Federated Learning-based Misbehaviour detection for the 5G-enabled Internet of Vehicles

Preeti Rani[1], Chandani Sharma[2], Janjhyam Venkata Naga Ramesh[3], Sonia Verma[4], Rohit Sharma[1], Ahmed Alkhayyat[5], Sachin Kumar[6]

Abstract **– The concept of federated learning (FL) is becoming increasingly popular as a method for training collaborative models without loss the sensitive information. The term has become ubiquitous due to the extensive development of autonomous vehicles. Vehicular Networks and the Internet of Vehicles (IoV) enable cooperative learning through federated learning. It is still necessary to address several technical challenges. In recent years, Federated Learning (FL) has attracted significant interest in various sectors, including smart cities and transportation systems. FL-enabled attack detection for IoVs are still in its infancy. However, to determine the main challenges of deployment in real-world scenarios, there needs to be research efforts from various areas. Performance metrics are used to evaluate the effectiveness of the proposed FL framework. According to experiments, the proposed FL approach detected attacks in IOV networks with a maximum accuracy of 99.72%. In addition to precision, recall, and F1 scores, 99.70%, 99.20%, and 99.26% were achieved. A comparison of the proposed model with the existing model shows that the proposed model is more accurate.**

Keywords – Federated Learning, Internet of Vehicles (IOV), intelligent transportation system (ITS), 5G, vehicular ad-hoc networks (VANETs)

I. Introduction

IOV-derived VANETs are an important research area in IoT. As sensor technologies grow rapidly, IOV generates large volumes of data. Transportation systems must meet the IOV requirements to be effective. Different companies in different countries have introduced intelligent transportation systems based on IOVs (Gope & Sikdar, 2019; Hamid et al., 2019; Hussain et al., 2022; L.-L. Wang et al., 2020). The most common IOV applications are traffic control systems, traffic flow monitoring, toll plazas and intelligent vehicle control.

[1]*Department of Electronics & Communication Engineering, Faculty of Engineering and Technology,*
SRM Institute of Science and Technology, Delhi-NCR Campus, Delhi-Meerut Road, Modinagar, Ghaziabad, Uttar Pradesh, India
[2]*Associate professor, Department of Computer Science & Engineering, MMICTBM (MCA), Maharishi Markandeshwar (Deemed to be University), Mullana-Ambala, Haryana, India*
[3]*Assistant Professor, Dept.of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram-522502, Guntur Dist., Andhra Pradesh, India*
[4]*Assistant Professor, Department of Computer Science, ABES Engineering College, Ghaziabad, India -201204*
[5]*College of technical engineering, The Islamic University, Najaf, Iraq,* ahmedalkhayyat85@gmail.com
preetiresearcher1@gmail.com, chandani19nov@gmail.com, jvnramesh@gmail.com, sonverma@gmail.com, rohitapece@gmail.com
[6]*Big Data and Machine Learning Laboratory, South Ural State University, Chelyabinsk, Russia,* sachinagnihotri16@gmail.com

Corresponding- Rohit Sharma, Preeti Rani

Cellular networks' fifth generation aims to improve reliability, throughput, delay, and connectivity while improving service quality (Hassan et al., 2019). The Internet of Things (IoT), smart houses, Intelligent Transportation Systems (ITS), and health monitoring are some of the developing applications of fifth generation (5G) (I. A. Alablani & Arafah, 2021). Physical things are associated with the Internet through the IOTs, an emerging revolution (I. Alablani & Alenazi, 2019). As part of the IoT, the Internet of Vehicles (IOVs) connects vehicles to the Internet and allows them to communicate (Fabian et al., 2021; Rehman et al., 2019). IoV technology and Intelligent Transportation Systems (ITS) are evolutions towards vehicle-to-everything (V2X) tech. The V2X program aims to make roads safer, communication more reliable, and traffic flow more efficient (Chen et al., 2017; Raza et al., 2018). Figure 1 shows four kinds of vehicles-to-vehicles (V2V), vehicles-to-infrastructures (V2I), vehicles-to-pedestrians (V2P), and vehicles-to-networks (V2N) communication—most ITS use machine learning techniques to provide comfort and safety to end users (Sirohi et al., 2020).

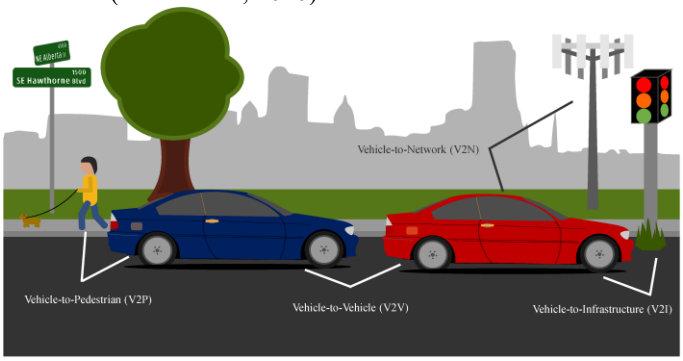

Figure 1: Internet of the vehicle and its communications.

The IoV has developed as a new standard in the field of ITS thanks to the advancement of IoT. As a result of the development of the traditional vehicle ad-hoc network (VANET) network, and road transport equipment, the IoV has evolved with the IoV (Babun et al., 2021). In addition to driving assistance services, traffic notifications, speed advisory services, accident notifications, and emergency vehicle notifications, the IoV promises to redefine how users live and offer them better opportunities (Puri et al., 2020). In the IoVs, a complex vehicular network connected to the Internet, sensors are installed on vehicles and roads to collect data. Vehicles, personal devices, and remote sensing units are

connected to cloud networks through wireless communication networks based on various wireless access technologies. There are different preferences and requirements for each IoV and network device. Vehicles and other devices utilize smart cloud services based on their application requirements. Safety, traffic management, and commercial applications are different IoV applications. Smart cloud infrastructure processes ITS applications like safety, navigation, real-time traffic information management, and parking; the results are fed back to IOVs.

With the ubiquity of sensors in vehicular networks, it is feasible to capture more data and train ML models. Machine Learning based models are generally applied to both vehicle management and traffic management (Tan et al., 2020). The dynamic nature of the surroundings limits the current autonomous driving decisions as the training is carried out offline. FL can rescue such situations by online training vehicles from different geographical locations, facilitating accurate labelling of the features. Similarly, a large amount of data is required for traffic flow prediction techniques. Still, most data is divided among various organizations and cannot be exchanged to protect privacy (Liu et al., 2020). To address such situations also, we can deploy FL methods.

1. In the paper proposed Federated Learning-based Misbehaviour detection for the 5G-enabled Internet of Vehicles (IOVs) using federated distillation (FD) training scheme with FL baselines.
2. In this paper, Four network attack detection datasets are evaluated, namely ISCXIDS2012 (Shiravi et al., 2012), CIC-110 IDS2017 (Panigrahi & Borah, 2018), CSE-CICIDS2018 (Leevy & Khoshgoftaar, 2020), and Car-hacking (Seo et al., 2018).
3. As a benchmark, other attack detections FL baselines are compared to evaluate the effectiveness of this method.

The rest of the paper is structured as follows: Section 2 explains the literature review based on the latest attack detection techniques using Federated Learning in the field of IOV. Section 3 presents the proposed Federated learning-based framework with complete mathematics. Section 4 presents the experimental setup, datasets, result discussion, and comparative analysis with the FL baselines. Section 5 presents the conclusion.

II. Literature Review

A rapidly growing market, especially the Internet of Things (IoT), is moving toward 6G networks (Mumtaz et al., 2022). This raises security concerns significantly (Tang et al., 2020). The Internet of Things and its Values reshape networking and communication (Aman et al., 2020). The Internet of Things/Internet of Vehicles can be defined differently (Youm, 2017), usually encompassing physical objects like sensors and vehicles. (Mahmood et al., 2022). In addition, the meaning has broadened from local to global, encompassing both local area networks (LANs) and next-generation 6G networks. The reasons for this are that 5G technology cannot handle the IoT/IoV applications requirements of the near future (Kim, 2021). IoT/IoV networks are experiencing much greater traffic as more devices are added. In one sense, IoT/IoV networks enrich our everyday lives and industrial equipment in the other. The number of malware threats is also increasing across all layers and phases of the life cycle of the network (Peng et al., 2021; Radoglou-Grammatikis et al., 2022). Due to the heterogeneity and dynamic nature of IoT and IoV, security has become one of the most important research areas (Aman et al., 2020). 6G networks are expected to support many heterogeneous devices and infrastructures, surpassing 5G networks in many ways (J. Wang et al., 2021). Table 1 shows some common vehicular cyber-attacks.

Table 1: The description of cyber-attacks in the IOV network

Attack	Description
DoS	In a DoS attack, malicious actors attempt to disable a device's normal operation to prevent its users from using it.
Jamming	An example of a jamming attack is when a malicious node interferes with a network to disrupt legitimate communication.
Spoofing	Spoofing is used in cyber security to describe the act of impersonating another entity to gain trust, gain access to systems, steal data, steal money, or distribute malware.
Acoustic	Computers and smart devices are vulnerable to acoustic attacks.
Eavesdropping	A technique of passively listening to network conversations is called eavesdropping, in which an attacker obtains private information such as node identifiers, routing changes, and application-sensitive information.
Blackhole	Routers that delete all messages supposed to be forwarded are said to be committing a blackhole attack. A router can be set incorrectly to give a zero-cost route to every destination on the Internet occasionally.
Cloaking	Users' credentials are often obtained through social engineering attacks on malicious websites impersonating well-known businesses. It is common practice for certain websites to hide hazardous information from search engine crawlers while displaying it to users/client browsers, known as cloaking.
Replay	Hackers typically perform replay attacks by intercepting and fraudulently delaying or resending data from a secure network connection to trick the receiver into doing what they want.

Federated Learning was introduced by Google in 2017 (McMahan et al., 2017), which enables machine learning models to be trained with data collected across multiple