DT-FU: Digital Twin-Driven Federated Unlearning for Resilient Vehicular Networks in the 6G Era

Wathsara Daluwatta, Shehan Edirimannage, Charitha Elvitigala, Ibrahim Khalil, and Mohammed Atiquzzaman

Abstract

In the evolving landscape of federated learning management systems (FLMS), ensuring robust security against adversarial attacks is paramount. However, traditional methods that filter harmful updates before aggregation often fail in already compromised networks. This research introduces a pioneering approach that integrates digital twin (DT) technology with an FLMS, bolstering the network's resilience and safeguarding communication among distributed networks in the 6G era. At the heart of our methodology is the novel federated unlearning techniques, designed to mitigate the influence of malicious or poisonous clients within an already compromised network. By leveraging the DTs of clients, the system can effectively tackle such threats while ensuring the integrity of the network. Federated learning enables collaborative model training across decentralized clients, while DTs provide a virtual representation of physical entities, allowing for accurate monitoring and analysis. The proposed system enhances the resilience of networked systems against adversarial attacks, ensuring dependable and secure communication among devices and infrastructure. This methodology can be applied to applications like vehicular networks, enhancing their robustness and security in adversarial conditions. The effectiveness of the proposed system is demonstrated through real-world experiments and simulations, showcasing its potential for enhancing the security and performance of vehicular networks in dynamic environments in the 6G era.

Introduction

Today's roads are actively filled with vehicles, traffic infrastructure, and pedestrian devices constantly exchanging information. They are on the brink of a significant technological leap. The existing ecosystem of shared information, facilitated by current network technologies, receives a substantial enhancement. This advancement fully realizes the potential of autonomous driving, pushing it to new heights of efficiency and safety. The next wave of network technology will refine the interactions within vehicular networks by significantly reducing latency and increasing data throughput. This improvement aims to optimize traffic flow, improve safety measures, and reduce environmental impact by enabling more intelligent and efficient vehicle operations. It is paving the way for a future where road travel is smoother, safer, and more sustainable.

As we enter the 6G [1] era, the potential to fulfill and exceed these requirements becomes tangible. 6G technology, with its promise of ultra-high speeds, near-zero latency, and unprecedented reliability, is uniquely positioned to unlock the full capabilities of vehicular networks. This next-generation connectivity not only supports the self-sustaining and proactive facets of wireless systems but also paves the way for more efficient, safe, and environmentally friendly transportation solutions.

At the heart of utilizing the capabilities of 6G within vehicular networks [2] is the application of federated learning (FL) [3]. This decentralized machine learning approach enables the collaborative development of models across the network without compromising individual data privacy. In the context of vehicular networks, FL facilitates the real-time, adaptive management of traffic flows, safety measures, etc, learning from vast amounts of data generated across the network while ensuring that sensitive information remains within local domains.

Furthermore, deploying Digital Twins (DT) [4] within this network offers a comprehensive solution that incorporates self-sustaining capabilities and proactive analytics. DT serves as virtual replicas of physical components of vehicular networks, enabling predictive maintenance, scenario simulation, and the optimization of network resources in harmony with FL algorithms. Together, these technologies forge a path toward intelligent, efficient, and secure vehicular networks, fully leveraging the transformative potential of 6G connectivity to meet and surpass the demanding requirements of modern applications.

However, as networks become more advanced, they also become vulnerable to new forms of cyber threats, including data poisoning attacks [5]. Data Poisoning Attacks are notably harmful, leveraging the decentralized architecture of FL to compromise the model's integrity, efficacy, or privacy of its participants. These attacks, through the insertion of maliciously altered or completely fabricated data into the training sets, aim to disrupt the learning process, leading to inaccurate model predictions or

Wathsara Daluwatta, Shehan Edirimannage, Charitha Elvitigala, and Ibrahim Khalil are with RMIT University, Australia; Mohammed Atiquzzaman is with The University of Oklahoma, USA.

Digital Object Identifier: 10.1109/MCOM.001.2400229

0163-6804/25/$25.00 © 2025 IEEE

decisions. Such vulnerabilities are especially critical in the context of vehicular networks, where they can severely compromise the FL model's reliability, effectiveness, and security.

Identifying and addressing poisoning attacks early is essential to protect the overall quality of the global model in FLMS. However, traditional methods [6], which focus on detecting and removing harmful updates before aggregating them, often fall short when dealing with networks that are already compromised. This limitation highlights a significant weakness in existing security approaches within FL, particularly in areas such as vehicular networks. Developing advanced techniques to inspect and correct the influence of poisonous data is crucial for enhancing the resilience of FLMS against cyber threats. Strengthening the resilience of FL is crucial for protecting the network's reliability and ongoing operations, ensuring user safety in a digitally connected vehicular ecosystem.

This article introduces a novel framework DT-FU designed to enhance the security of FLMS by integrating Federated Unlearning. Unlike traditional federated unlearning methods that require explicit identification of problematic clients, our framework autonomously detects and mitigates the presence of harmful clients. By incorporating Federated Unlearning [7] with Digital Twins (DTs), DT-FU is ingeniously designed to integrate seamlessly with both new and existing FLMS. It offers advanced detection and remediation capabilities, effectively removing harmful effects from a compromised global model at any operational stage. DT-FU safeguards the FL process while ensuring the operational efficacy, resource management, and safety standards vital for advanced applications in vehicular networks.

Preliminaries

This section presents a concise background on the concepts of Vehicular Networks, Federated Learning, Federated Unlearning, and Digital Twin.

Vehicular Network is an advanced ecosystem of interconnected components, including Roadside Units (RSUs), smart vehicles [8], cameras, traffic lights, etc, designed to enhance road safety, efficiency, and sustainability. At the heart of its functionality and optimization lies in Artificial Intelligence, crucial for processing and analyzing the data generated within these networks. AI's role is instrumental, enabling smart vehicles to make autonomous decisions in real time, facilitating adaptive traffic light control, and aiding RSUs in managing traffic flows effectively. The integration of AI within vehicular networks marks a pivotal step forward, driving immediate improvements in traffic management and laying the groundwork for future innovations in intelligent transportation systems.

Federated Learning introduces a paradigm shift in machine learning by decentralizing the training process across a federation of distributed nodes. Unlike traditional centralized machine learning methodologies, where data aggregation and model training occur in a singular, centralized location, FL ensures that training data remains at its source. Each node in the network trains a local model on its dataset, thus preserving data privacy and reducing data centralization risks. After local training, each node computes model updates, which encapsulate the learned parameters, and transmits these updates to a central aggregator. This aggregator performs a critical function within the FL framework: it aggregates these local updates into a global model using algorithms designed to synthesize the distributed knowledge without requiring access to the raw training data. The global model, representing the collective learning outcome of all participating nodes, is then distributed back to the nodes for further local refinements. The iterative process of local training, model update aggregation, and global model dissemination continues until the global model converges to an optimal state. Convergence in this context implies that the global model has reached a state of maximal learning efficacy, as informed by the collective data and learning experiences of all nodes in the federation. Figure 1 illustrates the concept of FL in a heterogeneous vehicular network.

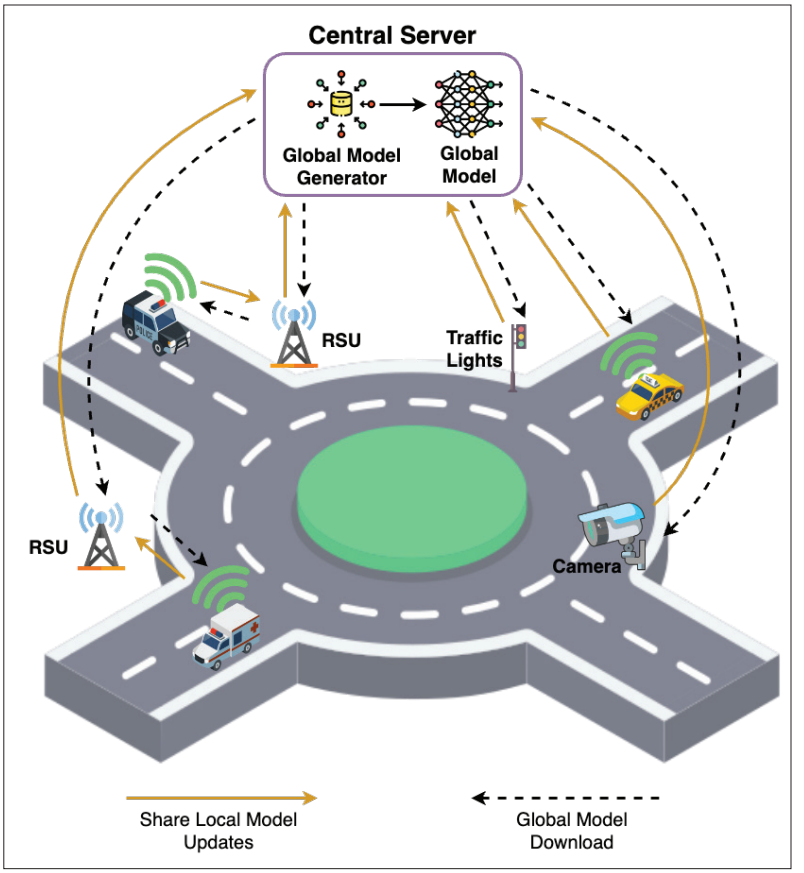

FIGURE 1. An illustration of the concept of FL in a vehicular network.

6G, the sixth generation of wireless technology, follows 5G and promises transformative benefits for communication systems with phenomenal advancements in speed, reliability, and data handling capabilities, offering even lower latency at ultra-low levels and significantly higher data rates that surpass the already impressive speeds of 5G. This advancement in technology is poised to improve the performance of FL systems. By leveraging 6G's robust infrastructure, FL systems can achieve more efficient data synchronization across vast networks of distributed nodes, such as those in vehicular networks. This ensures quicker convergence of learning models and more timely updates, which are crucial for real-time decision-making processes. Additionally, 6G will facilitate the handling

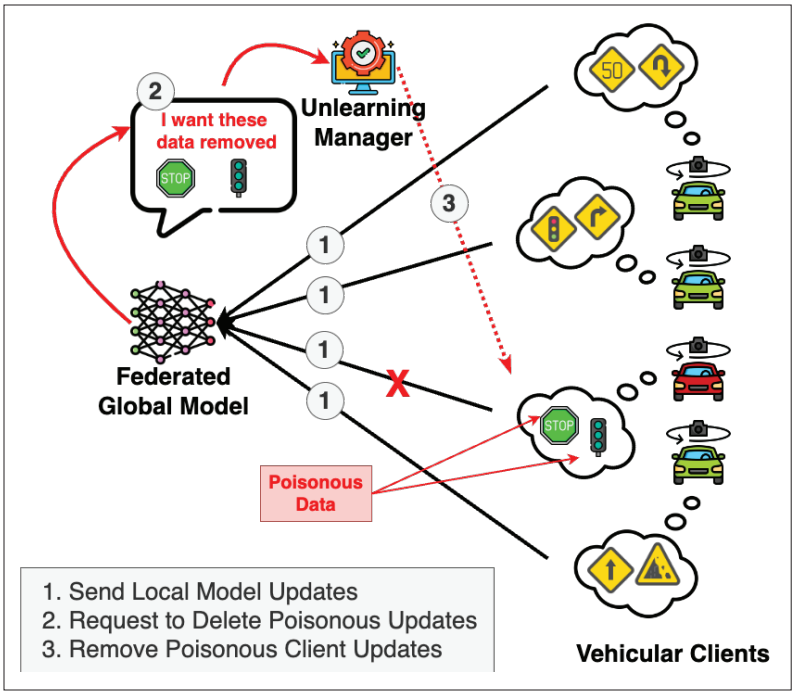

FIGURE 2. An illustration of the concept of FUL in a vehicular network.

of higher volumes of data with enhanced security measures, significantly improving the scalability and reliability of FL systems.

Federated Unlearning (FUL), an emergent concept within the domain of machine learning, particularly in the context of FL, addresses the need for dynamic and secure removal of a client's influence from a global model without compromising the integrity and performance of the system. This process is critical in scenarios where data privacy laws, such as the General Data Protection Regulation (GDPR) [9], mandate the right to be forgotten, or in instances where a client's data has been identified as malicious or corrupt. Figure 2 illustrates the concept of federated unlearning in a vehicular network.

Digital twin (DT) technology, pivotal in driving digital transformation, has attracted significant interest across industries for its ability to mirror physical objects in the real world as digital equivalent in the digital realm. Creating such virtual replicas enables the addressing of real-world challenges and the optimization of physical systems through the simulation or prediction of future scenarios. In the context of vehicular networks, the integration of digital twin technology with FL emerges as a particularly powerful combination. FL, in tandem with digital twins, allows for the collective intelligence of the network to inform the predictive analytics and simulations of digital twins. This synergy enables the optimization of traffic flows, predictive maintenance, and enhanced safety measures, all while safeguarding user data.

Data poisoning attacks in FL are a significant security concern, where malicious participants deliberately introduce false data into their updates, aiming to compromise the integrity of the global model. These attacks can be particularly insidious because they exploit the decentralized nature of FL, where data remains on the client side and is not fully visible to other participants or the central server. The attacker can subtly manipulate the training process by altering the data samples or labels ultimately leading to a degraded or biased model.

Moreover, conventional anomaly detection and robust aggregation strategies frequently struggle to address data poisoning in already compromised global models efficiently [6]. To rectify these systems, FUL can be leveraged. Existing algorithms such as FedEraser [10] and unlearning with knowledge distillation [11] require the server to retain all client weights, which is impractical in many scenarios due to the substantial storage and management overhead involved. The study in [12] introduces a novel framework that selectively forgets specific classes or categories. This method highlights the evolving need for FUL solutions to address more sophisticated requirements, such as targeted unlearning of specific data segments, rather than removing data associated with certain clients. FedRecovery [13] intensifies these storage demands by requiring the preservation of the global model from each round complicating the scalability and efficiency of FL systems. Existing FUL methods often assume the identities of malicious clients are known in advance and fail to identify and remove malicious clients adaptively. In response to these limitations, we propose a novel framework that aligns with the fundamental goals of model unlearning and refinement and eliminates the need for extensive storage. DT-FU, which utilizes a combination of gradient ascent and gradient descent techniques, does not require the server to retain all client weights. Additionally, methods predominantly focus on eliminating the influence of low-quality data from the global model [14], which might not adequately address scenarios where a client's dataset is tainted with sophisticated, hard-to-detect malicious inputs. In response to these limitations, to enhance the security and integrity of our federated learning model against data poisoning attacks, we adopt a novel framework for removing all clients that exhibit malicious behavior or have been compromised.

System Architecture

This section outlines the architecture and key components of the proposed framework (DT-FU), which comprises clients, a central server, and the Unlearning Service.

Clients

As depicted in Fig 3, the framework incorporates a vehicular network, with all components (such as RSUs, vehicles, and traffic lights) designated as clients. Each client is tasked with training a local machine-learning model using its dataset. After this training phase, clients send their model updates to a central server, aggregating them to create a cohesive global model.

Unlearning Agent (UA): The primary role of this agent is to execute the unlearning algorithm using parameters shared by the unlearning service. Specifically, this agent employs the gradient ascent method to maximize the loss associated with the client's data, thereby effectively unlearning the client's data. This process, called client-level unlearning, ensures that the learned features of the specified data are diminished within the model. Upon completion of the unlearning tasks, the unlearning agent sends the unlearned model back to the unlearning service, to integrate the changes into the FL system.

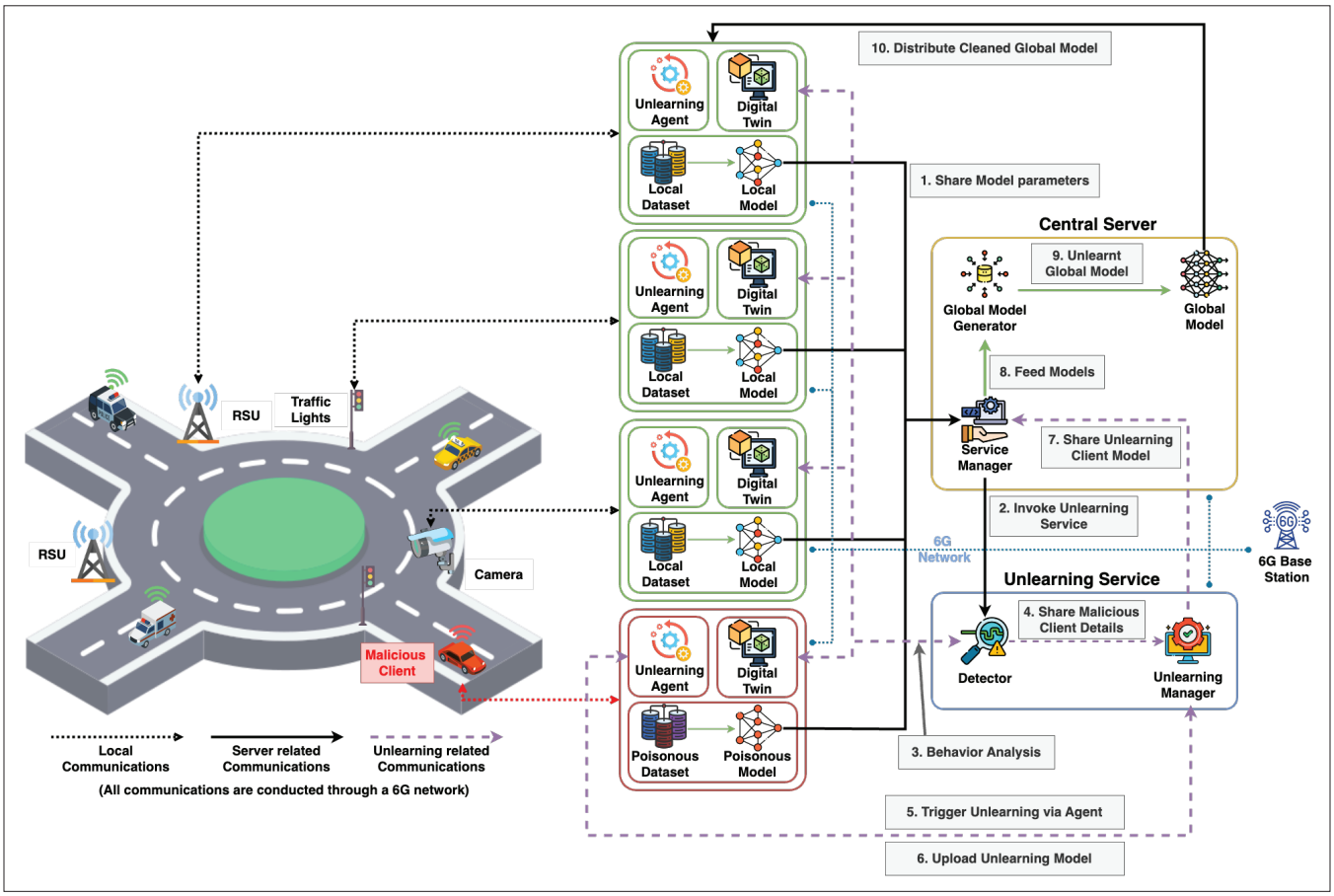

FIGURE 3. Overview of the DT-FU Framework: This illustration showcases the integration process of DTs and FUL within vehicular networks, with all communications occurring over a 6G network. It highlights the framework's communication with DTs to identify malicious clients accurately. Following identification, the framework employs FUL techniques to effectively remove these harmful influences, thereby significantly enhancing the security and reliability of the global model in a vehicular network.

Central Server

In a FL system, the central server stands as the critical nexus for model aggregation and distribution.

Model Aggregator: This function within the central server is responsible for aggregating the model updates received from various clients. It combines these updates to refine and update the global model, ensuring that the learning achieved by individual clients contributes to the collective intelligence of the system.

Service Manager: This component facilitates communication between the unlearning service. It is instrumental in managing the flow of model updates from clients to the unlearning service and ensuring that the outcomes of the unlearning process, namely the unlearned models, are accurately conveyed to the model aggregator.

Unlearning Service

Within the FL ecosystem, the Unlearning Service emerges as a pivotal component, designed with the primary objective of managing the unlearning process to enhance the security and integrity of the system:

Detector: This element of the Unlearning Service plays the crucial role of monitoring client behaviors through communication with their Digital Twins. By analyzing these interactions, the Detector identifies potentially malicious clients whose data or behavior could compromise the system's learning process or overall security.

Unlearning Manager (UM): Following the identification of malicious clients, the UM takes charge. It sets operational thresholds to govern unlearning processes, ensuring that the process halts at an optimal point to prevent over-unlearning. Additionally, it generates parameters such as learning rate, batch size, and epoch count based on the nature and extent of the client's contributions, which are crucial for the execution of the unlearning algorithm within the UA. Upon completion of this process, the UM ensures that the unlearned models are securely transmitted back to the central server, maintaining the integrity and security of the FL network.

System Overview

As illustrated in Fig. 3, the proposed framework details the process of identifying and eliminating malicious clients, with all communications between clients, the central server, and the unlearning processes occurring over a 6G network. The following steps further explain the framework's approach to maintaining security within the FL ecosystem:

1. To start, clients get a global model from the FL Management System. Then, they modify this model with their own data sets, which makes it better at handling their own data needs. After all the clients have trained their models, they send them to the central server, where they are combined to make a better model.

In the evolving landscape of vehicular networks, image classification emerges as a cornerstone application, leveraging machine learning to revolutionize how vehicles understand and interact with their surroundings. From distinguishing traffic signs to identifying potential hazards on the road, image classification serves as the digital eyes for autonomous and connected vehicles, making navigation safer and more efficient.

2. The central server's Service Manager plays a proactive role by collaborating with the unlearning service detector. Using the DTs of the clients, this intelligent scanner utilizes anomaly detection techniques to monitor behavior patterns and detect potential threats from clients. Specific behaviors analyzed include unusual patterns in model performance trends, Inconsistencies in learning rates, and inconsistencies in client updates compared to established baselines. These behaviors are analyzed to identify any clients potentially compromised with bad data. By doing so, it prevents any malicious actions from infiltrating the system unnoticed.
3. After identifying the malicious client, the automated detective system notifies the UM, allowing for a tailored and effective response to address the specific issues caused by these malicious clients.
4. Then the UM steps in, commanding a UA nestled within the target clients to reverse all changes. This agent meticulously performs client-level unlearning to reverse the influence of the client's data and returns a cleansed, unlearned model to the central server via UM.
5. After receiving the client-level unlearned model, it is proactively incorporated into the global model by a centralized Global Model Aggregator. This critical step reverses the adverse effects of malicious data and enhances the global model's integrity.
6. Finally trustworthy clients engage in a few more additional training rounds with the comprehensive model. This phase is critical for reinforcing the model's foundation, significantly enhancing its overall effectiveness.

Results

Example Application

In the evolving landscape of vehicular networks, image classification emerges as a cornerstone application, leveraging machine learning to revolutionize how vehicles understand and interact with their surroundings. From distinguishing traffic signs to identifying potential hazards on the road, image classification serves as the digital eyes for autonomous and connected vehicles, making navigation safer and more efficient.

Dataset and Neural Network Structure: To showcase the effectiveness of DT-FU, we established an experimental setup focusing on three diverse datasets: MNIST, Fashion-MNIST, and CIFAR10. MNIST and Fashion-MNIST are datasets used for digit and fashion item classification, respectively. Both datasets feature 28x28 pixel grayscale images. We employ the LeNet-5 model for these datasets for its efficiency in handling similar input sizes. The LeNet-5 model includes two convolutional layers and two pooling layers, followed by three fully connected layers, using tanh activations and optimized with stochastic gradient descent (SGD). CIFAR-10, consisting of 32x32 pixel color images in 10 classes, demands a more complex model due to its higher complexity and color information. We use the VGG-11 model, which features eight convolutional layers with small 3x3 filters and three fully connected layers. This model uses ReLU activations and is also optimized with SGD.

This experimental setup showcases the versatility and adaptability of our framework, highlighting its potential to enhance image classification tasks within vehicular networks. Through rigorous testing, we demonstrate how our framework advances machine learning applications in the automotive domain, paving the way for safer and smarter vehicular technologies.

Experimental Setup

In our study, we designed an experimental setup to illuminate FL's practical applications and effectiveness within a simulated environment. In our setup, we simulated the realistic and unpredictable nature of data generation in vehicular networks, we randomly distributed the data among clients in a manner that ensures no two nodes receive identical datasets. This approach accounts for node activity and availability variations, mirroring real-world scenarios. Additionally, to examine the effects of potentially compromised nodes, we introduce a specific pattern sized 3×3 pixels into the dataset of a selected client to create a "backdoor." This manipulation, facilitated using the Adversarial Robustness Toolbox [15], a reputable toolkit for evaluating the security of machine learning models against adversarial threats creates a backdoor that renders the global FL model vulnerable to manipulated data inputs, thereby simulating a real-world adversarial scenario.

A pivotal aspect of our experiment was the creation of digital twins for each client, modeled using Python and PyTorch. To introduce a challenge akin to encountering a security threat, we intentionally modified a subset of the data for one randomly selected client. This modification involved embedding backdoor images into the dataset, distinguished by a specific pattern, to simulate a potential attack vector within the FL process. This alteration allowed us to detect and analyze unusual patterns in model performance trends, enabling a detailed behavior analysis to identify and understand anomalies that could indicate potential security vulnerabilities.

Throughout the simulation, our FL system was actively engaged in training, with a keen focus on the behavior analysis facilitated by the digital twins. This continuous monitoring allowed us to pinpoint the malicious client harboring the backdoor images. Upon identifying the compromised client, we initiated the unlearning process. This crucial step aimed to surgically excise the influence of the backdoor images from the global model, effectively neutralizing the threat they posed.

Comparison Method

To validate the effectiveness of our proposed service, we conducted comprehensive experiments comparing it against established methods. Retrain [10], where the global model is trained from scratch excluding the malicious client, serves as a baseline. Federated Averaging (FedAvg), which synthesizes a global model by aggregating local updates, is employed to demonstrate the impact of data removal. Additionally, we use FedEraser and FedRecovery to assess the unlearning efficiency of our method. These well-known FUL methods help illustrate the robustness and efficiency of our approach in real-world scenarios.

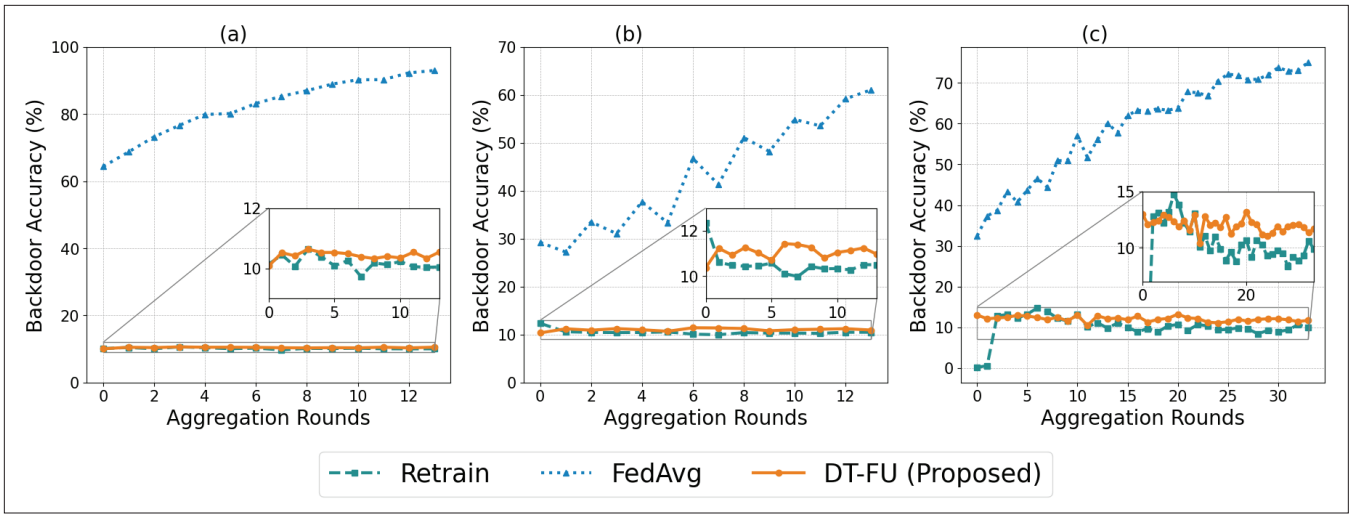

FIGURE 4. Backdoor Accuracy vs. Aggregation Rounds: Illustrating the Effectiveness of the DT-FU. This figure compares three scenarios: (1) the FedAvg algorithm demonstrating how the backdoor accuracy increases if the FL system retains the malicious client, (2) Retraining which establishes the baseline by showing the backdoor accuracy without the malicious client, & (3) DT-FU which effectively restores the backdoor accuracy to match the baseline level. The comparison highlights DT-FU's capability to neutralize the impact of malicious influences and align the system's performance with secure baseline conditions: a) MNIST; b) Fashion-MNIST; c) CIFAR-10.

Evaluation and Comparison Metrics

In the domain of FL, particularly when integrating an FUL, the dual objectives of removing targeted malicious influences without affecting the model's overall learning achievements are paramount. To assess our framework's effectiveness, we used two key metrics known for their relevance and insight.

Backdoor Accuracy: This metric showcases the model's performance in handling adversarial data that has been subtly altered by an adversary to manipulate model predictions toward a specific, erroneous target label. The evaluation of backdoor accuracy is instrumental in understanding the extent to which the model has successfully "Forgotten" the manipulations introduced by the adversary. A decline in backdoor accuracy indicates effective FUL, showcasing the model's resilience in neutralizing the intended manipulation and thereby reducing its vulnerability to such attacks.

Clean Accuracy: Beyond the adversarial resilience, it is essential to maintain the model's overall accuracy, which is the ability to predict or classify across the entire dataset. This metric evaluates the global model's performance and its iteration following the unlearning process. After FUL, a stable or improved clean accuracy signifies the ability to selectively eliminate harmful data influences without compromising the model's general learning capabilities.

Experimental Results

Our experimental findings, illustrated in Fig 4, underscore the efficacy of the proposed framework in mitigating backdoor vulnerabilities within the FL setup. Specifically, we observed that DT-FU can reduce the backdoor accuracy to baseline levels, similar to those seen in scenarios where the model undergoes retraining. This is a significant achievement, highlighting the framework's capacity to neutralize adversarial influences effectively. Moreover, the comparison with the FedAvg method reveals, that without the unlearning process, the global model's vulnerability to malicious patterns increases, as evidenced by the rise in backdoor accuracy.

Figure 5 demonstrates the superiority of DT-FU over the conventional retraining method in terms of efficiency and effectiveness in restoring clean accuracy. In contrast to the traditional retraining approach, which typically necessitates a prolonged series of aggregation rounds to achieve high accuracy, DT-FU capitalizes on the existing global model. By precisely adjusting it to diminish the influence of the malicious client, our method significantly enhances clean accuracy within fewer aggregation rounds than retraining. This not only underscores the effectiveness of our algorithm but also emphasizes its efficiency in utilizing computational resources.

Figure 6 showcases a clear representation of the performance comparison on the DT-FU method across two datasets. The FedAvg method exhibits high clean accuracy but is notably susceptible to backdoor attacks. FedRecovery offers significant security improvements over FedAvg by reducing backdoor accuracy. However, it achieves slightly lower clean accuracy compared to the other strategies. FedEraser also shows comparable performance in clean accuracy and an effective reduction in backdoor vulnerabilities. Our DT-FU framework maintains clean accuracy close to the Retrain baseline and excels in minimizing backdoor risks better than both FedEraser and FedRecovery. This comprehensive visual comparison underscores DT-FU's capability to sustain high accuracy and significantly improve security measures, thereby advancing the protective measures within FL systems.

Collectively, these results highlight the proposed system's dual strengths: its adeptness at excising adversarial influences and resource efficiency. By ensuring the learning model's integrity and optimizing computational expenditure, DT-FU stands out as a robust solution for maintaining system security and operational efficiency in the face of adversarial threats in FL environments.

Impact of 6G

The advent of 6G technology promises a transformative shift, primarily by significantly reducing com-

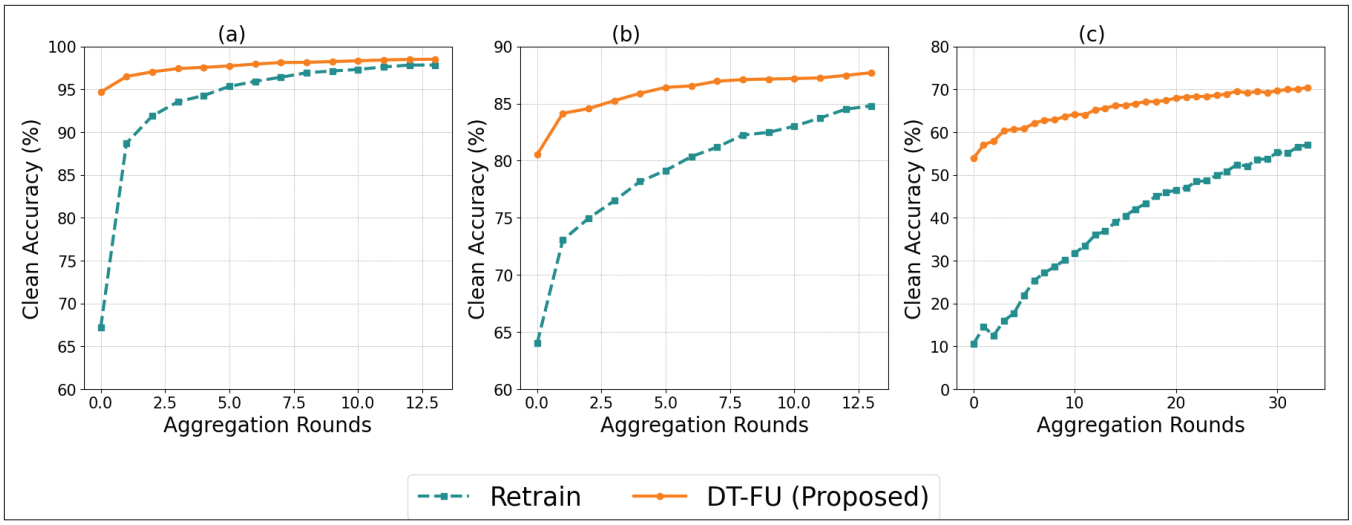

FIGURE 5. Clean Accuracy vs. Aggregation Rounds: Demonstrating the Efficiency of the Proposed Framework. This figure displays the starting accuracy of our proposed framework immediately after the aggregation of the client-level unlearned models. It effectively illustrates the efficiency with which the global model converges in subsequent training rounds, showcasing the framework's capability to achieve high accuracy levels across multiple aggregation rounds rapidly: a) MNIST; b) Fashion-MNIST; c) CIFAR-10.

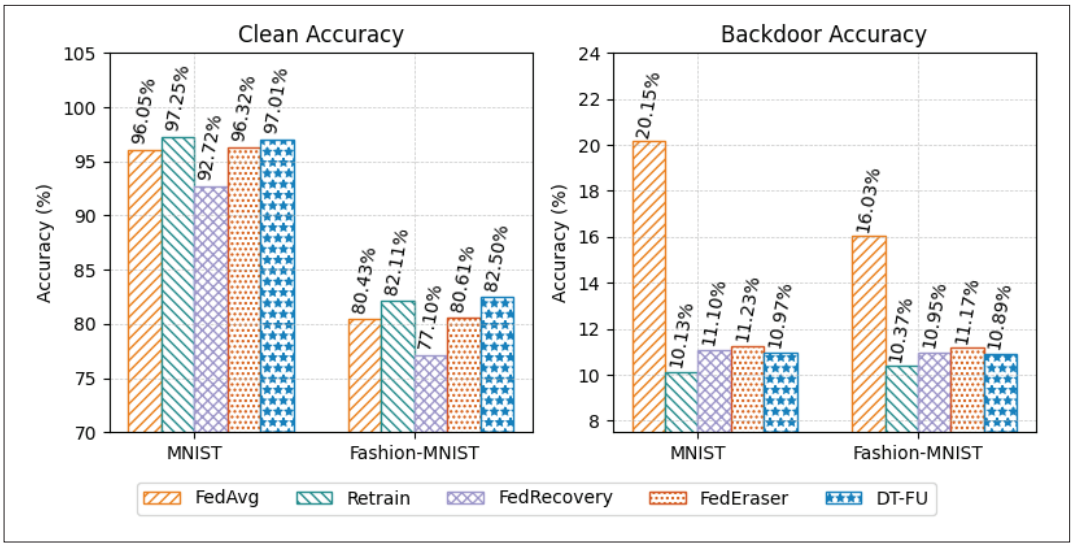

FIGURE 6. Comparative Analysis of the DT-FU Against FedAvg, Retrain, FedEraser, and FedRecovery.

munication costs. This reduction facilitates quicker detection of malicious activities and accelerates the FUL processes necessary for maintaining model integrity. Notably, 6G is expected to enhance several key features that are critical for vehicular networks: increased bandwidth and lower latency will enable faster communication, allowing for more rapid identification and removal of malicious clients. Furthermore, the enhanced reliability offered by 6G minimizes transmission errors and connection drops improve the efficient convergence of FL models. The integration of 6G significantly boosts real-time performance capabilities, enhancing the overall effectiveness of the systems. This enables quicker responses to security threats and more robust protection in Vehicular Networks.

Conclusion and Future Works

This article presented a novel framework (DT-FU) that leverages FUL within vehicular networks to cleanse the global model of distortions introduced by malicious entities. Central to our framework is the innovative use of digital twins and the integration of 6G technology, which provide a robust foundation for securing and optimizing vehicular networks. Through empirical analysis and real-world simulations, our research has demonstrated the framework's effectiveness in neutralizing data poisoning attacks, a prevalent threat in distributed learning environments. By addressing and mitigating adversarial impacts, DT-FU marks a significant advance in applying vehicular networks to remain resilient and secure.

In future developments of our DT-FU framework, we aim to enhance its effectiveness and adaptability in FL environments. A key area for enhancement is the refinement of our FUL algorithm to utilize DT more effectively. Currently, DT-FU uses DTs to identify malicious clients. Moving forward, we plan to evolve this approach by enabling the DTs to perform unlearning processes autonomously, without direct interaction with the actual clients. This capability will be crucial for safeguarding against emerging cyber-attacks in a manner that minimizes the risk of compromising client operations or data integrity.

Acknowledgment

This work is supported by the Australian Research Council Discovery Project (DP220100215).

References

[1] M. Giordani *et al.*, "Toward 6G Networks: Use Cases and Technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, 2020, pp. 55–61.
[2] J. Posner *et al.*, "Federated Learning in Vehicular Networks: Opportunities and Solutions," *IEEE Network*, vol. 35, no. 2, 2021, pp. 152–59.
[3] T. Li *et al.*, "Federated Learning: Challenges, Methods, and Future Directions," *IEEE Signal Processing Mag.*, vol. 37, no. 3, 2020, pp. 50–60.
[4] X. Lin *et al.*, "6G Digital Twin Networks: From Theory to Practice," *IEEE Commun. Mag.*, vol. 61, no. 11, 2023, pp. 72–78.
[5] G. Xia *et al.*, "Poisoning Attacks in Federated Learning: A Survey," *IEEE Access*, vol. 11, 2023, pp. 10,708–22.
[6] A. Uprety and D. B. Rawat, "Mitigating Poisoning Attack in Federated Learning," *Proc. 2021 IEEE Symposium Series on Computational Intelligence*, 2021, pp. 01–07.
[7] L. Wu *et al.*, "Federated Unlearning: Guarantee the Right of Clients to Forget," *IEEE Network*, vol. 36, no. 5, 2022, pp. 129–35.
[8] W. Collier and R. Weiland, "Smart Cars, Smart Highways," *IEEE Spectrum*, vol. 31, no. 4, 1994, pp. 27–33.
[9] B. Custers *et al.*, E*U Personal Data Protection in Policy and Practice, 1st ed.*, ser. Information Technology and Law Series, T.M.C. Asser Press, 2019; available: https: //www.springer.com/gp/book/9789462652811.
[10] G. Liu *et al.*, "Federaser: Enabling Efficient Client-Level Data Removal From Federated Learning Models," Proc. 2021 *IEEE/ACM 29th Int'l. Symposium on Quality of Service*, 2021, pp. 1–10.
[11] C. Wu, S. Zhu, and P. Mitra, "Federated Unlearning With Knowledge Distillation," ArXiv, vol. abs/2201.09441, 2022.
[12] J. Wang *et al.*, "Federated Unlearning via Class-Discriminative Pruning," *Proc. ACM Web Conf. 2022*, ser. WWW '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 622–32; available: https://doi.org/10.1145/3485447.3512222.
[13] L. Zhang *et al.*, "Fedrecovery: Differentially Private Machine Unlearning for Federated Learning Frameworks," *IEEE Trans. Information Forensics and Security*, vol. 18, 2023, pp. 4732–46.
[14] P. Wang *et al.*, "Server-Initiated Federated Unlearning to Eliminate Impacts of Low-Quality Data," *IEEE Trans. Services Computing*, vol. 17, no. 03, may 2024, pp. 1196–1211.
[15] M.-I. Nicolae *et al.*, "Adversarial Robustness Toolbox v1.2.0," *CoRR*, vol. 1807.01069, 2018; available: https://arxiv.org/pdf/1807.01069.

Biographies

Wathsara Daluwatta is a Ph.D. candidate at RMIT University in Melbourne, Victoria, Australia.

Shehan Edirimannage is currently pursuing a Ph.D. at RMIT University in Melbourne, Victoria, Australia.

Charitha Elvitigala is currently pursuing a Ph.D. at RMIT University in Melbourne, Victoria, Australia.

Ibrahim Khalil is a Professor at RMIT University in Melbourne, Australia. His research interests include Federated Learning, Privacy, Blockchain Secure AI & Data Analytics.

Mohammed Atiquzzaman [SM], Edith Kinney Gaylord Presidential Professor at the University of Oklahoma, serves as Editor-in-Chief of the *Journal of Networks and Computer Applications and Vehicular Communications*. He is also involved with various IEEE journals and has co-chaired numerous IEEE international conferences, including IEEE Globecom.