Federated Learning for Intrusion Detection Systems in Internet of Vehicles

Bayu Sulistiyanto Ipung Sutejo[1], Fajar Pratama[2]
[1] Department of computer science, Esa Unggul University, Indonesia; r.bayoesutejo@gmail.com
[2] Department of computer science, Esa Unggul University, Indonesia; fajarpratama0115@gmail.com

ABSTRACT The Internet of Vehicles needs Intrusion Detection Systems. With Federated Learning, models can perform local training on selected vehicles while maintaining anonymity through joint learning; thus redundant knowledge is shared and the overall resilience of the connected vehicular network in terms of security also increases. This article introduces Federated Learning (FL) as a privacy-preserving machine learning framework for Intrusion Detection Systems in the Internet of Vehicles. The types of IoV IDS detection technique, data source and deployment location are categorized as taxonomies. This article stresses applications of FL in optimizing real-time threat detection, data sharing without privacy risk and cooperation among vehicles. It bring possibilities of future challenge along with potential ways for improving and integrating it-such as 5G, edge computing etc.

KEYWORDS Federated learning, Intrusion system, anomaly detection, edge computing

I. INTRODUCTION

Federated Learning(FL) is an innovative machine learning paradigm that allows model training across decentralized devices, so long as data remains local. In FL, models are trained on local devices. The updates of the aggregate model are uploaded to a central server without revealing user data privacy issues. It has become increasingly popular in a variety of fields, providing answers where data privacy matters [1].

The Internet of Vehicles (IoV) is a networked, integrated environment where vehicles interact with one another and their surroundings. At this point, the security of the IoV is significant [2]. These Intrusion Detection Systems (IDS) are largely responsible for detecting possible threats and their neutralization, protecting the entire vehicular network in terms of both integrity and safety. As the IoV develops, more powerful IDS protection is needed to protect against all kinds of cyber-attacks and guarantee that intelligent transport runs smoothly [3].

This article's main goal is to develop a complete taxonomy of Federated Learning for Intrusion Detection Systems within the Internet of Vehicles. This taxonomy will organize different types of approaches, methodologies and components concerning the integration of FL with IDS into a systematic framework for researchers and users.

Finally, the article will examine how Federated Learning can strengthen Intrusion Detection Systems that are related to the Iov. This exploration will examine situations where FL may help in the detection, prevention and reaction to information security threats within vehicular networks. These applications are essential to realizing the full potential of FL in protecting IoV.

Finally, the article will explore some of possible directions and bottlenecks at this crossroads between FL & IDS for IoV. The article attempts to point the way towards future directions for secure and privacy-preserving vehicular communications by pinpointing areas in need of improvement, catching emerging trends early on, or finding gaps between problems needing solving and solutions yet discovered. The insights below will help inform the discussion about how FL can be best harnessed to strengthen Intrusion Detection Systems in response to ever changing Internet of Vehicles.

II. Federated Learning in Intrusion Detection Systems

In Federated Learning (FL), model training is done locally on each individual device, and only the trained model updates are transmitted to a central server [4]. With this privacy-preserving paradigm, devices can work together to learn a global model without disclosing raw data. With respect to Intrusion Detection Systems (IDS) [5], FL provides a disruptive new way to improve security in systems that are interconnected with each other, such as the Internet of Vehicles (IoV). Fundamentally, then, FL allows intrusion detection models to be strong and flexible without impairing the privacy of sensitive data.

There are several advantages to the use of Federated Learning in Intrusion Detection Systems. Secondly, because the raw data never leaves local devices, FL makes privacy easy to preserve. This is especially true of IoV, where data generated by vehicles sometimes involves sensitive

information and must be protected [6] Thirdly, FL enables distributed learning, making it ideal for the IoV's dynamic and diffuse environments. It allows for individual cars or parts of the network to learn and respond to local threats, contributing to systemwide security. Finally, FL encourages collaborative training[7]. The IDS can better take advantage of the different data sources available throughout the IoV ecosystem to create more robust and accurate intrusion detection models.

While Federated Learning for Intrusion Detection Systems in the Internet of Vehicles has many advantages, implementing it involves its own issues and considerations. One major obstacle is making sure the federated learning process itself is secure. Since feeds are shuttled back and forth between local devices and the central server, they must be guarded against attackers [8]. Besides, there is the problem of model convergence and performance when heterogeneous or dynamic data from different vehicles come into play. Besides, the problem of non-IID (non-Independently and Identically Distributed) data [9]. Different cars may be subject to different forms of attack. We must proceed gingerly so that the whole federated learning model operates effectively connects sentences

The application of Federated Learning for intrusion detection systems in the Internet of Vehicles has tremendous potential. The privacy-protecting properties of FL along with the ability to adapt to changing and dispersed situations make it an attractive option for improving the security between connected systems. Yet tackling the problems of security, heterogeneity and non-IID data is necessary if FL-based IDS are to be implemented in the IoV. The result will be a more secure and private model of intrusion detection.

III. Taxonomy of Intrusion Detection Systems in Internet of Vehicles

A. Classification of IDS Based on Detection Techniques

Intrusion Detection Systems (IDS) in the Internet of Vehicles (IoV) employ various techniques to identify and mitigate potential threats. This taxonomy categorizes IDS based on detection techniques:

1. Signature-based IDS

A signature-based IDS uses predefined patterns of known types of cyber threats. In recognizing established attack types, this approach works well. It is a valuable part of IoV security. Yet, as for its limitation, it is that it cannot detect a novel or hitherto unseen threat [10].

2. Anomaly-based IDS

In anomaly-based IDS, the focus is on observing deviations from normal system behavior. Establishing a normal baseline of activity means that anything untoward can be viewed as suspicious behavior. While anomaly-based IDS is very effective in detecting new or changing threats, it may produce false positives in complex and everchanging vehicular environments [11].

3. Hybrid IDS

Hybrid IDS employs both signature-based and anomaly-based detection techniques. This approach seeks to combine the best of each method, creating a better equipped and more flexible intrusion detection system. A hybrid IDS, which combines signature-based precision with anomaly-based flexibility, provides a powerful answer to the multiple kinds of threat facing the IoV [12].

B. Taxonomy Based on Data Sources

The data sources utilized by IDS significantly influence their effectiveness. This taxonomy classifies IDS based on the origin of data for intrusion detection:

1. In-vehicle Sensors

Sensor-rich vehicles provide plenty of data for IDS. They have accelerometers, GPS modules, and sometimes even cameras. By analyzing in-vehicle sensor data, anomalies in vehicles are detected. Cyber-physical threats can then be identified by the system [13].

2. Communication Networks

IDS is able to monitor and scan the communications network within the IoV. It filters data packets, looking for traces of malicious behavior. Anomalies in network traffic, communication patterns and protocol deviations provide the system with a method to identify threats from V2V (vehicle-to-vehicle) or V2I (vehicle-to infrastructure) communication[14].

3. Cloud-based Data

Utilizing cloud-based data for intrusion detection involves aggregating and analyzing information from various vehicles within a cloud infrastructure. This approach enables centralized monitoring and the identification of global attack patterns, enhancing the overall security of the IoV [15].

C. Categorization by Deployment Locations

The deployment location of an IDS within the IoV architecture is a crucial aspect of its effectiveness. This taxonomy categorizes IDS based on their deployment locations:

1. Edge-based IDS

Edge-based IDS are deployed on individual vehicles or at the edge of the vehicular network. These systems analyze data locally, providing real-time intrusion detection and response. Edge-based IDS are advantageous for minimizing latency and ensuring timely threat mitigation. [16]

2. Cloud-based IDS

Cloud-based IDS centralizes intrusion detection processes in cloud servers. This approach enables comprehensive analysis of data from multiple vehicles, facilitating the identification of global attack patterns and trends. Cloud-based IDS are effective for scalability and collaborative threat intelligence [17].

3. Hybrid Deployment Models

Hybrid deployment models combine both edge-based and cloud-based IDS components. This approach optimizes the advantages of local detection and centralized analysis, providing a balanced and adaptive intrusion detection solution for the diverse and dynamic IoV environment [18].

In conclusion, this taxonomy provides a structured framework for understanding the diverse landscape of Intrusion Detection Systems in the Internet of Vehicles. By classifying IDS based on detection techniques, data sources, and deployment locations, researchers and practitioners can navigate the complexities of designing, implementing, and optimizing intrusion detection solutions tailored to the unique challenges of the IoV.

IV. Applications of Federated Learning in Internet of Vehicles IDS

A. Real-time Threat Detection

Federated Learning (FL) proves instrumental in enhancing real-time threat detection capabilities within the Internet of Vehicles (IoV) Intrusion Detection Systems (IDS). By employing FL, individual vehicles equipped with edge-based IDS can collaboratively analyze local data to identify emerging threats in real-time. The decentralized nature of FL ensures timely detection of anomalies and potential cyber-attacks, contributing to the overall security and safety of the vehicular network [19].

B. Privacy-preserving Data Sharing

Privacy is a huge concern in the IoV, where vehicles generate sensitive data. FL addresses this concern by allowing model training to occur locally on individual vehicles without sharing raw data. Privacy-preserving data sharing ensures that sensitive information remains on the device, with only aggregated model updates being transmitted to a central server. This application of FL in IoV IDS not only enhances security but also fosters user trust by prioritizing data privacy [20].

C. Collaborative Learning Among Vehicles

FL facilitates collaborative learning among vehicles within the IoV, enabling them to collectively improve their intrusion detection capabilities. Each vehicle contributes to the learning process by sharing insights gained from its unique local data. This collaborative approach allows the entire vehicular network to benefit from the diverse experiences and perspectives of individual vehicles, leading to a more robust and adaptive IDS [21].

D. Adaptive Model Updating for Evolving Threats

The dynamic nature of cyber threats in the IoV requires IDS to adapt continually. FL excels in this aspect by enabling adaptive model updating. As new threat patterns emerge, the FL model can be updated collaboratively across vehicles, ensuring that the IDS remains effective against evolving cyber threats. This adaptability is crucial for staying ahead of sophisticated attacks and maintaining the resilience of the IoV security infrastructure [22].

E. Case Studies or Examples Illustrating Successful FL Applications in IoV IDS

Several case studies and examples highlight the successful application of Federated Learning in IoV IDS:

1. Traffic Anomaly Detection

In a scenario where multiple vehicles share insights about traffic anomalies, FL-based IDS can collectively identify irregularities such as sudden stops, traffic congestions, or abnormal traffic patterns. This collaborative approach improves the accuracy of anomaly detection, enhancing overall traffic management and safety [23].

2. Malicious Communication Identification

FL in IoV IDS can be applied to identify patterns of malicious communication between vehicles. By analyzing communication data locally and collaboratively updating models, the system can quickly detect and mitigate potential cyber-attacks targeting the vehicular communication network [24].

3. Privacy-preserving Incident Reporting

FL enables vehicles to report security incidents without compromising user privacy. By locally processing and sharing encrypted incident data, vehicles contribute to a collective knowledge base for threat detection without revealing sensitive information.

These case studies showcase the versatility and effectiveness of Federated Learning in IoV IDS, emphasizing its ability to address diverse security challenges while preserving user

privacy and fostering collaborative security measures within the vehicular network [25].

V. Future Challenges

Integration with new technologies The future of Federated Learning (FL) in Internet of Vehicles (IoV) Intrusion Detection Systems (IDS) looks promising. Several key areas include:

1. 5G Networks and Edge Computing

In the IoV, FL infrastructure will be greatly enhanced with the arrival of 5G networks and advances in edge computing capabilities. Faster model updates through lower latency and increased bandwidth will let intrusion detection be faster, more effective, and more responsive [26].

2. Explainable AI (XAI)

Given that the complexity of FL models is increasing, XAI becomes a necessity. As future FL-based IDS in the IoV, such models will be transparent and interpretable, providing insights into decision making processes. This boosts users ' trust as well that of other stakeholders [27].

3. Secure Hardware areas

Using secure hardware enclaves such as Trusted Execution Environments (TEEs) will improve the security of FL in IoV IDS. This technology is designed to have the sensitive operations carried out in isolated and secure environments, protecting against possible attacks upon this very learning process [28].

Gaps in Research and Areas for Improvement

While FL in IoV IDS shows great potential, several research gaps and areas for improvement warrant exploration:

1.Robustness to Adversarial Attacks

Robustifying FL models against adversarial attacks is still an important open research problem. The development of detection and preventive mechanisms in future work should also take into account the features of decentralization and dynamism unique to IoV environments [29].

2. Energy Efficiency

For resource-limited IoV devices, optimizing the energy efficiency of FL algorithms is essential. Sustainable and large-scale deployment depends on Lightweight models, compression techniques, and energy-aware learning strategies [30].

3. Cross-domain Knowledge Transfer

This is particularly true when vehicles must switch from one environment to another of different threat structure, for example. Under such conditions it will be necessary to find ways of knowledge transfer effective in the different environments. One way to improve model adaptability is transfer learning and domain adaptation techniques.

VI. CONCLUSIONS

To sum up, the work on Federated Learning (FL) for Intrusion Detection Systems (IDS) in Internet of Vehicles IoV has provided valuable results. A transformative solution emerges: FL offers privacy-protected security and decentralized collaboration among vehicles. Other key findings are that data privacy problems have been resolved through local model training, intrusion detection accuracy is improved by collaborative learning, and ICTs can get around communication overhead and heterogeneity points.

The implications for future development of IoV IDS with FL are deep. For the IoV, this integration marks a step toward robust Cybersecurity infrastructure, which enhances user trust and adoption as well as overall resource utilization. Future implementations of IoV IDS, made up of emerging technologies such as 5G networks, edge computing and explainable AI, have the promise of improving communication or responsiveness and interpretability.

For this reason, further research and innovation in this area is urgently needed. The claim here is that researchers and practitioners must face evolving threats, conduct interdisciplinary collaborations, test solutions in real-world environments, and catch up to technological developments. With this determination to constantly innovate, a set of robust, flexible and privacy-preserving FL-based IDS solutions that can keep up with the changing situation are sure to emerge.

References

[1] AbdulRahman, S., Tout, H., Ould-Slimane, H., Mourad, A., Talhi, C., & Guizani, M. (2020). A survey on federated learning: The journey from centralized to distributed on-site learning and beyond. IEEE Internet of Things Journal, 8(7), 5476-5497.

[2]Ji, B., Zhang, X., Mumtaz, S., Han, C., Li, C., Wen, H., & Wang, D. (2020). Survey on the internet of vehicles: Network architectures and applications. IEEE Communications Standards Magazine, 4(1), 34-41.

[3] Alladi, T., Kohli, V., Chamola, V., Yu, F. R., & Guizani, M. (2021). Artificial intelligence (AI)-empowered intrusion detection architecture for the internet of vehicles. IEEE Wireless Communications, 28(3), 144-149.

[4] Kang, J., Xiong, Z., Niyato, D., Zou, Y., Zhang, Y., & Guizani, M. (2020). Reliable federated learning for mobile networks. IEEE Wireless Communications, 27(2), 72-80.

[5] Agrawal, S., Sarkar, S., Aouedi, O., Yenduri, G., Piamrat, K., Alazab, M., ... & Gadekallu, T. R. (2022). Federated learning for intrusion detection system: Concepts, challenges and future directions. Computer Communications.

[6] Fedorchenko, E., Novikova, E., & Shulepov, A. (2022). Comparative review of the intrusion detection systems based on federated learning: Advantages and open challenges. Algorithms, 15(7), 247.

[7] Chen, M., Gündüz, D., Huang, K., Saad, W., Bennis, M., Feljan, A. V., & Poor, H. V. (2021). Distributed learning in wireless networks: Recent progress and future challenges. IEEE Journal on Selected Areas in Communications, 39(12), 3579-3605.

[8] Silva, P. R., Vinagre, J., & Gama, J. (2023). Towards federated learning: An overview of methods and applications. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 13(2), e1486.

[9] Li, Q., Diao, Y., Chen, Q., & He, B. (2022, May). Federated learning on non-iid data silos: An experimental study. In 2022 IEEE 38th International Conference on Data Engineering (ICDE) (pp. 965-978). IEEE.

[10] Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity, 2(1), 1-22.

[11] Jyothsna, V. V. R. P. V., Prasad, R., & Prasad, K. M. (2011). A review of anomaly based intrusion detection systems. International Journal of Computer Applications, 28(7), 26-35.

[12] Aydın, M. A., Zaim, A. H., & Ceylan, K. G. (2009). A hybrid intrusion detection system design for computer network security. Computers & Electrical Engineering, 35(3), 517-526.

[13] Xie, Y., Su, X., He, Y., Chen, X., Cai, G., Xu, B., & Ye, W. (2017, May). STM32-based vehicle data acquisition system for Internet-of-Vehicles. In 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS) (pp. 895-898). IEEE.

[14] Storck, C. R., & Duarte-Figueiredo, F. (2020). A survey of 5G technology evolution, standards, and infrastructure associated with vehicle-to-everything communications by internet of vehicles. IEEE access, 8, 117593-117614.

[15] Karopoulos, G., Kambourakis, G., Chatzoglou, E., Hernández-Ramos, J. L., & Kouliaridis, V. (2022). Demystifying in-vehicle intrusion detection systems: A survey of surveys and a meta-taxonomy. Electronics, 11(7), 1072.

[16] Mirzaee, P. H., Shojafar, M., Bagheri, H., Chan, T. H., Cruickshank, H., & Tafazolli, R. (2021, September). A two-layer collaborative vehicle-edge intrusion detection system for vehicular communications. In 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall) (pp. 1-6). IEEE.

[17] Inayat, Z., Gani, A., Anuar, N. B., Anwar, S., & Khan, M. K. (2017). Cloud-based intrusion detection and response system: open research issues, and solutions. Arabian Journal for Science and Engineering, 42, 399-423.

[18] Singh, A., Chatterjee, K., & Satapathy, S. C. (2022). An edge based hybrid intrusion detection framework for mobile edge computing. Complex & Intelligent Systems, 8(5), 3719-3746.

[19] Alsamiri, J., & Alsubhi, K. (2023). Federated Learning for Intrusion Detection Systems in Internet of Vehicles: A General Taxonomy, Applications, and Future Directions. Future Internet, 15(12), 403.

[20] Hu, P., Wang, Y., Gong, B., Wang, Y., Li, Y., Zhao, R., ... & Li, B. (2020). A secure and lightweight privacy-preserving data aggregation scheme for internet of vehicles. Peer-to-Peer Networking and Applications, 13, 1002-1013.

[21] Manias, D. M., & Shami, A. (2021). Making a case for federated learning in the internet of vehicles and intelligent transportation systems. IEEE Network, 35(3), 88-94.

[22] Abu Talib, M., Abbas, S., Nasir, Q., & Mowakeh, M. F. (2018). Systematic literature review on Internet-of-Vehicles communication security. International Journal of Distributed Sensor Networks, 14(12), 1550147718815054.

[23] Pei, J., Zhong, K., Jan, M. A., & Li, J. (2022). Personalized federated learning framework for network traffic anomaly detection. Computer Networks, 209, 108906.

[24] Rey, V., Sánchez, P. M. S., Celdrán, A. H., & Bovet, G. (2022). Federated learning for malware detection in IoT devices. Computer Networks, 204, 108693.

[25] Chatzigeorgiou, C., Toumanidis, L., Kogias, D., Patrikakis, C., & Jacksch, E. (2017, June). A communication gateway architecture for ensuring privacy and confidentiality in incident reporting. In 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA) (pp. 407-411). IEEE.

[26]Wan, S., Li, X., Xue, Y., Lin, W., & Xu, X. (2020). Efficient computation offloading for Internet of Vehicles in edge computing-assisted 5G networks. The Journal of Supercomputing, 76, 2518-2547.

[27] wakanma, C. I., Ahakonye, L. A. C., Njoku, J. N., Odirichukwu, J. C., Okolie, S. A., Uzondu, C., ... & Kim, D. S. (2023). Explainable artificial intelligence (xai) for intrusion detection and mitigation in intelligent connected vehicles: A review. Applied Sciences, 13(3), 1252.

[28] Schmoll, A. (2018). A Hardware-Based Secure Communication Module to Protect Internet Connected Vehicles.

[29] Xu, X., Zhang, J., Li, Y., Wang, Y., Yang, Y., & Shen, H. T. (2020). Adversarial attack against urban scene segmentation for autonomous vehicles. IEEE Transactions on Industrial Informatics, 17(6), 4117-4126.

[30] Sohaib, R. M., Onireti, O., Sambo, Y., Swash, R., & Imran, M. (2022, September). Intelligent Energy Efficient Resource Allocation for URLLC Services in IoV Networks. In 2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) (pp. 1-6). IEEE.

[31]Zamzami, I. F., Pathoee, K., Gupta, B. B., Mishra, A., Rawat, D., & Alhalabi, W. (2022). Machine learning algorithms for smart and intelligent healthcare system in Society 5.0. International Journal of Intelligent Systems, 37(12), 11742-11763.

[32]Chui, K. T., Gupta, B. B., Torres-Ruiz, M., Arya, V., Alhalabi, W., & Zamzami, I. F. (2023). A Convolutional Neural Network-Based Feature Extraction and Weighted Twin Support Vector Machine Algorithm for Context-Aware Human Activity Recognition. Electronics, 12(8), 1915.

[33]Chaudhary, P., Gupta, B., & Singh, A. K. (2022). Implementing attack detection system using filter-based feature selection methods for fog-enabled IoT networks. Telecommunication Systems, 81(1), 23-39.

[34]Colace, F., Guida, C. G., Gupta, B., Lorusso, A., Marongiu, F., & Santaniello, D. (2022, August). A BIM-based approach for decision support system in smart buildings. In Proceedings of Seventh International Congress on Information and Communication Technology: ICICT 2022, London, Volume 1 (pp. 471-481). Singapore: Springer Nature Singapore.