Received: 10 May 2022 | Revised: 21 June 2022 | Accepted: 12 July 2022

DOI: 10.1111/exsy.13103

ORIGINAL ARTICLE

Expert Systems WILEY

Privacy-preserving federated learning cyber-threat detection for intelligent transport systems with blockchain-based security

Tarek Moulahi[1] | Rateb Jabbar[2] | Abdulatif Alabdulatif[3] | Sidra Abbas[4] | Salim El Khediri[1] | Salah Zidi[5] | Muhammad Rizwan[6]

[1]Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia

[2]Department of International Affairs, College of Arts and Science, Qatar University, Doha, Qatar

[3]Department of Computer Science, College of Computer, Qassim University, Buraydah, Saudi Arabia

[4]ASET LAB, Islamabad, Pakistan

[5]Institut Superieur des Systemes Industriels de Gabes (ISSIG), Gabes University, Gabes, Tunisia

[6]Secure Cyber Systems Research Group, WMG, University of Warwick, Coventry, UK

Correspondence
Muhammad Rizwan, Secure Cyber Systems Research Group, WMG, University of Warwick, Coventry CV4 7AL, UK.
Email: muhammad.rizwan.1@warwick.ac.uk

Abstract

Artificial intelligence (AI) techniques implemented at a large scale in intelligent transport systems (ITS), have considerably enhanced the vehicles' autonomous behaviour in making independent decisions about cyber threats, attacks, and faults. While, AI techniques are based on data sharing among the vehicles, it is important to note that sensitive data cannot be shared. Thus, federated learning (FL) has been implemented to protect privacy in vehicles. On the other hand, the integrity of data and the safety of aggregation are ensured by using blockchain technology. This paper applied classification approaches to VANET and ITS cyber-threats detection at the vehicle. Subsequently, by using blockchain and by applying an aggregation strategy to different models, models from the previous step were uploaded in a smart contract. Lastly, we returned the updated models to the vehicles. Furthermore, we conducted an experimental study to measure the effectiveness of the proposed prototype. In this paper, the VeReMi data set was distributed in a balanced manner into five parts in the experimental study. Thus, classification techniques were executed by each vehicle separately, and models were generated. Upon the aggregation of the models in blockchain, they were returned to the vehicles. Lastly, the vehicles updated their decision functions and accessed the precision and accuracy of cyber-threat detection. The results indicated that the precision and accuracy decreased by 7.1% on average with comparable *F*1-score and recall. Our solution ensures the privacy preservation of vehicles whereas blockchain guarantees the safety of aggregation technique and low gas consumption.

KEYWORDS

blockchain technology, cyberthreat, data privacy, federated learning, intelligent transport systems, VANET

1 | INTRODUCTION

It is estimated by Fortune Business Insights (2022) that the global smart transportation market size will significantly increase in size over the 2021–2028 period. More precisely, the size of this market is anticipated to increase from USD 98.74 billion to USD 206.80 billion worldwide. In particular, the Intelligent Transportation System Application (ITS) segment is expected to increase particularly. Already in 2020, this segment

dominated the market with a share of more than 30%, and this domination is likely to continue. The ITS refers to the traffic management solutions aimed at improving traffic flows and mobility. To be more precise, these solutions are based on real-time data that are analysed with an aim to promptly respond to emergencies. It is necessary to implement effective traffic management systems because long travel times and related fuel consumption lead to traffic congestion, which in turn, results in considerable financial losses. In addition, the ITS also contributes to better the operational performance and reliability of road networks. Furthermore, the initially proposed vehicular ad-hoc networks (VANETs) (Zeadally et al., 2012) were substantially changed through cutting-edge computation and communication technologies into the Internet of Vehicles (IoV) (Obaidat et al., 2020). The IoV interconnects smart vehicles on the Internet and consequently enables the functioning of ITS. As reported by the U.S. Department of Transportation (DOT) (Lamssaggad et al., 2021), the IoV has been proven to be especially efficient in reducing crashes committed by unimpaired drivers. It is estimated that IoV can help avoid approximately 79% of such crashes because it establishes the effective communication and collaboration among vehicles as well as the interconnection of roadside infrastructure, pedestrians, and bicycles. The exchange of messages on traffic conditions, accidents, and safety can not only resolve traffic jams and reduce accident rates but also contribute to a reduction in environmental pollution. Consequently, overall safety, comfort, and convenience are improved.

Nevertheless, the integration of intelligent transportation systems (ITS) and consequent high connectivity has resulted in considerable challenges such as how to guarantee the security of all participants in intelligent transportation systems (Hassan et al., 2022; Rehman Javed et al., 2020) because it is prone to malicious attacks. Furthermore, the exchange of sensitive data can lead to difficulties in ensuring privacy. VANETs have experienced numerous attacks such as Man-In-The-Middle (MITM) which have resulted in compromised safety (Ahmad et al., 2017; Aloqaily et al., 2019; Javed et al., 2020, 2021; Mittal et al., 2021; Sheikh et al., 2019). To illustrate, Hunt (2018), successfully controlled features of Nissan LEAFs against an attack in 2016. Subsequently, in June 2016, a Mitsubishi Outlander PHEV (2016), via a smartphone application with remote control through the hacked the controlling of non-driving essential components of the vehicle systems (e.g., to lock or unlock doors, to turn on/off air conditioning, and to change charger settings). In addition, Keen Security Lab, a Chinese cybersecurity company, managed to hack Tesla (Finkle, 2016). The scientists were able to control the components of Model S Tesla (e.g., mirrors and brakes) from a distance of 20 km away.

These examples demonstrate that it is necessary to propose more complex solutions are needed to improve data transmission. For instance, the payment systems must be enhanced to become efficient, secure, and integrated (Naeem et al., 2021). Following the introduction of Bitcoin, blockchain technology (Nofer et al., 2017) has substantially changed digital currencies (Reyna et al., 2018). Blockchain represents a distributed ledger keeping an immutable log of transactions taking part in a network. Blockchain can be used in our context to ensure the safety of decisions performed by a machine learning algorithm.

The convergence of machine learning and BIoV is likely to lead to novel services and applications that have been enabled to learn from training data, reach data-driven conclusions, assess the improvement of the network performance, and guarantee decision support (Jabbar et al., 2022). However, the application of machine learning in mobile networks (e.g., vehicular networks) is primarily limited because of data that takes the form of isolated islands. Consequently, the integration of different data sources in vehicular networks is difficult as it requires increasing the number of vehicles, which in turn increases the data size and requires heavy computation (Du et al., 2020). Google developed 'federated learning' (FL) (Hamdi et al., 2021; Qi et al., 2021) in 2016 to establish data models distributively and solve this problem. In this type of architecture, data are kept on participating devices (e.g., smartphones or vehicles). The devices then download and train the model using their data.

In this way, they are going to learn from training data and reach data-driven conclusions, ensure decision support, and estimate how to improve the network performance (Jabbar et al., 2022). In this paper, we use FL to protect vehicles' privacy while blockchain is used to ensure the integrity of data and the safety of aggregation (Rehman et al., 2022). We apply classification approaches for VANET and ITS cyber-threats detection at the vehicle. Then, we upload each model, resulting from the previous step, in a smart contract using blockchain to apply an aggregation strategy between different models. Finally, the updated models are sent back to the vehicles.

The rest of this paper is organized as follows: Section 2 outlines the related work. Section 3 gives an overview of the problem. Section 4 presents the proposed framework. Section 5 shows and discuss the experimental result of the proposed method. The conclusion is given in Section 6.

2 | RELATED WORK

We subdivide this section into two subsections. First, we discuss first the adaptation of blockchain technology in ITS and VANET. Next, we outline the use of FL to deal with some issues in VANET and ITS.

2.1 | Blockchain technology for ITS and VANET

It is paramount to guarantee security in vehicular networks and to prevent adverse impacts on the users. To illustrate, it is likely that cyberattacks, hacker threats of hackers, and failures in ensuring security would lead to financial losses, roads accidents, and vehicle immobilization. Moreover,

disclosure of sensitive data may result in endangering the safety of road users. Accordingly, numerous studies have focused on enhancing credibility, transparency, identification, accessibility, integrity, confidentiality, immutability, reputation, resilience to attacks, trust, authentication, anonymity, and privacy. Yang et al. (2019) developed a highly credible and trustworthy distributed blockchain-based platform for vehicular systems that is highly credible and trustworthy. In this solution, vehicles use a Bayesian inference model to validate messages from vehicles in the vicinity. Depending on the validation results, messages from source vehicles are rated. In addition, roadside units (RSUs) assess the trust value offsets of participating vehicles, which are subsequently stored in blocks. While this solution is aimed at ensuring privacy and security, and its performance is assessed in terms of these properties, cost and execution time are not taken into account. Furthermore, Li et al. (2018) investigated Creditcoin. Its architecture is based on the privacy-preserving announcement. Accordingly, identities are preserved during the message broadcasting. The incentive mechanism based on blockchain and anonymous message aggregation protocol enables the users to sign and forward messages. It is significant that Bitcoin, which is typically employed for purchasing and selling commodities in a marketplace that is pseudo-anonymous and secure, was implemented into this solution. Therefore, Bitcoin cannot be employed to generate smart contracts and programing features that can be used to resolve computational problems and allow the transferring of sensitive data. In addition, Yang et al. (2017) used blockchain to create a novel reputation system aimed at assessing data credibility. This solution enables cars to assess the transmitted data by observing the system. Subsequently, the shared data receives a value kept in the block. The reputation of the sender and the credibility of the data are verified by vehicles on the basis of the score. Malik et al. (2018) created a smart car authentication and revocation framework employing blockchain to update the status of participating vehicles. Accordingly, when compared with solutions based on a trusted-authority central architecture, the cost of processing and communication is minimized. Furthermore, Labrador and Hou (2019) employed blockchain technology to ensure and enhance the authentication of vehicle identity and shared data. This solution includes a mechanism that can identify data and vehicles by analysing transmitted data packets. In addition, Reimers et al. (2019) created a prototype based on blockchain. This prototype is aimed at ensuring transparency in a system based on the IoT. This solution includes a traceability information system that gathers data (e.g., sensor data from IoT devices). Moreover, end-users can use the data. Mostafa (2019) developed a trustworthy vehicle platform based on blockchain to identify malicious and misbehaving vehicles. In this way, vehicles decide if other vehicles are trustworthy by validation of packets, VANET environmental data, and how the participating vehicles create and share blocks. In Zhang et al. (2014), this platform was able to identify Sybil attacks. Rathee et al. (2019) also proposed a transparent and secure framework for CAVs using blockchain for extracting and storing data. This solution minimizes the risk of exchanging information with fake participants, receiving corrupted data, and compromising vehicles' smart sensors. Noh et al. (2020) used an authentication mechanism based on blockchain with a message authentication code and asymmetric keys to authenticate messages for VANET and enhance privacy. PoW and PBWT were employed to ensure a message authentication consensus. Furthermore, Nadeem et al. (2019) developed a fog node-based distributed blockchain cloud architecture scheme to improve vehicular data privacy and prevent attacks. This solution allows for the efficient management of enormous amounts of created data owing to the extraordinary computational performance at the edge of the network. Successful vehicular systems require secure communication among vehicles. Thus, Singh and Kim (2018) used blockchain to develop a trustworthy, smart communication protocol for a cloud vehicular system. An incentive mechanism was applied to improve the trustworthiness of the notes. Consequently, the system achieved successful exchanges.

Furthermore, Jabbar, Fetais, et al. (2020) used blockchain to create a Decentralized IoT Solution for Vehicles Communication (DISV). More precisely, the DISV (Jabbar, Krichen, Kharbeche, et al., 2020) is a real-time application specification created to ensure secure communication among all transportation system users. Considering the limitations, it is not possible for this solution to be used by many vehicles in the vicinity, and accordingly, it lacks scalability. The scalability problem was resolved by via a blockchain-based solution that ensures secure payment and communication (PSEV) (Jabbar et al., 2022; Rateb, 2021). Nkenyereye et al. (2020) created a hybrid 5G and cloud vehicle network aimed at ensuring the privacy of users' sensitive information and establishing a communication protocol for warning messages. Furthermore, Singh et al. (2020) proposed a hybrid architecture to ensure vehicle-to-infrastructure and vehicle-to-vehicle communications. In addition, Kumar et al. (2020) developed for a blockchain-based vehicular network an access control scheme hash-based storage for managing traffic data. There is an innovative, emerging market of smart transport applications for vehicular systems and the IoV. Various APIs have been developed and tested by researchers which are aimed at providing in-vehicle entertainment, ensuring traffic safety, and preventing congestion. In addition, different mobility services have been created to locate, unlock, and read the odometers of cars across brands. Researchers have significantly contributed to the development of transport applications for vehicle systems. These solutions primarily comprise the IoV security layer to ensure privacy and security. Various scientists have developed platforms based on blockchain for establishing trusted multiparty insurance (Demir et al., 2019), training and learning autonomous cars (Gandhi & Salvi, 2019), car leasing (Obour Agyekum et al., 2018), and smart car parking services (Hu et al., 2019; Jabbar, Fetais, et al., 2020; Jabbar, Krichen, Shinoy, et al., 2020; Jabbar et al., 2021). Moreover, Masoud et al. (2019) developed the transparent dissemination of the users' motor history for trading, whereas Zhang et al. (2019) used blockchain to ensure secure purchasing and selling of second-hand vehicles.

2.2 | Blockchain-based FL for ITS and VANET

The purpose of introducing blockchain to FL is to respond to specific limitations. To illustrate, the problem of the dependency of the resiliency of an aggregator being dependent on the robustness of the centre of charge of the FL network can be overcome. Similarly, vulnerability to malicious

clients is reduced, especially for those clients that attach to the network via poisonous models. Lu et al. (2020) used FL to create a solution for reducing the transmission load and ensuring users privacy through a hybrid blockchain architecture. This solution includes the local directed acyclic graph and allows blockchain to enhance model parameters' security and reliability. This solution also selected the node using DRL and accordingly created an asynchronous FL scheme to enhance efficacy. The integration of learned models into the blockchain system and the implementation of a two-stage verification can ensure the reliability of shared data. Furthermore, the numerical results demonstrated that faster convergence and higher learning accuracy are achieved through this data-sharing scheme. However, this paper did not address critical performance metrics in machine learning, including *F*1 score, sensitivity, and precision. In addition, Chai et al. (2021) created a hierarchical FL algorithm and a hierarchical blockchain framework allowing knowledge sharing. Machine learning methods are employed in this process to allow vehicles to learn environmental data and share the learned knowledge. Shared data can be reliable if learned models are incorporated in blockchain and if a two-stage verification is implemented. Furthermore, it was data-sharing scheme that can improve learning accuracy and convergence. Privacy requirements and the distributed pattern of IoVs can be met using the hierarchical FL algorithm. Moreover, it is possible to apply this hierarchical blockchain framework to vehicular networks at a large scale. The proposed solution was demonstrated to have an approximately 10% higher accuracy than conventional FL algorithms. Nevertheless, it is still necessary to explore the overhead and transaction throughput of the proposed hierarchical blockchain framework. In addition, Otoum et al. (2020) integrated blockchain and FL and created an innovative solution that ensures data privacy and network security maintenance. On end devices, this framework decentralizes the mutual machine learning models. A consensus solution based on blockchain is employed as a second line of protecting privacy to make sure that the shared cloud training can be trusted. Centralized coordination and training of data coordination are not required in this model for enabling in this model because it is conducted via a consensus method in blockchain. Nevertheless, it is necessary to perform more tests must be performed which take into account various FL algorithms.

Abdel-Basset et al. (2022) developed a federated deep learning-based intrusion detection framework (FED-IDS). In this framework, the learning process is offloaded from servers to distributed vehicular edge nodes; and consequently, attacks are detected efficiently. A context-aware transformer network is incorporated into FED-IDS. This network can classify various attack categories by learning spatial–temporal representations of vehicular traffic flows. However, the interpretability of classification decisions in IDS as an essential requirement for security communities, needs to be investigated in Smart Transportation Systems.

Maaroufi and Pierre introduced the BCOOL (Maaroufi & Pierre, 2021). This solution includes innovative, hybrid, and dynamic blockchain Fog-based Distributed Trust Contract Strategy (FDTCS). Its purpose is to manage massages and assess the trustworthiness of vehicles. According to the simulation results, the BCOOL significantly outperforms similar strategies. More precisely, it is 80% more reliable and 100% more efficient in terms of responding to data congestion environments.

Cui et al. (2018) developed an edge-computing concept to enhance message authentication efficiency. In the proposed solution, the authentication tasks of vehicles are shared by roadside units (RSU) and edge computing vehicles (ECVs). Accordingly, the burden on the RSU is reduced. Alfadhli et al. (2020) confirmed that CPPA schemes developed by Cui et al. (2018) could not provide enough driving safety for vehicles in critical areas. In addition, vehicles are vulnerable considering VANET's management, communication, and computational efficiency. Furthermore, an SD2PA authentication scheme using a hash function was proposed to prevent non-safe driving and to promote safer driving while also preserving the vehicles' security and privacy. Blockchain acted as a Trusted Authority (TA) enabling the ledger management to store information about the vehicle. Furthermore, using the intermediate node, vehicles must carry out mutual authentication with the TA. Nevertheless, this study highlighted low computing overhead (Xu, Liang, et al., 2021). Kim et al. (2021) focused on studying common centralized protocols, including request/reply and publish/subscribe, and decentralized P2P computing in constrained IoT environments. First, the authors introduced cost modelling of network protocols aimed at analysing the basic network performance. Furthermore, a comparative analysis considering push, pull, and P2P-based IoT data transmission approaches was conducted. In this experiment, latency and availability created by the request message and determined factors influencing different IoT networks were evaluated. Xu, Li, et al. (2021) proposed a computationally efficient and safe key agreement scheme and authentication for the IoV, allowing the RSU to authenticate vehicles. The authors conducted the analysis using the simulation tool ProVerif and the Real-or-Random model and confirmed that the proposed scheme was secure. In comparison to previous schemes, this one enhanced authentication efficiency and reduced energy consumption. This scheme, however, took into account only the key agreement and authentication between the vehicle and the TA. Therefore, in the future, studies should focus on exploring the key agreement and mutual authentication between different vehicles. Hasan et al. (2022) proposed an artificial intelligence (AI)-based synchronization scheme aimed at resolving smart grid timing issues. The back propagation neural network was employed as the AI method using the timing estimations and error corrections to ensure precise performance. Furthermore, the innovative AIFS scheme takes into account radio communication functionalities to relate to the external timing server. The simulation results confirmed this scheme's efficacy and substantial contributions such as no occurrence of synchronization errors do not occur if the external clock receives increments daily in the Wide Area Monitoring application (WAM) of the smart grid system. In addition, Hasan et al. (2020) developed an innovative resource-efficient Flow-enabled Distributed Mobility Anchoring (FDMA) framework. This solution improved the functionalities of mobility entities and centralized network entities. The mathematical analysis revealed that the packed tunnelling cost and the signalling overhead cost of the FDMA are lower than those of PNEMO (Ryu et al., 2013) and NBSP (Ernest et al., 2014).

3 | PROBLEM OVERVIEW

In this section, we give an overview of the safety provided by the adaptation of BC in our context. Next, we outline the goals of FL. Finally, we present the advantages of joining BC and FL in VANET and ITS for cyberthreats detection.

3.1 | Using the blockchain for safety

Studies have primarily focused on blockchain in the financial sector; however, the research focus has recently shifted to the Internet of Things (IoT) (Reyna et al., 2018). The IoT is widely explored and employed because it is able to create a safe, trustworthy, and decentralized environment. Owing to blockchain, new high technology has emerged in various sensitive and active sectors. To be more specific, the IoV solutions can ensure reliable information due to consensus, that the records are immutable along with the transparency of transactions. Most importantly, blockchain considerably improves security and trust. Moreover, smart contracts have enabled optimizing and automatizing of the managing process of information, resulting in substantially lower costs (Ali et al., 2021). In comparison to traditional centralized architectures, this technology has many benefits. However, there are significant limits to this technology of storage space, inflexibility, and expensiveness. Therefore, the integration of ITS and blockchain technology brings extraordinary advantages such as efficient management of the ITS, big data storage, and improved intelligence and security. Blockchain-based ITS solutions (BITS) have been investigated by academics and practitioners. Other emerging technologies, including 5G, big data, and machine learning, contribute to their additional advancement (Mollah et al., 2021).

3.2 | Overview on FL

The basics of FL can be described mathematically as follows (Qi et al., 2021):

Let:

K: the number of clients participating in FL,
w_t: the global model parameters,
w_t^k: the client local model,
D_k: the local dataset of the client k ($k \in K$),
$n_k = |D_k|$: the data in the client k,
l: indicates the loss function, and
∇w_i^k: describes the gradient symbol of w_t^k.

The FL process starts when the server distributes w_t to w_t^k. Then computes the gradient is computed by the following equation:

$$g_t^{k(1)} = \nabla w_i^k \sum_{(x_i, y_i) \in D_k} l\left(w_i^k, (x_i, y_i)\right). \tag{1}$$

The next step is to collect all gradient at the server and applying the weighted average in order to update the parameters. In the next subsection we give more details on FL.

3.3 | FL illustration

As evident in Figure 1, a classical FL system's operational process is ideally comprised of four stages:

1. The MEC server that is deployed at the edge cloud selects a computing task in addition to the task requirements and the learning parameters, such as selection of vehicular traffic analytics, task classification or prediction, and learning rate, respectively. Subsequently, it designates a subset of devices as learning clients for collaborative learning. Practically, the MEC server is able to designate various subsets of devices throughout various rounds of updating through the use of a fitting client scheduling mechanism for the purpose of ameliorating the quality of its training (Masoud et al., 2019).
2. The server initiates a training model which is then sent to each client to set up a new training round. At this stage, each client trains the model through the use of its own dataset and calculates a new update.

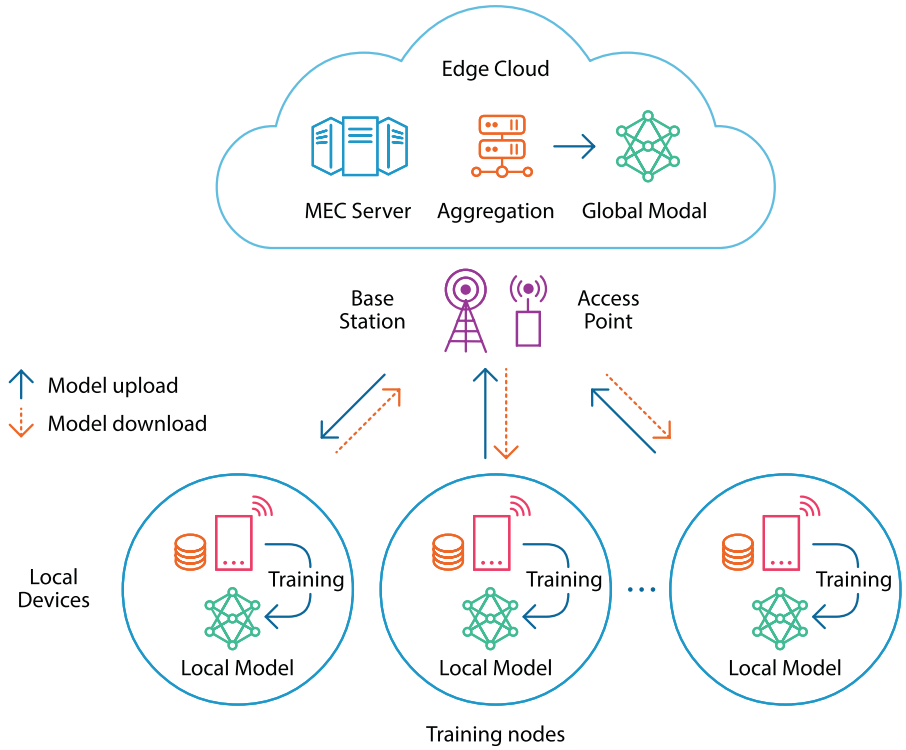

FIGURE 1 Classical federated learning framework

3. All clients upload the update they computed to the MEC server to perform a global computation. A novel global model is constructed by the server by means of aggregating all of the clients' updates in a way that decreases the loss function (Nguyen et al., 2021; Sattler et al., 2019).
4. The global model that has been previously computed is broadcasted by the server towards each client for the subsequent training round. The FL process is repeated until either the global loss function converges or the sought-after accuracy is attained. Based on the learning procedure, it is apparent that the classical FL is dependent on a centralized server; namely, an MEC server for model aggregation that has malfunction as well as scalability problems in common wireless networks.

Furthermore, this type of centralized FL architecture is not able to attract devices which are far away from the server to be trained, which in turn hinders the overall system's learning performance. The primary limitation of applying machine learning in mobile networks (e.g., vehicular networks) is the presence of data in the form of isolated islands. More specifically, it is highly challenging to integrate various data sources in vehicular networks because of necessary heavy computation, an increasing number of vehicles, and enormous data size (Du et al., 2020). To resolve this issue, 'federated learning' (FL) (Qi et al., 2021) was developed by Google in 2016 for establishing data models distributive. In FL architecture, data are stored on participating devices (e.g., smartphones or vehicles) that download the model and train the model by their data.

3.4 | Joining FL to blockchain

The operational process of blockchain-based FL in edge computing, as shown in Figure 2, is formulated through the following steps: First, the devices send a participant request to the server to be integrated into the network. Then, there is the selection and the review of the appropriate device by the server. After this stage of authentication, each terminal calculates its appropriate hashes, and then the data, including hashes and model, is uploaded to the blockchain for providing local model update training with private data. The offloads of data from the blockchain to servers perform to ensure the training update task. Finally, the server aggregates and updates parameters to be provided to the global model.

4 | PROPOSED FRAMEWORK

The proposed framework is given in Figure 3. The workflow is performed in eight steps that can be grouped into four phases. Before starting the process, we divide the VeReMi dataset into five balanced sub-datasets. We suppose that each vehicle is working only with its dataset without knowing each other. Table 1 summarizes the list of phases and steps.

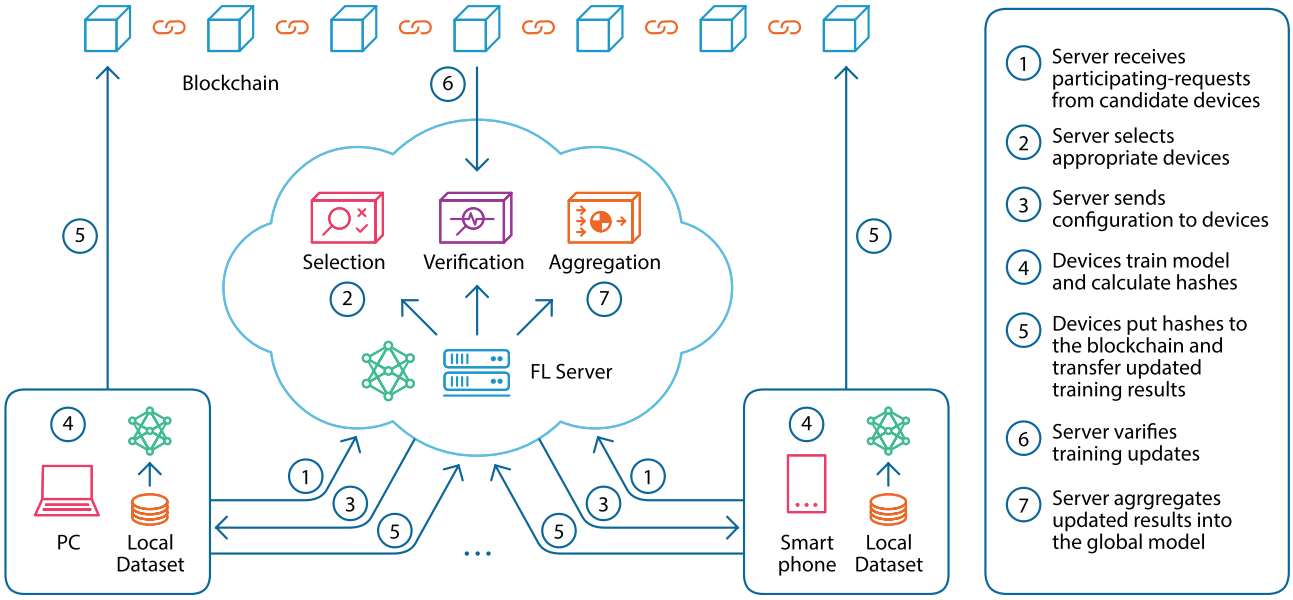

FIGURE 2 Federated learning and BC

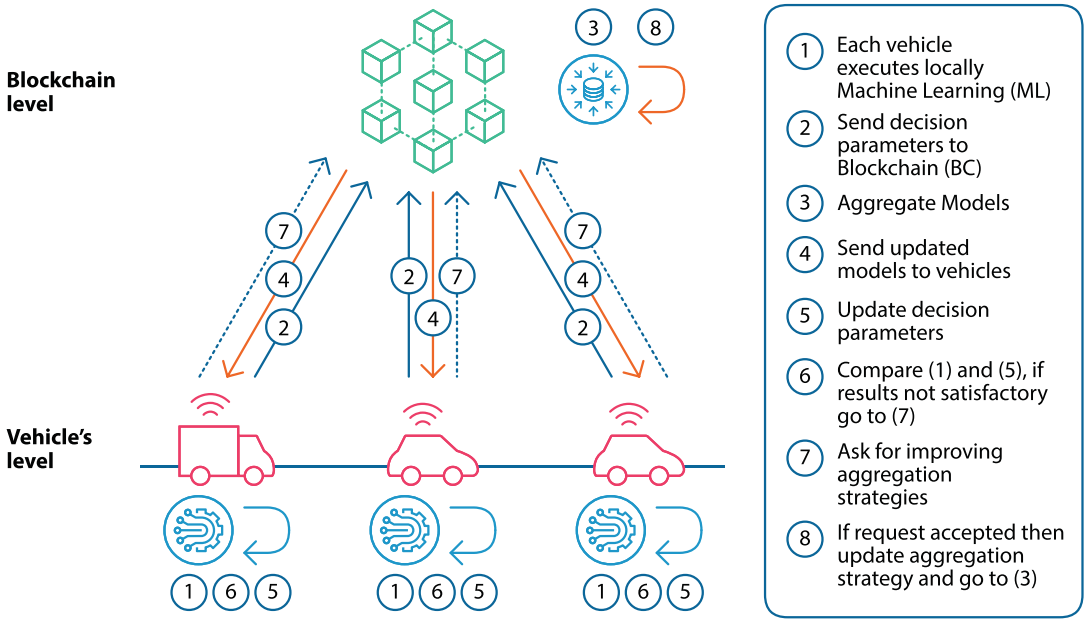

FIGURE 3 In the proposed framework steps 1–8 on the right side of the figure describe the list of tasks from executing ML locally, followed by the aggregation and the improvement process, and finally, the sharing of final results

4.1 | Phase 1: Local training

During the first step of this phase, each vehicle trains locally using four different models: Support vector machines (SVMs), Random Forest (RF), Naïve Bayesian (NB), and k-nearest neighbours' algorithm (KNN) algorithm. The models are trained only on the local dataset issued from the initial VeReMi dataset. In the second step, each vehicle extracts the decision parameters belonging to each model and sends them securely to the blockchain.

The decision function for each model, which is based on the entered vector of features, and the list of parameters for each model and can be described by the following equation:

TABLE 1 The proposed framework's phases and steps

Level	Phase	Step	Step description
Vehicle level	Phase 1	Step 1	Each vehicle executes ML locally
		Step 2	Send decision parameters to BC
Blockchain level	Phase 2	Step 3	Aggregate models
		Step 4	Send updated models to vehicles
Vehicle level	Phase 3	Step 5	Update decision parameters
		Step 6	Compare (1) and (5). If the results are not satisfactory, go to (7)
Blockchain level	Phase 4	Step 7	Request improvement of aggregation strategies
		Step 8	If the request is accepted, then update the aggregation strategy and go to (3)

$$\mathrm{DF}_{\mathrm{model}}(v_i, p_i) = \begin{cases} 1 & : \text{Abormal behavior} \\ -1 & : \text{Normal behavior} \end{cases},$$

where: model = {SV M, KNN, NB or RF}. v_i: The features of the entered vector. p_i: The decision parameters of the model.

4.2 | Phase 2: Aggregate model

This phase is performed at the blockchain level and consists of two steps. In step 3, the aggregation of decision parameters is done for each model. In step 4, the newly updated decision parameters are sent back to each vehicle.

4.3 | Phase 3: Update training models

During this phase, in step 5, each vehicle updates its decision parameters depending on the received values from the previous step. If the results of the predictions are not satisfactory (step 6), then the vehicles can go to step 7 by sending a request sent to the blockchain to improve the aggregation strategies.

4.4 | Phase 4: Improve models

This optional phase is optional and is executed only in the case of non-satisfactory classification results in the previous phase. Many possible actions can be taken to improve the aggregation strategy, such as excluding some vehicles that appear to be fraudsters. This can be decided if the decision parameters are not close to other vehicles' parameters. The degree of closeness can be defined depending on the deviation compared to mean values.

5 | EXPERIMENT AND RESULTS

In this section, we present the experimental study and the evaluation of the results. Finally, we discuss the obtained results.

5.1 | Dataset description

The VeReMi dataset is available in VeReMi (2022). In Table 2, we outline the different types of attacks and faults in this dataset.

5.2 | Evaluation parameters and environment setup

In this subsection, we describe the environment of simulation and different versions of the used systems. Next, we outline the parameters of evaluation.

5.2.1 | Environment setup

The environment setup and different system version are given in Table 3.

5.3 | Evaluation parameters

Accuracy: describes a ratio of correctly predicted attacks among all attacks. Accuracy is defined by the following equation:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}. \tag{2}$$

Precision: shows the ratio of correctly predicted positive attacks among the total predicted positive attacks. A high precision relates to the low false-positive rate. It is defined by the following equation:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{3}$$

Recall, also called sensitivity: describes the ratio of correctly predicted positive attacks to all attacks in the actual class. It is defined by the following equation:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{4}$$

TABLE 2 Cyberthreats in VeReMi dataset

ID	Cyberthreats	Type	Description
1	Constant position	Fault	The vehicle broadcasts a defective constant position (*X*, *Y*). Every time a new faulty node is introduced, the fixed values are discarded
2	Constant position offset	Fault	The vehicle broadcasts its true position with coordinates (*X*, *Y*)
15	Constant speed	Fault	The vehicle broadcasts a defective constant speed controller
16	Constant speed offset	Fault	The vehicle broadcasts a defective restricted number of random positions
3	Random position	Fault	The vehicle broadcasts a defective restricted number of random positions
4	Random position offset	Fault	With a small random offset, the vehicle diffuses its true position
18	Random speed	Fault	The vehicle broadcasts a malfunctioning limited random speed
19	Random speed offset	Fault	The vehicle broadcast its real speed with a limited random offset
17	Delayed messages	Attack	The vehicle sends out the right messages; however, they are sent out following a delay (*t*)
13	Eventual stop	Attack	The malicious node disseminates a fixed position and a null speed, simulating an abrupt stop
5	DoS	Attack	The car sends messages at a higher rate than what is specified in IEEE or ETSI standards. As a result of the overhead on the transmission channel, other cars may be unable to use it
6	Dos Random	Attack	The vehicle launches a DoS attack while also generating random data for all V2X messages
7	DoS random Sybil	Attack	To avoid detection, the attacker does a DoS Random by changing her identity in any forward message
10	Disruptive	Attack	The malicious vehicle replicates previously received communications that were sent by a random neighbour. This may prevent real messages from being broadcast and cause the network to become overburdened
12	DoS disruptive	Attack	At the same time, the attacker launches DoS and disruptive attacks
11	DoS disruptive Sybil	Attack	The malicious vehicle launches a Dos Disruptive assault while also changing its identity in all forwarding data
8	Data Replay	Attack	The attacker delivers the previously received data to a distinct destination neighbour. The attacker's certificate is used to sign this message
9	Data Replay Sybil	Attack	The attacker uses a Data Replay attack, changing their pseudonym on each new destination

TABLE 3 Environment setup

OS	Windows 10
CPU	i7-7500U CPU @ 2.70 GHz
RAM	8 GB
Language	Python and Solidity
Compiler of solidity	Solc 0.5.16
Test framework	Truffle 5.0.5
Local ethereum	Ganache v 5.4

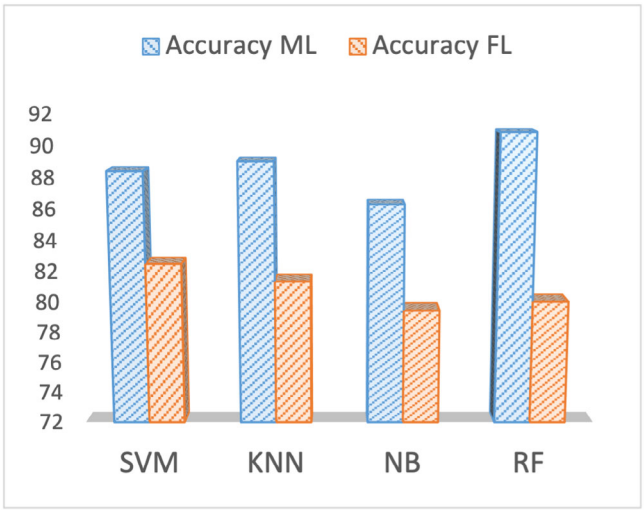

FIGURE 4 Comparison of accuracy between ML approach and the developed FL approach

F1-score describes the weighted average of Precision and Recall. It is defined by the following equation:

$$F1\text{Score} = \frac{2^{*}(\text{Recall}^{*}\text{Precision})}{(\text{Recall} + \text{Precision})}. \quad (5)$$

6 | RESULTS DISCUSSION

The machine learning (ML) results that were executed on the whole VeReMi dataset are taken directly from Sharma and Liu (2021). The results of our contribution are called FL results. The levels of accuracy are shown in Figure 4. The best accuracy for FL is provided by SVM with (82.45%). This accuracy level was (88.38%) with an ML approach. The use of the FL approach is very beneficial in terms of privacy preservation; however, it can decrease the level of accuracy due to the distrusted behaviour of FL. In this case, the rate of accuracy reduction is (5.93%). The decrease in levels of accuracy is (7.68%) for KNN, (6.85%) for NB, and (10.83%) for RF.

Figures 5 outlines the precisions results of ML approach compared to our FL approach. The precision shows the rate of detected true attack. The best level of precision in our approach is provided again by SVM at (91.02%). However, this precision level is (5.4%) less than that of ML approach. This decrease is due to the choice of favouring the privacy preservation. The decrease in level of precision is 8.22% for KNN, 8.36% for NB, and 6.72% for RF.

Figure 6 shows the recall result of ML approach compared to our FL approach.

Figure 7 outlines the *F*1-score results. According to these results, we can see that there are no big differences between FL and ML. Recall and *F*1-score are almost for SVM in both approaches. For other models, the maximum difference is (1.57%) for RF. These results show that the loss in recall and *f*1-score are very acceptable compared to the advantage of our method with is essentially the privacy preservation.

The curve shown in Figure 8 outlines amount of gas to execute aggregation methods for each model. The amount of gas is linked to the number of instructions executed for each model. This final result is needed in case of similarity in the previous performance of accuracy. The

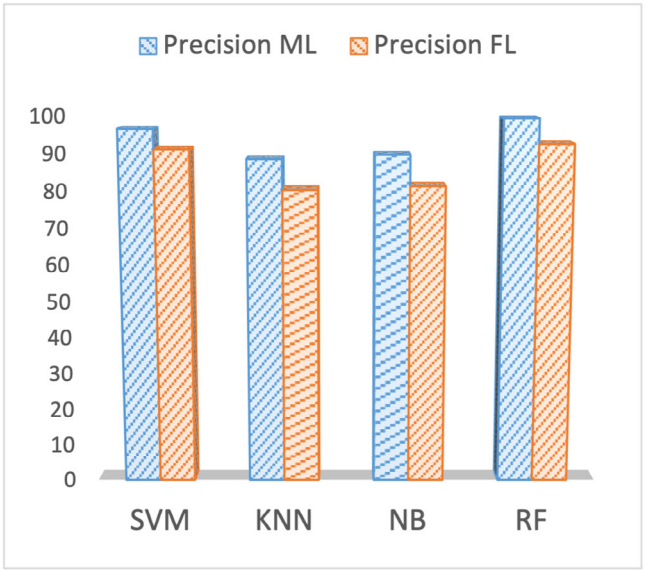

FIGURE 5 Precisions results of ML approach compared to our FL approach

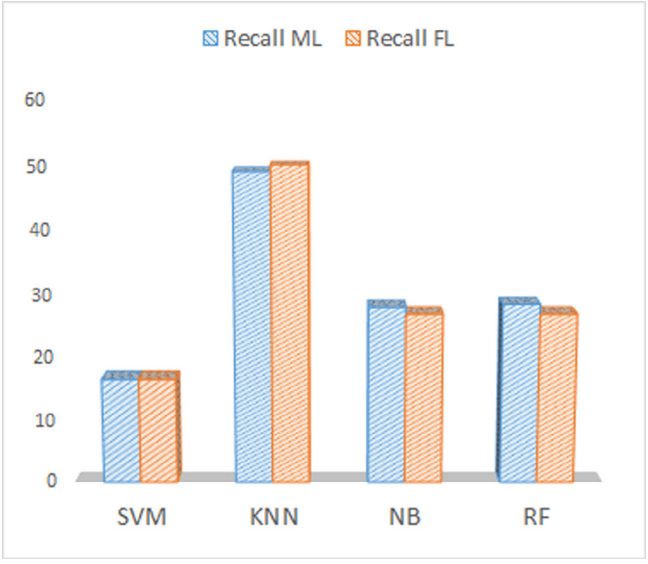

FIGURE 6 Recall results of ML approach compared to our FL approach

consumed gas can be used to select the best model to apply when the level of accuracy or precision is almost the same. In this case, it is preferable to apply a model that minimizes gas consumption. We can see that SVM is the lowest model in terms of consumed gas with only (5.2 gwei) while KNN is the highest model in terms of gas consumed reaching 10.1 gwei.

Figure 9 describes the elapsed time for aggregation according to the model. We can see that times results are confirm consumed energy results. Elapsed time can vary depending on the processing capacity of the machine where the simulation is one. The best time is achieved by SVM with less than 0.2 s while KNN takes the longest time.

7 | CONCLUSION

This paper developed a framework based on FL used to preserve vehicles' privacy and blockchain that ensured the safety of the aggregation strategy and data integrity. The classification models (SVM, KNN, NB, and RF) were applied to VANET and ITS cyber-threats detection and executed

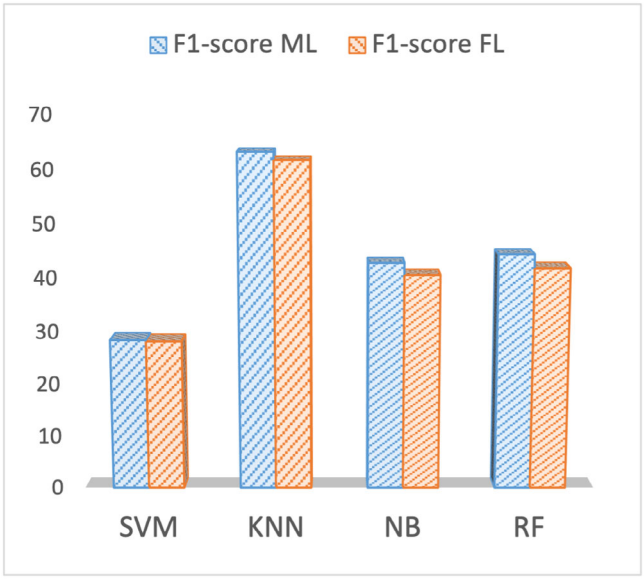

FIGURE 7 *F*1-score results of ML approach compared to our FL approach

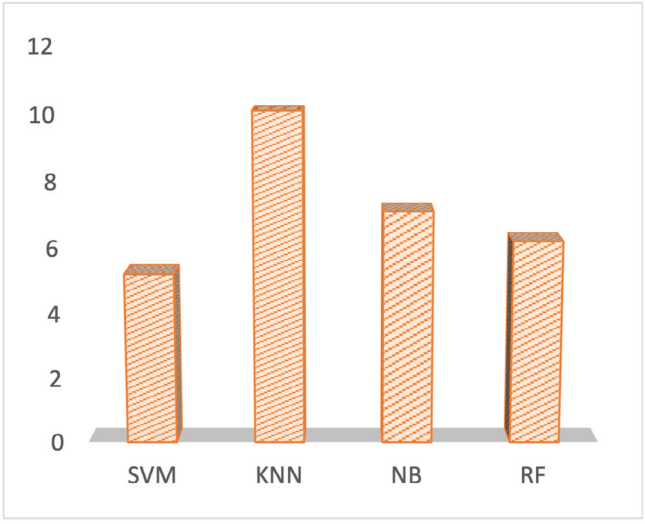

FIGURE 8 Consumed gas in (gwei) for aggregations according to models

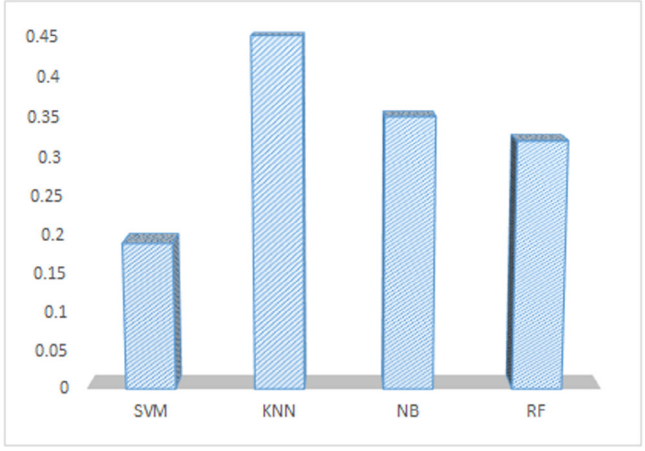

FIGURE 9 Elapsed time in (seconds) for aggregations according to models

locally on vehicles. Subsequently, we aggregated the models on the blockchain and returned them to vehicles. In comparison to previous studies that applied centralized ML on the same dataset, our approach was demonstrated to decrease precision and accuracy by 7.1% on average while maintaining the *F*1 score and the recall. The developed solution ensures the privacy preservation of vehicles whereas blockchain guarantees the safety of aggregation technique and a low consumption of gas. Future research should improve the definition of the most appropriate aggregation strategies and the selection of vehicles for participating in the FL process while removing malicious ones. It is critical to assess the behaviour of the vehicle in a prepossessing phase by classifying the vehicles before including them in the FL process.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Vehicular Reference Misbehavior (VeReMi) at https://veremi-dataset.github.io/.

ORCID

Muhammad Rizwan https://orcid.org/0000-0002-4408-4934

REFERENCES

Abdel-Basset, M., Moustafa, N., Hawash, H., Razzak, I., Sallam, K. M., & Elkomy, O. M. (2022). Federated intrusion detection in blockchain-based smart transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, *23*(3), 2523–2537. https://doi.org/10.1109/TITS.2021.3119968

Ahmad, F., Hall, J., Adnane, A., & Franqueira, V. N. L. (2017). Faith in vehicles: A set of evaluation criteria for trust management in vehicular ad-hoc network. In *2017 IEEE international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData), Exeter, United Kingdom, June 21–23, 2017* (pp. 44–52). IEEE Computer Society. https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2017.13.

Alfadhli, S. A., Lu, S., Fatani, A., Al-Fedhly, H., & Ince, M. (2020). SD2PA: A fully safe driving and privacy-preserving authentication scheme for VANETs. *Human-centric Computing and Information Sciences*, *10*(1), 1–25.

Ali, O., Jaradat, A., Kulakli, A., & Abuhalimeh, A. (2021). A comparative study: Blockchain technology utilization benefits, challenges and functionalities. *IEEE Access*, *9*, 12730–12749. https://doi.org/10.1109/ACCESS.2021.3050241

Aloqaily, M., Otoum, S., Ridhawi, I. A., & Jararweh, Y. (2019). An intrusion detection system for connected vehicles in smart cities. *Ad Hoc Networks*, *90*, 101842. https://doi.org/10.1016/j.adhoc.2019.02.001

Chai, H., Leng, S., Chen, Y., & Zhang, K. (2021). A hierarchical blockchain-enabled federated learning algorithm for knowledge sharing in internet of vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *22*(7), 3975–3986. https://doi.org/10.1109/TITS.2020.3002712

Cui, J., Wei, L., Zhang, J., Xu, Y., & Zhong, H. (2018). An efficient message-authentication scheme based on edge computing for vehicular ad hoc networks. *IEEE Transactions on Intelligent Transportation Systems*, *20*(5), 1621–1632.

Demir, M., Turetken, O., & Ferworn, A. (2019). Blockchain based transparent vehicle insurance management. In *6th International conference on software defined systems, SDS 2019, Rome, Italy, June 10–13, 2019* (pp. 213–220). IEEE. https://doi.org/10.1109/SDS.2019.8768669

Du, Z., Wu, C., Yoshinaga, T., Yau, K. L. A., Ji, Y., & Li, J. (2020). Federated learning for vehicular internet of things: Recent advances and open issues. *IEEE Open Journal of the Computer Society*, *1*, 45–61.

Ernest, P. P., Chan, H. A., Falowo, O. E., Magagula, L. A., & Céspedes, S. (2014). Network-based distributed mobility management for network mobility. In *2014 IEEE 11th consumer communications and networking conference (CCNC)* (pp. 417–425).

Finkle, J. (2016). *Tesla fixes security bugs after claims of model S hack*. https://www.reuters.com/article/us-tesla-cyber-idUSKCN11Q2SD

Fortune Business Insights. (2022). *Smart transportation market size, share & covid-19 impact analysis, by solution (traffic management system, integrated supervision system, parking management system, tcketing management system), by service (business-, professional, cloud) and regional forecast 2021–2028*. https://www.fortunebusinessinsights.com/smart-transportation-market-105736

Gandhi, G. M. & Salvi. (2019). Artificial intelligence integrated blockchain for training autonomous cars. In *2019 Fifth international conference on science technology engineering and mathematics (ICONSTEM)* (Vol. 1, pp. 157–161).

Hamdi, R., Chen, M., Said, A. B., Qaraqe, M., & Poor, H. V. (2021). Federated learning over energy harvesting wireless networks. *IEEE Internet of Things Journal*, *9*(1), 92–103.

Hasan, M. K., Ahmed, M. M., Hashim, A. H. A., Razzaque, A., Islam, S., & Pandey, B. (2020). A novel artificial intelligence based timing synchronization scheme for smart grid applications. *Wireless Personal Communications*, *114*(2), 1067–1084.

Hasan, M. K., Islam, S., Memon, I., Ismail, A. F., Abdullah, S., Budati, A. K., & Nafi, N. S. (2022). A novel resource oriented DMA framework for internet of medical things devices in 5G network. *IEEE Transactions on Industrial Informatics*, 1.

Hassan, M. A., Javed, A. R., Hassan, T., Band, S. S., Sitharthan, R., & Rizwan, M. (2022). Reinforcing communication on the internet of aerial vehicles. *IEEE Transactions on Green Communications and Networking*.

Hu, Q., Fong, S., Qin, P., Guo, J., Zhang, Y., Xu, D., Chen, Y., & Yen, J. (2019). Intelligent car parking system based on blockchain processing reengineering. In *Advances in E-business engineering for ubiquitous computing, proceedings of the 16th international conference on e-business engineering, ICEBE 2019, Shanghai, China, 12–13 October 2019* (Vol. 41, pp. 265–273). Lecture Notes on Data Engineering and Communications Technologies, Springer. https://doi.org/10.1007/978-3-030-34986-8_19

Hunt, T. (2018). *Controlling vehicle features of Nissan leafs across the globe via vulnerable apis*. https://www.troyhunt.com/controlling-vehicle-features-of-nissan/

Jabbar, R., Dhib, E., Ben Said, A., Krichen, M., Fetais, N., Zaidan, E., & Barkaoui, K. (2022). Blockchain technology for intelligent transportation systems: A systematic literature review. *IEEE Access*, *10*, 20995–21031.

Jabbar, R., Fetais, N., Kharbeche, M., Krichen, M., Barkaoui, K., & Shinoy, M. (2021). Blockchain for the internet of vehicles: How to use blockchain to secure vehicle-to-everything (V2X) communication and payment ? *IEEE Sensors Journal*, *21*, 15823. https://hal.archives-ouvertes.fr/hal-03154122

Jabbar, R., Kharbeche, M., Al-Khalifa, K., Krichen, M., & Barkaoui, K. (2020). Blockchain for the internet of vehicles: A decentralized IoT solution for vehicles communication using ethereum. *Sensors*, *20*(14), 3928. https://doi.org/10.3390/s20143928

Jabbar, R., Krichen, M., Kharbeche, M., Fetais, N., & Barkaoui, K. (2020). A formal model-based testing framework for validating an IoT solution for blockchain-based vehicles communication. In R. Ali, H. Kaindl, & L. A. Maciaszek (Eds.), *Proceedings of the 15th international conference on evaluation of novel approaches to software engineering, ENASE 2020, Prague, Czech Republic, May 5–6, 2020* (pp. 595–602). Scite Press. https://doi.org/10.5220/0009594305950602

Jabbar, R., Krichen, M., Shinoy, M., Kharbeche, M., Fetais, N., & Barkaoui, K. (2020). A model-based and resource-aware testing framework for parking system payment using blockchain. In *16th international wireless communications and mobile computing conference, IWCMC 2020, Limassol, Cyprus, June 15–19, 2020* (pp. 1252–1259). IEEE. https://doi.org/10.1109/IWCMC48107.2020.9148212

Javed, A. R., Ur Rehman, S., Khan, M. U., Alazab, M., & Reddy, T. (2021). CANintelliIDS: Detecting in-vehicle intrusion attacks on a controller area network using CNN and attention-based GRU. *IEEE Transactions on Network Science and Engineering*, *8*(2), 1456–1466.

Javed, A. R., Usman, M., Rehman, S. U., Khan, M. U., & Haghighi, M. S. (2020). Anomaly detection in automated vehicles using multistage attention-based convolutional neural network. *IEEE Transactions on Intelligent Transportation Systems*, *22*(7), 4291–4300.

Kim, M., Kim, K., & Kim, J. H. (2021). Cost modeling for analyzing network performance of IoT protocols in blockchain-based IoT. *Human-centric Computing and Information Sciences*, *11*.

Kumar, A., Yadav, A. S., & Kushwaha, D. S. (2020). VChain: Efficient blockchain based vehicular communication protocol. In *2020 10th international conference on cloud computing, data science engineering (confluence)* (pp. 762–768).

Labrador, M. & Hou, W. (2019). Implementing blockchain technology in the internet of vehicle (IoV). In: *2019 International conference on intelligent computing and its emerging applications (ICEA)* (pp. 5–10).

Lamssaggad, A., Benamar, N., Hafid, A. S., & Msahli, M. (2021). A survey on the current security landscape of intelligent transportation systems. *IEEE Access*, *9*, 9180–9208. https://doi.org/10.1109/ACCESS.2021.3050038

Li, L., Liu, J., Cheng, L., Qiu, S., Wang, W., Zhang, X., & Zhang, Z. (2018). CreditCoin: A privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *19*(7), 2204–2220. https://doi.org/10.1109/TITS.2017.2777990

Lu, Y., Huang, X., Zhang, K., Maharjan, S., & Zhang, Y. (2020). Blockchain empowered asynchronous federated learning for secure data sharing in internet of vehicles. *IEEE Transactions on Vehicular Technology*, *69*(4), 4298–4311. https://doi.org/10.1109/TVT.2020.2973651

Maaroufi, S., & Pierre, S. (2021). BCOOL: A novel blockchain congestion control architecture using dynamic service function chaining and machine learning for next generation vehicular networks. *IEEE Access*, *9*, 53096–53122. https://doi.org/10.1109/ACCESS.2021.3070023

Malik, N., Nanda, P., Arora, A., He, X., & Puthal, D. (2018). Blockchain based secured identity authentication and expeditious revocation framework for vehicular networks. In *17th IEEE international conference on trust, security and privacy in computing and communications/12th IEEE international conference on big data science and engineering, TrustCom/BigDataSE 2018, New York, NY, USA, August 1–3, 2018* (pp. 674–679). IEEE. https://doi.org/10.1109/TrustCom/BigDataSE.2018.00099

Masoud, M. Z., Jaradat, Y., Jannoud, I., & Zaidan, D. (2019). CarChain: A novel public blockchain-based used motor vehicle history reporting system. In: *2019 IEEE Jordan international joint conference on electrical engineering and information technology (JEEIT)* (pp. 683–688).

Mitsubishi Outlander PHEV. (2016). *Hackers can remotely disable car alarm on Mitsubishi Outlander PHEV suvs*. https://securityaffairs.co/wordpress/48114/hacking/mitsubishi-outlander-phev-hacking.html

Mittal, M., Iwendi, C., Khan, S., & Rehman Javed, A. (2021). Analysis of security and energy efficiency for shortest route discovery in low-energy adaptive clustering hierarchy protocol using Levenberg-Marquardt neural network and gated recurrent unit for intrusion detection system. *Transactions on Emerging Telecommunications Technologies*, *32*(6), e3997.

Mollah, M. B., Zhao, J., Niyato, D., Guan, Y. L., Yuen, C., Sun, S., Lam, K. Y., & Koh, L. H. (2021). Blockchain for the internet of vehicles towards intelligent transportation systems: A survey. *IEEE Internet of Things Journal*, *8*(6), 4157–4185. https://doi.org/10.1109/JIOT.2020.3028368

Mostafa, A. (2019). VANET blockchain: A general framework for detecting malicious vehicles. *The Journal of Communication*, *14*(5), 356–362. https://doi.org/10.12720/jcm.14.5.356-362

Nadeem, S., Rizwan, M., Ahmad, F., & Manzoor, J. (2019). Securing cognitive radio vehicular ad hoc network with fog node based distributed blockchain cloud architecture. *International Journal of Advanced Computer Science and Applications*, *10*(1). https://doi.org/10.14569/IJACSA.2019.0100138

Naeem, A., Javed, A. R., Rizwan, M., Abbas, S., Lin, J. C. W., & Gadekallu, T. R. (2021). DARE-SEP: A hybrid approach of distance aware residual energy-efficient SEP for WSN. *IEEE Transactions on Green Communications and Networking*, *5*(2), 611–621.

Nguyen, D. C., Ding, M., Pham, Q., Pathirana, P. N., Le, L. B., Seneviratne, A., Li, J., Niyato, D., & Poor, H. V. (2021). Federated learning meets blockchain in edge computing: Opportunities and challenges. *IEEE Internet of Things Journal*, *8*(16), 12806–12825. https://doi.org/10.1109/JIOT.2021.3072611

Nkenyereye, L., Tama, B. A., Shahzad, M. K., & Choi, Y. (2020). Secure and blockchain-based emergency driven message protocol for 5G enabled vehicular edge computing. *Sensors*, *20*(1), 154. https://doi.org/10.3390/s20010154

Nofer, M., Gomber, P., Hinz, O., & Schiereck, D. (2017). Blockchain. *Business and Information Systems Engineering*, *59*(3), 183–187. https://doi.org/10.1007/s12599-017-0467-3

Noh, J., Jeon, S., & Cho, S. (2020). Distributed blockchain-based message authentication scheme for connected vehicles. *Electronics*, *9*(1). https://www.mdpi.com/2079-9292/9/1/74

Obaidat, M., Khodjaeva, M., Holst, J., & Ben Zid, M. (2020). *Security and privacy challenges in vehicular ad hoc networks* (pp. 223–251). Springer International Publishing. https://doi.org/10.1007/978-3-030-36167-9_9

Obour Agyekum, K. O. B., Xia, Q., Boateng Sifah, E., Amofa, S., Nketia Acheampong, K., Gao, J., Chen, R., Xia, H., Gee, J. C., Du, X., & Guizani, M. (2018). V-chain: A blockchain-based car lease platform. In *2018 IEEE international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData)* (pp. 1317–1325).

Otoum, S., Ridhawi, I. A., & Mouftah, H. T. (2020). Blockchain-supported federated learning for trustworthy vehicular networks. In: *IEEE global communications conference, GLOBECOM 2020, virtual event, Taiwan, December 7–11, 2020* (pp. 1–6). IEEE. https://doi.org/10.1109/GLOBECOM42002.2020.9322159

Qi, Y., Hossain, M. S., Nie, J., & Li, X. (2021). Privacy-preserving blockchain-based federated learning for traffic flow prediction. *Future Generation Computer Systems*, *117*, 328–337.

Rateb, J. (2021). *Blockchain for the internet of vehicles: A decentralized IoT solution for vehicles communication and payment using ethereum* [PhD thesis]. Paris: HESAM.

Rathee, G., Sharma, A., Iqbal, R., Aloqaily, M., Jaglan, N., & Kumar, R. (2019). A blockchain framework for securing connected and autonomous vehicles. *Sensors*, *19*(14), 3165. https://doi.org/10.3390/s19143165

Rehman, A., Razzak, I., & Xu, G. (2022). Federated learning for privacy preservation of healthcare data from smartphone-based side-channel attacks. *IEEE Journal of Biomedical and Health Informatics*.

Rehman Javed, A., Jalil, Z., Atif Moqurrab, S., Abbas, S., & Liu, X. (2020). Ensemble adaboost classifier for accurate and fast detection of botnet attacks in connected vehicles. *Transactions on Emerging Telecommunications Technologies*, e4088.

Reimers, T., Leber, F., & Lechner, U. (2019). Integration of blockchain and internet of things in a car supply chain. In *IEEE international conference on decentralized applications and infrastructures, DAPPCON 2019, Newark, CA, USA, April 4–9, 2019* (pp. 146–151). IEEE. https://doi.org/10.1109/DAPPCON.2019.00028

Reyna, A., Martín, C., Chen, J., Soler, E., & Díaz, M. (2018). On blockchain and its integration with IoT. Challenges and opportunities. *Future Generation Computer Systems*, *88*, 173–190. https://doi.org/10.1016/j.future.2018.05.046

Ryu, S., Choi, J. W., & Park, K. J. (2013). Performance evaluation of improved fast PMIPv6-based network mobility for intelligent transportation systems. *Journal of Communications and Networks*, *15*(2), 142–152.

Sattler, F., Wiedemann, S., Müller, K., & Samek, W. (2019). Robust and communication-efficient federated learning from non-IID data. *CoRR. abs/1903.02891*. http://arxiv.org/abs/1903.02891

Sharma, P., & Liu, H. (2021). A machine-learning-based data-centric misbehavior detection model for internet of vehicles. *IEEE Internet of Things Journal*, *8*(6), 4991–4999. https://doi.org/10.1109/JIOT.2020.3035035

Sheikh, M. S., Liang, J., & Wang, W. (2019). A survey of security services, attacks, and applications for vehicular ad hoc networks (VANETs). *Sensors*, *19*(16), 3589. https://doi.org/10.3390/s19163589

Singh, M. & Kim, S. (2018). Trust Bit: Reward-based intelligent vehicle commination using blockchain paper. In *4th IEEE world forum on internet of things, WF-IoT 2018, Singapore, February 5–8, 2018* (pp. 62–67). IEEE. https://doi.org/10.1109/WF-IoT.2018.8355227

Singh, P., Khanna, P., & Kumar, S. (2020). Communication architecture for vehicular ad hoc networks, with blockchain security. In *2020 International conference on computation, automation and knowledge management (ICCAKM)* (pp. 68–72).

Veremi. (2022). *Veremi dataset*. https://veremi-dataset.github.io/

Xu, Z., Li, X., Xu, J., Liang, W., & Choo, K. K. R. (2021). A secure and computationally efficient authentication and key agreement scheme for internet of vehicles. *Computers & Electrical Engineering*, *95*, 107409.

Xu, Z., Liang, W., Li, K. C., Xu, J., & Jin, H. (2021). A blockchain-based roadside unit-assisted authentication and key agreement protocol for internet of vehicles. *Journal of Parallel and Distributed Computing*, *149*, 29–39.

Yang, Z., Yang, K., Lei, L., Zheng, K., & Leung, V. C. M. (2019). Blockchain-based decentralized trust management in vehicular networks. *IEEE Internet of Things Journal*, *6*(2), 1495–1505. https://doi.org/10.1109/JIOT.2018.2836144

Yang, Z., Zheng, K., Yang, K., & Leung, V. C. M. (2017). A blockchain-based reputation system for data credibility assessment in vehicular networks. In *28th IEEE annual international symposium on personal, indoor, and mobile radio communications, PIMRC 2017, Montreal, QC, Canada, October 8–13, 2017* (p. 1–5). IEEE. https://doi.org/10.1109/PIMRC.2017.8292724

Zeadally, S., Hunt, R., Chen, Y., Irwin, A., & Hassan, A. (2012). Vehicular ad hoc networks (VANETS): Status, results, and challenges. *Telecommunication Systems*, *50*(4), 217–241. https://doi.org/10.1007/s11235-010-9400-5

Zhang, J., Zhao, H., Yang, Y., Yan, J. (2019). Towards transparency and trustworthy: A used-car deposit platform based on blockchain. In *19th IEEE international conference on software quality, reliability and security companion, QRS companion 2019, Sofia, Bulgaria, July 22–26, 2019* (pp. 46–50). IEEE. https://doi.org/10.1109/QRS-C.2019.00022

Zhang, K., Liang, X., Lu, R., & Shen, X. (2014). Sybil attacks and their defenses in the internet of things. *IEEE Internet of Things Journal*, *1*(5), 372–383. https://doi.org/10.1109/JIOT.2014.2344013

AUTHOR BIOGRAPHIES

Tarek Moulahi received the Ph.D. degree from the University of Franche-Comté, Besançon, France, in March 2015, and the Sfax National School of Engineering, Tunisia. He is currently an Assistant Professor with the Mathematics and Computer Science Department, Faculty of Science and Technology of Sidi Bouzid (FSTSB), University of Kairouan, Tunisia, and with the Department of Information Technology, College of Computer, Qassim University, Saudi Arabia. His research interests include wireless sensor networks, vehicular ad hoc networks (VANET), the Internet of Things (IoT), and blockchain. He received the 2019 IEEE Sensors Council Sensors Journal Best Paper Runner-Up Award.

Rateb Jabbar received the Ph.D. degree in Computer Science from Hautes Écoles Sorbonne Arts et Métiers (HESAM) University, Paris, France, in 2021. He worked as a Senior Software Engineer for 10 years specialist in web, cloud, machine learning, and blockchain technologies. He is a Postdoctoral Fellow with the Department of Computer Science and Engineering, Qatar University. He is a Microsoft Certified Professional and a Microsoft Certified Technology Specialist in developing ASP.NET MVC 4 web applications and Microsoft Azure cloud service and web services.

Abdulatif Alabdulatif is an assistant professor at the School of Computer Science & IT, Qassim University, Saudi Arabia. He completed his Ph. D. degree in Computer Science from RMIT University, Australia in 2018. He received his B.Sc. degree in Computer Science from Qassim University, Saudi Arabia in 2008 and his M.Sc. degree in Computer Science from RMIT University, Australia in 2013. His research interests include applied cryptography, cloud computing, data mining, and remote healthcare.

Sidra Abbas received her BSCS from COMSATS University, Islamabad, Pakistan. She has more than 5 years of experience in the industry. She has published more than 10 international publications. Currently, her areas of research include Machine Learning, the Internet of Things, Deep Neural Networks, Blockchain, and Computer Vision.

Salim El Khediri is an Associate Professor at the Department of Information Technology, Qassim University, Saudi Arabia. He received his Ph. D. in Computer Science from the University of Sfax in collaboration with CNAM Paris, his Master's degree in Computer Science from Luminy University in Marseille, France, he also received his Bachelor's degree in Computer Science from Luminy University. He has authored over 63 referred papers. His research interests include Internet of Things, Artificial Intelligence, and Network Security.

Salah Zidi received the Ph.D. degree from the University of Lille, France, with a focus on regulation and reconfiguration of multimodal transportation systems, in July 2007, and the HDR degree from the University of Lille1, France, in 2017. He is currently an Assistant Professor with the MIS Department, College of Business and Economics, Qassim University, Saudi Arabia, and an Associate Professor with the University of Gabes, Tunisia. His research interests include optimization, artificial intelligence, machine learning, feature extraction, and data analysis for automation systems and complex systems. He received the 2019 IEEE Sensors Council Sensors Journal Best Paper Runner-Up Award.

Muhammad Rizwan received the M.Sc. degree from PUCIT, Lahore, Pakistan, in 2006, the MS degree from CIIT, Lahore, Pakistan, in 2012, and the Ph.D. degree from HUST, Wuhan, China, in 2017. He was an Assistant Professor with the Department of Computer Science, Kinnaird College for Woman University, Lahore, Pakistan for four years. He is currently an Assistant Professor in Cyber Security and member of the Secure Cyber Systems Research Group, WMG, University of Warwick, UK. He is the author of more than 70 research articles published in international conferences and journals. His research interests include the areas of Cybersecurity, Digital Forensics, Industrial Cyber Physical Systems, Artificial Intelligence, Health Informatics, Mobile Cloud Computing & IoT.

How to cite this article: Moulahi, T., Jabbar, R., Alabdulatif, A., Abbas, S., El Khediri, S., Zidi, S., & Rizwan, M. (2023). Privacy-preserving federated learning cyber-threat detection for intelligent transport systems with blockchain-based security. *Expert Systems*, *40*(5), e13103. https://doi.org/10.1111/exsy.13103