IEEE Open Journal of Vehicular Technology

Received 25 May 2025; accepted 1 June 2025. Date of publication 4 June 2025;
date of current version 18 June 2025. The review of this article was coordinated by Editor Rongxing Lu.

*Digital Object Identifier 10.1109/OJVT.2025.3576366*

# Explainable Federated Framework for Enhanced Security and Privacy in Connected Vehicles Against Advanced Persistent Threats

**SUDHINA KUMAR G K [1], KRISHNA PRAKASHA K [1], BALACHANDRA MUNIYAL [1] (Member, IEEE), AND MUTTUKRISHNAN RAJARAJAN [2] (Senior Member, IEEE)**

[1]Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India
[2]Department of Engineering, School of Science and Technology, City University of London, Northampton Square, EC1V 0HB London, U.K.

CORRESPONDING AUTHORS: KRISHNA PRAKASHA K; BALACHANDRA MUNIYAL (e-mail: kkp.prakash@manipal.edu; bala.chandra@manipal.edu).

**ABSTRACT** The increasing adoption of autonomous and intelligent vehicles within ground transportation systems faces new security challenges. This shift from human-controlled operations opens up a broader attack surface for malicious players. As the interconnected Internet of Things (IoT) become ubiquitous in vehicles, they continuously generate and exchange a large amount of data. This tendency creates vulnerabilities that attackers can exploit using sophisticated techniques, such as Advanced Persistent Threats (APT). Detecting APTs in IoT-enabled vehicular environments is crucial. These APTs demand advanced detection mechanisms. The critical need for vehicular data privacy restricts traditional centralized Machine Learning (ML) approaches. Furthermore, the absence of publicly available APT datasets in the vehicular domain complicates model development and validation, creating a significant gap in cybersecurity capabilities for this evolving vehicular domain. This research proposes a novel Federated Deep Neural Network (FDNN) framework with a privacy-preserving technique to address these concerns. This study presents the key challenges in the APT detection phase and outlines the novel contributions to the body of knowledge. The research questions guiding the investigation are addressed and discussed. The features of the UNSW-NB15, Edge-IIoTset, and CSE-CIC-IDS2018 datasets are aligned with different stages of APT attacks. Using these datasets, the developed framework is analyzed and evaluated. For the mentioned datasets, the framework without privacy-preserving technique shows high APT detection accuracies of 97.32% , 96.81% and 98.06% , respectively. However, with the privacy-preserving technique, the framework shows 95.62% , 96.11% and 95.63% accuracies, respectively. All results with other evaluation metrics, such as Precision, False positive rate, F1 score etc., are tabulated. The developed framework is subjected to "Shapley Additive explanations (SHAP)," analysis to filter the considerably influential features in APT detection. This research establishes the efficacy of a novel framework for detecting APTs in distributed vehicular environments. The framework achieves superior performance by minimizing the number of data and reducing the number of features, which is demonstrated through rigorous experimentation on multiple benchmark datasets. The potential of the developed framework to detect the APTs in the cross-domain is discussed in future works.

**INDEX TERMS** Advanced persistent threats, cyber security, ground transport, Internet of Things (IoT), privacy preserving, XAI and V2X.

## I. INTRODUCTION

Integrating Internet of Things (IoT) technologies into the vehicular infrastructure of critical systems represents a significant advancement in modern transportation, offering potential improvements in efficiency, real-time monitoring, and resource management. However, this integration also exposes these vital sectors to unprecedented cyber threats, necessitating robust security measures. Infrastructures encompass

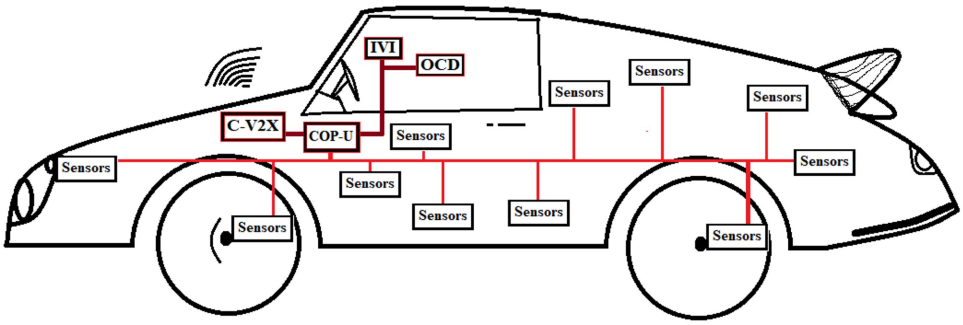


**FIGURE 1.** The complex interactions between the different units and components in the vehicular system.

sectors such as vehicles, energy, transportation, healthcare, supply chain, etc., [1], [2], which increasingly rely on IoT devices for data collection, analysis, and system control.

IoT-enabled vehicles encompass centrally orchestrated processing units (COP-U) and software elements crafted to handle hardware functionalities, interact with other entities using vehicle-to-everything (v2x), collect data and logs, and communicate with the internal devices and sensors via In-Vehicle Infotainment (IVI) System and other Connected Devices (OCD) as in Fig. 1. The technological paradigm comprises multiple facets with the primary objectives of enhancing consistency, identifying avenues for progress, and harnessing untapped potential within vehicular frameworks. Cyber-Physical Systems (CPS) represent a convergence of computation, communication, and physical processes, where integrated devices monitor and control physical environments through real-time feedback. Incorporating IoT sensors, data integration mechanisms, analytics, and ML techniques into CPS can enhance the interoperability and coordination among heterogeneous systems to handle challenge of detecting APTs in IoT networks primarily due to the extreme rarity of APT attacks within normal network traffic, leading to highly imbalanced datasets and the lack of comprehensive public datasets encompassing all APT attack types, which hinders robust model training and evaluation [3]. These sensors play a pivotal role in acquiring data logs from various equipment and relaying them to system analyzers for subsequent processing. Subsequently, ML algorithms utilize these data logs to gain insights and enhance the system's operations for optimal efficiency—the schematic depiction of the interactions in the vehicular system as in Fig. 1.

One of the most pressing concerns associated with IoT-enabled infrastructures is the vulnerability to cyber-attacks [4], [5]. These infrastructures rely on interconnected sensors, actuators, and control systems, as in the Fig. 1, that malicious actors can exploit. Cyber-attacks targeting these systems can lead to devastating consequences, including service disruptions, environmental disasters, and loss of life. Comprehensive security frameworks must be established to eliminate these risks, encompassing encryption, access controls, intrusion detection systems, and continuous monitoring.

The attack surface expands as the number of connected devices within these vehicular infrastructures grows, increasing the complexity of the cyber security efforts. Collaborative efforts among governments, industries, and security experts are imperative to address these evolving threats effectively. One of the primary concerns in IoT security is the sheer diversity of devices, ranging from smart thermostats to industrial sensors, each potentially serving as an entry point for malicious actors. These attackers employ advanced techniques such as zero-day vulnerabilities, malware propagation and encrypted communication to infiltrate networks undetected. Sophisticated cyber-attacks on IoT networks often compromise device credentials, leading to unauthorized access, data breaches, and even full-scale network control. Attackers may employ tactics like Distributed Denial of Service (DDoS) attacks, leveraging IoT botnets to disrupt services and extort victims. Furthermore, these attacks may exploit supply chain vulnerabilities, compromising devices at the manufacturing or distribution stage. Attackers can then manipulate these devices for espionage, data theft, or cyber-physical harm.

Some of the challenges [6] encountered while addressing APTs are

- Complex execution strategies and zero-day vulnerabilities are hard to detect in the traditional detection mechanism.
- Rapid addition of heterogeneous devices to the network will harden the process of flagging the APTs.
- Extreme rarity of APT attacks within normal network traffic, leads to highly imbalanced datasets.
- Since most data are not available in the public domain, adopting large pre-trained models to understand the APTs is challenging.
- Due to victim organisations' privacy concerns, they can't share the entire original data related to the APT attacks for the study.

As scope of the experiment is to detect APTs in the vehicular system, this study addresses some of the earlier challenges by investigating the research inquiries which are as listed in the Table 1 .

Robust security measures are imperative to counter the issues mentioned above. The data is the key, and finding the correct data to train the model to detect APTs in the IoT environment is crucial to avoid a biased model. However, synthetic and semi-synthetic data are available to some extent. One of the solutions to getting actual data is to find the real-world

**TABLE 1. List of Research Questionnaires (RQ)**

| Questions |
|---|
| 1) How are intrusion-related datasets, such as CSE-CICIDS2018 Edge-IIoTset and UNSW-NB15, used as APT-related datasets? |
| 2) How do the datasets' features contribute to the detection of APTs? |
| 3) How can deep learning techniques be leveraged to detect APTs? |
| 4) Does this study preserve the privacy of vehicular data? |

attack datasets in an IoT environment and try to map the possible APT stages from the data. Also, it is essential to address the privacy issues that arise during the data-sharing process.

This research presents the FDNN framework, which effectively addresses two critical aspects: the detection of APTs and the preservation of privacy in VIoT. Contribution to the body of knowledge as follows.

*Contributions:*

1) A Novel Privacy preserved Federated Framework to Detect Advanced Persistent Threats in IoT-enabled connected Vehicules (PF-DAPTIV) is developed to detect the APTs in the IoT enabled vehicular environment. PF-DAPTIV also addresses the fundamental concern: ensuring privacy preservation. This framework operates efficiently on varied datasets, minimizing communication overhead during computations and fortifying vehicular-IoT networks against potential data breaches and privacy vulnerabilities on the server side.
2) As the availability of the APT datasets in VIoT is rare, here the feature of the three datasets are aligned with different stages of APT attacks and utilized. Identify the subset of features that most accurately portray the characteristics of APT attacks, exhibiting commendable model performance within the framework using three distinct datasets.
3) While APT detection is not an unknown subject, the swift evolution of transport infrastructures and the proliferation of diverse IoT devices within vehicular environments have propelled the demand for extensive exploration within this newly formed and intricate domain. The study addresses Research Questions (RQs), listed in Table 1, to provide deeper insights and comprehension of APTs within vehicular-IoT environments, catering to researchers and system defenders alike. The developed PF-DAPTIV framework underwent rigorous Analysis using three distinct publicly accessible datasets, "UNSW-NB15", "Edge-IIoTset" and "CSECIC- IDS2018". A comprehensive ablation study was conducted, meticulously examining the results and drawing pertinent inferences.

The solution is to develop a robust, privacy-preserving, and resource-efficient framework capable of accurately detecting sophisticated attacks like APTs across their entire lifecycle within IoT-enabled Connected Vehicles. At the same time, surmounting the critical challenges posed by the lack of available data and privacy sensitivity. This research addresses these pressing issues by proposing a novel framework, PF-DAPTIV, meticulously designed for the VIoT context.

The subsequent sections of this article are organised as follows: Section II furnishes an overview encompassing the concepts of APT and Non-APTs. Section III delineates the background regarding utilising ML and DL methodologies in the domain of attack detection within the IoT, complemented by a comprehensive review of relevant literature. Section IV expounds the study's methodology, encompassing intricate details on the architecture, framework, experimental configurations, dataset application, and mapping of APT attack stages. Section V comprehensively presents and scrutinises the PF-DAPTIV framework, encompassing 11 parameters, alongside employing eXplainable Artificial Intelligence (XAI) techniques, particularly SHAP, on three distinct datasets. Furthermore, Section VI furnishes a condensed summary of this research endeavour, tabulates the obtained results, addresses the posed Research Questions (RQs), and culminates in Section VII, which deliberates upon future research avenues and concluding remarks.

## II. ADVANCED PERSISTENT THREATS (APTS)

Sophisticated cyber-attacks on IoT networks represent a growing and alarming threat in the digital age. These attacks, characterized by their complexity and strategic nature, exploit vulnerabilities in the interconnected web of IoT devices, posing substantial risks to individuals, organizations, and critical infrastructures. One such form of the sophisticated cyber-attack is APTs [7]. APT attacks are well-planned and carried out with a very high success rate.

APTs are usually misunderstood in the industries and are used as an escape answer for any regular breaches or attacks. If a given attack scenario could have been handled in more than one way or the methods used are very typical, these are targeted attacks, not APTs. There is a silver line between these attacks; some fall into both categories, including "WannaCry". This comes under the ransomware of malware category and APT [8].

These APTs may have many purposes; some of them are:

- Goal is to steal sensitive data
- Intention is to remain inside the network for a very long period
- To observe the system and its behaviours

One of the main reasons for using APT is to steal data, and Alshamrani et al., [6] modelled this process carefully, as in Fig. 2. These stages need not happen in every APT; some models are more generic, and Ussath et al., [9] described the most miniature stages of three in number, a specific case of APT. In general, there are five stages, which are (1)" Initial Compromise", (2)" Foothold", (3)" Lateral Movement", (4)" Data ex-filtration", and (5)" Post-ex-filtration." These stages are self-explanatory. The stages mentioned in Fig. 2 could be elaborated.

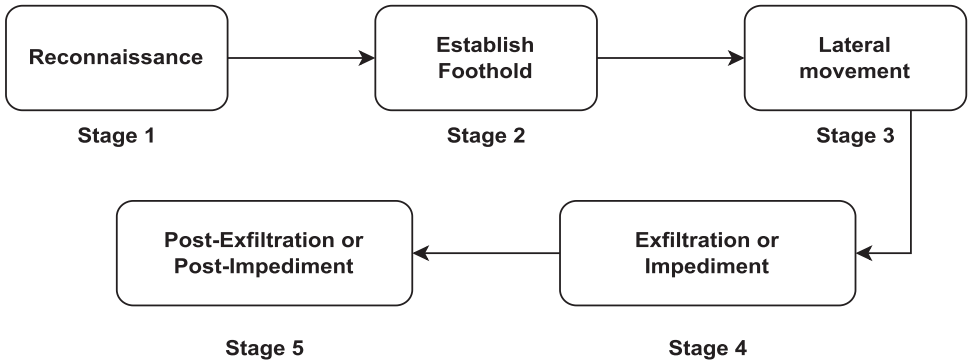


**FIGURE 2.** Generalized APT attack stages on IoT systems, as APTs silently navigate through IoT environments, aiming to infiltrate, exploit, and maintain prolonged access while minimizing detection.

"Reconnaissance" is one of the first steps in many successful APTs. The attacker will monitor and learn the target system and attempt to grasp the IT infrastructures to get access to the target network. Here, threat actors gather comprehensive information about their target. This process involves the identification of potential vulnerabilities, key personnel, and critical systems. Chen et al., [10] found that it encompasses activities such as open-source intelligence (OS-INT) research, the scrutiny of publicly available data, and probing target systems for weaknesses.

"Establish Foothold" emerges as the immediate subsequent stage following the acquisition of crucial information required for initiating an attack. Among the prevalent methods employed, the exploitation of vulnerabilities stands prominent. Motoyama et al., [11] highlighted the presence of these vulnerabilities within extensively documented datasets such as the "Common Vulnerabilities and Exposures (CVE)" alongside circulation in various hacker forums. Notably, most APT attacks have leveraged known vulnerabilities [9] rather than zero-day vulnerabilities for their execution.

"Lateral Movement" Once inside the network, APT operatives navigate and progress laterally to reach other systems and resources. Their objectives encompass privilege escalation, the acquisition of access to sensitive data, and consolidating control over a broader spectrum of assets.

"Ex-filtration or Impediment" refers to APT operations as the systematic collection of valuable information. Here, Target may encompass intellectual property, financial data, credentials, or any other form of sensitive information of interest. To maintain stealth and avoid detection, APT actors employ clandestine techniques, such as data ex-filtration through encrypted channels. After acquiring the desired data, APT attackers clandestinely extract it from the compromised network. This process is conducted covertly, using encrypted and obfuscated communication channels as a common practice.

"Post Ex-filtration or Post Impediment" is to evade detection and sustain their access, APT actors meticulously erase or manipulate logs and any traces of their presence within the compromised network. This step holds critical importance in ensuring their continued undetected existence.

### A. WHAT IS NOT AN APT?

Many misunderstand APTs, often employing them in the industry to explain why organizations struggle to defend against targeted attacks. The specific tactics used by attackers and the precise characteristics of these attacks have spurred the security community to advocate for a reevaluation of what constitutes an APT, encompassing different domains with diverse attack objectives. Alshamrani et al., [6] outlined a set of criteria as listed below. If any of these criteria are met, it suggests that the particular assault may not fit the definition of an APT.

- *Attacks can be Prevented through multiple ways:* Considering the attack process and the targeted environment, if the attack was foreseeable and could have been averted with minimal countermeasures and existing security protocols.
- *Minimal adaptation by attackers:* If the attackers' efforts to achieve their objectives did not necessitate significant adaptations in response to the defender's measures.
- *Lack of novelty in attack variants:* The success of an APT often hinges on the novelty of its methods or techniques. Suppose an attack demonstrates no innovative methods or techniques.

## III. BACKGROUND

Recognizing and precisely classifying APT within cyber Systems poses complex challenges [12] that numerous researchers and facilitators encountered to address.

### A. USE OF MACHINE LEARNING IN ATTACK DETECTION

ML is critical in bolstering attack detection and security within the IoT realm [13]. Through the adept application of ML algorithms and techniques, IoT systems can be fortified to identify and respond to a spectrum of cyber security threats with heightened precision and efficiency [14]. This section delineates several critical avenues through which ML augments the detection of attacks in IoT environments as in Fig. 3.

Connected vehicles offer a glimpse into the future of transportation, promising improved traffic management, enhanced safety features, and, ultimately, autonomous driving. However, this increased connectivity introduces new attack

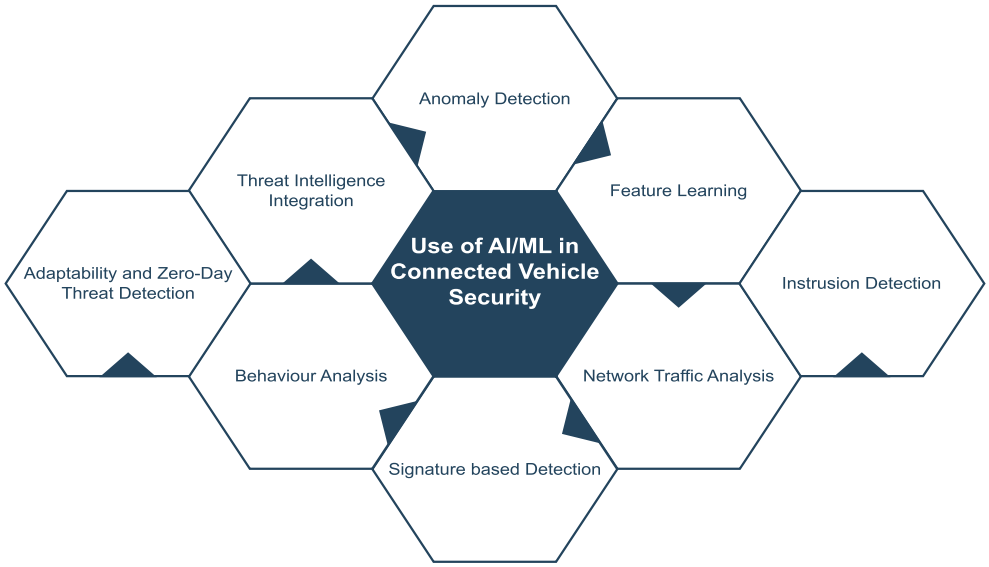


**FIGURE 3.** The portrayal of prevalent applications of AI/ML within the realm of cybersecurity primarily serves as a defence mechanism aimed at comprehending and thwarting various types of cyber attacks on CAVs.

surfaces that malicious actors can exploit. AI/ML techniques are emerging as a powerful tool to address these security challenges. This section reviews the state of the art at the intersection of security and machine learning for connected vehicles.

### B. ANOMALY DETECTION IN VEHICLES

Anomaly detection is crucial in identifying unusual behaviour that might indicate security threats or malfunctions within a connected and autonomous vehicle (CAV). These threats can range from cyberattacks targeting a vehicle's control systems to mechanical failures in critical components. Fang et al., [15] provided a comprehensive survey of anomaly detection techniques designed for CAVs. Their work categorizes existing methods based on the data they utilize (sensor data, communication data, etc.) and the algorithms they employ (statistical methods, machine learning, deep learning). This survey is a valuable resource for researchers developing novel anomaly detection solutions for CAV security. However, even with this extensive categorization, critical challenges still need to be addressed in anomaly detection for CAVs. A crucial challenge is achieving real-time detection. CAVs operate in dynamic environments, and timely identification of anomalies is essential for ensuring safety and security. Additionally, the explainability of deep learning models, often used for anomaly detection due to their ability to handle complex patterns, is another crucial aspect. Understanding why a model flags a particular event as an anomaly can be vital for troubleshooting and mitigating security risks.

Several recent works explore various techniques to address these challenges. Prathiba et al., [16] propose a hybrid deep sensor anomaly detection system for autonomous vehicles. Their system combines the power of deep learning for complex pattern recognition with traditional statistical techniques. This approach could improve the overall accuracy of anomaly detection. However, the computational demands of deep learning algorithms pose a challenge for real-time implementation on resource-constrained vehicles. Another approach that leverages federated learning for anomaly detection is proposed by Zhang et al., [17]. They present a federated graph neural network for fast anomaly detection in controller area networks (CANs), a standard communication protocol within vehicles. This method allows for distributed learning across CAVs while preserving privacy. However, the effectiveness of this approach can be limited by the quality of the local models on individual vehicles. The exploration of anomaly detection techniques for CAVs is an ongoing area of research. As we will see in the next section, researchers continuously develop novel approaches that leverage advancements in machine learning and consider the unique constraints of the CAV environment.

### C. SECURITY FRAMEWORKS AND INTRUSION DETECTION

The ability to detect anomalies paves the way for robust security frameworks and intrusion detection systems (IDS) within CAVs. These systems are critical for identifying and mitigating cyberattacks that target a vehicle's control systems, potentially compromising passenger safety and causing significant disruption to transportation networks. Sedjelmaci et al., [18] proposed a secure attack detection framework designed explicitly for hierarchical 6G-enabled Networks of Vehicles (IoVs). Their approach considers the potential cascading effects of attacks within such networks, where a successful attack on one vehicle can create vulnerabilities for others. This highlights the interconnected nature of CAV security and the need for comprehensive frameworks that

address individual vehicles and the broader transportation ecosystem.

Similar concerns regarding trust and security in future-generation IoVs motivate the work of Rathee et al., [19]. They propose a trust management solution that leverages the capabilities of 5 G networks to establish trust between vehicles and other network entities. Building trust is essential for secure communication and collaboration within CAV networks. Focusing on individual vehicles, Qiu et al., [20] delve into intelligent security authentication for CAVs. Their work explores various authentication mechanisms and analyzes potential attack vectors that could exploit vulnerabilities in these mechanisms. Understanding these vulnerabilities is crucial for developing effective defence strategies. Haddaji et al., [21] take a step further by proposing a framework for enforcing security within a vehicle's internal network. Their approach compartmentalizes critical systems from non-critical applications, creating an additional layer of defence against cyberattacks. This compartmentalization can limit the potential damage caused by a successful intrusion.

Several recent works focus on specific intrusion detection techniques for CAVs, such as DivaCAN [22], a system designed to detect in-vehicle intrusion attacks. DeepSecDrive, a deep learning framework for real-time intrusion detection with explainability features [23]. The explainability aspect of DeepSecDrive is precious as it allows security personnel to understand the reasoning behind the system's decisions. Additionally, Zhou and Che et al. address the challenge of detecting stealthy attacks on autonomous vehicles, which are designed to evade traditional detection methods [24]. Developing security frameworks and intrusion detection systems for CAVs is an active area of research. As we explore in the next section, researchers are also investigating the potential of machine learning and communication security to enhance the safety and security of CAVs further.

*Zero-Day Threat Detection:* ML models can pinpoint previously undisclosed threats by discerning deviations from established behavioural norms, even when no prior signature exists. This capability assumes paramount importance in thwarting attackers exploiting undiscovered vulnerabilities.

*Edge Computing:* Deploying ML models at the periphery of IoT networks enables real-time data analysis without relaying it to a centralized server. This curtails latency and bolsters responsiveness, making it suitable for IoT applications.

*Adaptive Security:* ML-powered security systems exhibit an innate adaptability to evolving threats. They acquire knowledge from fresh data and calibrate their detection capabilities to ensure perpetual protection.

*Threat Intelligence Integration:* ML fosters the seamless integration of threat intelligence feeds, thereby Improving the detection of known threat indicators and indicators of compromise (IoCs) within IoT environments. Enhancing IoT Attack Detection with Deep Neural Networks (DNNs) [25] though ML has manifested efficacy across various cyber security applications [26], including detecting multiple attack types [27], it encounters limitations when detecting advanced attacks within IoT networks. The subsequent section delves into the confines of traditional ML in this context and elucidates how DNN adeptly surmount these challenges.

*Feature Learning:* DNNs boast the innate capability to acquire hierarchical representations of data autonomously. This learning enables them to extract complex and abstract features from raw IoT data, a part of profound significance within IoT, where feature engineering can be inherently intricate.

### D. LITERATURE REVIEW

The ever-expanding landscape of connected vehicles, encompassing traditional internet-connected features and integration with IoT devices, presents a rapidly evolving attack surface for cybercriminals. The timeline of attacks demonstrates a concerning progression from manipulating entertainment systems to potentially hijacking critical driving functions. This highlights the urgent need for robust and multi-layered vehicle cyber defence systems. These systems must safeguard core software and encompass the security of connected infotainment systems, telematics units, and the growing number of IoT devices integrated into modern vehicles. Failure to implement such advanced cyber defences leaves vehicles vulnerable to unauthorized access, potentially endangering passenger safety, causing widespread road disruption, and exposing sensitive driver and vehicle data to theft.

Zidi et al. [28] proposed a hybrid machine learning (ML) and Hierarchical Temporal Memory (HTM) approach for fault prediction and mitigation in vehicular environments. Their hierarchical framework learns spatiotemporal patterns, enabling robust fault prediction and adaptive recovery through real-time data integration. Awan et al. [29] introduced a federated learning-based, privacy-aware location prediction model for vehicular systems. This decentralized approach protects user privacy by training locally on each vehicle and sharing only anonymized model updates. Differential privacy and encryption further enhance security, balancing privacy and prediction accuracy for secure location-based applications.

The convergence of electric vehicles (EVs) and Vehicle-to-Everything (V2X) communication is poised to transform transportation [30]. As data hubs, coupled with real-time V2X data exchange, EVs promise safer ecosystems through enhanced situational awareness and more efficient systems via reduced congestion and energy consumption. This synergy has the potential to enable autonomous driving and a future of sustainable, intelligently connected transportation.

Zhang et al. [31] proposed an intrusion detection model for the Internet of Vehicles (IoV) using GRIPCA for feature extraction and OWELM for adaptive anomaly detection, addressing the challenges of high-dimensional data and evolving network conditions. Similarly, Sharma et al. [32] presented a machine learning-based approach for misbehaviour detection in VANETs, analyzing consecutive Basic Safety Messages (BSMs) with SVM and Random Forest algorithms to identify anomalous patterns and adapting to new attack strategies.

Lv et al. [33] proposed a novel misbehaviour detection framework for VANETs leveraging privacy-preserving FL

and blockchain. Addressing limitations in traditional methods regarding privacy and scalability, the framework enables collaborative model training across vehicles without raw data sharing. Blockchain secures and verifies model update integrity. VANET clustering under trusted authorities facilitates local federated learning with subsequent aggregation and secure blockchain storage. This approach enhances trustworthiness and transparency while safeguarding privacy.

In a recent study [34] investigated optimal defence strategies against APTs in dynamic networks, addressing the challenge of long-term, stealthy attacks. They developed an "optimal repair strategy" to minimize APT impact in constantly evolving networks. Combining game theory and network dynamics, they analyze various scenarios and defences, considering attacker capabilities, network structure, and resource constraints, providing actionable insights for robust, adaptive defence mechanisms.

The 2FLIP authentication scheme, while initially robust, is vulnerable to relay and key exposure attacks, jeopardizing VANET communication security. Current research explores countermeasures such as secure message authentication, time-based authentication, and enhanced key management [35] to mitigate these vulnerabilities. However, the dynamic nature of VANETs necessitates ongoing refinement to address evolving security challenges.

Focusing on satellite navigation, the study [36] explores countermeasures against worst-case spoofing attacks, a critical vulnerability that can cause intense disruptions and safety hazards. These attacks abuse receivers with deceptive signals, leading to inaccurate position and velocity calculations. Researchers are actively exploring methods to enhance system resilience, often combining hardware and software solutions for spoofing signal detection and rejection while safeguarding legitimate navigation data. Advancements in encryption and authentication protocols are also crucial for defence. However, the evolving landscape of cyber threats necessitates continuous research and development to stay ahead of potential attackers, ensuring these critical systems' continued reliability and trustworthiness.

Asensio-Garriga et al., [37] presented a novel approach to securing Vehicle-to-Everything (V2X) communication in Multi-access Edge Computing (MEC) networks against Distributed Denial of Service (DDoS) attacks. The system leverages a Zero-Second-Microservice (ZSM) architecture to manage and allocate security slices across the MEC network dynamically. This combination allows for rapid deployment and adjustments of security measures, ensuring real-time responses to threats like DDoS attacks. By prioritizing proactive DDoS protection in latency-critical V2X environments, the solution enhances communication resilience and optimizes resource allocation within the MEC infrastructure. This dynamic adaptation of security measures based on real-time threat data and network conditions improves the overall reliability and security of V2X communication, paving the way for secure and reliable connected and autonomous vehicles.

### E. MACHINE LEARNING AND COMMUNICATION SECURITY

Secure and reliable communication between CAVs and other network entities [38] is paramount for ensuring intelligent transportation systems' overall safety and effectiveness. Machine learning and robust communication security protocols are vital in achieving this goal. Alqahtani and Kumar [39] provide a concise survey on the use of machine learning for enhancing transportation security in electric and flying vehicles. Their work highlights the potential of machine learning algorithms to identify patterns and anomalies that might indicate security threats. This knowledge can be used to develop intelligent intrusion detection systems and predictive maintenance models for CAVs. Machine learning extends beyond anomaly detection within individual vehicles. Miao et al., [40] propose a deep-meta-heuristic system for unmanned aerial vehicle (UAV) intrusion detection. Similar to CAVs, UAVs operate in dynamic environments and require robust security measures. This work demonstrates the applicability of machine learning for securing various components within intelligent transportation systems.

Several studies address communication security in connected and autonomous vehicles (CAVs). Choi et al. [41] introduce a computationally efficient intrusion detection method using fuzzy hashing. Jangam et al. [42] propose abnormal traffic prediction for enhanced security in IoT-based automated vehicle systems. Prasad et al. [43] highlight the need to protect against data breaches, sensor spoofing, and denial-of-service attacks in wireless communication infrastructure. Chen et al. [44] specifically address DDoS attacks in 6 G V2X networks. While promising, integrating intelligent cyber-physical systems and deep learning, as explored by Aleisa et al. [45], faces challenges in real-time implementation due to CAV resource constraints.

The interplay between machine learning and communication security is crucial for safeguarding CAVs and ensuring the smooth operation of intelligent transportation systems. As research in these areas continues to evolve, we can expect even more robust and efficient security solutions for the future of connected and autonomous vehicles.

### F. FEDERATED LEARNING FOR SECURE COMMUNICATION

The growing reliance on data for CAV development necessitates addressing privacy concerns surrounding sensitive data like driving behaviour and sensor information [20]. Federated learning (FL) offers a promising solution, allowing CAVs to train a central model collaboratively without directly sharing raw data [46], [47]. Each CAV trains the model locally using its dataset and transmits only model updates, significantly reducing privacy risks. This approach is particularly advantageous for CAV security applications such as anomaly detection and intrusion detection systems (IDS).

However, FL for CAVs faces challenges. The central model's effectiveness relies on the quality and participation of individual vehicles [48]. Malfunctioning vehicles or those dropping out can negatively impact accuracy [47]. Secure

**TABLE 2. Summary Table of the Literature Review**

| Papers | Focused Areas | Key Contribution/Method | Addressed Security Concerns | Techniques/Technologies Used |
|---|---|---|---|---|
| [30] | Fault Prediction & Mitigation | Hybrid ML and HTM for historical data analysis and fault prediction; Adaptive fault recovery using real-time sensor data. | Fault prediction, proactive maintenance | Machine Learning, Hierarchical Temporal Memory (HTM) |
| [31] | Location Prediction | Federated Learning-based privacy-aware location prediction. | Data privacy in IoVT location prediction | Federated Learning, Differential Privacy, Encryption |
| [32] | EV & V2X Convergence | Explores the synergy of EVs as data hubs and V2X communication for safer and more efficient transportation. | Traffic congestion, energy consumption, safety | Electric Vehicles (EVs), Vehicle-to-Everything (V2X) |
| [33] | Intrusion Detection (IoV) | GRIPCA and OWELM for feature extraction and adaptive anomaly detection. | Intrusion detection in high-dimensional IoV data | OWELM and Group Incremental Principal Component Analysis (GRIPCA) |
| [34] | Misbehavior Detection (VANETs) | Consecutive BSM analysis and ML (SVM, Random Forests) for misbehavior detection. | Misbehavior detection in dynamic VANETs | Machine Learning (SVM, Random Forests), Basic Safety Messages (BSM) |
| [35] | Misbehavior Detection (VANETs) | Privacy-preserving FL and blockchain for collaborative misbehavior detection. | Data privacy, scalability in misbehavior detection | FL and Blockchain |
| [36] | APT Defense | Optimal repair strategies against APTs in dynamic networks using game theory. | Advanced Persistent Threats (APTs) in dynamic networks | Game Theory, Network Dynamics |
| [37] | 2FLIP Authentication | Explores vulnerabilities and countermeasures for the 2FLIP authentication scheme. | Relay attacks, key exposure in 2FLIP | Secure message authentication, time-based authentication, key management |
| [38] | Spoofing Attacks (Satellite Navigation) | Explores countermeasures against worst-case spoofing attacks. | Spoofing attacks in satellite navigation | Hardware/software solutions, encryption, authentication |
| [39] | DDoS Protection (V2X in MEC) | Zero-Second-Microservice (ZSM) architecture for dynamic security slice management against DDoS attacks. | DDoS attacks in V2X communication within MEC networks | Zero-Second-Microservice (ZSM), Multi-access Edge Computing (MEC) |

communication channels between CAVs and the central server are also crucial [49].

[50] proposed a CNNFL-based APT detection system, leveraging FL for distributed alert analysis. Centralized training of several models, including an ensemble, peaked at 88.15% accuracy, while the distributed CNNFL model achieved 90.18% accuracy with a low false alarm rate.

Despite these challenges, FL holds immense potential for enhancing CAV security while preserving privacy. The effectiveness of FL for anomaly detection can also be limited by data quality and quantity [51]. Furthermore, integrating FL with deep learning for communication and anomaly detection might introduce real-time implementation challenges due to computational demands.

The rapid evolution of federated learning presents exciting possibilities for CAV security, complementing advances in machine learning and communication security. A critical gap in research on detecting advanced persistent threats (APTs) targeting vehicles emerges from the literature review.

This gap compels the proposal of a novel framework that leverages a federated deep learning model with a privacy-preserving mechanism. This approach aims to bridge the identified gap by facilitating efficient APT detection on CAVs, ultimately contributing valuable knowledge to CAV security.

The summary of literature in Table 2. The following section delves into the methodology of the proposed framework. The system architecture, the federated learning algorithm, and the privacy-preserving mechanisms will be detailed. This methodology aims to bridge the gap in APT detection and pave the way for a new chapter in CAV security.

## IV. METHODOLOGY

### A. OVERVIEW

The exponential growth from standalone systems to interconnected vehicles in the ground transportation system necessitates robust security measures to counter sophisticated attacks targeting these platforms. This proposed explainable framework employs a privacy-preserving mechanism to shield the data generated by in-vehicle IoT devices from unauthorized access and utilizes techniques to defend against advanced cyber attacks like APTs. The following subsections delve into the technical details of the proposed framework's architecture, implementation, and functionalities.

### B. DATASET DESCRIPTION AND MAPPING OF APT STAGES

The inherent secrecy surrounding Advanced Persistent Threat (APT) attacks presents a significant challenge in acquiring realistic datasets for research purposes. Victim organizations are understandably reluctant to disclose their logs due to the potential exposure of underlying IT infrastructure vulnerabilities. This underscores the critical need for developing

and public availability of large, comprehensive, and attack-specific datasets within the cybersecurity research community.

Given the limited availability of real-world APT attack datasets in the vehicular field, this study adopts a semi-synthetic approach. Three publicly available Internet of Things (IoT)-related datasets, "CSE-CIC-IDS2018" and "UNSW NB-15", along with the "Edge-IIoTset" datasets, are leveraged. These datasets are chosen to approximate real-world scenarios by representing network traffic commonly encountered in vehicular IoT environments. A feature mapping process bridges the gap between generic network activity and APT attack stages. This process involves mapping relevant network traffic features within the chosen datasets to the various stages of the APT lifecycle model. This mapping aims to create a more realistic and APT-relevant training environment for subsequent experimentation.

Although the datasets are ultimately utilized in CSV format, their transformation from raw network data (where applicable) to the final feature set involves several crucial steps.

**Specific data pipeline is detailed as follows:**

*Raw Data Acquisition:* For datasets like UNSW-NB15 and CSE-CIC-IDS2018, which originate from raw PCAP files, the initial step involves converting these packet captures into network flow records and extracting higher-level features.

Tools Utilized - CICFlowMeter: For the CSE-CIC-IDS2018 dataset, the publicly available CICFlowMeter tool was used to process the raw PCAP files. This tool extracts over 80 network flow features (e.g., flow duration, bytes per second, packet length statistics) by analyzing TCP, UDP, and ICMP traffic. For the UNSW-NB15 dataset, where raw PCAPs were a source, tools conceptually similar to Argus or Zeek (used by the original dataset creators) were applied, and flow data were obtained in CSV.

*Custom Python Scripts:* For the Edge-IIoTset dataset, which is typically distributed in feature-extracted CSV format, and for any additional feature engineering on UNSW-NB15 or CSE-CIC-IDS2018 features, custom Python scripts were developed. These scripts handled parsing, aggregation, and the calculation of derived metrics relevant to APT detection.

*Feature Engineering and Selection:* The features initially extracted from raw traffic (e.g., by CICFlowMeter) are often numerous. This stage involves cleaning. Handling missing values, encoding categorical features (e.g., one-hot encoding for protocol types), and normalizing numerical features to ensure consistency.

*Derivation of APT-Specific Features:* as certain features critical for APT detection are not always directly present in the raw output of generic flow tools, we derived them. For instance, traffic Patterns are calculated statistics over time stamp features. While direct payload content was not used due to privacy, features like payload size variance, non-standard port usage, or specific flag combinations were derived to indicate potential malicious activity. Anomaly deviation from expected behaviour, such as changes in inter-arrival times for IIoT sensor data, was computed based on a statistical analysis of baseline data.

*Labelling and Final CSV Generation:* The processed features were then aligned with their corresponding attack labels. Multi-stage APT detection involves mapping individual flow and event labels from the original datasets to the broader APT stages, which are tabulated as in the Table 3, 4 and 5.

**TABLE 3. Mapping of CSE-CIC-IDS2018 Dataset Contents to APT Stages**

| APT Stage | Dataset Features |
|---|---|
| Reconnaissance | Network traffic logs, packet captures, metadata showing scanning, probing activities, and reconnaissance attempts. |
| Initial Compromise | Records of attempted unauthorized access, exploitation of vulnerabilities, phishing attempts, or suspicious connection requests. |
| Establishing Foothold | Indicators of persistent presence such as unusual user account creations, backdoor installations, or unexpected system changes. |
| Privilege Escalation | Logs showing attempts to gain elevated access or brute-force attacks on privileged accounts. |
| Lateral Movement | Anomalies indicating movement across hosts, unusual data access, or atypical communication patterns. |
| Data Exfiltration | Anomalies suggesting large data transfers, unusual file access, or traffic indicating the movement of sensitive data. |
| Covering Tracks | Evidence of log manipulation, timestamp alterations, or attempts to erase traces to evade detection. |

**TABLE 4. Mapping of UNSW-NB15 Dataset Contents to APT Stages**

| APT Stage | Dataset Features |
|---|---|
| Reconnaissance | Network flow statistics, Payload properties and Traffic patterns |
| Initial Compromise | Unauthorized access records, Connection request anomalies, Exploitation attempt indicators |
| Establishing Foothold | Persistent presence indicators, Foothold establishment anomalies, Backdoor installation evidence |
| Privilege Escalation | Privilege elevation attempts, Anomalies related to account access, Brute-force attack indicators |
| Lateral Movement | Host movement anomalies, Unusual data access, Traffic patterns across segments |
| Data Exfiltration | Large data transfer indicators, Unusual file access, Outbound traffic anomalies |
| Covering Tracks | Not extensively covered in network traffic data |

**TABLE 5. Mapping of Edge-IIoTset Dataset Contents to APT Stages**

| APT Stage | Dataset Features |
|---|---|
| Reconnaissance | Network Traffic Volume (Sudden Increase), Network Traffic Frequency (Unusual Spikes/Patterns), Network Traffic Destination - Suspicious IPs or Unusual Locations. |
| Initial Compromise | Sensor Data Anomalies (Deviations from Expected Readings - Temperature, Pressure, etc.), Time-based Features (Changes in Inter-arrival Times, Missing/Delayed Readings). |
| Establishing Foothold | System Resource Usage (CPU Spikes, Memory Increase, Storage Usage Growth), File System Changes (New Files, Modifications to Critical System Files). |
| Data Exfiltration | Network Traffic Analysis (Patterns - Persistent Connections, Unusual Destinations), Frequency-domain Features (Identifying Hidden Communication Patterns). |
| Covering Tracks | Sensor Data Manipulation (Significant Deviations Beyond Normal Ranges). |

#### 1) CSE-CIC-IDS2018 DATASET

The CSE-CIC-IDS2018 dataset [52] acts as a cornerstone in cybersecurity research, specifically for developing and evaluating intrusion detection systems (IDS) and ML models. Created by the Canadian Institute for Cybersecurity (CIC) within their network environment, this extensive dataset provides network flow features and raw packet captures collected over ten days. It meticulously records benign traffic and diverse attack scenarios, mirroring the intricate complexities of modern network ecosystems by incorporating traffic from various applications and services. The dataset's original storage includes raw PCAP files and corresponding CSV files, where comprehensive features were extracted using their proprietary CICFlowMeter tool. This dataset, alongside others like UNSW-NB15 and Edge-IIoTset, is characterized by its labelled network traffic, real-world captures, and rich feature sets, making it invaluable for diverse experimental analyses. In the context of our study, a key focus is on minimizing these extensive feature sets to enable more efficient diagnosis of Advanced Persistent Threats (APTs).

#### 2) UNSW NB-15 DATASET

UNSW-NB15 [53], [54], [55], [56], [57] is a pivotal and comprehensive network security and intrusion detection research resource. It was meticulously generated by the UNSW Cyber Centre utilizing the IXIA PerfectStorm tool in the UNSW-NB15 Lab, simulating a realistic network environment encompassing contemporary normal activities and various modern attack behaviours over an extended period. Primarily comprising network traffic features, this extensive collection provides a holistic view of network dynamics and potential threats. The dataset's original form includes raw network packet captures (PCAP files), from which flow-level features were subsequently derived and stored in CSV format using specialized tools such as Argus and Bro-IDS (now Zeek). UNSW-NB15 meticulously emulates various network conditions, communication protocols, and attack vectors across millions of network connections, offering multiple features per connection.

#### 3) EDGE-IIOTSET DATASET

The Edge-IIoTset [58] serves as a benchmark dataset, for advancing intrusion and other attack detection systems within Edge and Industrial IoT (IIoT) domains. This dataset is designed explicitly for edge-of-network IIoT security, encompassing network traffic features and sensor readings for IIoT devices. It meticulously emulates authentic scenarios by integrating sensor data—including parameters such as temperature, pressure, and specialized readings—alongside pertinent network traffic. The dataset encapsulates unaltered sensor readings and deliberately tampered data streams, simulating cyber assaults such as data injection or denial-of-service attempts directed at IoT devices across various systems (e.g., Modbus, MQTT, DNS attacks). Conceptually originating from raw network traffic and device telemetry, it is typically provided as CSV files with pre-extracted features. This emphasis on real-world sensor data and its manipulation during cyber offensives renders Edge-IIoTset an invaluable asset, facilitating a deeper comprehension of the cyber domain and empowering the creation of real-time Advanced Persistent Threat (APT) detection frameworks, particularly relevant for IoT-enabled systems like Connected Autonomous Vehicles (CAVs) leveraging federated deep learning models trained to discern malicious deviations by analyzing extracted features like sensor readings, statistical attributes, and temporal characteristics.

### C. SYSTEM ARCHITECTURE

The main goal of this proposed architecture, established on the federated idea of distributed entities, is to minimize data transactions and protect the privacy of the components and the associated data.

The framework assumes an honest but curious central server, colluding clients as potential adversaries. The central server is supposed to follow the protocol honestly but may attempt to infer private information from the aggregated model updates. Other clients are also considered potential threats if they try to reconstruct individual data contributions from shared model parameters. The experiment is continued here, where the IoT devices in the vehicles are grouped and treated as separate nodes. Compared to traditional DNN architecture, here the weights are transmitted between server and clients as in the Fig. 16. Where the local group/client will run the model and update the global model.

**Algorithm 1:** Server-Side Computation in the Proposed Framework With DP.

**Initialization:**
Initialize global model parameters $\theta$ on the server representing the IoT device behavior.
Privacy budget $\varepsilon$ for the training process.
**Differential privacy induced Federated Learning process at central orchestration:**
**for** $r = 1$ to $R_{nd}$, where $R_{nd}$ is the total number of rounds **do**
  Initialize $\theta_{\text{aggregated}}$ as zero.
  **for** $t = 1$ to $T_l$, where $T_l$ is the total number of epochs **do**
    Receive encrypted gradients $\Delta\theta_i$ from $N$ clients.
    Decrypt received gradients: $\Delta\theta_{\text{decrypted}} = \text{Decrypt}(\Delta\theta_i)$.
    Apply DP mechanism (here, Laplace noise injection): $\Delta\theta_{\text{noisy}} = \Delta\theta_{\text{decrypted}} + \text{LaplaceNoise}(\varepsilon/(2NR_t))$, where $NR_t$ is the total number of gradients received so far (accounting for all clients across all rounds).
    Aggregate noisy gradients: $\theta_{\text{aggregated}} = \theta_{\text{aggregated}} + \Delta\theta_{\text{noisy}}$.
  **end for**
  Compute the average of aggregated gradients: $\Delta\theta_{\text{avg}} = \frac{1}{N}\theta_{\text{aggregated}}$.
  Update global model: $\theta \leftarrow \theta - \eta \cdot \Delta\theta_{\text{avg}}$, where $\eta$ is the learning rate.
**end for**

In this scenario, each node trains the designed DNN model, and the only information communicated with the central orchestration to update the global model is the weights of the DNN model.

The received DNN weights are then subjected to the computation and re-distributed to the nodes to update the local model. This architecture was designed and represented in Fig. 4. The architecture shows the pooling of weight information from different nodes without the data movement between the devices and hence provides an extra layer of privacy; the data generated at the node level are cleaned, processed, DP is applied as in

$$\Pr[M(D_1) \in S] \leq \exp(\epsilon) \cdot \Pr[M(D_2) \in S] + \delta \qquad (1)$$

where:

- $M$ is a randomized mechanism.
- $D_1$ and $D_2$ are neighboring datapoints differing by at most one element.
- $S$ is any possible set of outputs.
- $\epsilon$ is the privacy parameter.
- $\delta$ is the failure probability.

**Algorithm 2:** Client-Side Computation in the Proposed Framework With DP.

**Initialization:**
Each IoT device (client) initializes with local data $D_i$ and a local model $\theta_i$ based on $\theta$.
**Training Iterations:**
**for** $t = 1$ to $T_l$, where $T_l$ is the total number of epochs **do**
  Receive global model $\theta$ from the server.
  Compute local gradients: $\Delta\theta_i = \nabla_\theta \mathcal{L}_i(\theta, D_i)$, where $\mathcal{L}_i(\theta, D_i)$ is the local loss function for the $i$th IoT device.
  Apply Laplace noise with DP: $\Delta\theta_{i,noisy} = \Delta\theta_i + \text{LaplaceNoise}(\sigma)$, where $\sigma$ is the noise scale.
  Encrypt gradients: encrypted_gradients = $\text{Encrypt}(\Delta\theta_{i,noisy})$.
  Transmit encrypted gradients to the server.
**end for**

and data is fed into the local DNN model at the node level, and only the weight parameters are collected and processed by the aggregator. The general workflow as in the Fig. 5.

### D. FRAMEWORK AT NODE LEVEL

This innovative framework safeguards data privacy within IoT (Internet of Things) systems through a multi-tiered approach while concurrently identifying potential malicious activities targeting these systems. The solution integrates numerous mechanisms and harnesses a fusion of practical techniques to deliver a robust security solution, all while upholding the integrity of sensitive data privacy. PF-DAPTIV framework integrates diverse methods for feature selection and reduces data processing, tailoring it to various data sources. Additionally, it is validated using the existing datasets "CSE-CIC-IDS2018", Edge-IIoTset and "UNSW-NB15". The developed framework at the node level, as in Fig. 6.

### E. FEDERATED SET-UP

This experiment uses "Horizontal Federated Learning (HFL)", an ML paradigm where multiple "Independent and Identically Distributed (IID)" datasets are held across decentralized devices. In this context, each device, maintaining ownership of its local data, collaborates in training a shared model without sharing raw data. Here, the calculated noise is added to the data at the node level to preserve privacy. The process involves initializing a global model and individual local model training on respective datasets. After local training, model parameters are aggregated at a central server, generating an updated global model. This updated model is then redistributed to participating devices, iterating the process; the following are achieved.

- *Privacy is preserved:* by the fact that data stays on the devices and differential privacy is used at the node level.

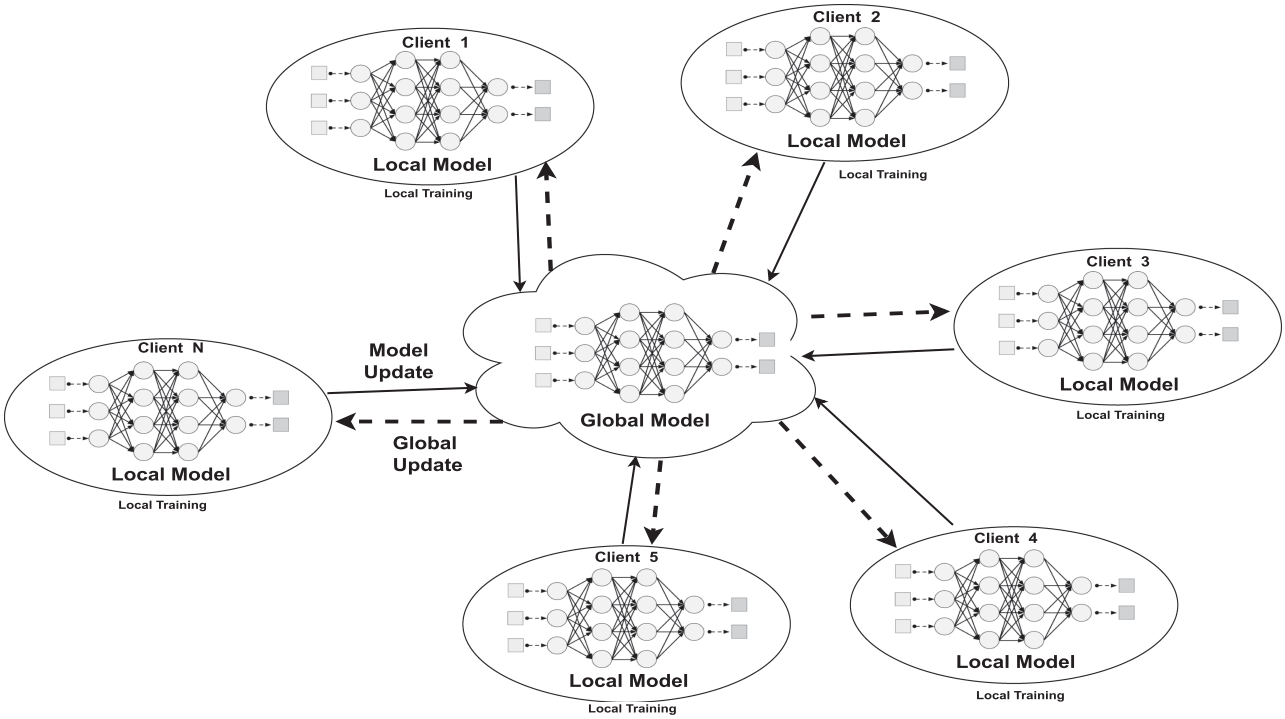


**FIGURE 4.** Architectural Diagram of proposed PF-DAPTIV Framework with a centrally orchestrated global model and aggregator, communicating with N nodes (here in this experiment N=6). The developed DL model is distributed to six client nodes to get trained on locally available IoT datasets and communicates only the weights with the aggregator to form the final model to detect APT attacks.

- *Data Diversity:* The model's training process is enhanced by the possibility that each device's local dataset represents distinct facets of the global dataset.
- *Reduced Communication Overhead:* Communication expenses are drastically reduced when model updates are sent rather than raw data.

### F. EXPERIMENTAL SETUP

The experimental configuration for this proposed study involves developing, training and evaluating a privacy-preserved federated DL model. The research concentrates on utilizing privacy-preserving, Federated Learning and DL methodologies to detect APTs in a distributed manner while alleviating the computational load in centralized systems. Terms and symbols used in the algorithms at client and Server-side are as in the Table 6. The entire experiment is conducted on a cloud infrastructure, the specifications of which are detailed in Table 7.

All three datasets are further categorized into six groups each, and these groups are considered as nodes in each system. After the feature set understanding, at every node, only the selected components are used to train the developed model locally, and the weights are shared to update the global model. These updates are performed based on the loss calculation, and the communication rounds between the local and global models are determined. Evaluation metrics such as Accuracy, precision, Recall, F1 score, etc., are calculated and tabulated.

Accuracy denotes the proportion of correctly classified instances compared to the total count within the test set. Precision represents the proportion of accurately classified instances of the intended outcome relative to the total number of cases classified as that specific outcome. Recall is the ratio of correctly classified instances of the intended outcome to the overall count of target records. The F1-score, employing the harmonic mean instead of the arithmetic mean, concurrently assesses precision and recall, providing a consolidated measure of both metrics.

- *Sensitivity (True Positive Rate, Recall):*

$$\text{Sensitivity} = \frac{\mathcal{TP}}{\mathcal{TP} + \mathcal{FN}} \tag{2}$$

- *Specificity (True Negative Rate):*

$$\text{Specificity} = \frac{\mathcal{TN}}{\mathcal{TN} + \mathcal{FP}} \tag{3}$$

- *Precision (Positive Predictive Value):*

$$\text{Precision} = \frac{\mathcal{TP}}{\mathcal{TP} + \mathcal{FP}} \tag{4}$$

- *Negative Predictive Value (NPV):*

$$\text{NPV} = \frac{\mathcal{TN}}{\mathcal{TN} + \mathcal{FN}} \tag{5}$$

- *False Positive Rate (FPR):*

$$\text{FPR} = \frac{\mathcal{FP}}{\mathcal{FP} + \mathcal{TN}} \tag{6}$$

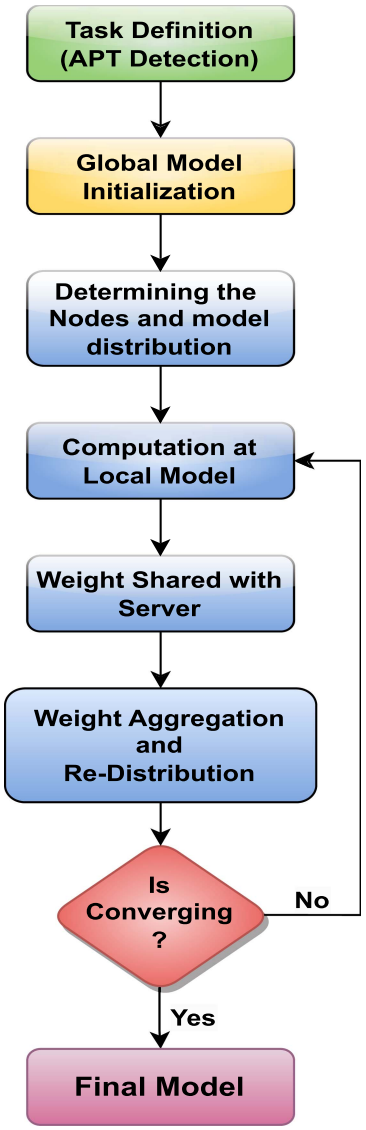


**FIGURE 5.** The operational workflow of the proposed PF-DAPTIV framework's concise overview of its functioning methods in detecting APTs across federated environments.

- *False Discovery Rate (FDR):*

$$\text{FDR} = \frac{\mathcal{FP}}{\mathcal{FP} + \mathcal{TP}} \tag{7}$$

- *False Negative Rate (FNR, Miss Rate):*

$$\text{FNR} = \frac{\mathcal{FN}}{\mathcal{FN} + \mathcal{TP}} \tag{8}$$

- *Accuracy:*

$$\text{Accuracy} = \frac{\mathcal{TP} + \mathcal{TN}}{\text{Total Population}} \tag{9}$$

- *F1 Score:*

$$\text{F1 Score} = 2 \times \frac{\mathcal{P}\text{recision} \times \mathcal{R}\text{ecall}}{\mathcal{P}\text{recision} + \mathcal{R}\text{ecall}} \tag{10}$$

**TABLE 6.** Terms and Symbols Used in the Algorithms at Client and Server-Side

| Symbol | Description |
|---|---|
| $D_i$ | Local data set on client $i$ |
| $\theta$ | Global model parameters representing IoT device behavior |
| $\theta_i$ | Local model on client $i$ (copy of global model) |
| $T_l$ | Total number of epochs (training iterations) within each round |
| $R_{nd}$ | Total number of rounds in federated learning process |
| $N$ | Total number of participating clients in the system |
| $\eta$ | Learning rate for model update |
| $\varepsilon$ | Privacy budget for the entire training process |
| $\Delta\theta_i$ | Encrypted gradients received from client $i$ |
| $\Delta\theta_{\text{decrypted}}$ | Decrypted gradients received from client |
| $\Delta\theta_{\text{noisy}}$ | Noisy gradients after applying differential privacy |
| $\theta_{\text{aggregated}}$ | Accumulated noisy gradients from all clients within a round |
| $\Delta\theta_{\text{avg}}$ | Average of aggregated noisy gradients |
| $\sigma$ | Noise scale for Laplace noise (controls privacy) used on client-side (not shown in server-side algorithm) |

**TABLE 7.** Details of Cloud-GPU and Experimental Set-up

| Attribute | Details |
|---|---|
| CPU Cores | 32 |
| Processor | AMD EPYC |
| Graphics Card | Nvidia A100 |
| GPU Memory | 40GB |
| Operating System | Ubuntu 22.0 |
| Programming Language | Python3 |
| Libraries Used | Pandas, NumPy, Tensorflow 2.0 |
| Learning Techniques | Deep Learning, Federated Learning |

- *Matthews Correlation Coefficient (MCC):*

$$\frac{\mathcal{TP} \times \mathcal{TN} - \mathcal{FP} \times \mathcal{FN}}{\sqrt{(\mathcal{TP} + \mathcal{FP})(\mathcal{TP} + \mathcal{FN})(\mathcal{TN} + \mathcal{FP})(\mathcal{TN} + \mathcal{FN})}} \tag{11}$$

where: $\mathcal{TP} = \mathcal{T}\text{rue } \mathcal{P}\text{ositives}$, $\mathcal{TN} = \mathcal{T}\text{rue } \mathcal{N}\text{egatives}$, $\mathcal{FP} = \mathcal{F}\text{alse } \mathcal{P}\text{ositives}$, $\mathcal{FN} = \mathcal{F}\text{alse } \mathcal{N}\text{egatives}$.

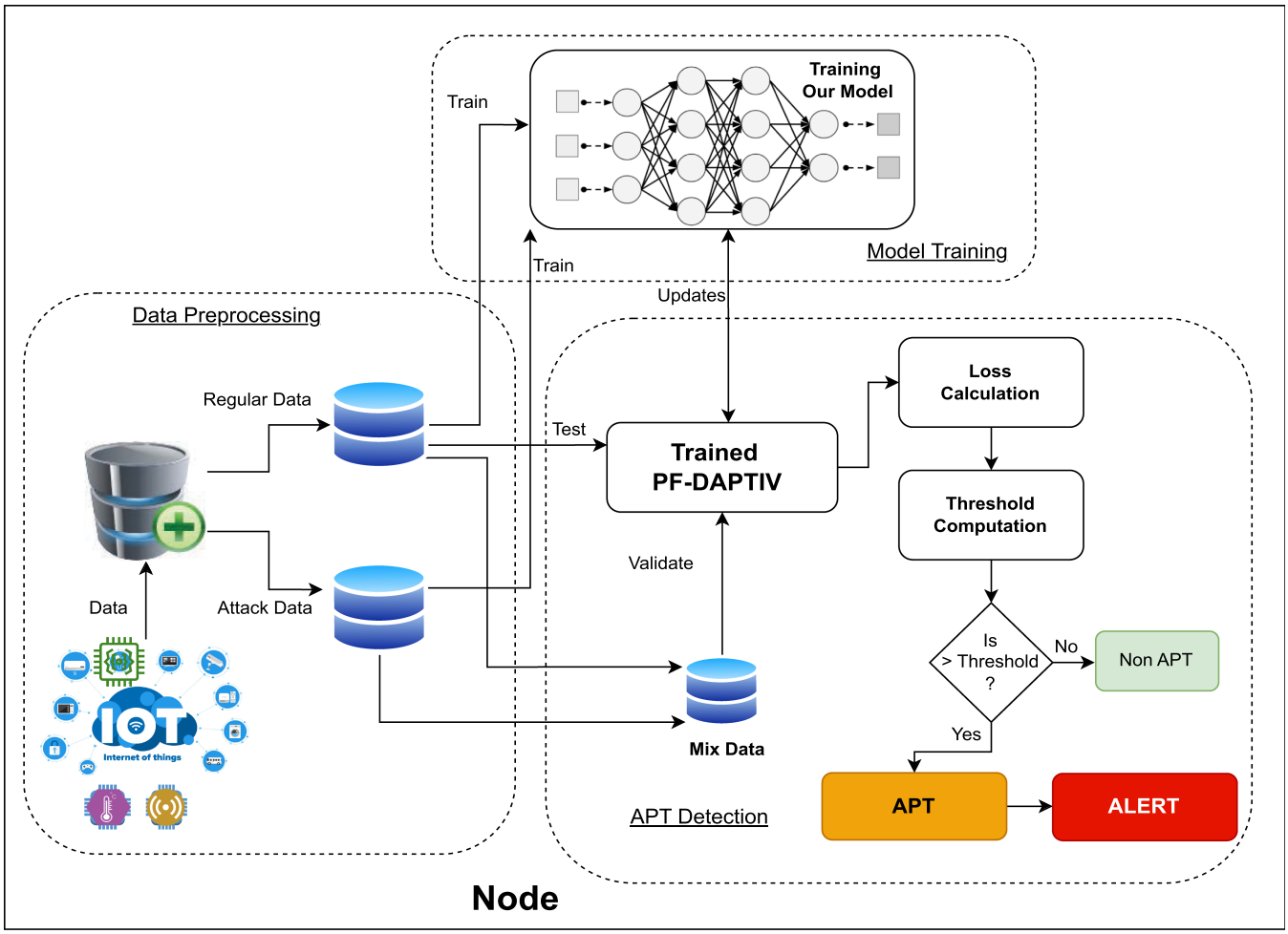


**FIGURE 6.** The intricate architectural diagram of the proposed PF-DAPTIV framework at the node level.

## V. ANALYSIS OF PF-DAPTIV FRAMEWORK

Framework Analysis stands as a cornerstone in assessing the viability and functionality of frameworks across various domains. This analytical process holds substantial importance, offering critical insights into a framework's performance, scalability, and adaptability within real-world scenarios. A crucial aspect of evaluating frameworks involves conducting ablation studies on datasets to meticulously analyze their impact on the framework's performance. Ablation studies systematically modify or remove specific components or features within a dataset to gauge their influence on the framework's functionality. By selectively altering dataset elements, such as removing certain parts or instances, to assess the framework's robustness and sensitivity to various data conditions. Insights garnered from these studies aid in pinpointing critical components that significantly contribute to the framework's efficacy while identifying redundant or less influential elements. Such insights provide invaluable guidance for refining the framework, optimizing performance, and enhancing adaptability to diverse data environments.

Tools like SHAP enrich the evaluation process, offering interpretability and understanding of feature importance. Here, SHAP aids in explaining the developed model outputs by attributing predictions to input features. Within ablation studies, SHAP facilitates a deeper understanding of the impact of individual elements or components removed from the dataset on the framework's performance. It enables identifying influential features, their contributions to model predictions, and the consequences of their absence on the overall model behaviour.

Integrating SHAP in the ablation process enhances result interpretability and makes informed decisions about feature selection, model refinement, and framework enhancement. SHAP is used on the UNSW-NB15, Edge-IIoTset and CI-CIDS2018 Datasets to understand the features contributing more to the detection process.

### A. PRIVACY ANALYSIS

The core privacy mechanism integrated into this framework is the strategic injection of Laplace noise into the gradients, providing robust privacy guarantees for sensitive vehicular IoT data.

#### 1) CLIENT-SIDE PRIVACY (ALGORITHM 2)

Each IoT device, acting as a client, fortifies its local privacy by adding carefully calibrated Laplace noise to its locally

computed gradients ($\Delta\theta_i$) prior to both encryption and transmission to the central server. This crucial step ensures local differential privacy (LDP). LDP offers strong privacy guarantees, making it computationally infeasible for the server or any potential eavesdropper to infer an individual client's raw data from its transmitted gradients, even in scenarios where the entire local model update might theoretically be compromised. The noise scale, denoted by $\sigma$, effectively masks the precise contribution of any single data record within the local dataset ($D_i$). This mechanism is paramount for protecting highly sensitive IoT data, such as real-time sensor readings or unique vehicular behavioral patterns.

#### 2) NOISE SCALE ($\sigma$) FOR CLIENT-SIDE LAPLACE MECHANISM

For the client-side Laplace mechanism, the noise scale $\sigma$ is critically determined by the sensitivity of the gradients and the allocated privacy budget $\epsilon$. In general, for a function $f$ with an $L_1$-sensitivity $\Delta f$, the noise scale is often set as $\sigma = \Delta f/\epsilon$. In the context of gradient-based methods, particularly when dealing with the $L_2$-norm of gradients (e.g., after clipping), the choice of $\sigma$ must align with the clipping threshold ($C$). If the gradients are clipped to an $L_2$-norm of $C$, their $L_2$-sensitivity is $C$. While the Laplace mechanism conventionally uses $L_1$-sensitivity, a conversion from $L_2$-sensitivity or the direct application of a Gaussian mechanism might be considered for gradient perturbation. The $\sigma$ in LaplaceNoise($\sigma$) reflects the maximum possible change an individual data point can induce in a single dimension of the gradient component, ensuring the required level of privacy.

#### 3) SERVER-SIDE AGGREGATION (ALGORITHM 1)

Beyond client-side privacy, the server further enhances data protection by applying additional Laplace noise during the aggregation of the decrypted gradients. This operation contributes to global differential privacy (GDP). While LDP at the client level safeguards individual contributions, the server-side noise provides an extra layer of obscurity to the aggregated information. This makes it substantially more challenging for an adversary to reconstruct characteristics of the collective dataset or to infer information about individuals from the overall model updates.

#### 4) NOISE SCALE FOR SERVER-SIDE LAPLACE MECHANISM

The noise scale for the server-side aggregation, as detailed in Algorithm 1 (Line 11), is dynamically determined by the expression $\epsilon/(2NR_t)$. Here, $\epsilon$ represents the total global privacy budget allocated for the entire federated learning process, and $NR_t$ is a cumulative term reflecting the total number of gradients received across all participating clients and training rounds up to time $t$ . The factor of 2in the denominator is often incorporated in certain DP implementations to manage the distribution of the privacy budget across different privacy-preserving layers or to account for the cumulative privacy loss in multi-round learning settings. This dynamic adjustment of the noise scale ensures that the overall $\epsilon$ budget is systematically consumed throughout the federated learning process, rather than being expended prematurely.

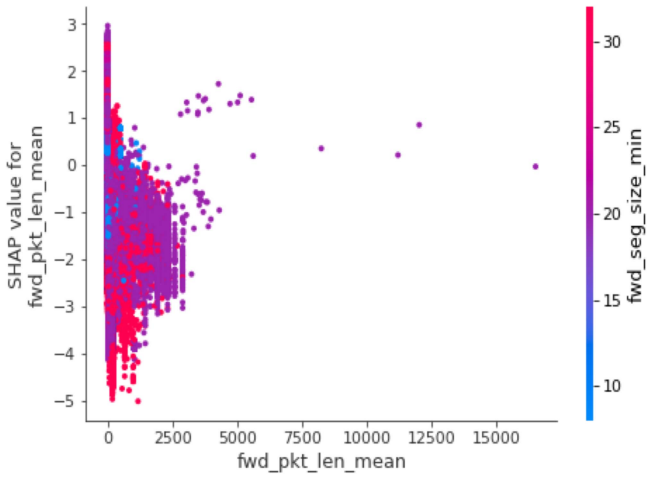


**FIGURE 7.** SHAP value of selected feature fwd_pkt_len_mean in CICIDS2018.

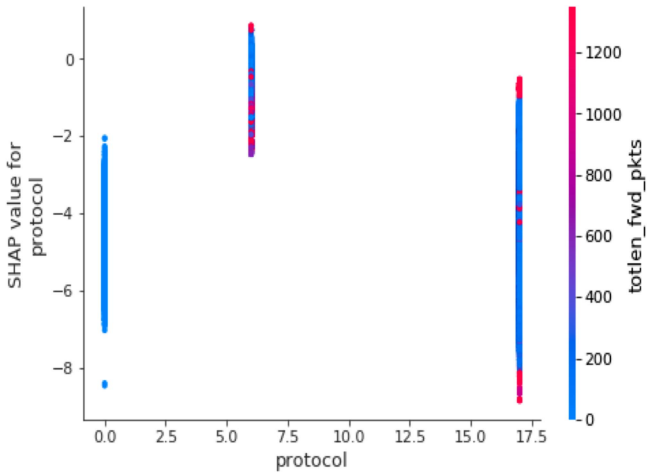


**FIGURE 8.** SHAP value of selected feature protocol in CICIDS2018.

### *B. ABLATION STUDY BASED ON CSE-CIC-IDS2018, UNSW NB-15 AND EDGE-IIOTSET DATASETS ON PF-DAPTIV FRAMEWORK*

The study employed a feature ablation technique on a deep learning model trained on three network traffic datasets: CI-CIDS2018, UNSW-NB15, and Edge-IIoTset. This involved iteratively removing individual features or feature groups to assess their contribution to the model's performance in identifying abnormal activity within the PF-DAPTIV framework, as depicted in Figs. 7 to 17. The analysis aimed to quantify the importance of each feature for the model's overall classification accuracy and identify potential redundancies or irrelevant features within the datasets.

The waterfall model in Fig. 17 comprehensively depicts vital insights regarding the developed DL model's predictions. Commencing with the Base Value, serving as the starting

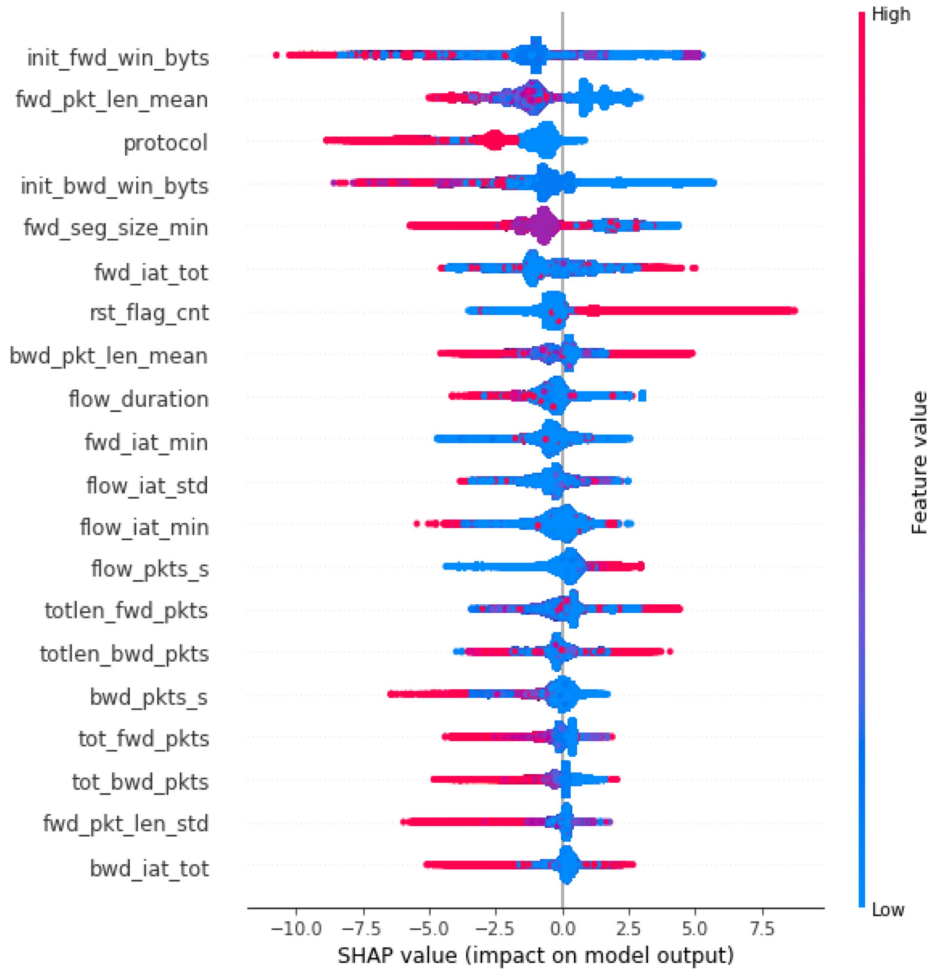


**FIGURE 9.** High impact features which influence model output in CICIDS2018 dataset.

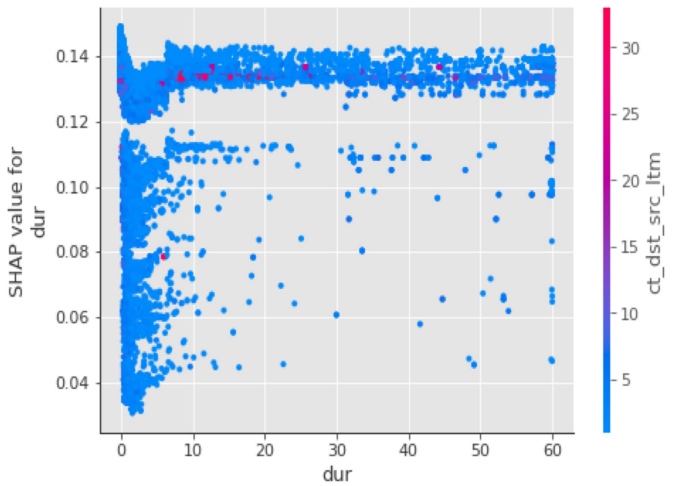


**FIGURE 10.** Distribution of dur and ct_dst_src_ltm feature in UNSW-NB15 dataset.

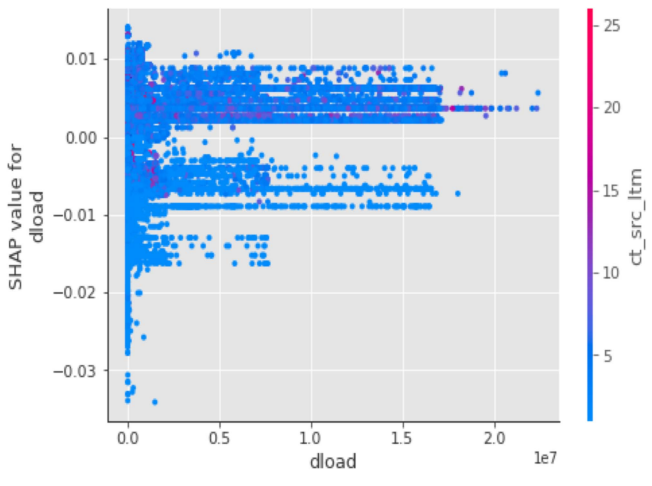


**FIGURE 11.** Distribution density of dload and ct_src_ltm feature in UNSW-NB15 dataset.

point, it signifies the model's baseline prediction or the average forecast across all instances. Contributions from each feature are visually represented as stacked bars within the plot, with positive contributions elevating predictions and negative ones reducing them. The length of each bar delineates the Feature Impact, thereby illustrating the magnitude of influence that each feature holds in altering the prediction; longer bars signify features wielding more significant influence in

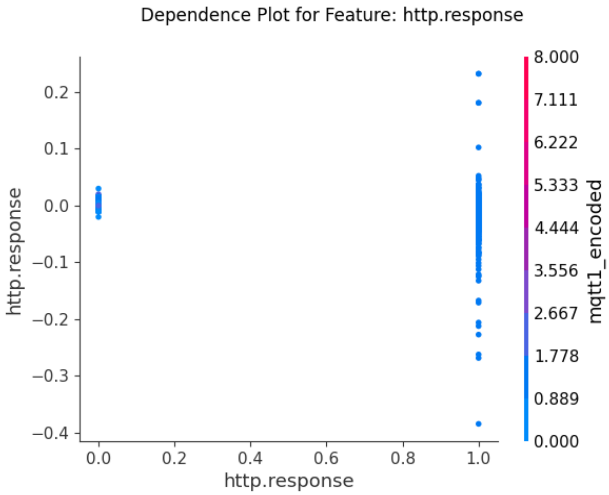


**FIGURE 12.** SHAP value of selected feature http.response having less contribution towards model output in Edge-IIoTset dataset.

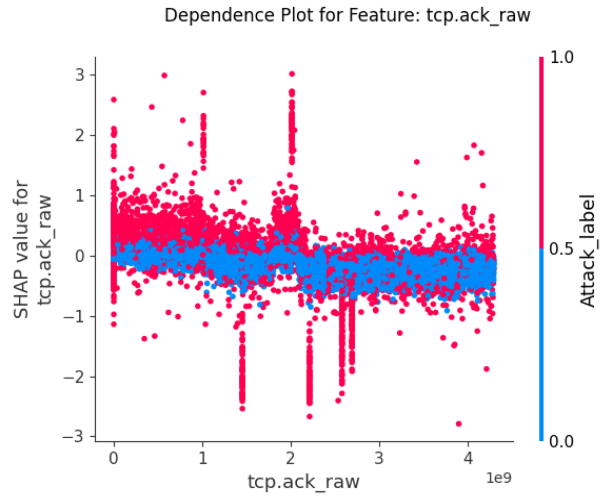


**FIGURE 13.** Distribution density of major contributers tcp.ack_raw and Attack types features in Edge-IIoTset dataset.

deviating predictions from the baseline. Ultimately, the plot culminates in the Total Prediction, denoting the comprehensive prediction derived from the combined contributions of all features, encapsulating the model's collective decision-making process.

The value f(x)=0.757 signifies the specific prediction generated by the DL model for a distinct instance x. This indicates that, for the given set of input features or data point x, the machine learning model predicts an output of 0.75. On the other hand, E[f(x)]=0.111 represents the expected value of the model's output (f(x)) calculated across a distribution or range of instances. This expected value E[f(x)] embodies the anticipated prediction made by the model when considering various instances or scenarios, resulting in an average prediction of 0.111 across these diverse scenarios. These numerical values offer insights into both the specific prediction made by the model for a particular instance (f(x)=0.75) and the average prediction derived from multiple instances (E[f(x)]=0.111), providing an understanding of both specific and collective predictive behaviours of the model.

## VI. RESULTS AND DISCUSSION

The Loss Vs. Rounds for the different Batch sizes between the Global and Client models have been analysed to determine the efficient number of communication rounds required. It is observed that at 100 rounds, the Loss is minimal, as in the Fig. 18.

The performance evaluation of the PF-DAPTIV framework involved three publicly available datasets: CSE-CIC-IDS2018, Edge-IIoTset, and UNSW-NB15. Each dataset was evaluated under two conditions: one where the model was trained directly on the raw data, without privacy preservation technique and another where the PF-DAPTIV framework's privacy-preserving techniques were employed during training. The results, documented in separate tables as in Table 8, 10, and 12 for without privacy preserving techniques and Table 9, 11, and 13 in PF-DAPTIV with privacy preserving technique. Tables revealed the expected trade-off between privacy and accuracy.

In all three datasets (CSE-CIC-IDS2018, Edge-IIoTset, and UNSW-NB15), the model trained directly on the raw data achieved slightly higher accuracy in APT detection compared to the model trained with PF-DAPTIV's privacy-preserving techniques. This is understandable as the raw data provides the model with the most complete and unobscured information for learning.

However, the slight decrease in accuracy with PF-DAPTIV is outweighed by the significant benefit of maintaining data privacy at the vehicular level. The framework ensures that the data remains confidential throughout training, eliminating the need to transfer raw data to a central server. This decentralized approach protects the anonymity of participating nodes and safeguards sensitive information. The DP techniques employed by PF-DAPTIV further enhance this vehicle data protection, ensuring that even the aggregated model updates used for training do not reveal sensitive details about individual nodes.

Therefore, while there is a slight accuracy trade-off, the PF-DAPTIV framework offers a valuable solution for balancing APT detection effectiveness with robust data privacy in connected vehicle systems.

The developed PF-DAPTIV framework compared with state-of-the-art and tabulated as in the Table 16.

*Discussion:*

The presented framework, PF-DAPTIV, extends beyond a direct implementation by incorporating several critical advancements warranting detailed discussion within the context of FL for VIoT security.

A notable aspect of this work is the innovative strategy to mitigate the inherent scarcity of dedicated APT datasets within the VIoT domain. Rather than relying on a singular, potentially restrictive dataset, PF-DAPTIV leverages three distinct and comprehensive datasets—UNSW-NB15, CSE-CIC-IDS2018, and Edge-IIoTset—as heterogeneous clients

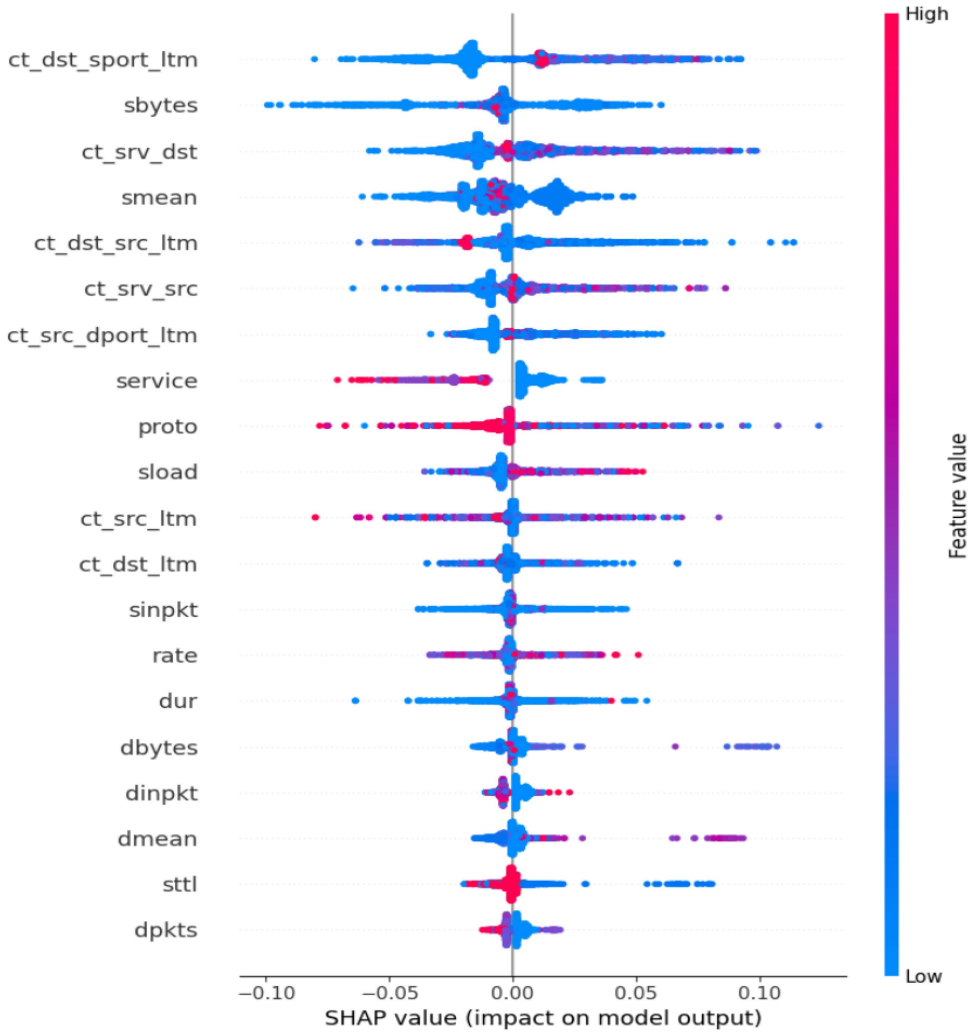


**FIGURE 14.** High impact features which influence model output in UNSW-NB15 Dataset.

**TABLE 8.** Results Obtained From the Developed Model at six Different Nodes Subjected to CICIDS2018 Dataset in Federated Set-up Without Privacy Preserving Mechanism

| Metrics | Node 1 | Node 2 | Node 3 | Node 4 | Node 5 | Node 6 |
|---|---|---|---|---|---|---|
| Sensitivity | 0.8911 | 1.0000 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| Specificity | 0.9969 | 0.9456 | 0.9535 | 0.9752 | 0.9093 | 0.9427 |
| Precision | 0.9999 | 0.9997 | 0.9981 | 0.9984 | 0.9976 | 0.9975 |
| Negative Predictive Value (NPV) | 0.1366 | 1.0000 | 0.9965 | 0.9978 | 0.9978 | 0.9968 |
| False Positive Rate | 0.0031 | 0.0544 | 0.0465 | 0.0248 | 0.0907 | 0.0573 |
| False Discovery Rate | 0.0001 | 0.0003 | 0.0019 | 0.0016 | 0.0024 | 0.0025 |
| False Negative Rate | 0.1089 | 0.0000 | 0.0001 | 0.0001 | 0.0001 | 0.0001 |
| Accuracy | 0.8929 | 0.9997 | 0.9980 | 0.9984 | 0.9976 | 0.9974 |
| F1 Score | 0.9424 | 0.9999 | 0.9990 | 0.9991 | 0.9987 | 0.9987 |
| Matthews Correlation Coefficient (MCC) | 0.3482 | 0.9723 | 0.9738 | 0.9856 | 0.9513 | 0.9681 |

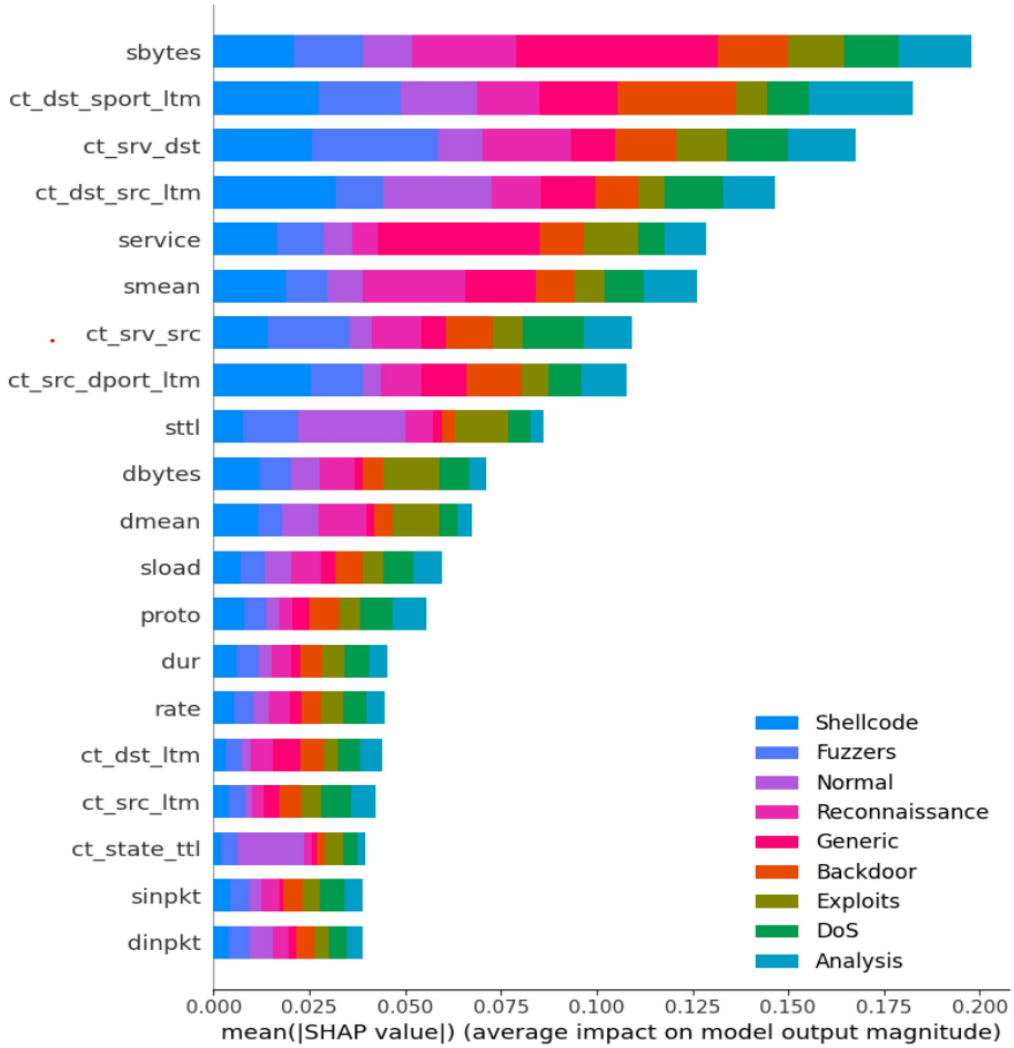


**FIGURE 15.** SHAP Summary plot showing the contribution of features to different types of attack categories present in the UNSW-NB15 dataset.

**TABLE 9.** Results Obtained From the Developed Model at six Different Nodes Subjected to CICIDS2018 Dataset in Federated Set-up With Privacy Preserving Mechanism

| Metrics | Node 1 | Node 2 | Node 3 | Node 4 | Node 5 | Node 6 |
|---|---|---|---|---|---|---|
| Sensitivity | 0.8451 | 0.9725 | 0.9836 | 0.9914 | 0.9972 | 0.9999 |
| Specificity | 0.9919 | 0.7781 | 0.8333 | 0.9425 | 0.7365 | 0.1706 |
| Precision | 0.9999 | 0.9985 | 0.9934 | 0.9963 | 0.9913 | 0.9644 |
| Negative Predictive Value (NPV) | 0.0590 | 0.1543 | 0.6667 | 0.8768 | 0.8966 | 0.9826 |
| False Positive Rate | 0.0081 | 0.2219 | 0.1667 | 0.0575 | 0.2635 | 0.8294 |
| False Discovery Rate | 0.0001 | 0.0015 | 0.0066 | 0.0037 | 0.0087 | 0.0356 |
| False Negative Rate | 0.1549 | 0.0275 | 0.0164 | 0.0086 | 0.0028 | 0.0001 |
| Accuracy | 0.8465 | 0.9712 | 0.9779 | 0.9885 | 0.9888 | 0.9646 |
| F1 Score | 0.9160 | 0.9853 | 0.9885 | 0.9939 | 0.9942 | 0.9818 |
| Matthews Correlation Coefficient (MCC) | 0.2220 | 0.3387 | 0.7343 | 0.9030 | 0.8071 | 0.4018 |

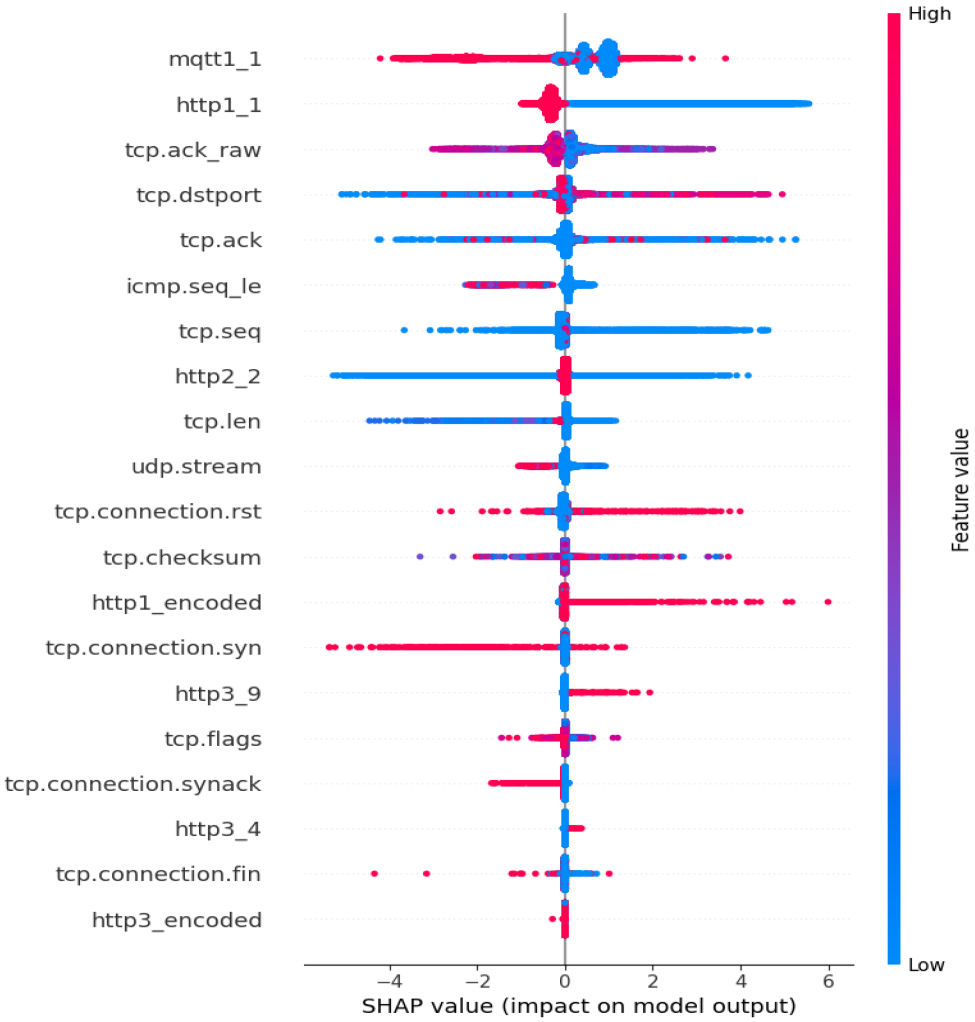


**FIGURE 16.** High impact features which influence model output in the Edge-IIoTset dataset.

**TABLE 10.** Results Obtained From the Developed Model at six Different Nodes Subjected to UNSW-NB15 Dataset in Federated Set-up Without Privacy Preserving Mechanism

| Metrics | Node 1 | Node 2 | Node 3 | Node 4 | Node 5 | Node 6 |
|---|---|---|---|---|---|---|
| Sensitivity | 0.9709 | 0.9697 | 0.9851 | 0.9950 | 0.9825 | 0.9804 |
| Specificity | 0.9899 | 0.9620 | 0.9583 | 0.9574 | 0.9583 | 0.9388 |
| Precision | 0.9901 | 0.9697 | 0.9706 | 0.9901 | 0.9655 | 0.9434 |
| Negative Predictive Value (NPV) | 0.9703 | 0.9620 | 0.9787 | 0.9783 | 0.9787 | 0.9787 |
| False Positive Rate | 0.0101 | 0.0380 | 0.0417 | 0.0426 | 0.0417 | 0.0612 |
| False Discovery Rate | 0.0099 | 0.0303 | 0.0294 | 0.0099 | 0.0345 | 0.0566 |
| False Negative Rate | 0.0291 | 0.0303 | 0.0149 | 0.0050 | 0.0175 | 0.0196 |
| Accuracy | 0.9802 | 0.9663 | 0.9739 | 0.9879 | 0.9714 | 0.9600 |
| F1 Score | 0.9804 | 0.9697 | 0.9778 | 0.9926 | 0.9739 | 0.9615 |
| Matthews Correlation Coefficient (MCC) | 0.9606 | 0.9317 | 0.9464 | 0.9604 | 0.9425 | 0.9206 |

**TABLE 11.** Results Obtained From the Developed Model at six Different Nodes Subjected to UNSW-NB15 Dataset in Federated Set-up With Privacy Preserving Mechanism

| Metrics | Node 1 | Node 2 | Node 3 | Node 4 | Node 5 | Node 6 |
|---|---|---|---|---|---|---|
| Sensitivity | 0.9709 | 0.9697 | 0.9706 | 0.9852 | 0.9655 | 0.9615 |
| Specificity | 0.9703 | 0.9268 | 0.9388 | 0.9375 | 0.9388 | 0.9388 |
| Precision | 0.9709 | 0.9412 | 0.9565 | 0.9852 | 0.9492 | 0.9434 |
| Negative Predictive Value (NPV) | 0.9703 | 0.9620 | 0.9583 | 0.9375 | 0.9583 | 0.9583 |
| False Positive Rate | 0.0297 | 0.0732 | 0.0612 | 0.0625 | 0.0612 | 0.0612 |
| False Discovery Rate | 0.0291 | 0.0588 | 0.0435 | 0.0148 | 0.0508 | 0.0566 |
| False Negative Rate | 0.0291 | 0.0303 | 0.0294 | 0.0148 | 0.0345 | 0.0385 |
| Accuracy | 0.9706 | 0.9503 | 0.9573 | 0.9761 | 0.9533 | 0.9505 |
| F1 Score | 0.9709 | 0.9552 | 0.9635 | 0.9852 | 0.9573 | 0.9524 |
| Matthews Correlation Coefficient (MCC) | 0.9412 | 0.8999 | 0.9121 | 0.9227 | 0.9059 | 0.9010 |

**TABLE 12.** Results Obtained From the Developed Model at six Different Nodes Subjected to Edge-IIoTset Dataset in Federated Set-up Without Privacy Preserving Mechanism

| Metrics | Node 1 | Node 2 | Node 3 | Node 4 | Node 5 | Node 6 |
|---|---|---|---|---|---|---|
| Sensitivity | 0.9902 | 0.9648 | 0.9781 | 0.9926 | 0.9744 | 0.9714 |
| Specificity | 0.9754 | 0.9565 | 0.9495 | 0.9485 | 0.9495 | 0.9307 |
| Precision | 0.9758 | 0.9648 | 0.9640 | 0.9877 | 0.9580 | 0.9358 |
| Negative Predictive Value (NPV) | 0.9900 | 0.9565 | 0.9691 | 0.9684 | 0.9691 | 0.9691 |
| False Positive Rate | 0.0246 | 0.0435 | 0.0505 | 0.0515 | 0.0505 | 0.0693 |
| False Discovery Rate | 0.0242 | 0.0352 | 0.0360 | 0.0123 | 0.0420 | 0.0642 |
| False Negative Rate | 0.0098 | 0.0352 | 0.0219 | 0.0074 | 0.0256 | 0.0286 |
| Accuracy | 0.9828 | 0.9611 | 0.9661 | 0.9841 | 0.9630 | 0.9515 |
| F1 Score | 0.9830 | 0.9648 | 0.9710 | 0.9901 | 0.9661 | 0.9533 |
| Matthews Correlation Coefficient (MCC) | 0.9657 | 0.9213 | 0.9303 | 0.9486 | 0.9255 | 0.9035 |

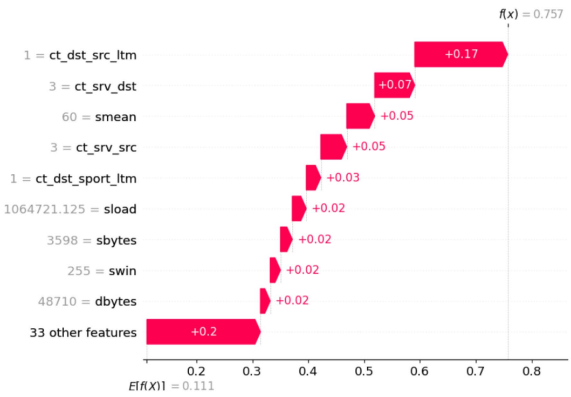


**FIGURE 17.** SHAP waterfall model for the UNSWNB-15 Dataset.

within the FL architecture. A mapping and alignment process was undertaken to correlate specific features from each dataset with distinct stages of an APT attack. This multi-dataset integration empowers PF-DAPTIV to detect a broad spectrum of APT attack patterns in VIoT. This approach fosters the development of a more robust and generalized APT detection.

This approach of cross-domain feature mapping and the utilization for comprehensive APT stage detection explicitly addresses the privacy challenge in VIoT. This shows the significant contribution. The framework rigorously integrates DP at the client side before gradient transmission, addressing a concern in the vehicular domain.

The privacy guarantees rely on the assumption that the encryption mechanism used for transmitting gradients is

**TABLE 13.** Results Obtained From the Developed Model at six Different Nodes Subjected to Edge-IIoTset Dataset in Federated Set-up With Privacy Preserving Mechanism

| Metrics | Node 1 | Node 2 | Node 3 | Node 4 | Node 5 | Node 6 |
|---|---|---|---|---|---|---|
| Sensitivity | 0.9758 | 0.9648 | 0.9640 | 0.9877 | 0.9580 | 0.9533 |
| Specificity | 0.9754 | 0.9506 | 0.9495 | 0.9388 | 0.9495 | 0.9307 |
| Precision | 0.9758 | 0.9600 | 0.9640 | 0.9853 | 0.9580 | 0.9358 |
| Negative Predictive Value (NPV) | 0.9754 | 0.9565 | 0.9495 | 0.9485 | 0.9495 | 0.9495 |
| False Positive Rate | 0.0246 | 0.0494 | 0.0505 | 0.0612 | 0.0505 | 0.0693 |
| False Discovery Rate | 0.0242 | 0.0400 | 0.0360 | 0.0147 | 0.0420 | 0.0642 |
| False Negative Rate | 0.0242 | 0.0352 | 0.0360 | 0.0123 | 0.0420 | 0.0467 |
| Accuracy | 0.9756 | 0.9584 | 0.9580 | 0.9782 | 0.9541 | 0.9423 |
| F1 Score | 0.9758 | 0.9624 | 0.9640 | 0.9865 | 0.9580 | 0.9444 |
| Matthews Correlation Coefficient (MCC) | 0.9512 | 0.9160 | 0.9135 | 0.9301 | 0.9075 | 0.8846 |

**TABLE 14.** Performance Comparison of the Developed Framework Without Integrated Privacy Mechanisms on Multiple Benchmark Datasets

| Dataset | Accuracy | Precision | F1 score | False Positive Rate |
|---|---|---|---|---|
| CSE-CIC-IDS2018 | 98.06 | 99.85 | 98.96 | 0.0461 |
| UNSW-NB15 | 97.32 | 97.15 | 97.59 | 0.0392 |
| Edge-IIoTset | 96.81 | 96.43 | 97.14 | 0.0483 |

**TABLE 15.** Performance Comparison of the Developed Framework With Integrated Privacy Mechanisms on Multiple Benchmark Datasets

| Dataset | Accuracy | Precision | F1 score | False Positive Rate |
|---|---|---|---|---|
| CSE-CIC-IDS2018 | 95.63 | 99.06 | 97.66 | 0.25785 |
| UNSW-NB15 | 95.96 | 95.77 | 96.41 | 0.0581 |
| Edge-IIoTset | 96.11 | 96.31 | 96.52 | 0.0509 |

secure and that the random noise generation process is truly random and correctly calibrated. As explicitly illustrated in Algorithm 2 (Client-side computation), carefully calibrated Laplace noise is systematically introduced to the local gradients $\Delta\theta_i^{\text{noisy}} = \Delta\theta_i + \text{LaplaceNoise}(\sigma)$ before their encryption and subsequent transmission to the server. This mechanism ensures strong privacy for each client's sensitive local data. In Algorithm 1 (Server-side computation), the server aggregates these noisy, encrypted gradients after decryption, thereby preventing the reconstruction of individual client contributions and further augmenting privacy. This meticulous application of DP shows that the global model benefits from collaborative learning. The privacy of individual IoT-enabled vehicles and their respective network patterns is maintained.

Overall, the proposed framework distinguishes itself from a mere client-side FL algorithm with DP primarily through its strategic multi-client, multi-dataset integration for comprehensive APT stage detection, specifically tailored to address the unique challenges of the VIoT domain. While standard FL with DP ensures privacy by perturbing individual client updates, this work's novelty lies in leveraging the inherent heterogeneity of multiple distinct, publicly available datasets (UNSW-NB15, CSE-CIC-IDS2018, and Edge-IIoTset) to simulate diverse "clients" or data sources in the VIoT environment. This method involves mapping features across these varied datasets to represent the full spectrum of APT attack stages, from reconnaissance to data exfiltration. This systematic cross-domain feature utilization, a significant departure from typical FL applications, allows the global model to learn from a much broader and more complex set of attack behaviours that are characteristic of multi-stage APTs often distributed across heterogeneous VIoT devices and networks, thereby overcoming the critical real-world problem of dedicated APT dataset scarcity in this VIoT system.

**TABLE 16. Comparing the PF-DAPTIV Framework With the State-of-the-art**

| Papers | Approach | Field | Dataset | Distributed Environment | Feature Extraction | Application of XAI | Privacy Preservation |
|---|---|---|---|---|---|---|---|
| [61] | DL model | IIoT | Malimg data | × | × | × | × |
| [62] | Game theory | CPS | NA | × | × | × | × |
| [63] | BERT scheme | IIoT | Own data | × | × | × | × |
| [64] | Correlations | IIoT | CICIDS2018 ICS data | × | ∼ | × | × |
| [65] | GAN based | CPS-IIoT | DAPT2020 Edge I-IoT | × | ∼ | × | × |
| [66] | FL model | Edge computing | UNSW-NB15 | ✓ | × | × | × |
| [67] | FL model | SDN network | NF-UQ NIDS | ✓ | × | × | × |
| [68] | FL model | 5G networks | system logs | ✓ | ∼ | × | × |
| **Proposed Framework Fed-DAPTIV** | **PP- DL model in Federated setup** | **Vehicular IoT CPS, CV** | **UNSW-NB15 CICIDS2018 Edge-IIoTset** | ✓ | ✓ | ✓ | ✓ |

Here, ✓ = Condition satisfied, × = condition not satisfied, ∼ = Partial

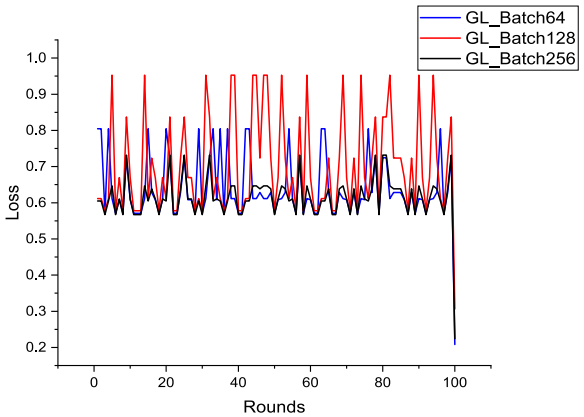


**FIGURE 18. Loss Vs number of rounds for the batch sizes 64, 128 and 256 respectively in PF-DAPTIV framework.**

Integration of SHAP, comprehending the rationale behind a framework's specific predictions, is indispensable for security analysts to respond to APTs effectively. Incorporating SHAP analysis provides essential interpretability by identifying the most influential features contributing to APT detection across different stages, even within the context of differentially private, aggregated updates. This capability to provide interpretability for a federated, DP-enabled model, particularly for multi-stage APT detection, offers invaluable insights often absent in black-box deep learning approaches and significantly enhances practical deployment in safety-critical vehicular systems.

The research is predicated upon specific Research Questions (RQs), as detailed in the contributions and enumerated in Table 1 of the manuscript, which directly address the unique and evolving threat landscape posed by APTs in vehicular-IoT environments. This targeted problem definition and analysis, including a comprehensive ablation study utilizing the selected datasets, provides profound insights into APTs within this VIoT domain. This focused inquiry into VIoT cybersecurity challenges through an FL with DP distinguishes this work from broader applications, providing valuable contributions for researchers and system defenders.

### A. RESEARCH QUESTIONARIES AND ANSWERS

The list of research questionnaires (RQ) formed is addressed in different parts of the experiment, which are summarized as below;

1) *How are intrusion-related datasets, such as CSE-CICIDS2018, Edge-IIoTset and UNSW-NB15, used as APT-related datasets?:* The UNSW-NB15, Edge-IIoTset and CSE-CICIDS2018 datasets primarily centre on intrusion data, offering a robust diversity in feature sets and encompassing various attack categories. These distinctive features and attack categories are the foundation for associating them with APT stages. Subsequently, the experiment is conducted using the derived dataset. The resultant mappings are systematically tabulated, as in Tables 3, 4, and 5.
2) *How do the datasets' features contribute to the detection of APTs? Does some feature influence model output more than other features in different environments?:* The features within the datasets assume a pivotal role in APT detection by serving as discriminative indicators or distinctive traits facilitating the identification of potentially anomalous activities. These features encompass diverse aspects pertinent to various stages of APTs and exhibit specific attributes indicative of attack stages. Employing the SHAP (SHapley Additive exPlanations)

to scrutinize the influence of these features on the model output has provided valuable insights into their significance and impact on the detection process.

3) *How can deep learning techniques be leveraged to detect APTs?:* Several methodologies exist to harness DL for APT detection, including employing DL for "feature learning," "sequence modeling," "transfer learning," and other techniques. Within this research, a DL-CNN (Deep Learning - Convolutional Neural Network) model has been constructed and integrated into the PF-DAPTIV framework, given the primary focus of the datasets on raw data. In the future, exploring and establishing a robust transfer learning model could offer additional valuable insights within this context.
4) *Does this study preserve the privacy of vehicular data?:* To ensure ironclad data privacy, this research trained their deep learning model using a federated learning approach, where data never leaves individual devices. Only model weights are exchanged, guaranteeing anonymity. In addition to anonymization, differential privacy is meticulously applied to further protect the confidentiality of data points while still facilitating collaborative model improvement through weight aggregation within a horizontal federated structure. Here, all devices contribute features of the same type.

## VII. CONCLUSION

The experimental endeavour has created a specialised privacy-preserved federated framework to detect APTs in IoT networks. Validation of the PF-DAPTIV framework has been conducted by utilising three distinct datasets, "CICIDS2018", "Edge-IIoTset", and "UNSW-NB15". A comprehensive analysis of all features within these datasets has been achieved, culminating in identifying and using a condensed subset of significant features as displayed in Fig. 9, Figure and Fig. 14 these set of features played crucial role in prominently detecting APT attacks and for model training to predict APT occurrences. Fig. 8, 10, 11, 12 and 13 shows the few selected feature's SHAP values. Eleven distinct metrics have been gathered at each network node to facilitate comparative assessments and subsequently organised into tabulated formats in the result section, Tables 8, 10, and 12 showcase the performance of the framework without the privacy preserving methods and Tables 9, 11, and 13 shows the results obtained, with privacy preserving technique.

Table 14 and 15 shows the performance comparison of the developed framework with and without integrated privacy mechanisms on multiple benchmark datasets. From these tables, it is very evident that the detection capacity of a system and the level of privacy offering are inversly proportional to each other; it is always the trade-off, and the researcher has to select the $\varepsilon$ value carefully as suitable to their field of application; Strategic selection of the $\varepsilon$ value is paramount, ensuring a judicious balance between robust privacy and high model accuracy tailored to the specific application domain. While increased noise enhances privacy, it inversely affects model accuracy. Conversely, diminished noise yields heightened accuracy at the expense of privacy.

This research successfully showcases the feasibility of APT detection within vehicular IoT environments, concurrently elucidating the underlying deep learning model's operational mechanisms through SHAP analysis. The developed deep learning model exhibits commendable efficacy in a decentralized setting, leveraging Horizontal Federated Learning with independently and identically distributed (IID) datasets. The proposed Privacy-preserving Federated Framework for APT Detection in Vehicular IoT (PF-DAPTIV) effectively safeguards node-level data privacy and mitigates communication overhead by optimizing data transfer between the central aggregator and client devices. While acknowledging the utilization of benchmark datasets due to the inherent challenges of accessing real-time vehicular network data, it is essential to emphasize that this methodological approach provides a solid foundation for evaluating the proposed framework. These datasets afford a controlled environment for rigorous performance assessment, facilitating the demonstration of the framework's potential.

## VIII. FUTURE WORK

The experiments were conducted under idealized network conditions, providing a baseline for performance evaluation. Consequently, future research may explore the framework's adaptability within dynamic and unpredictable network environments. To further enhance the model's applicability, integration with real-time data streams and local-level implementation could significantly contribute to advancing the comprehension of these systems. Furthermore, the proposed PF-DAPTIV framework possesses the potential for extension to related environments characterized by diverse IoT configurations. Specifically, applications within the Internet of Medical Things (IoMT) and Military Internet of Things (MIoT) domains, which feature distinct feature sets and real-life implications, represent a promising avenue for future investigation.

## ACKNOWLEDGMENT

The authors thank the Manipal Institute of Technology for facilitating Cloud GPU services and infrastructure to carry out the experiment.

## REFERENCES

[1] N. Mishra and S. Pandya, "Internet of Things applications, security challenges, attacks, intrusion detection, and future visions: A systematic review," *IEEE Access*, vol. 9, pp. 59353–59377, 2021.
[2] M. A. Khan and K. Salah, "IoT security: Review, blockchain solutions, and open challenges," *Future Gener. Comput. Syst.*, vol. 82, pp. 395–411, 2018.
[3] Z. Chen et al., "Machine learning-enabled IoT security: Open issues and challenges under advanced persistent threats," *ACM Comput. Surv.*, vol. 55, no. 5, 2022, Art. no. 105.

[4] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations," *IEEE Commun. Surv. Tut.*, vol. 21, no. 3, pp. 2702–2733, Thirdquarter 2019.

[5] A. Mosenia and N. K. Jha, "A comprehensive study of security of Internet-of-Things," *IEEE Trans. Emerg. Topics Comput.*, vol. 5, no. 4, pp. 586–602, Oct.–Dec. 2017.

[6] A. Alshamrani, S. Myneni, A. Chowdhary, and D. Huang, "A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities," *IEEE Commun. Surveys. Tuts.*, vol. 21, no. 2, pp. 1851–1877, Secondquarter 2019.

[7] B. Stojanović, K. Hofer-Schmitz, and U. Kleb, "APT datasets and attack modeling for automated detection methods: A review," *Comput. Secur.*, vol. 92, 2020, Art. no. 101734.

[8] R. Baksi and S. Upadhyaya, "A comprehensive model for elucidating advanced persistent threats (APT)," in *Proc. Int. Conf. Secur. Manage.*, 2018, pp. 245–251.

[9] M. Ussath, D. Jaeger, F. Cheng, and C. Meinel, "Advanced persistent threats: Behind the scenes," in *Proc. Annu. Conf. Inf. Sci. Syst.*, 2016, pp. 181–186.

[10] P. Chen, L. Desmet, and C. Huygens, "A study on advanced persistent threats," in *Proc. 15th Int. Conf. Commun. Multimedia Secur.*, 2014, vol. 8735, pp. 63–72.

[11] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. Voelker, "An analysis of underground forums," in *Proc. ACM SIGCOMM Conf. Int. Meas. Conf.*, 2011, pp. 71–80.

[12] A. Sharma, B. B. Gupta, A. K. Singh, and V. Saraswat, "Advanced persistent threats (APT): Evolution, anatomy, attribution and countermeasures," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 7, pp. 9355–9381, 2023.

[13] A. Gaurav, B. B. Gupta, and P. K. Panigrahi, "A comprehensive survey on machine learning approaches for Malware detection in IoT-based enterprise information system," *Enterprise Inf. Syst.*, vol. 17, no. 3, 2023, Art. no. 2023764.

[14] U. Sakthivelu and C. Vinoth Kumar, "Advanced persistent threat detection and mitigation using machine learning model," *Intell. Automat. Soft Comput.*, vol. 36, no. 3, pp. 3691–3707, 2023.

[15] Y. Fang et al., "Anomaly diagnosis of connected autonomous vehicles: A survey," *Inf. Fusion*, vol. 105, 2024, Art. no. 102223.

[16] S. B. Prathiba, G. Raja, S. Anbalagan, K. Arikumar, S. Gurumoorthy, and K. Dev, "A hybrid deep sensor anomaly detection for autonomous vehicles in 6G-V2X environment," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 3, pp. 1246–1255, May/Jun. 2023.

[17] H. Zhang, K. Zeng, and S. Lin, "Federated graph neural network for fast anomaly detection in controller area networks," *IEEE Trans. Inf. Forensics Secur.*, vol. 18, pp. 1566–1579, 2023.

[18] H. Sedjelmaci, N. Kaaniche, A. Boudguiga, and N. Ansari, "Secure attack detection framework for hierarchical 6G-enabled Internet of Vehicles," *IEEE Trans. Veh. Technol.*, vol. 73, no. 2, pp. 2633–2642, Feb. 2024.

[19] G. Rathee, A. Kumar, C. A. Kerrache, and C. T. Calafate, "A trust management solution for 5G-based future generation Internet of Vehicles," *Comput. Netw.*, vol. 248, 2024, Art. no. 110501.

[20] X. Qiu, J. Yu, W. Jiang, and X. Sun, "Intelligent security authentication for connected and autonomous vehicles: Attacks and defenses," *Electronics*, vol. 13, no. 8, 2024, Art. no. 1577.

[21] A. Haddaji, S. Ayed, and L. C. Fourati, "A novel and efficient framework for in-vehicle security enforcement," *Ad Hoc Netw.*, vol. 158, 2024, Art. no. 103481.

[22] M. H. Khan, A. R. Javed, Z. Iqbal, M. Asim, and A. I. Awad, "DivaCAN: Detecting in-vehicle intrusion attacks on a controller area network using ensemble learning," *Comput. Secur.*, vol. 139, 2024, Art. no. 103712.

[23] W. Ding, I. Alrashdi, H. Hawash, and M. Abdel-Basset, "DeepSecDrive: An explainable deep learning framework for real-time detection of cyberattack in in-vehicle networks," *Inf. Sci.*, vol. 658, 2024, Art. no. 120057.

[24] M. Zhou and X. Che, "Stealthy attack detection based on controlled invariant subspace for autonomous vehicles," *Comput. Secur.*, vol. 137, 2024, Art. no. 103635.

[25] H. K. Alkhpor and F. M. Alserhani, "Collaborative federated learning-based model for alert correlation and attack scenario recognition," *Electronics*, vol. 12, no. 21, 2023, Art. no. 4509.

[26] S. Salim, N. Moustafa, M. Hassanian, D. Ormod, and J. Slay, "Deep federated learning-based threat detection model for extreme satellite communications," *IEEE Internet Things J.*, vol. 11, no. 3, pp. 3853–3867, Feb. 2024.

[27] T. Bodström and T. Hämäläinen, "A novel deep learning stack for APT detection," *Appl. Sci.*, vol. 9, no. 6, 2019, Art. no. 1055.

[28] S. Zidi, B. Alaya, T. Moulahi, A. Al-Shargabi, and S. El Khediri, "Fault prediction and recovery using machine learning techniques and the HTM algorithm in vehicular network environment," *IEEE Open J. Intell. Transp. Syst.*, vol. 5, pp. 132–145, 2024.

[29] W. Ali, I. U. Din, A. Almogren, and J. J. Rodrigues, "Federated learning-based privacy-aware location prediction model for Internet of Vehicular Things," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 1968–1978, Feb. 2025.

[30] T. Aldhanhani, A. Abraham, W. Hamidouche, and M. Shaaban, "Future trends in smart green IoV: Vehicle-to-Everything in the era of electric vehicles," *IEEE Open J. Veh. Technol.*, vol. 5, pp. 278–297, 2024.

[31] K. Zhang et al., "Intrusion detection model for Internet of Vehicles using GRIPCA and OWELM," *IEEE Access*, vol. 12, pp. 28911–28925, 2024.

[32] A. Sharma and A. Jaekel, "Machine learning based misbehaviour detection in VANET using consecutive BSM approach," *IEEE Open J. Veh. Technol.*, vol. 3, pp. 1–14, 2022.

[33] P. Lv, L. Xie, J. Xu, X. Wu, and T. Li, "Misbehavior detection in vehicular ad hoc networks based on privacy-preserving federated learning and blockchain," *IEEE Trans. Netw. Serv. Manage.*, vol. 19, no. 4, pp. 3936–3948, Dec. 2022.

[34] Z. Wang, J. Li, Y. Wang, Z. Su, S. Yu, and W. Meng, "Optimal repair strategy against advanced persistent threats under time-varying networks," *IEEE Trans. Inf. Forensics Secur.*, vol. 18, pp. 5964–5979, 2023.

[35] M. A. R. Baee, L. Simpson, E. Foo, and J. Pieprzyk, "The security of "2FLIP" authentication scheme for VANETs: Attacks and rectifications," *IEEE Open J. Veh. Technol.*, vol. 4, pp. 101–113, 2023.

[36] L. Crosara, F. Ardizzon, S. Tomasin, and N. Laurenti, "Worst-case spoofing attack and robust countermeasure in satellite navigation systems," *IEEE Trans. Inf. Forensics Secur.*, vol. 19, pp. 2039–2050, 2024.

[37] R. Asensio-Garriga et al., "ZSM-based E2E security slice management for DDoS attack protection in MEC-enabled V2X environments," *IEEE Open J. Veh. Technol.*, vol. 5, pp. 485–495, 2024.

[38] K. Sharma and B. Gupta, "Multi-layer defense against Malware attacks on smartphone Wi-Fi access channel," *Procedia Comput. Sci.*, vol. 78, pp. 19–25, 2016.

[39] H. Alqahtani and G. Kumar, "Machine learning for enhancing transportation security: A comprehensive analysis of electric and flying vehicle systems," *Eng. Appl. Artif. Intell.*, vol. 129, 2024, Art. no. 107667.

[40] S. Miao, Q. Pan, D. Zheng, and G. Mohi-ud din, "Unmanned aerial vehicle intrusion detection: Deep-meta-heuristic system," *Veh. Commun.*, vol. 46, 2024, Art. no. 100726.

[41] M. J. Choi, I. R. Jeong, and H. M. Song, "Fast and efficient context-aware embedding generation using fuzzy hashing for in-vehicle network intrusion detection," *Veh. Commun.*, vol. 47, 2024, Art. no. 100786.

[42] K. R. Jangam, R. Kumudham, V. Rajendran, and M. R. Prabhu, "Enhancing secure communication in IoT-based automated vehicle systems through accurate prediction of abnormal traffic data," *Int. J. Eng. Trends Technol.*, vol. 72, no. 3, pp. 358–369, 2024.

[43] N. R. Prasad, B. Andersen, and D. K. Rubin-Grøn, "Overview of security challenges in wireless IoT infrastructures for autonomous vehicles," in *Proc. 7th Int. Conf. Saf. Secur. IoT*, 2024, pp. 63–82.

[44] X. Chen, W. Feng, Y. Chen, N. Ge, and Y. He, "Access-side DDOS defense for space-air-ground integrated 6G V2X networks," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 2847–2868, 2024.

[45] H. N. Aleisa et al., "Transforming transportation: Safe and secure vehicular communication and anomaly detection with intelligent cyber-physical system and deep learning," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 1736–1746, Feb. 2024.

[46] Z. Lin and J. Li, "FedEVCP: Federated learning-based anomalies detection for electric vehicle charging pile," *Comput. J.*, vol. 67, no. 4, pp. 1521–1530, 2024.

[47] C.-K. Tham, L. Yang, A. Khanna, and B. Gera, "Federated learning for anomaly detection in vehicular networks," in *Proc. IEEE 97th Veh. Technol. Conf.*, 2023, pp. 1–6.

[48] V. P. Chellapandi, L. Yuan, C. G. Brinton, S. H. Zak, and Z. Wang, "Federated learning for connected and automated vehicles: A survey of existing approaches and challenges," *IEEE Trans. Intell. Veh.*, vol. 9, no. 1, pp. 119–137, Jan. 2024.
[49] Y. Wang, D.-H. Zhai, D. Han, Y. Guan, and Y. Xia, "MITDBA: Mitigating dynamic backdoor attacks in federated learning for IoT applications," *IEEE Internet Things J.*, vol. 11, no. 6, pp. 10115–10132, Mar. 2024.
[50] H. K. Alkhpor and F. M. Alserhani, "Collaborative federated learning-based model for alert correlation and attack scenario recognition," *Electronics*, vol. 12, no. 21, 2023, Art. no. 4509.
[51] X. Han et al., "ADS-Lead: Lifelong anomaly detection in autonomous driving systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 1, pp. 1039–1051, Jan. 2023.
[52] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proc. 4th Int. Conf. Inf. Syst. Secur. Privacy*, 2018, pp. 108–116.
[53] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *Proc. Mil. Commun. Inf. Syst. Conf.*, 2015, pp. 1–6.
[54] N. Moustafa and J. Slay, "The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set," *Inf. Secur. J.: A Glob. Perspective*, vol. 25, no. 1–3, pp. 18–31, 2016.
[55] N. Moustafa, J. Slay, and G. Creech, "Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks," *IEEE Trans. Big Data*, vol. 5, no. 4, pp. 481–494, Dec. 2019.
[56] N. Moustafa, G. Creech, and J. Slay, "Big data analytics for intrusion detection system: Statistical decision-making using finite Dirichlet mixture models," in *Data Analytics and Decision Support for Cybersecurity: Trends, Methodologies and Applications*. Berlin, Germany: Springer, 2017, pp. 127–156.
[57] M. Sarhan, S. Layeghy, N. Moustafa, and M. Portmann, "NetFlow datasets for machine learning-based network intrusion detection systems," in *Proc. Big Data Technol. Appl.: 10th EAI Int. Conf., BDTA, 13th EAI Int. Conf. Wireless Internet, WiCON, Virtual Event*, Dec. 2020, pp. 117–135.
[58] M. A. Ferrag, O. Friha, D. Hamouda, L. Maglaras, and H. Janicke, "Edge-IIoTset: A new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning," *IEEE Access*, vol. 10, pp. 40281–40306, 2022.
[59] M. Azizjon, A. Jumabek, and W. Kim, "1D CNN based network intrusion detection with normalization on imbalanced data," in *Proc. Int. Conf. Artif. Intell. Inf. Commun.*, 2020, pp. 218–224.
[60] L. Huang and Q. Zhu, "A dynamic games approach to proactive defense strategies against advanced persistent threats in cyber-physical systems," *Comput. Secur.*, vol. 89, 2020, Art. no. 101660.
[61] K. Yu et al., "Securing critical infrastructures: Deep-learning-based threat detection in IIoT," *IEEE Commun. Mag.*, vol. 59, no. 10, pp. 76–82, Oct. 2021.
[62] A. Kumar and V. Thing, "RAPTOR: Advanced persistent threat detection in industrial IoT via attack stage correlation," in *Proc. 20th Annu. Int. Conf. Privacy, Secur. Trust*, 2023, pp. 1–12.
[63] S. Hussain et al., "APT adversarial defence mechanism for industrial IoT enabled cyber-physical system," *IEEE Access*, vol. 11, pp. 74000–74020, 2023.
[64] Z. Li, J. Chen, J. Zhang, X. Cheng, and B. Chen, "Detecting advanced persistent threat in edge computing via federated learning," in *Proc. 1st Int. Conf. Secur. Privacy Digit. Economy*, Oct./Nov. Quzhou, China, 2020, pp. 518–532.
[65] H. T. Thi, N. D. Hoang Son, P. T. Duy, and V.-H. Pham, "Federated learning-based cyber threat hunting for apt attack detection in SDN-enabled networks," in *Proc. 21st Int. Symp. Commun. Inf. Technol.*, 2022, pp. 1–6.
[66] X. Cheng, Q. Luo, Y. Pan, Z. Li, J. Zhang, and B. Chen, "Predicting the APT for cyber situation comprehension in 5G-enabled IoT scenarios based on differentially private federated learning," *Secur. Commun. Netw.*, vol. 2021, 2021, Art. no. 8814068.

**SUDHINA KUMAR G K** received the B.E. degree in information science and engineering from Visvesvaraya Technological University (VTU), Belgaum, Karnataka, India, in 2017, and M.Tech degree in computer networks and engineering with Manipal Institute of Technology (MIT), Manipal, Karnataka, India, in 2020. He is currently working toward the Ph.D. degree in information and communication technology with Manipal Institute of technology, Manipal, Karnataka, India. His research interests includes Privacy preserving, understanding different types of threats in cyber Space.

**KRISHNA PRAKASHA K** received the B.E. and M.Tech. degrees from Viswesvaraya Technological University, Belagavi, Karnataka, India, and the Ph.D. degree in network security from the Manipal Academy of Higher Education (MAHE), Manipal, Karnataka, India. He is currently an Assistant Director: Alumni, Public & International Relations, MIT and an Associate Professor with the Department of Information and Communication Technology, Manipal Institute of Technology, MAHE. He has more than 30 publications in national and international conferences and journals. His current research interests include information security, network security, algorithms, real-time systems, and wireless sensor networks.

**BALACHANDRA MUNIYAL** (Member, IEEE) received the B.E. degree in computer science and engineering from Mysore University, Mysore, Karnataka, India, and the M.Tech. and Ph.D. degrees in computer science and engineering from the Manipal Academy of Higher Education, Manipal, Karnataka, India. He carried out the M.Tech. project work in T-Systems Nova GmbH, Bremen, Germany. He was deputed to Manipal International University, Nilai, Malaysia, in 2014. He is Coordinator of Center of Excellence for Cybersecurity and Professor with the Department of Information & Communication Technology, Manipal Institute of Technology, Manipal. He has 27 years of teaching experience in various institutes. He has more than 50 publications in national and international conferences/journals. His research interests include network security, cryptography, and intrusion detection systems.

**MUTTUKRISHNAN RAJARAJAN** (Senior Member, IEEE) received the Ph.D. degree from the City University of London, London, U.K., in 2001. He is the founding Director of the institute for Cyber Security with City University of London and the CEO of Citydefend Limited. He is a Professor of Security engineering with the City University of London, U.K., and he is currently actively engaged in the U.K. governments Identity Assurance Programme (verify U.K.). He has authored or coauthored more than 350 articles, three books and holds two patents in the area of cloud data privacy. His research interests include mobile security, intrusion detection, and privacy techniques. He is a member of ACM and an Advisory board member of the The Institue of Information Security Professionals (IISP), U.K.