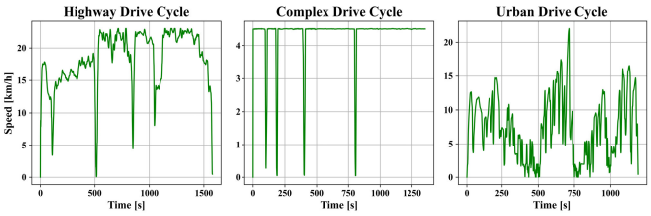


Fig. 4. Driving cycles under different road conditions.

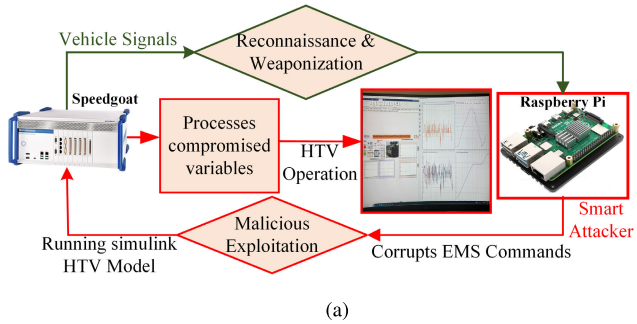


Fig. 5. Experimental setup for the proposed system. (a) Vulnerability assessment setup for HTV. (b) Complete running experimental setup, including vehicle simulation, attacker layer, and cyberattack detection components.

by generating subtle FDIAs. The attacker program's interface receives the observation o_t as its input and produces corrupt torque variables T_t^{je} as its output, which gets embedded into control decisions directed toward the motor inverter.

A 50 ms sampling time is adopted in our experimental setup, balancing computational efficiency with system responsiveness. This interval aligns with typical operational timescales of vehicle control systems and CAN bus message cycles, ensuring realistic simulation of attacker interactions with vehicle data and control commands [46], [47]. Furthermore, the attacker's design facilitates variable attack intensity, ranging from subtle to pronounced, to gauge system behavior. The

vehicle model is run under normal conditions for the three drive cycles to generate datasets for the detection method. A sampling rate of 50ms is considered for data collection. The dataset, consisting of 1600, 1354, and 1200 samples for the three respective drive cycles, includes the EMS output, vehicle speed, acceleration, and road grade. This data is leveraged to train the iForest network. An ensemble of 500 isolation trees (iTrees) and a contamination factor 0.01 is adopted for the iForest algorithm.

In setting up the DRL training, the DQN parameters are determined through experiments and testing. The complete training process spans 500 episodes. The ϵ, which governs the balance between exploration and exploitation in action selection, undergoes a linear decay of 0.00005 from an initial value of 1.0–0.1. This approach ensures that the agent explores a wide range of actions in the early stages of training, gathering diverse experiences and reducing the risk of overfitting to specific sequences. As training progresses, the agent gradually shifts toward exploiting the learned policy, fine-tuning its decision-making based on the acquired knowledge, as shown in Fig. A.1. The learning rate for Q-network updates is set to 0.001. The Q-network architecture consists of two LSTM blocks and four hidden neural network layers with 128, 64, 64, and 32 neurons, respectively. To mitigate overfitting, dropout regularization is applied after each fully connected layer with a dropout probability of 0.2, and L2 regularization is employed in the optimizer with a weight decay factor of 0.001. These regularization techniques help prevent the model from relying too heavily on specific patterns and promote generalizable learning. Executing one forward pass of our trained LSTM-DQN model on a Raspberry Pi 4 takes approximately 5–11 milliseconds, with a memory footprint of about 1.2–1.5 MB. While these requirements are potentially feasible for modern HTVs, real-time deployment would necessitate careful model optimization. Further improvements, such as model hardware acceleration options (e.g., GPUs and TPUs), adaptive model pruning, and quantization, can be applied to reduce inference time while maintaining accuracy. Tables A.3 and A.4 provide detailed information on the hyperparameters and model architecture, respectively.

VIII. Impact Analysis of Cyber Attacks

A. Attack Intensity I

The experimental results are shown in Figs. 6 and 7. The DRL agents target the EMS output signals to produce subtle, efficiency-degrading FDIAs on the four compromised torque references T_t^{je} across varied driving scenarios. It can be observed that when the stealthy attack vectors are initiated, the system deviates from the normal operating condition. Fig. 6(a)–(c) contrasts normal and attacked torque commands for highway, complex, and urban drive cycles. The comparison between the energy consumption profiles reveals that cyberattacks significantly affect energy efficiency while having a minimal impact on vehicle dynamics. Specifically, during positive torque periods, the attacks increase energy consumption by inducing excessive torque demands, forcing the engine and battery to supply more power than necessary. Conversely, during negative torque periods (regenerative braking), the attacks disrupt the energy recovery process, reducing the amount of energy that can be stored in the battery. It is important to note that energy losses due to internal resistance and other system inefficiencies also contribute to the overall energy deficit.

In the highway cycle, the impact on system deviation averaged between 2.69% and 8.50% across motors. Motor 4 shows a substantial increase in torque values compared to the normal scenario because it consistently operates in a positive torque region, providing propulsive power to the vehicle. This indicates that when corrupt torque references are injected, motor four is forced to bear a significant burden in compensating for the compromised behavior of the other motors. In the urban cycle, the rise ranged from 0.77% to 2.48% is recorded. A sharp torque increase was evident during the complex cycle, averaging between 4.27% and 5.07%. This observation is attributed to this drive cycle's complex positive road grade, which the agent exploits to escalate energy demands and increase energy consumption. Critically, the attack maintains the original signal patterns while executing coordinated manipulations across all motors. This preservation of signal patterns mimics normal operating conditions, making detection extremely difficult. Such stealthy attacks ensure that the deviations remain within plausible limits, effectively concealing the malicious intent while progressively degrading system energy efficiency.

The results of the impact on the supply-side performance are presented in Fig. 7(a)–(c). As observed in the figures, the overall energy efficiency of the vehicle is compromised, resulting in higher battery depletion and fuel consumption. In the highway cycle, the attack's impact is evident through a 0.93% surge in fuel consumption and a 0.61% increase in SoC depletion. In the urban cycle, a 0.3% rise in fuel consumption and a 0.86% increase in SoC depletion are recorded. Finally, given the significant torque manipulation in the complex cycle, a 4.52% spike in fuel consumption and a 1.99% modest deviation while maintaining believable discharge patterns. The minor surges in battery energy consumption observed during the experiment can be attributed to the constrained SoC range of 0.2–0.3. This constraint ensured the presence of distinct charge and discharge cycles during the specific drive cycles, providing more granular insights into the battery's behavior. A broader SoC range would have resulted in only a single charge operation, limiting our ability to observe variations in battery performance across multiple cycles.

In scenarios where the objective is undermining energy efficiency and battery health, stealthy cyberattacks are designed to hide in the system for days to months to achieve the intended attacker's goal, which is to degrade the vehicle's monetary value. The highway driving cycle was extended to operate for 500 cycles with the attack to capture this better. Fig. 8 reveals the attack's cumulative impact, showing that substantial efficiency losses occur over time if left undetected. Exponential projections suggest an extra 7-L fuel consumption, while polynomial projections point to 25 L. As for SoC, exponential predictions anticipate an added 0.1391 kW draw (considering an initial battery capacity of 53.5 kWh), with worst case scenarios exhausting an alarming 42.8 kW.

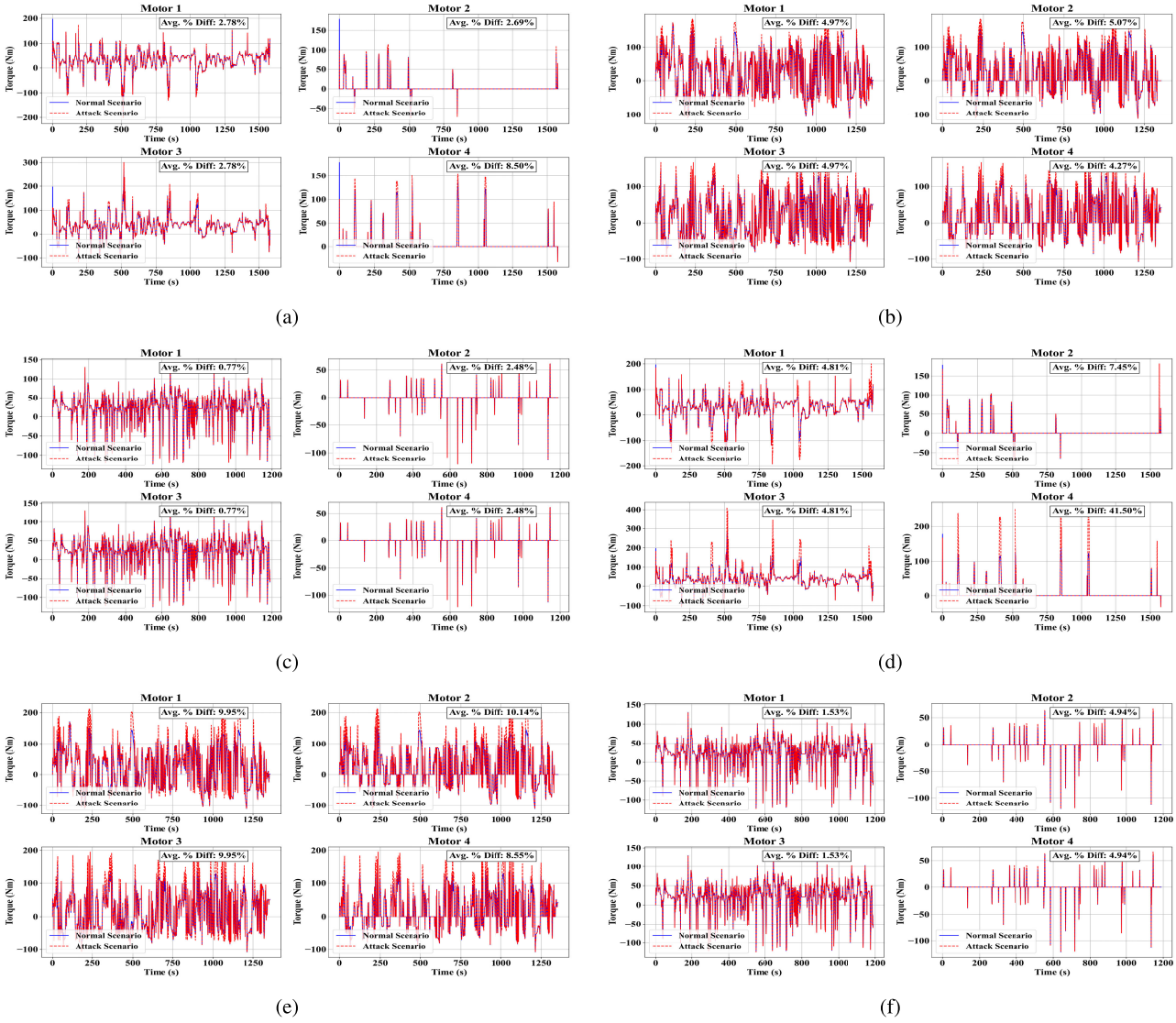


Fig. 6. Normal and compromises torque references for case study I (a) highway driving, (b) complex driving, and (c) urban driving. Case study II (d) highway driving, (e) complex driving, and (f) urban driving.

B. Attack Intensity II

This study amplifies the modification factor $\Omega_t^{(j)}$ leveraged in generating the attack vectors to assess the implications of a more powerful attacker. Fig. 7(d)–(f) demonstrates that these intensified cyberattacks can induce severe consequences on the battery, potentially causing drastic shifts in SoC and fuel consumption trajectories. The intensified attack led to significant spikes in energy consumption, reaching up to 35.54% for fuel and 6.02% for the battery. Fig. 6(d) presents an analysis of the average torque demand increase for each motor in the highway case. Motors 1 and 3 experienced a 4.81% increase, motor 2 saw a 7.45% rise, and motor 4 faced a substantial 41.50% surge in torque demand. The complex cycle observed in Fig. 6(e) displayed an increase varying from 8.55% to 10.14% across the four motors. In the urban cycle, as seen in Fig. 6(f), the energy request increased by 1.53% for motors 1 and 3 and 4.94% for motors 2 and 4. Comparing attack intensity I (stealthy attack) and II (powerful attack), it is evident that attack intensity II has a more pronounced detrimental effect on energy consumption despite sharing some similarities with attack intensity I.

The impacts on the supply sources are observed in Fig. 7(d)–(f). In the highway scenario, a 6.87% increase in fuel consumption and a 6.02% rise in SoC are noted. The urban cycle shows a 2.10% surge in fuel consumption and a 1.42% increase in SoC. The complex cycle is the most affected, with a massive 35.54% rise in fuel consumption and a 5.51% escalation in SoC. These results demonstrate that potent controller cyberattacks can dramatically impact energy consumption and probably send the vehicle off path. However, the amplified attack vectors make them easier to detect than stealthy attacks. Notably, the same attack strategy yielded nuanced outcomes across different driving cycles. The complex drive cycle amplified the output, making the cyberattacks particularly impactful. In contrast, the urban cycle's minimal impact highlights that specific driving conditions could potentially camouflage aggressive cyberattacks, rendering them more challenging to detect. These impacts are summarized in Table II.

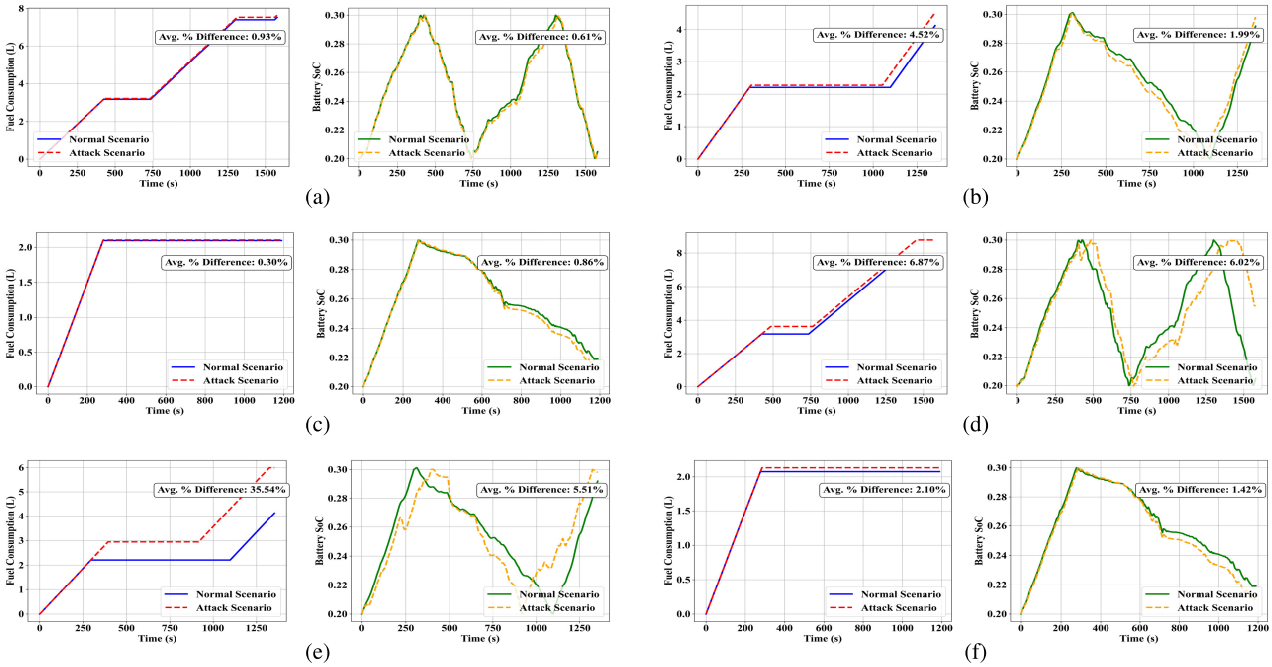


Fig. 7. Impact of cyberattacks for case study I (a) highway driving, (b) complex driving, and (c) urban driving. Case study II (d) highway driving, (e) complex driving, and (f) urban driving.

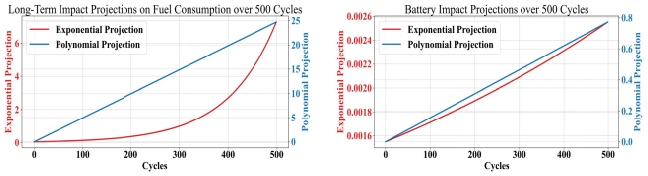


Fig. 8. Highway Terrain Impact Projections Under Stealthy Cyberattacks. (a) Fuel Consumption Increase Over 500 Cycles. (b) Battery Power Draw Projections Over 500 Cycles.

TABLE II

IMPACT OF CYBERATTACKS ON THE SYSTEM

Drive Cycles	Ave. % diff.			
	Stealthy Attack Vectors		Powerful Attack Vectors	
	Fuel (%)	Battery (%)	Fuel (%)	Battery (%)
Highway	0.93	0.61	6.87	6.02
Complex	4.52	1.99	35.54	5.51
Urban	0.30	0.86	2.10	1.42

IX. PERFORMANCE OF THE PROPOSED DETECTION MECHANISM UNDER DIFFERENT ROAD CONDITIONS

The proposed sliding window-based iForest algorithm and DTW distance provide a comprehensive approach for detecting stealthy sequential FDIAs in HTVs. The iForest algorithm is an unsupervised anomaly detection technique that excels at detecting previously unseen attack patterns by constructing an ensemble of isolation trees, where anomalies are more easily isolated closer to the tree's root. On the other hand, DTW distance is a similarity measure that captures the temporal alignment and distance between two-time series. By focusing on smaller segments of data and extracting relevant features, the scheme can effectively identify attacks localized in specific vehicle operating conditions. On a Raspberry Pi 4B, iForest detects these attacks in approximately 0.37 s per data window, while DTW processes one data window in around 38 ms. Figs. 9 and 10 display the performance of the proposed scheme, which combines the iForest algorithm and DTW distance with an adaptive threshold. The adaptive threshold is calculated based on the moving average of the DTW distance and helps distinguish between normal variations and actual attacks. Fig. 9 showcases scatter plots and anomaly scores generated by the iForest algorithm for three drive cycles (highway, complex, and urban) and two attack intensities (stealthy and powerful).

For the highway and urban drive cycles [see Fig. 9(a), (b), (d), and (e)], the iForest algorithm successfully identifies clusters of anomalies (red points) that deviate significantly from the normal operational data (blue points). However, several points in the middle region surpass the anomaly threshold, which could be false positives. This highlights the challenge of relying solely on the iForest algorithm for attack detection in tracked vehicles, as it may be sensitive to normal variations in the vehicle's operation, particularly in the context of tracked vehicles exhibiting high uncertainty and variability levels. For the complex drive cycle [see Fig. 9(c) and (f)], the iForest algorithm's performance varies depending on the attack intensity. During stealthy attacks, the anomaly scores of some attacked points are relatively close to the normal points, making it challenging to distinguish between normal variations and actual cyberattacks. However, for the powerful attack (amplified FDI intensity), there is a clearer separation between normal and attacked points, enabling the iForest algorithm to identify anomalies more effectively and accurately. From Fig. 9, it can be seen that the likelihood of false negatives in stealthy attacks scenario, where attacked points are misclassified as normal due to their anomaly scores being close to normal data points, is present.

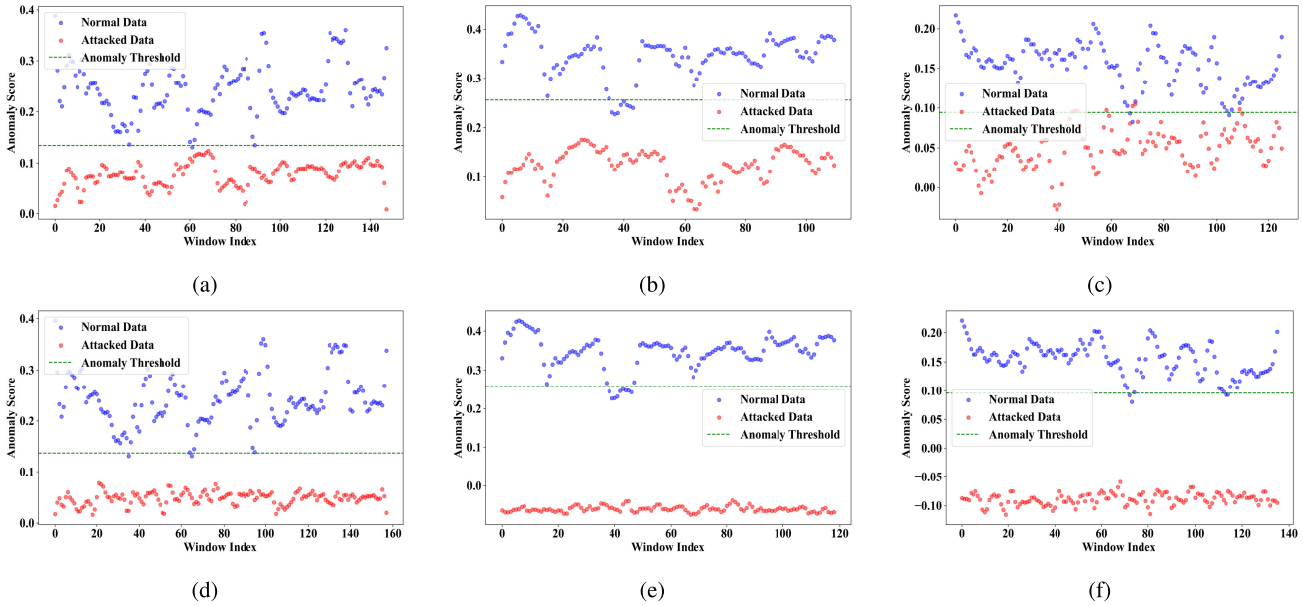


Fig. 9. Scatter plots for case I (a) highway driving, (b) urban driving, and (c) complex driving. Scatter plots for case II (d) highway driving, (e) urban driving, and (f) complex driving.

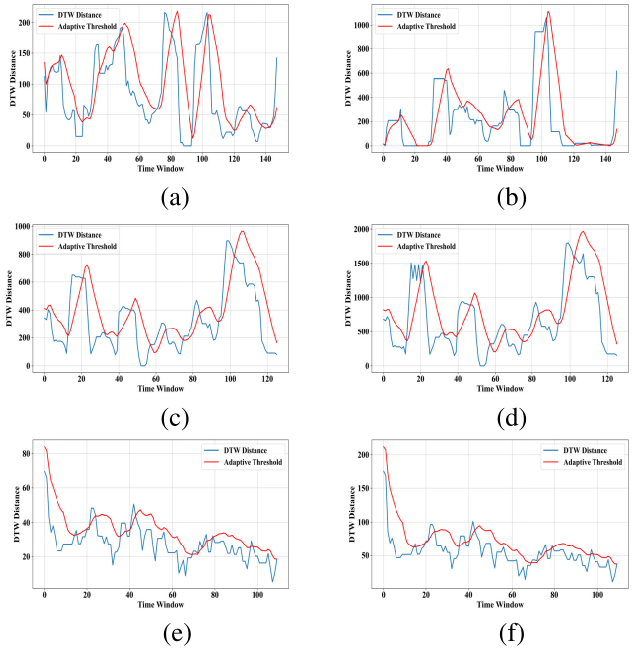


Fig. 10. DTW distances for highway terrain (a) attack intensity I, (b) attack intensity II, complex terrain (c) attack intensity I, (d) attack intensity II, urban terrain (e) attack intensity I, and (f) attack intensity II.

Fig. 10(a) and (b) presents the DTW distance with an adaptive threshold for the highway drive cycle to identify the occurrence of consecutive outliers. As observed in Fig. 10(b), the DTW distance shows spikes during time windows 50, 80, and 100. The adaptive threshold (red line), which adjusts based on recent DTW distance, helps distinguish between normal variations and potential attacks. Significant and consecutive spikes in the DTW distance above the threshold indicate the presence of compromised windows associated with stealthy FDIAs. The DTW distance captures anomalous patterns and sequences that point-wise anomaly detection methods like iForest may miss by considering the temporal context and the relationship between consecutive data points. Fig. 10(c)–(f) presents the DTW distances for stealthy and powerful attacks in the other drive cycles. As observed in Fig. 10(d), DTW distance exhibits noticeable spikes, particularly around time windows 15, 45, and 90, indicating the presence of anomalous windows corresponding to periods where the attacks are most active. When the DTW distance exceeds the adaptive threshold, it strongly suggests the presence of attacks and their prolonged impact on the system. The DTW distance provides valuable insights into the temporal aspects of the stealthy attack by capturing anomalous patterns and sequences.

Combining the iForest and DTW algorithms creates a robust attack detection scheme. The iForest excels at detecting point anomalies, while the DTW distance uncovers anomalous sequences that may be more subtle and gradual. By correlating the detections from both techniques and filtering for agreement between them, the proposed scheme can effectively reduce false positives and improve the overall accuracy of attack detection in HTVs across different driving conditions and attack intensity levels. By leveraging the strengths of both iForest and DTW distance and triggering alarms based on either method's detections, the system ensures high sensitivity to potential stealthy attacks (that seamlessly hide within normal system data) while minimizing false positives and negatives. The proposed defense framework carefully considers the tradeoffs between detecting fine-grained discordances and persistent point anomalies. By balancing these two aspects, the framework provides reliable decision support to ensure secure control in HTVs.

X. Conclusion

This study investigated the vulnerabilities of HTVs to stealthy sequential cyberattacks aimed at degrading the battery capacity and energy efficiency of the hybrid powertrain system. By leveraging the POMDP model, sophisticated learning-based cyberattacks were designed to escalate the

TABLE A.1
OFF-ROAD SERIES HTV SPECIFICATIONS [41]

Description	Units	Value
Mass	[kg]	2900
Effective frontal area	[m^2]	3.6
Drag coefficient	[-]	0.8
Sprocket radius	[m]	0.19
Coefficient of rolling friction	[-]	0.045
Driveline efficiency	[-]	0.90
Battery specifications	[kWhr, V]	50 kWhr, 800V
Silent mode specifications	[km]	50
Traction/Generator motor specifications	[kW]	100 continuous, 200 peak
Max/continuous power per track	[kW]	400 max, 200 continuous
ICE specifications	[kW]	194 kW 3.0V6 FCA EcoDiesel
Top speed	[kmph, mph]	72 kmph (45 mph)

TABLE A.2
COMPARISON OF LSTM-BASED DQN APPROACH WITH RELATED WORKS

Features	LSTM-DQN (Proposed)	Q-learning [10]	DDPG [20]
Network Architecture	Uses recurrent units to capture long-term temporal dependencies	Tabular representation, no neural network architecture	Actor-critic architecture with separate policy and value networks
Handling Sequential Data	Excels at learning from sequential/time-series data patterns	Limited sequence handling, treats observations independently	Can handle continuous action spaces, but limited temporal modeling
Overfitting Risks	Moderate risk, mitigated by techniques like data augmentation, regularization, dropout and experience replay	High risk of overfitting due to tabular representation	Moderate risk, mitigated by target networks and experience replay
Training Requirements	Higher computational demands due to recurrent computations	Low computational requirements, but suffers from the curse of dimensionality	High computational requirements due to separate networks
Computational Latency	Increased inference latency from sequential computations	Low latency for inference, but may not scale well	Moderate latency for inference
Convergence Speed	LSTM converges faster due to better temporal modeling capability	Slow convergence, especially in large state-action spaces	Faster convergence compared to Q-learning, but slower than DQN

vehicle's energy consumption while evading detection subtly. The execution of the cyberattack employed a DRL agent that learns from experiences to identify and exploit the vehicle's vulnerabilities. Specific metrics are developed to evaluate the severity of the designed cyberattacks on the vehicle. The results highlighted that stealthy cyberattacks can seamlessly blend within normal system data, potentially causing a 4.52% increase in the vehicle's energy consumption. In contrast, aggressive cyberattacks were discovered to be more discernible despite having a more pronounced impact of reducing battery capacity by 6.02% and increasing fuel consumption by 35.54%. A sliding window-based controller monitoring scheme combining iForest and DTW algorithms was proposed to address the challenge of detecting these stealthy FDIAs. This detection module demonstrated promising results, achieving 98.47% average detection accuracy for the evaluated drive cycles. The iForest algorithm's ability to effectively identify point anomalies and DTW's capability to capture discordant temporal patterns proved to be a robust combination for real-time detection of stealthy sequential FDIAs in HTVs.

While the LSTM-DQN model successfully simulates sophisticated attack strategies, its computational intensity poses challenges for resource-constrained vehicular environments. To address this, future work will focus on optimizing the model using techniques such as pruning and quantization, with the goal of reducing computational demands while retaining its key capabilities. These optimizations aim to bridge the gap between the theoretical framework and the practical constraints of real-world implementations.

TABLE A.3
DQN HYPERPARAMETERS

Hyperparameter	Value
Discount factor (γ)	0.8
Initial exploration rate (ϵ)	1.0
Learning rate (l)	0.001
Input dimensions	6
Number of actions	5
Memory size	25000
Minimum exploration rate (ϵ_{min})	0.1
Batch size	64
Target network update frequency	200
Exploration rate decay (ϵ_{decay})	0.00005
L2 regularization decay factor	0.001

In addition, we plan to explore transfer learning, distributed computing, and model-based RL to reduce sample complexity while maintaining learning accuracy. These approaches will help accelerate training and enhance the model's adaptability to various scenarios. Furthermore, we intend to develop digital twins to facilitate proactive countermeasures and resilience strategies tailored specifically to vehicular systems. This digital twin approach will improve the robustness of HTV control systems, enhancing their defenses against emerging cyber–physical threats.

APPENDIX

See Table A.1 for the off-road series HTV specifications, Table A.2 for a comparison of the LSTM-based DQN approach with related works, Table A.3 for the DQN hyperparameters,

TABLE A.4
DQN MODEL ARCHITECTURE

Layer
LSTM(input_size=6, hidden_size=128, num_layers=2, batch_first=True)
Linear(128, 128)
Dropout(p=0.2)
Linear(128, 64)
Dropout(p=0.2)
Linear(64, 64)
Dropout(p=0.2)
Linear(64, 32)
Dropout(p=0.2)
Linear(32, 5)

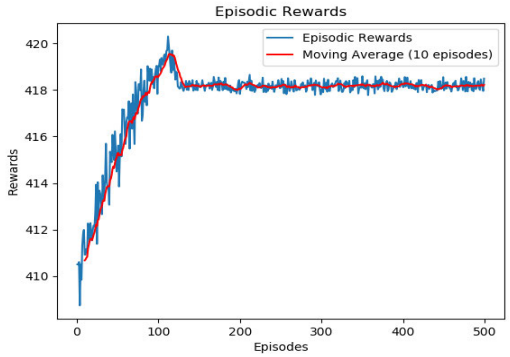


Fig. A.1. Episodic rewards during the training process (complex model).

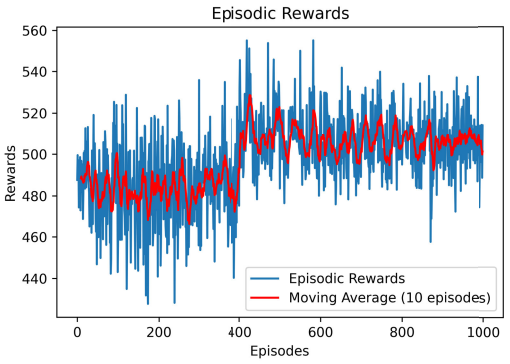


Fig. A.2. Episodic rewards during the training process (simplified model).

and Table A.4 for the DQN model architecture, respectively. Figs. A.1 and A.2 illustrate the episodic rewards of the DRL agent during the training process for two different model configurations.

A. Episodic Rewards for Different Model Configurations

This section provides a comparison of the episodic rewards observed during the training process for two configurations of the DRL agent: one using a complex model with two LSTM layers and another using a simplified model with one LSTM layer.

Fig. A.1 shows the episodic rewards of the DRL agent during the training process using a model with two LSTM layers. This configuration, discussed in the main body of the manuscript, demonstrates faster convergence and more stable performance. The stability of these rewards reflects the stronger learning capacity of the more complex architecture. On the other hand, Fig. A.2 displays the episodic rewards during training for a simplified version of the model, where the number of LSTM layers was reduced to one. The goal of this exploration was to examine the tradeoff between model complexity and computational efficiency, which is crucial for deployment in resource-constrained environments. As shown, the rewards exhibit more variability, and the learning process is slower compared to the complex model. However, the moving average (shown in red) still demonstrates a stable upward trend, indicating that the model is learning, albeit with increased noise.

ACKNOWLEDGMENT

The authors would like to express their gratitude for the valuable resources and facilities provided by the Secure Energy and Automation Laboratory (SEAL) and Real Time Controls and Optimization Laboratory (RTCOOL). Distribution Statement A: Approved for public release; distribution is unlimited (OPSEC 8147).

REFERENCES

[1] S. Carberry. (2022). *Army Rolls Ahead With Ground Vehicle Modernization*. National Defense Magazine. [Online]. Available: https://www.nationaldefensemagazine.org/articles/2022/10/11/army-rolling-ahead-with-ground-vehicle-modernization

[2] Q. Zhu et al., "Development of a series hybrid electrified powertrain for a high speed tracked vehicle based on driving cycle simulation," *SAE Int. J. Adv. Current Practices Mobility*, vol. 4, no. 4, pp. 1403–1412, Mar. 2022.

[3] L. Guo, J. Ye, and L. Du, "Cyber–physical security of energy-efficient powertrain system in hybrid electric vehicles against sophisticated cyber-attacks," *IEEE Trans. Transport. Electrific.*, vol. 7, no. 2, pp. 636–648, Jun. 2021.

[4] A. Arsalan et al., "Cyber attack detection and classification for integrated on-board electric vehicle chargers subject to stochastic charging coordination," *Transp. Res. Proc.*, vol. 70, pp. 44–51, Jan. 2023.

[5] B. Yang, L. Guo, F. Li, J. Ye, and W. Song, "Vulnerability assessments of electric drive systems due to sensor data integrity attacks," *IEEE Trans. Ind. Informat.*, vol. 16, no. 5, pp. 3301–3310, May 2020.

[6] J.-W. Wang and L.-L. Rong, "Cascade-based attack vulnerability on the U.S. power grid," *Saf. Sci.*, vol. 47, no. 10, pp. 1332–1336, Dec. 2009.

[7] Z. Zhang, S. Gong, A. D. Dimitrovski, and H. Li, "Time synchronization attack in smart grid: Impact and analysis," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 87–98, Mar. 2013.

[8] J. Yan, H. He, X. Zhong, and Y. Tang, "Q-learning-based vulnerability analysis of smart grid against sequential topology attacks," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 200–210, Jan. 2017.

[9] L. Guo et al., "Systematic assessment of cyber-physical security of energy management system for connected and automated electric vehicles," *IEEE Trans. Ind. Informat.*, vol. 17, no. 5, pp. 3335–3347, May 2021.

[10] Y. Chen, S. Huang, F. Liu, Z. Wang, and X. Sun, "Evaluation of reinforcement learning-based false data injection attack to automatic voltage control," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2158–2169, Mar. 2019.

[11] Y. Li, J. Yan, and M. Naili, "Deep reinforcement learning for penetration testing of cyber-physical attacks in the smart grid," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2022, pp. 1–9.

[12] R. Gottumukkala, R. Merchant, A. Tauzin, K. Leon, A. Roche, and P. Darby, "Cyber-physical system security of vehicle charging stations," in *Proc. IEEE Green Technol. Conf. (GreenTech)*, Apr. 2019, pp. 1–5.

[13] S.-K. Kang and J.-T. Seo, "An analysis of the security threats and security requirements for electric vehicle charging infrastructure," *J. Korea Inst. Inf. Secur. Cryptol.*, vol. 22, no. 5, pp. 1027–1037, 2012.

[14] J. Ye et al., "Cyber–physical security of powertrain systems in modern electric vehicles: Vulnerabilities, challenges, and future visions," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 9, no. 4, pp. 4639–4657, Aug. 2021.

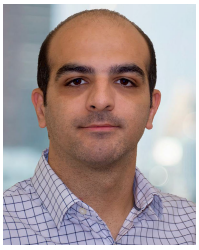
[15] G. Muriithi et al., "Impact analysis of cyberattacks in electric propulsion systems for hybrid tracked vehicles," in *Proc. SAE Tech. Paper Ser.*, vol. 1, Sep. 2024, pp. 1–13.
[16] Y. Huang, L. Huang, and Q. Zhu, "Reinforcement learning for feedback-enabled cyber resilience," *Annu. Rev. Control*, vol. 53, pp. 273–295, Jan. 2022.
[17] Y. L. Yuan, Z. Y. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 382–390, Jun. 2011.
[18] Z.-H. Yu and W.-L. Chin, "Blind false data injection attack using PCA approximation method in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1219–1226, May 2015.
[19] Y. Zhu, J. Yan, Y. Tang, Y. L. Sun, and H. He, "Resilience analysis of power grids under the sequential attack," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 12, pp. 2340–2354, Dec. 2014.
[20] A. J. Abianeh, Y. Wan, F. Ferdowsi, N. Mijatovic, and T. Dragicevic, "Vulnerability identification and remediation of FDI attacks in islanded DC microgrids using multiagent reinforcement learning," *IEEE Trans. Power Electron.*, vol. 37, no. 6, pp. 6359–6370, Jun. 2022.
[21] L. Guo, J. Ye, and B. Yang, "Cyberattack detection for electric vehicles using physics-guided machine learning," *IEEE Trans. Transport. Electrific.*, vol. 7, no. 3, pp. 2010–2022, Sep. 2020.
[22] Y. Wan and T. Dragicevic, "Data-driven cyber-attack detection of intelligent attacks in islanded DC microgrids," *IEEE Trans. Ind. Electron.*, vol. 70, no. 4, pp. 4293–4299, Apr. 2023.
[23] E. Gumrukcu et al., "Impact of cyber-attacks on EV charging coordination: The case of single point of failure," in *Proc. 4th Global Power, Energy Commun. Conf. (GPECOM)*, Jun. 2022, pp. 506–511.
[24] T. T. Nguyen and V. J. Reddi, "Deep reinforcement learning for cyber security," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 8, pp. 3779–3795, Nov. 2021.
[25] J. Giraldo et al., "A survey of physics-based attack detection in cyber-physical systems," *ACM Comput. Surv.*, vol. 51, no. 4, pp. 1–36, Jul. 2019.
[26] Y. Mo, R. Chabukswar, and B. Sinopoli, "Detecting integrity attacks on SCADA systems," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 4, pp. 1396–1407, Jul. 2014.
[27] S. Sridhar and M. Govindarasu, "Model-based attack detection and mitigation for automatic generation control," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 580–591, Mar. 2014.
[28] S. Mishra, Y. Shoukry, N. Karamchandani, S. N. Diggavi, and P. Tabuada, "Secure state estimation against sensor attacks in the presence of noise," *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 1, pp. 49–59, Mar. 2017.
[29] X.-K. Liu, C. Wen, Q. Xu, and Y.-W. Wang, "Resilient control and analysis for DC microgrid system under DoS and impulsive FDI attacks," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 3742–3754, Sep. 2021.
[30] G. Muriithi, B. Papari, A. Moghassemi, A. Arsalan, G. Ozkan, and C. S. Edrington, "Security enhancement of cyber-physical DC ship power system using scalable deep learning method," in *Proc. IEEE Electr. Ship Technol. Symp. (ESTS)*, Aug. 2023, pp. 520–527.
[31] L. Yang, A. Moubayed, and A. Shami, "MTH-IDS: A multitiered hybrid intrusion detection system for Internet of Vehicles," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 616–632, Jan. 2022.
[32] X. Duan, H. Yan, D. Tian, J. Zhou, J. Su, and W. Hao, "In-vehicle CAN bus tampering attacks detection for connected and autonomous vehicles using an improved isolation forest method," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 2, pp. 2122–2134, Feb. 2023.
[33] A. M. M. Sizkouhi, M. Rahimifard, and R. R. Selmic, "Covert attack and detection through deep neural network on vision-based navigation systems of multi-agent autonomous vehicles," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2022, pp. 2583–2590.
[34] J. Kandasamy, S. Arunagirinathan, P. Sivaraj, M. Pameela, G. T. Subham, and R. Nagarajan, "Detection of cyber attack in electric vehicles using ALSTM based machine learning," in *Proc. 4th Int. Conf. Inventive Res. Comput. Appl. (ICIRCA)*, Sep. 2022, pp. 596–600.
[35] H. Zhang, K. Zeng, and S. Lin, "Federated graph neural network for fast anomaly detection in controller area networks," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 1566–1579, 2023.
[36] B. Yang, J. Ye, and L. Guo, "Fast detection for cyber threats in electric vehicle traction motor drives," *IEEE Trans. Transport. Electrific.*, vol. 8, no. 1, pp. 767–777, Mar. 2022.
[37] P. Darby and R. Gottumukkala, "Decentralized computing techniques in support of cyber-physical security for electric and autonomous vehicles," in *Proc. IEEE Green Technol. Conf.(GreenTech)*, Apr. 2019, pp. 1–5.
[38] M. Hanselmann, T. Strauss, K. Dormann, and H. Ulmer, "CANet: An unsupervised intrusion detection system for high dimensional CAN bus data," *IEEE Access*, vol. 8, pp. 58194–58205, 2020.
[39] Y. Gao, S. Yang, X. Wang, W. Li, Q. Hou, and Q. Cheng, "Cyber hierarchy multiscale integrated energy management of intelligent hybrid electric vehicles," *Automot. Innov.*, vol. 5, no. 4, pp. 438–452, Nov. 2022.
[40] L. Tang, G. Rizzoni, and S. Onori, "Energy management strategy for HEVs including battery life optimization," *IEEE Trans. Transport. Electrific.*, vol. 1, no. 3, pp. 211–222, Oct. 2015.
[41] A. Sundar, A. Vannarath, R. Prucka, Q. Zhu, V. Korivi, and Y. Ruan, "The influence of cooling air-path restrictions on fuel consumption of a series hybrid electric off-road tracked vehicle," in *Proc. SAE Tech. Paper Ser.*, Oct. 2023, pp. 1–13.
[42] G. Muriithi and S. Chowdhury, "Deep Q-network application for optimal energy management in a grid-tied solar PV-battery microgrid," *J. Eng.*, vol. 2022, no. 4, pp. 422–441, Apr. 2022.
[43] B. Bakker, "Reinforcement learning with long short-term memory," in *Proc. Adv. neural Inf. Process. Syst.*, vol. 14, 2001, pp. 1–8.
[44] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Proc. 8th IEEE Int. Conf. Data Min.*, Sep. 2008, pp. 413–422.
[45] G. Cao, W. Gu, G. Lou, W. Sheng, and K. Liu, "Distributed synchronous detection for false data injection attack in cyber-physical microgrids," *Int. J. Electr. Power Energy Syst.*, vol. 137, May 2022, Art. no. 107788.
[46] R. I. Davis, A. Burns, R. J. Bril, and J. J. Lukkien, "Controller area network (CAN) schedulability analysis: Refuted, revisited and revised," *Real-Time Syst.*, vol. 35, no. 3, pp. 239–272, Apr. 2007.
[47] S. Onori, L. Serrao, and G. Rizzoni, *Hybrid Electric Vehicles: Energy Management Strategies*, vol. 13. Cham, Switzerland: Springer, 2016.

Grace Muriithi (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya, in 2018, and the M.Sc. degree in electrical engineering from the University of Cape Town, Cape Town, South Africa, in 2022. She is currently pursuing the Ph.D. degree in automotive engineering with Clemson University's International Center for Automotive Research (CU-ICAR), Greenville, SC, USA, specializing in the resilience and security of cyber–physical systems.

She has authored or co-authored several peer-reviewed papers in high-impact journals and international conferences. Her research focuses on hybrid electric vehicles (HEVs) and electric vehicle charging stations (EVCSs), with an emphasis on vulnerability assessment, machine learning-based detection, and robust control mechanisms. Her research interests include cyber–physical security, energy management, power systems, reinforcement learning, and anomaly detection in advanced control systems. Her work contributes to advancing the cybersecurity of connected and autonomous vehicles in the evolving landscape of electrified transportation. In addition to her research, she has been a volunteer reviewer for multiple journals and conferences.

Behnaz Papari (Senior Member, IEEE) received the B.S. degree in electrical engineering from the University of Bushehr, Bushehr, Iran, in 2008, the M.Sc. degree in electrical engineering from K. N. Toosi University of Technology, Tehran, Iran, in 2011, and the Ph.D. degree in electrical and computer engineering from Florida State University, Tallahassee, FL, USA, in 2018.

Prior to joining the Department of Automotive Engineering and Electrical and Computer Engineering, Clemson University, Greenville, SC, USA, in Spring 2022, she was an Assistant Professor at the Department of Marine Engineering Technology, Texas A&M University, College Station, TX, USA, and an Assistant Professor of Practice at the Energy Production and Infrastructure Center (EPIC) for three years at UNC Charlotte, Charlotte, NC, USA, where she taught at undergraduate and graduate levels. Her background is in power systems controls with an emphasis on modeling, analysis, planning, and AI-based optimization and her current research interests are in controls, cyber/physical security, secure control frameworks under uncertainty, and hardware in the loop (HIL) VV&A experimentation, applications on the energy systems, integration of renewable energy, storage, and stochastic optimization, electric vehicle, including electric, and hybrid-electric.



Ali Moghassemi (Graduate Student Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from Islamic Azad University, South Tehran Branch, Tehran, Iran, in 2012 and 2015, respectively. He is currently pursuing the Ph.D. degree with the Real-Time COntrol and Optimization Laboratory (RT-COOL), Holcombe Department of Electrical and Computer Engineering, Clemson University, Clemson, SC, USA.

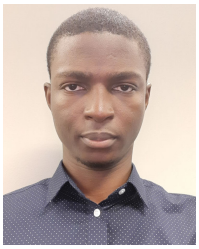
He is also a Graduate Research Assistant in electrical engineering under the supervision of Dr. Christopher S. Edrington with RT-COOL, Holcombe Department of Electrical and Computer Engineering, Clemson University. He authored or co-authored several peer-reviewed papers in high-quality journals and conferences, two book chapters, and one book. Currently, he is working on electro-thermal modeling, simulation, and control of power electronics building blocks (PEBBs) in all-electric ships (AESs). He has also been serving as a volunteer reviewer for several peer-reviewed journals/conferences over the years.

Anirudh Sundar received the master's degree in thermal engineering from IIT Madras, Chennai, India, in 2022, and the master's degree in automotive engineering from Clemson University's International Center for Automotive Research (CU-ICAR), Greenville, SC, USA, where he is currently pursuing the Ph.D. degree in optimal controls in energy and thermal management.

His research primarily focuses on optimal thermal management for batteries under extreme operating conditions and integrated energy and thermal management for modern and emerging powertrain solutions. His work integrates both experimental and simulation-based methodologies to advance this field.

Ali Arsalan (Graduate Student Member, IEEE) is currently pursuing the Ph.D. degree in automotive engineering with Clemson University, Greenville, SC, USA.

After his bachelor's studies, he was a Laboratory Engineer at Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Rahim Yar Khan, Pakistan, for five and a half years. He is currently a Graduate Research Assistant at the International Center for Automotive Research (ICAR), Greenville. His research interests include optimization, microgrids, energy management, and automotive cybersecurity.



Elutunji Buraimoh received the B.Tech. degree (Hons.) in electronic and electrical engineering from Ladoke Akintola University of Technology, Ogbomosho, Nigeria, in 2012, the M.Sc. degree in electronic and electrical engineering from Obafemi Awolowo University, Ife, Nigeria, in 2016, and the D.Eng. degree from Durban University of Technology, Durban, South Africa, in 2021. He is currently pursuing the Ph.D. degree in electrical engineering with Clemson University, Clemson, SC, USA.

His current research focus is the geographically distributed real-time co-simulation of power systems, communication latency compensation, and energy management. His previous studies focused on microgrids, renewable energy, power systems control, and FACTS devices.

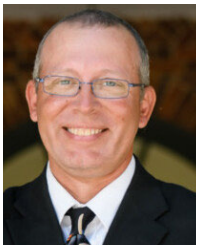
Mr. Buraimoh is a registered Engineer of the Council for the Regulation of Engineering in Nigeria (COREN), a Corporate Member of Nigerian Society of Engineers (NSE), a member of Society of Automotive Engineers (SAE), and a candidate for Engineering Council of South Africa (ECSA).

Laxman Timilsina (Member, IEEE) received the B.E. degree in electrical engineering from the Institute of Engineering, Tribhuvan University, Nepal, India, in 2016, the M.Sc. degree in distributed generation in electrical engineering from the Institute of Engineering in 2019, and the Ph.D. degree in electrical engineering from Clemson University, Greenville, SC, USA, in May 2024, under the mentorship of Dr. Christopher S. Edrington.

Transitioning to the next phase of his academic journey, he enrolled with Clemson University, in Fall 2021, to pursue the Ph.D. degree. Concurrently, he was a Graduate Research Assistant with the Real-Time COntrol and Optimization Laboratory (RT-COOL), Holcombe Department of Electrical and Computer Engineering (ECE), Clemson University, where he is currently a Post-Doctoral Researcher Fellow. His research interests include electrified vehicles, integration of renewable energy, ship power systems, and smart grids.

Gokhan Ozkan (Senior Member, IEEE) received the B.Sc. degree from the Electric Education Department, Marmara University, Istanbul, Turkey, in 2006, and the B.Sc. and M.Sc. degrees from the Energy System Engineering Department, Erciyes University, Kayseri, Turkey, in 2014 and 2016, respectively, and the Ph.D. degree from the Electrical and Computer Engineering Department, Florida State University, Tallahassee, FL, USA, in 2019.

In 2020, he joined Clemson University, Greenville, SC, USA, as a Post-Doctoral Fellow, where he has been a Research Assistant Professor with the Energy Innovation Center since 2022. He is working as a Research Lead with the Energy Innovation Center, North Charleston, SC, USA, and the Manager of Real-Time COntrol and Optimization Laboratory (RT-COOL). His main research interests are power systems and DER integration, model predictive control in power converters and drives, active thermal management and fault-tolerant control, distributed controls, power and energy management, transportation electrification, real-time power distribution system modeling, simulation, co-simulations, and hardware-in-the-loop instantiation. In addition to his research, he develops and executes standard and custom test cases for utility-scale power system equipment.



Christopher Edrington (Senior Member, IEEE) received the B.S. degree in engineering from Arkansas State University, Jonesboro, AR, USA, in 1999, and the M.S. and Ph.D. degrees in electrical engineering from Missouri University of Science and Technology, Rolla, MO, USA, in 2001 and 2004, respectively.

From 2004 to 2007, he was an Assistant Professor of electrical engineering with the College of Engineering, Arkansas State University. From 2007 to 2019, he held the ranks of Assistant Professor, Associate Professor, and Full Professor of electrical and computer engineering with the FAMU-FSU College of Engineering, Tallahassee, FL, USA, and was the Lead for the energy conversion and integration thrust with the Center for Advanced Power Systems, Florida State University, Tallahassee. In 2019, he joined Clemson University, Greenville, SC, USA, as the Warren H. Owen Distinguished Professor of electrical and computer engineering, where he leads the Real-Time COntrol and Optimization Laboratory (RT-COOL). He is also the Associate Director for industry relations with VIPR-GS Center, Clemson University. His research interests include modeling, simulation, optimization, and control of electromechanical drive systems, applied power electronics, integration of renewable energy, storage, and pulse power loads, and hardware in the loop (HIL) VV&A experimentation.

Dr. Edrington was a fellow of Graduate Assistance in Areas of National Need (GAANN) and Integrative Graduate Education and Research Traineeship (IGERT).