Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

Next-Generation Protection: Leveraging Federated Learning and Blockchain for Intrusion Detection in Smart Vehicle Network

Javaid Ahmad Malik[1], Sagheer Abbas[2], Altaf Hussain[3], Muhammad Saleem[4], Rahat Qudsi[5]

1. School of Computer Science, National College of Business Administration and Economics, Lahore 54000, Pakistan
Javed_ahmad2016@outlook.com

2. Department of Computer Science, Bahria University Lahore Campus, Lahore 54000, Pakistan
jamsagheer@gmail.com

3. Institute of Computing, MNS University of Agriculture Multan
Altaf.hussain@mnsuam.edu.pk

4. Minhaj University Lahore, Pakistan
msaleemkalru@gmail.com

5. Institute of Computing, MNS University of Agriculture Multan
noveenaabid@gmail.com

Correspondence:
Javaid Ahmad Malik
School of Computer Science, National College of Business Administration and Economics, Lahore 54000, Pakistan Email: javed_ahmad2016@outlook.com

Abstract:- Smart Car's era ushers in new challenges, some of which are in the field of information security such as wise cyber-attack prevention. Traditional IDS systems would lose everything in such a dynamic environment with their centralized architectures, and this approach could create single points of failure and privacy issues. Along with interconnectedness, cyber security becomes an inevitable problem as smart vehicles are incorporated into a daily life. Traditional security mechanisms usually lack scalability and privacy, which brings about the need to develop alter-nate or innovative methods. This research demonstrates a mixed security system that combines both federated learning and blockchain technologies to improve intrusion detection in smart vehicular networks. We evaluated the effectiveness of this framework using four machine learn-ing models as respectively; Support Vector Machine (SVM), Decision Tree, Neural Network, and Random Forest. Empirical results show that SVM had the highest accuracy of both 93.88% in training and 91.84% in validation, which is higher than Decision Tree, Neural Network, and Random Forest models. These findings evidently demonstrate that the federated learning and blockchain are a strong solution for the plausible security of smart vehicle networks; with SVM being employed mostly in complex security scenarios.

Keywords: Next-Generation Protection, Federated Learning, Blockchain, Intrusion Detection, Smart Vehicle Network

Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

1. Introduction

The concept of smart vehicle networks is, in fact, the biggest leap in automotive tech-nologies that has seen vehicles integrate advanced computational power, communication systems, and artificial intelligence for vehicles which can at the same time be aware of their surrounding area and other vehicles[1]. This marijuana is driven by the interplay of IoT (Internet of Things) technologies, the wireless communication, and analytics that gen-erate, sort, and share accurate data at the moment of need. The greatest objective of such contribution is to upgrade safety, lower traffic congestion, lessen carbon footprint and de-rives the driving experience[2]. The core of the hoped for transport future is the interaction of smart vehicles' networks among themselves and their vehicles' surrounding infrastruc-ture. Now, due to this connectivity, vehicles should be able not only to plan the optimal routes but also to do proper speed and safety management. This, in turn, can serve as an effective pragmatic measure that will reduce the risk of accidents and improve the traffic flow[3]. With the passage of time, these networks are steadily getting more integrated into the stream of daily routines thereby making the data they possess and consume more classy and thought-provoking, in turn, rendering major security and privacy issues. Be-sides, such computerized transport is capable of interacting and deliberating on data without humans, which casts difficulty in valuing the security of those communications from being intercepted or modified[4]. Cybersecurity in a smart vehicle network is not lim-ited only to repellent the onboard systems from the unwanted access; it encompasses the sharing of the data which should remain protected in integrity and confidentiality. The fact these online interactions are through limited by manufactures and service providers makes the universal safety system and protocols to be complex[5]. Overcoming these ob-stacles is the most important issue for the public acceptance and prosperity of smart au-tomobile systems and therefore officials should accept the holistic view that involves techno-regulatory, ethics and societal considerations[6].

The security of smart vehicle networks connotes the utmost importance because they not only rely on interoperable systems but they also handle large amounts of data. Net-works, which are by nature complex in structure, have to deal with a myriad of issues that can be attributed to their complexity and to the crucial role they play in ensuring safety and security[7]. The most crucial worry that arises is susceptibility to cyber-attacks. Rather than with the usual computers networks, in case of smart vehicle networks the stakes are very high, especially when the hackers manage to access the interface, a crash can occur. Vehicles are not only able to communicate with each other but also with smart traffic sys-tems, consequently, several attack points for the attackers have been created[8]. They can vary from a penetration of the vehicle's operations system to bombardment of any kind on the traffic network segment, with the both effects of catastrophic nature. The other notable challenge is the maintaining of data privacy and data integrity as well. Such critical in-formation, like fleet location, time spent at certain locations, and even driving styles are seen to be exposed to possible data fraud when shared

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

with other fleet vehicles[9]. And it simultaneously reveals the questions of personal data privacy and also possibility to inter-fere with traffic flow or the operation of vehicles through the manipulation of data. The heterogeneity of the technologies coupled with the trend adds yet another layer of com-plexity[10]. Smart road networks integrate a host of different technologies, standards, and protocols, which provides a wide opening for the attackers, making it challenging to estab-lish the uniform security measures. With the increase of diversity, a security framework that is flexible and adaptable in nature is the need of the hour as this security system can keep threats under check and flawlessly connect different systems and devices when needed[11].

The advent of federated learning marks a disruptive breakthrough in the field of ma-chine learning that involves a decentralized training mechanism across the devices net-work as part of the preservation of data privacy. Different from the centralized teaching techniques, with federated learning, the models are trained locally on the devices of indi-viduals, therefore the minimum amount of raw data needs to be exchanged between them[12]. In a nutshell, this trend is of the most significance in cases where privacy is of high priority, say in smart vehicle networks, where highly sensitive information is gener-ated and shared. Federated learning assembles the model training without revealing the details of user data, since it sends the model updates, not the raw data, to the central serv-er and to the other devices. This method can not only improve privacy protection but also reduce security and regulatory compliance concerns[13]. Through federated learning de-vices spread their collective intelligence and this enables the designers to create stronger and more versatile machine learning models that can work more precisely under diversity real-world data. With the further development of smart vehicle networks, federated learn-ing may therefore become an important tool for the progress of machine learning and en-sure privacy of data and integrity of users' data[14].

Blockchain technology signifies as the latest advent and has transformed into a vigi-lant and immutable ledger system for the applications such as smart vehicle networks that act as a traditional source of trust and transparency. In essence, blockchain is a data-base that is kept in a democratized way where the information is stored cryptographically across the computer network and it is impossible to be falsified. Through the utilization of crypto methods and consensus algorithmic decisions, blockchain allows for trustless in-teraction among transacting parties by eliminating the necessity for intermediaries and central authorities[15]. In a blockchain-enabled smart vehicle network blockchain is posi-tioning itself to become a foundational technology for securing data and transaction ex-changes and communication. The adoption of the smart contracts in blockchain as a means of inter-party contract administration and the fulfilment of the same involves creat-ing a smooth, transparent and efficient transactional environment. The embedded trans-parency and auditability of the blockchain technology help ensure reliability and tracea-bility as accountability and decorrelation attributes. These characteristics are vital for maintaining the integrity of data as well as

Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

operations within smart vehicle networks[16]. Blockchain's decentralized architecture encompasses a collection of nodes which is more resilient against single points of failure or unauthorized access; in addition, it further strengthens the security of smart vehicle ecosystems. Blockchain technology which is at the forefront of innovation is destined to be a crucial factor in encouraging trusted, trans-parent and secure connections as smart vehicle networks grow and expand. In the same manner, blockchain will ultimately enable connected and self-driven vehicles to truly real-ize their potential[17].

The integration of Federated Learning with Blockchain entails a new way of ensuring the security in major distributed applications such as self-driving vehicle networks where the privacy-preserving nature of Federated Learning is combined with the trust and trans-parency character of Blockchain technology. Collaborative model training across a net-work of devices in distributed environment happens while sensitive data remains locked in devices so the privacy risks are minimized. Through the use of blockchain based tech-nology process-integrity and transparency of the federated learning procedure is aug-mented. Blockchain stores model updates in an unalterable ledger that no one can alter or formulate without authority[18]. The smart contracts can be used to automate the func-tioning of the federated learning protocols which are aimed at providing enhanced securi-ty communication and coordination among involved devices. Blockchain's decentraliza-tion consensus process strengthens the robustness and resilience of the federated learning framework which in turn minimizes the risk of failure as well as malicious attacks[19]. Through this integrated scheme, not only the security of smart vehicle networks is forti-fied, but also confidence among the stakeholders is built up that is based on really existing assurances of confidentiality, integrity, and fairness of the federated learning algorithm. The fusion of federated learning with blockchain in self-driving and autonomous vehicles of the future is well on its way as the former technology continues to enhance the security and privacy of distributed and autonomous systems[20].

Network attacks are the ones where the cyber assailants capitalize on flaws and weaknesses within the networks and the systems. The forms of the attack may vary, in-cluding illegal access, data interception, and services disruption. They undermine a secre-cy, integrity, and availability of the network resources, in consequence causing breaches of data, financial losses and operational disruptions. The types of network attacks are DDoS, malware propagation, phishing and SQL injection capable of exploiting vulnerabilities in network architecture, software, and human behavior. This Table 1 shows various catego-ries of network attack types.

Table 1. Classes of Network Attacks

Sr. No.	Types of Attack	Explanation
1	Buffer Overflow	This action involves targeting the buffer's bounds and subsequently overwriting the memory region.
2	Denial of service	The occurrence of a security event aimed at disrupting network services. The process is initiated by performing a forced reset on

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

		the targeted computers. The inability of users to establish a connection with the system is attributed to the lack of service availability.
3	Common Gateway Interface Scripts	The assailant exploits CGI scripts to orchestrate an assault by transmitting unauthorized inputs to the web server.
4	Traffic Flooding	Critiques the constrained capacity of Network Intrusion Detection Systems (NIDS) in managing substantial volumes of network traffic and conducting investigations into potential intrusions. Suppose an individual engaged in cybercriminal activities can induce network congestion. In that case, it will increase the workload of Network Intrusion Detection Systems (NIDS) as they analyze the influx of network traffic.
5	Information Gathering	The act of gathering information or identifying vulnerabilities in computer systems or networks is accomplished through sniffing or searching.
6	User to Root (U2R) attack	Initially, the cybercriminal assumes the role of an ordinary user, then elevating their privileges to that of a super-user. This progression enables them to potentially exploit many vulnerabilities within the system.
7	Remote to Local (R2L) attack	The cybercriminal transmits packets to a remote system by establishing a network connection without possessing a user account on that system.
8	Probe	Identifying legitimate IP addresses involves a network scan to collect host data packets.

The research integrates distributed learning or federated learning with blockchain to enhance the security of smart vehicle networks. It focuses on the development of a novel protocol that integrates federated learning protocols with blockchain-based security data to protect privacy and reliability. This research aims to assess the efficacy of this framework through simulations and practical deployments. This ensures the development of cybersecurity for connected and autonomous vehicles. It is a comprehensive solution that renders smart vehicle networks more secure, private, and resilient, which in turn, implies safer transportation ecosystems in the future.

2. Literature Review

The smart vehicle network literature dives deep enough to reveal the architectures, functionalities, and target features employed in such advanced automotive systems. Accordingly, these networks, labeled usually as connected vehicle networks, utilize modern technologies such as IoT, wireless communications, and data analytics to bring such real-time communication between vehicles, road-related infrastructure, and external services. In this way, the digital platforms offer data and information regarding traffic conditions, unsafe road sections, weather details and vehicle performance to help in safety, efficiency and convenience

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

of drivers and their passengers[21]. The key elements of smart cars communicating systems are on-board sensors, communication modules, central processing module, and backend servers, which are connected and forms intelligence transportation system. The wide-spread development of such communications has brought forth a great deal of innovation in the form of autonomous driving, predictive maintenance, and personalized services which are turning the way that the automotive industries work inside out. As well, along with increasing number of positive uses of smart vehicle networks that provide diverse advantages on road transportation, new challenges rise including security, privacy, and interoperability which can be available by creating effective solutions to ensure continuous smoothness of smart vehicle networks[22].

The concerns that the security breaches play in complexty of smart vehicle network cause serious problems because of this network interconnectedness and sensitive data retention. They include those cyber threats which target vehicles in the form of hacking their systems, or denial of services mounted on vulnerable networks. Maintaining the confidentiality and validity of data during the constant flow of transmission in a network is of absolute importance[23]. A wide range of different technologies is the complex thing that makes it impossible to upgrade security with a uniform set of rules. The security vulnerabilities of smart networks in telecommunication, which are embedded in vehicles, become larger in the management of security protocols and updates. This complex issue is addressed by a comprehensive approach that encapsulates high-end encryption, discreet communication protocols, and continuous observation for a possible danger[24].

Decentralized Machine learning is based on federated learning, which enables learning by machines across many decentralized entities, while protecting data privacy at the same time. This paradigm is the base for the machine learning named model training on local data and this without sharing the raw data. Citing its relevance in healthcare, finance, and IoT systems, such as smart vehicle networks, the employment of federated learning has become imminent because of its future applications in various fields. Variable model training on fragmented data resources is what is facilitated by federated learning[25]. Privacy-preserving machine learning is thus implemented, where there is a problem of information confidentiality and safety. Through federated learning, it becomes possible to build models that are both more personalized and adaptive, whereas models could accommodate any diversity in real-world data that is found across a variety of distributed devices. These pillars of federated learning embed privacy, model aggregation, and communication efficiency, which make it an appealing option for training machine learning models in decentralized set-up[26].

The blockchain technology has been receiving significant spotlight for its possible uses in the automotive systems, driven by an existing system that is tamper-proof and decentralized that also provides a transparent, trustworthy and secure network. Within automotive systems, blockchain can be applied to providing secure, confirma¬tory, and immutable record of

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

transactions like vehicle sales, ownership transfer, and maintenance history, leading to restoration of trust and integrity of data. Blockchain technology is the feature in smart contracts that allow for automated transactions to occur between the parties involved in automotive process like buying a car or leasing a rental car without the need for human interaction. Blockchain-centered technologies can improve automotive industry supply chain by building real-time data about the movement of goods and component which will lead to faster distribution and fewer errors along the way[27]. The reason why immutable (i.e. unable to be changed) and verifiable (i.e. can be checked using methods) nature of blockchain makes it ideal for enhancing traceability and authentication in the automotive aftermarket where consumers may verify the source and record of the spare parts and accessories through blockchain[28]. As with auto systems that are gradually turning to be more connected and autonomous, blockchain technology is reckoning to be a main revolution heading to eradicate the various aspects of the auto sector for example vehicle ownership, supply chain management, and aftermarket services which will ultimately bring about efficiency, transparency, and trust.[29]

The existing approach to security in smart vehicles’ networks had been primarily aimed at traditional cybersecurity methods like encryption, authentication, and intrusion detection (IDS) systems. These approaches are targeted at ensuring that the vehicle systems remain immune to unauthorized access and malicious attacks thus protecting communication channels and enhancing systems to detect abnormal activity[30]. The measures of network segmentation and the systems of access control are employed to reduce the extension of the security breach and stop the attacker’s lateral movement within smart vehicle networks[31]. Hardware-based protection systems, such as dedicated secure elements and trusted platform modules (TPMs), are deployed to secure critical elements and the integrity of the vehicles. These existing security safeguards run into problems in dealing with the unconventional aspects of smart networks of vehicles which comprise such dynamic and heterogeneous kind of device as well as such large amount of data which is generated and exchanged all the time./ Thus, the employment of novel security techniques has become the key to successful protection of the systems against future cyber-attacks in connected vehicles, for example, the combination of federated learning and blockchain may prove effective in that they allow for distributed and privacy-preserving security in large-scale ecosystems[32].

Assessing the present-day literature and research gaps is worthwhile as it steers the progress of security measures in autonomous vehicle systems. Major gaps here include a missing detailed research on integrating the new technologies such as a federated learning as well as blockchain in the network security. The application and functionality of security frameworks, as well as the socio-technical issues including user views and the regulatory implications, are given less grounding in research. Scalability and interoperability issues are overlooked as well As a result, bridging these gaps needs the cooperation of researchers with different disciplines, empirical

Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

studies, and engagement of stakeholders, which can give birth to adapted security solutions[33].

3. Proposed Work

The strategy for implementing blockchain-based intelligent intrusion detection system is proposed by the integration of the detection systems into the existing systems. This covers concurrently integrating the best of signature, anomaly, and machine learning detecting techniques, and wisely mixing-up their distinct strengths. Signature-based detection method is useful for identifying the avoided trends while the anomaly-based detection method uses the established baseline system behavior as a base. At the same time, the introduction of the machine learning enables the system to determine in real time additional and new threats. This stratum of dynamic decision-making provides the platform for this fusion process. The system powered by machine learning algorithms comes up with adaptive weights for each detection method to each detection method with the decision based on history and context.

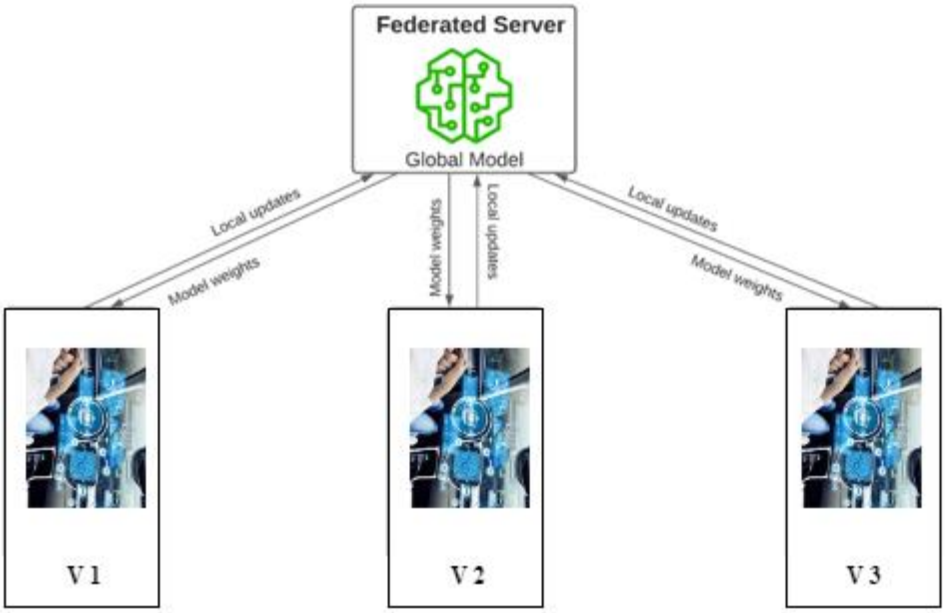

Figure 1. Federated Leaning Architecture

As shown in Figure 1, federated learning protocols encompass a decentralized ML architecture, which allows the training of the algorithms in multiple devices, or servers, without sharing raw data across the network. This scheme is designed to de-emphasize the privacy and security concerns, as only model updates are shared with the central server, whereas data remains on the devices of origin. This process guarantees that the risk of deep-data windowing is minimized making federated learning fit perfectly for such closed, confidential data scenarios,

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

where privacy is a must. Assembling models derived from the knowledge obtained from integrated diverse environmental data points disposed of across different locations can very well achieve robust and generalized models that have overcome the barriers posed by isolated data silos. It is not only this which provides more opacy and security, but also the reduced bandwith needed for moving over large datasets, allowing machine learning to be more easily available and efficient in many different sectors.

Network architectures are the backbone of an advanced communication system in automobiles of the modern era that make it possible to interconnect all the onboard features. The Controller Area Network (CAN) functions as a basic protocol for intra-car communication and enables real-time data transfer between electronic control units (ECUs) that control the engine, transmission and brake systems. In addition to CAN, the Local Interconnect Network (LIN) offers a low cost alternative for simple tasks like interior lights and window controls. The Global Positioning System (GPS) occupies a key position in navigation and location-based services, transmitting precise positioning data to features as lengthy as route guidance and vehicle tracking. Media Oriented Systems Transport (MOST) technology ensures low latency and reduced bandwidth for multimedia communication within the vehicle, thus enabling entertainment systems such as audio and video playback. Due to the Internet of Things (IoT) systems, vehicles can now communicate with external networks, leading to the assets like remote diagnostics, over-the-air software updates, and vehicle to vehicle (V2V) communication for improved safety and efficiency. Sensor fusion and communication technologies are utilized by ADAS to increase driver safety and comfort by offering features such as adaptive cruise control, lane-keeping assist and collision avoidance systems. Smart security devices use internet-based sensors and surveillance cameras to monitor the vehicle surroundings and prevent it from being stolen or accessed without authority. Infotainment systems, in turn, provide a variety of entertainment and communication options, such as internet connectivity, smartphone integration, and online streaming media services. Automatic cars constitute the epitome of such network architectures; they allow for the integrated services onboard as well as offboard to make the driving experience smart, interactive and interactive, meanwhile prepare the ground for the future technologies in the autonomous driving and mobility services.

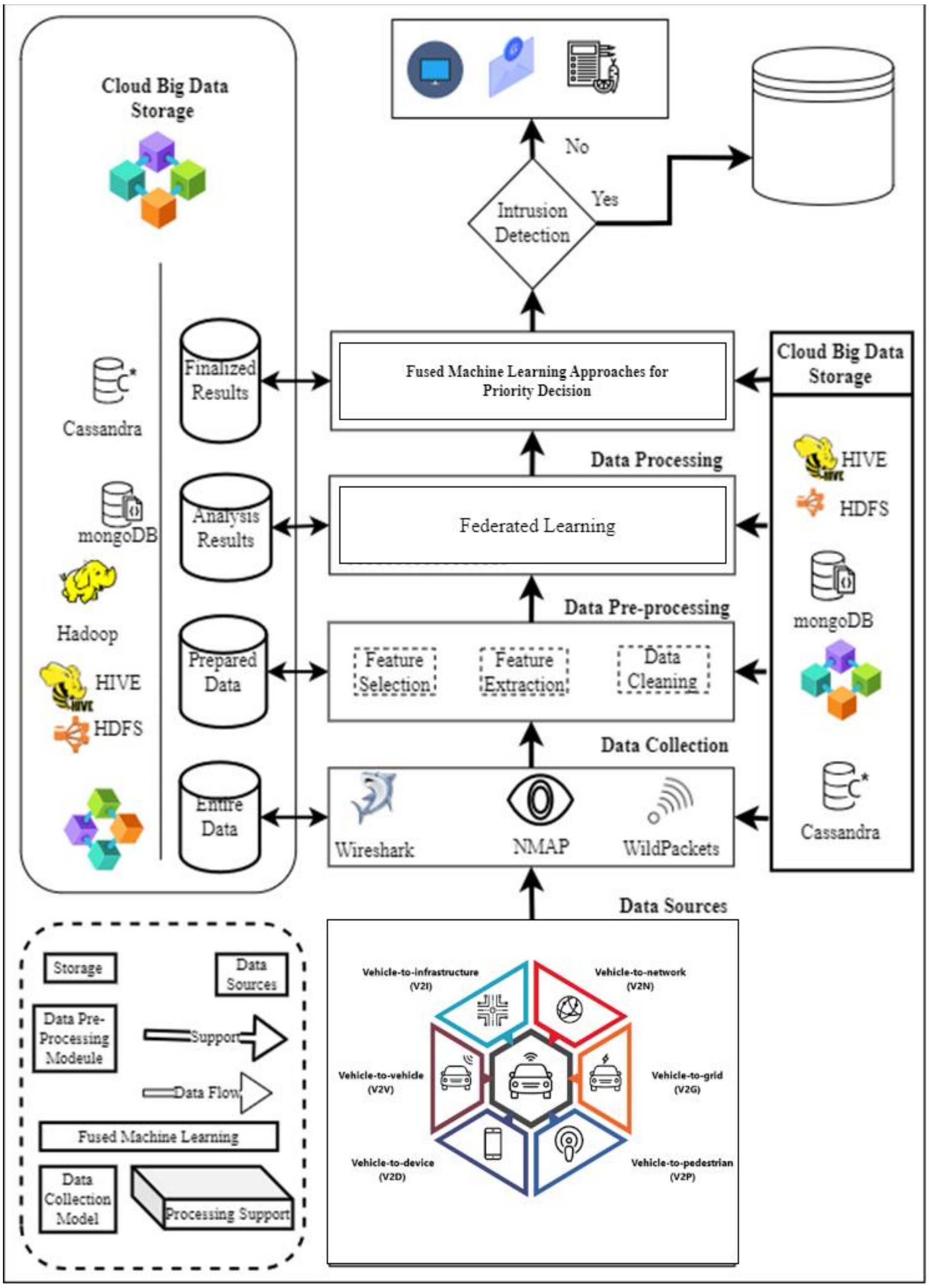

Figure 2. Proposed Model for Intrusion Detection in Smart Vehicles using Block-Chain and Federated Learning

Figure 2 shows the proposed blockchain-driven model for intrusion detection in smart Vehicle networks and clarifies the concluded coherent sequence of steps, each intricately connected to a blockchain-driven cloud database for transparent execution. Initiating with the Data Sources stage, data is gathered from a dataset as input sensors in smart Vehicles.

Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

Let Xt be the set of samples at node t. At each internal node t, a feature f is chosen to split the data into two subsets CapX,θ=argmaxf,0 Criterion (Xtf,θ) (1)

Where θ is the threshold for feature f, thedata is then split and child nodes tleft and tright are created.

At each node, the algorithm selects the best feature to split the data based on a specific criterion, often Gini impurity or entropy. Let p(i,t) be the proportion of samples of class i at node t.. The Impurity is defined as :

$\text{Impurity }(t)=1-\ \sum_{(i=1)}^{c}(p(i,t)))\ 2$ (2)

$\text{Defined as }(t)=-\sum_{(i=1)}^{c}(p(i,t)\ \log2\ (p(i,t))$ (3)

The algorithm splits nodes recursively until a stopping condition is met, often a maximum depth M or a minimum number of samples per leaf. At the leaf node, the majority class Ypred among the samples is assigned as the predicted class label:

$Ypred = argmaxi \sum_{i€Xt}〖p(i,t)〗$ (4)

Given a training dataset:

$X=\{x1,x2,\ldots xN\}$ (5)

where Xi represents the input feature vector of sample i.

$y=\{y1,y2,\ldots yN\}$ (6)

where yi is the binary class label for sample i (yi €{ -1,+1}).

SVM aims to find a hyperplane that maximizes the margin between the two classes while minimizing classification errors. The decision function of a linear SVM can be represented as:

$f(x)=sign\ (\sum_{(i=1)}^{N}〖aiyiK(x,xi)+b〗$ (7)

Where:

ai are the Lagrange multiplier obtained through optimization.

K (x,xi) is the kernel function used are the linear kernel

$K\ (x,xi)=xTxi=exp(-\ ɣ,x-xi)2$ (8)

For non-linear separation.

Where C is the regularization parameter that controls the trade-off between maximizing the margin and minimizing the classification error

3.1 Dataset Descriptions

During this interval, data is transmitted bidirectionally between a source IP address and a target IP address, following a clearly defined protocol. Furthermore, it is important to note that each

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

link is explicitly classified as either normal or an attack, with just one distinct attack type assigned to it. Each connection record is comprised of around 100 bytes. 41 quantitative and qualitative features are extracted from normal and attack data for every TCP/IP connection. These features consist of 3 qualitative aspects and 38 quantitative features. The class variable comprises two distinct groups, namely Normal and Anomalous.

Table 2. Dataset attributes

Data Type	Features	Data Type	Features
Duration	integer	is_guest_login	integer
protocol_type	nominal	Count	integer
Service	nominal	srv_count	integer
Flag	nominal	serror_rate	float
src_bytes	integer	srv_serror_rate	float
dst_bytes	integer	rerror_rate	float
Land	integer	srv_rerror_rate	float
wrong_fragment	integer	same_srv_rate	float
Urgent	integer	diff_srv_rate	float
Hot	integer	srv_diff_host_rate	float
num_failed_logins	integer	dst_host_count	float
logged_in	integer	dst_host_srv_count	float
num_compromised	integer	dst_host_same_srv_rate	float
root_shell	integer	dst_host_diff_srv_rate	float
su_attempted	integer	dst_host_same_src_port_rate	float
num_root	integer	dst_host_srv_diff_host_rate	float
num_file_creations	integer	dst_host_serror_rate	float
num_shells	integer	dst_host_srv_serror_rate	Float
num_access_files	integer	dst_host_rerror_rate	Float
num_outbound_cmds	integer	dst_host_srv_rerror_rate	Float
is_host_login	integer		

Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

Table 2 elaborates on the dataset that is to be subjected to an audit that has been submitted, encompassing a diverse range of simulated intrusions within a military network setting. The system was designed to facilitate the collection of unprocessed TCP/IP dump data from a network by emulating a standard local area network (LAN) configuration commonly found in the United States Air Force. The Local Area Network (LAN) was simulated to resemble a genuine environment and subjected to numerous attacks. A connection refers to a series of Transmission Control Protocol (TCP) packets that initiate and terminate within a specified time interval.

The dataset is widely acknowledged as the benchmark test set for network Intrusion Detection Systems (IDSs) (kaggle.com). The dataset is partitioned into two independent components: the training and validation datasets. A clear and identifiable identity characterizes the training dataset, but the test dataset lacks any specific identification. The test dataset includes attack types do not present in the training dataset. As a result, this improves the accuracy and dependability of the system's identification process. Preventive data processing was undertaken to eliminate differences in the data and protect against errors. The data classification process necessitates a relatively short duration, often quantified in milliseconds.

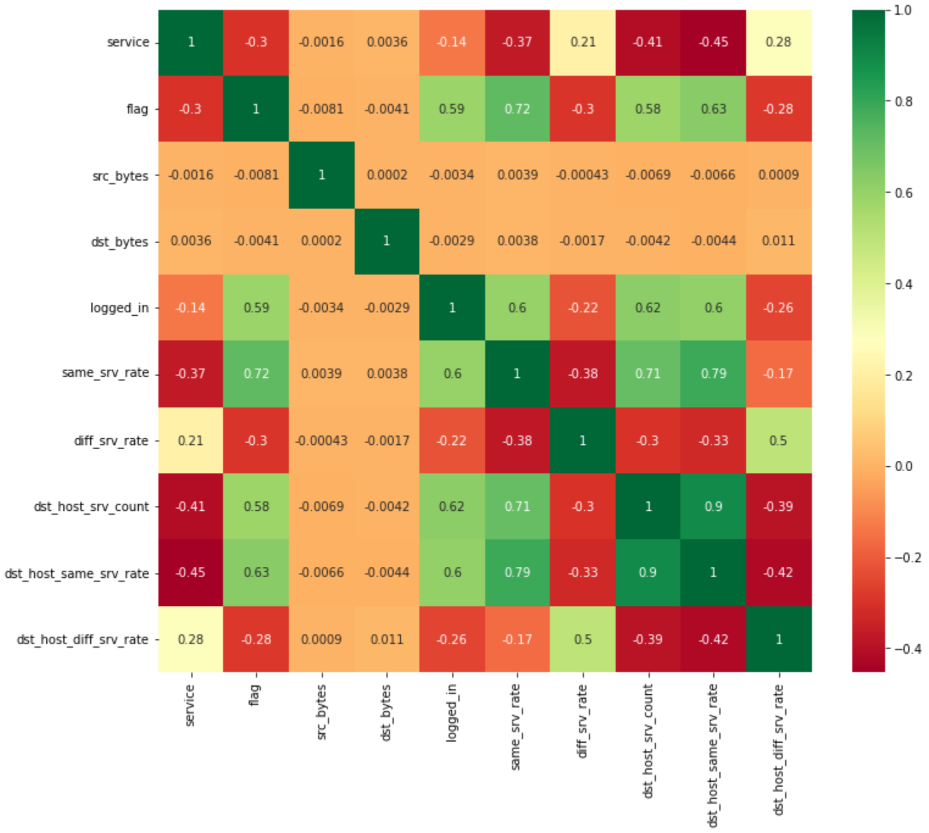

Figure 3. Distribution of Network Intrusion Detection Attributes by Data Type and Features

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

The collected information is then transmitted to the Data Collection layer, which securely exists in raw form within the cloud database, facilitated by wireless communication. The data preprocessing stage refines the raw data to minimize noise and optimise it for further analysis. Forwarding to the Data Processing layer, two machine learning algorithms (SVM and decision tree) are applied to predict trained patterns, and their outcomes are fed into the Fused Machine Learning stage.

The collected information is then transmitted to the Data Collection layer, which securely exists in raw form within the cloud database, facilitated by wireless communication. The data preprocessing stage refines the raw data to minimize noise and optimise it for further analysis. Forwarding to the Data Processing layer, two machine learning algorithms (SVM, Decision Tree, Neural Network and Random Forest) are applied to predict trained patterns, and their outcomes are fed into the Fused Machine Learning stage.

3.2 Simulation and Result

This research proposed a fused machine learning algorithm for Smart Vehicle Network networks to address intrusion detection, which is one of the critical and main concerns of recent times. The proposed approach is applied to a dataset, with 70% of the dataset used for training and 30% for validation to compare the performance matrices given in Table 3.

Table 3. Prediction and Classification Layout Parameter

Sensitivity (TPR)	$\frac{TP}{TP + FN}$	(10)
Specificity (TNR)	$\frac{TN}{TN + FP}$	(11)
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	(12)
Miss Rate (FNR)	$\frac{FN}{FN + TP}$	(13)
Fall out (FPR)	$\frac{FP}{FP + TN}$	(14)
Positive Ratio (LR +)	$\frac{TPR}{FPR}$	(15)

Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

Negative Ratio (LR –)	$\frac{\text{FNR}}{\text{TNR}}$	(16)
Precision (PPV)	$\frac{\text{TP}}{\text{TP}+\text{FP}}$	(17)
Negative Predicted Value (NPV)	$\frac{\text{TN}}{\text{TN}+\text{FN}}$	(18)

Table 3 demonstrated various evaluation criteria that were used for the assessment of a recommended model. This index is the gold standard of academic investigations and is known as specificity, which can be demonstrated as the True Positive Rate (TPR). Specificity as well as "True Negative Rate (TNR) " are widely used in scholarly literature - representation of correctly identified negative results. The accuracy of model to identify negatives is often measured as the number of wrong answers with the concept "False Negative Bias (FNR)" meaning the perfection of rate in classify negatives also called as miss-rate Academic literature supposes a term "False Positive Rate (FPR)" and a tropism call fallout at the same time. In academic writing, the term thesis" is often found to be positive predictive value (PPV). The confusion matrix is one methodology that is used for evaluating the accuracy of the classifier. The passage offers a concise summary of the comparison between expected outcomes and actual outcomes, categorizing the results into four distinct categories: it will involve the calculation of true positives (true events correctly predicted), true negatives (true events correctly predicted), false positives (incorrectly predicted negative events) and false negatives (incorrectly predicted positive events). The confusion matrix with performance metrics like accuracy, precision, recall, etc will show how the model could perform. This table uses symbols to indicate the true and false categories and their corresponding percentages in the confusion matrix. The method of measuring the quality of a model classifier is thus enabled by this.

Table 4. Training of Proposed Model using Decision Tree and Support Vector Machine

Number of Samples (17,635)	Results (Output)	
Expected Output	Predicted Positive	Predicted Negative
Positive (9,699)	TP (8,993)	FP (705)
Negative (7,583)	FN (352)	TN (7,230)

Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

Table 5. Validation of Proposed Model using Decision Tree and Support Vector Machine

Number of Samples (7,557)	Results (Output)	
Expected Output	Predicted Positive	Predicted Negative
Positive (4,156)	TP (3,778)	FP (377)
Negative (3,249)	FN (226)	TN (3,022)

In tables 4 and 5, respectively, the machine learning approach has been applied to a dataset consisting of 25,193 records. The dataset is partitioned into two subsets: a training set including 70% of the total samples (17,635 samples) and a validation set consisting of 30% of the samples (7,557 samples) from each class. This division is done to facilitate the aims of training and validation.

Table 5 presents the proposed model for intrusion detection on the server during the "Training Phase (TP)". During the training, a comprehensive 17,635 samples were employed. The dataset consists of 9,699 positive samples and 7,583 negative samples. 8,993 samples classified as "True Positive" demonstrate correct predictions, with no instances of intrusion being detected. A total of 705 records were mistakenly predicted as negative, indicating the detection of an intrusion. Similarly, a total of 7,583 samples are collected, with negative outcomes indicating intrusion detection. A total of 7,230 samples have been successfully predicted as negative, hence signifying the presence of intrusion. A total of 352 samples have been inaccurately classified as positive, suggesting that no intrusion has been detected despite the presence of an intrusion on the smart vehicles network.

During the validation process in Table 5, a total of 7,557 samples are applied. The dataset consists of 4,156 positive and 3,249 negative samples, resulting in 7,557 samples. The samples have been accurately recognized as "True Positive" in 3,778 instances, signifying the absence of any incursion. 377 data have been incorrectly classified as negatives, suggesting the presence of an incursion. In a similar vein, 3,249 samples are collected, with a negative outcome indicating the presence of infiltration. 3,022 samples have been successfully predicted as negative, suggesting the presence of infiltration. In conclusion, 226 samples have been inaccurately classified as positive, indicating the absence of any detected incursion despite the presence of the network.

Table 6. Statistical Measurement of Training and Validation using Support Vector Machine and Decision Tree

Training Model	Phases	Accuracy %	Sensitivity (TPR)	Specificity (TNR)	Miss Rate (FNR)	Precision (PPV)
SVM	Training	93.88	0.9623	0.9111	0.0377	0.9273
	Validation	91.84	0.9434	0.8889	0.0566	0.9091
Decision Tree	Training	92.85	0.9541	0.8986	0.0459	0.9168
	Validation	90.28	0.9310	0.8699	0.0690	0.8932
Neural Network	Training	90.77	0.9213	0.8911	0.0877	0.8973
	Validation	88.84	0.9134	0.8589	0.0966	0.8891
Random Forests	Training	88.88	0.8923	0.8911	0.0897	0.8873
	Validation	85.76	0.8734	0.8689	0.0976	0.8681

The assessment of the presented model is concerned with the SVM, Decision Tree, Neural Network, Random Forest algorithms and is detailed below in Table 6. Thus we undergo the analysis of the metrics, like accuracy of detection, sensitivity, specificity, miss extent and precision. It's carried at both training and validation steps. The SVM (Support Vector Machine) training performance was characterized by 93.88% of accuracy, 0.9623 of sensitivity, 0.9111 of specificity, 0.0377 of the rate of misses, and 0.9273 of precision. The entire validation procedure goes through the performance of the model, which shows up in the form of various metric performance, including detection accuracy, sensitivity, specificity, miss rate and precision. These metrics are at 91,84, 0.1941, 0.0889, 0.0566, and 0.9091. At the training stage, the Decision Tree demonstrated detection accuracy as 92.85% and showed sensitivity, specificity, and miss rate values of 0.9541, 0.8986, and 0.0459, correspondingly, and the precision value was 0.9168. Two more algorithms – Neural Network and Random Forest – evaluated with this method is also part of the table 7. The validation model described throughout the process is characterized by the performance-metrics its shows. Indicators of the system's performance incorporate detection accuracy,sensitivity, specificity, miss rate and accuracy. That's why the specific values for these metrics are 90.28%, 0.9310, 0.8699, 0.0690, and 0.8932.

During the validation step, the test data retained in the database is retrieved from the edge database, together with the explained patterns. These data and patterns are then utilised in machine learning techniques to make predictions on the presence of Intrusion. If the response is negative, the procedure is disregarded; conversely, if the response is affirmative, the

Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

notification signifies the presence of intrusion. The fusion strategy, which utilises machine learning techniques, involves developing and applying to enhance the performance of classification algorithms. Machine learning approaches such as Decision Trees, Support Vector Machines (SVM), Neural Network and Random Forest generate logical structures.

Besides the output of the proposed technique is obtained by using the fusion of SVM and Decision tree, where 17 out of the 18 random samples are classified accurately by the simulation as per the human decision-making. On the other hand, the value once deemed as low turned out to be in line with the new system encompassing the fused ML approach but not proper. It is determined that the performance accuracy of the proposed approach under study by fused ML techniques is 94.44% with a miss rate of 5.56%.

Table 7. Comparison Result of the Proposed Model

Sr. No.	Literature	ML Techniques	Security and Privacy	Miss Rate	Accuracy
1	Gao et al., 2019 [34]	Multi Tree	No	15.77%	84.23%
2	Latah et al., 2020 [35]	KNN+ELM	No	15.71%	84.29
3	Wu et al., 2018 [36]	CNN	No	20.52%	79.48%
4	Tavallaee et al., 2009 [37]	NB Tree	No	33.84%	66.16%
5	Ingre et al., 2015 [38]	ANN	No	18.8%	81.2%
6	Aggarwal et al., 2015 [39]	Random Tree	No	16.96%	83.04%
7	Ambusaidi et al., 2016 [40]	LSSVM-IDS	No	21.14%	78.86%
8	Al-Qatf et al., 2018 [41]	SAE_SVM	No	15.04%	84.96%
9	Proposed fused ML Approach	Decision Tree + SVM	Yes	5.56%	94.44%
10	Proposed fused ML Approach	Neural Network + Random Forest	Yes	10.18%	89.83%

Table 7 elaborates on the comparison of the proposed approach with previously published approaches using the fused ML approach, showing 94.44% accuracy and 5.56% miss-rate for

Power System Technology
ISSN:1000-3673

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

Decision Tree and SVM, 89.83% accuracy and 10.18% miss rate for Neural Network and Random Forest, which are better as compared to others.

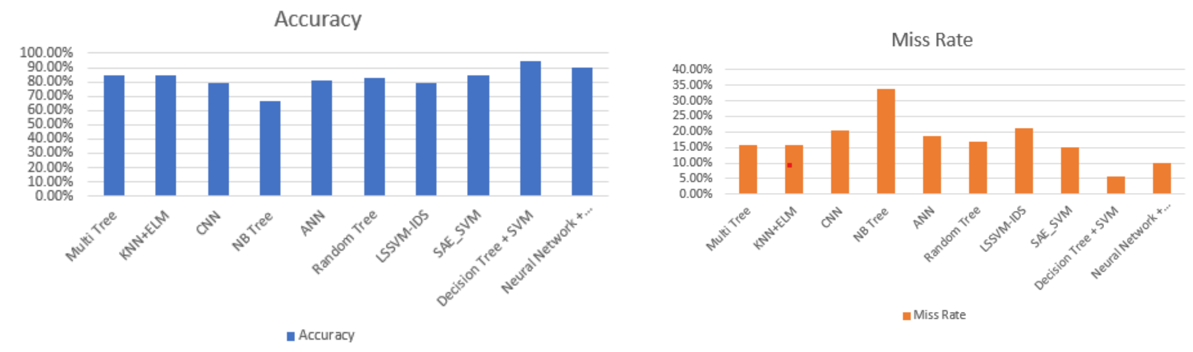

Figure 3. (a) Accuracy Comparison Figure 3. (b) Miss Rate Comparison

The proposed approach characterized a significant improvement in the power of intrusion detection to make the solution appropriate for smart vehicles network security.

4. Conclusion

In the ever-growing field of Smart Vehicle Network, traditional intrusion detection systems rarely have sufficient features to overcome new and advanced cyber security threats in dynamic and smart technologies. This research considers a blockchain-driven fused ML approach for IDS detection to address the intricate challenges of smart vehicles internal network security with the constantly evolving cyber threats. The proposed research is pursuing a solution to close the gaps in model interpretation, robust decision making and transparency, and to respond to the dynamic threats during the process.

The integration of blockchain technology and smart fusion techniques in intrusion detection is applied in this research. The proposed approach using fused ML demonstrates significant performance, attaining a training accuracy of 94.44 as well as a miss-rate of 5.56%. This outcome highlights the system's robust learning and classification proficiencies, revealing the promising implications of combining blockchain and intelligent fusion for enhancing cybersecurity strategies—the system's integration of performance as compared to the previous published approaches.

Author Contributions: Conceptualization, methodology, software, writing—original draft preparation, J.A.M. and S.A.; writing—review and editing, F.S and J.B.; supervision, A.H. and F.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: Not applicable

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

Acknowledgments: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number RGP.2/379/45.

Conflicts of Interest: The authors declare no conflicts of interest.

References

[1] Abou El Houda, Z., et al., Blockchain-Enabled Federated Learning for Enhanced Collaborative Intrusion Detection in Vehicular Edge Computing. IEEE Transactions on Intelligent Transportation Systems, 2024.

[2] Ahmad, J., et al., Machine learning and blockchain technologies for cybersecurity in connected vehicles. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2024. 14(1): p. e1515.

[3] Amari, H., Smart models for security enhancement in the internet of vehicles. 2023, Normandie Université; Université de Sfax (Tunisie).

[4] Aouedi, O., et al., Handling privacy-sensitive medical data with federated learning: challenges and future directions. IEEE Journal of Biomedical and Health Informatics, 2022. 27(2): p. 790-803.

[5] Asad, M., et al., Secure and Efficient Blockchain-Based Federated Learning Approach for VANETs. IEEE Internet of Things Journal, 2023.

[6] Boualouache, A., et al., On-demand security framework for 5GB vehicular networks. IEEE Internet of Things Magazine, 2023. 6(2): p. 26-31.

[7] Chellapandi, V.P., et al., Federated learning for connected and automated vehicles: A survey of existing approaches and challenges. IEEE Transactions on Intelligent Vehicles, 2023.

[8] Ebrahim, M., A. Hafid, and E. Elie, Blockchain as privacy and security solution for smart environments: A Survey. arXiv preprint arXiv:2203.08901, 2022.

[9] Ganesan, P. and S.K. Jagatheesaperumal, Revolutionizing Emergency Response: The Transformative Power of Smart Wearables Through Blockchain, Federated Learning, and Beyond 5G/6G Services. IT Professional, 2023. 25(6): p. 54-61.

[10] Gebremariam, G.G., J. Panda, and S. Indu, Blockchain-based secure localization against malicious nodes in IoT-based wireless sensor networks using federated learning. Wireless communications and mobile computing, 2023. 2023.

[11] Gebremariam, G.G., J. Panda, and S. Indu, Research Article Blockchain-Based Secure Localization against Malicious Nodes in IoT-Based Wireless Sensor Networks Using Federated Learning. 2023.

[12] Goethals, T., B. Volckaert, and F. De Turck, Enabling and leveraging ai in the intelligent edge: A review of current trends and future directions. IEEE Open Journal of the Communications Society, 2021. 2: p. 2311-2341.

Received: 05-04-2024 Revised: 28-04-2024 Accepted: 02-05-2024

[13] Hazra, A., et al., Federated-learning-aided next-generation edge networks for intelligent services. IEEE Network, 2022. 36(3): p. 56-64.

[14] Issa, W., et al., Blockchain-based federated learning for securing internet of things: A comprehensive survey. ACM Computing Surveys, 2023. 55(9): p. 1-43.

[15] Javeed, D., et al., Quantum-Empowered Federated Learning and 6G Wireless Networks for IoT Security: Concept, Challenges and Future Directions. Quantum, 2023. 96: p. 1.

[16] Liu, H., et al., Blockchain and federated learning for collaborative intrusion detection in vehicular edge computing. IEEE Transactions on Vehicular Technology, 2021. 70(6): p. 6073-6084.

[17] Mao, B., et al., Security and privacy on 6g network edge: A survey. IEEE communications surveys & tutorials, 2023.

[18] Movahedian Attar, M., Adaptive Model Aggregation for Decentralized Federated Learning in Vehicular Networks. 2023.

[19] Neto, H.N.C., et al., Securing Federated Learning: A Security Analysis on Applications, Attacks, Challenges, and Trends. IEEE Access, 2023.

[20] Pham, Q.-V., et al., Fusion of federated learning and industrial Internet of Things: A survey. arXiv preprint arXiv:2101.00798, 2021.

[21] Rahman, M.A., et al., Secure and provenance enhanced internet of health things framework: A blockchain managed federated learning approach. Ieee Access, 2020. 8: p. 205071-205087.

[22] Saraswat, D., et al., Blockchain-based federated learning in UAVs beyond 5G networks: A solution taxonomy and future directions. IEEE Access, 2022. 10: p. 33154-33182.

[23] Sethi, P., Reinforcement Learning assisted Adaptive difficulty of Proof of Work (PoW) in Blockchain-enabled Federated Learning. 2023, Virginia Tech.

[24] Sood, K., et al., Intrusion detection scheme with dimensionality reduction in next generation networks. IEEE Transactions on Information Forensics and Security, 2023. 18: p. 965-979.

[25] Truong, V.T. and L.B. Le, MetaCIDS: Privacy-preserving collaborative intrusion detection for metaverse based on blockchain and online federated learning. IEEE Open Journal of the Computer Society, 2023.

[26] Truong, V.T. and L.B. Le, Security for the Metaverse: Blockchain and Machine Learning Techniques for Intrusion Detection. IEEE Network, 2024.

[27] Zhu, C., et al., Blockchain-enabled federated learning for UAV edge computing network: Issues and solutions. Ieee Access, 2022. 10: p. 56591-56610.

[28] Xu, H., et al., A survey on digital twin for industrial internet of things: Applications, technologies and tools. IEEE Communications Surveys & Tutorials, 2023.

[29] Chen, Y., et al. Federated learning for metaverse: A survey. in Companion Proceedings of the ACM Web Conference 2023. 2023.

[30] Vaiyapuri, T., et al., Deep learning approaches for intrusion detection in IIoT networks–opportunities and future directions. International Journal of Advanced Computer Science and Applications, 2021. 12(4).

[31] Haider, N., M.Z. Baig, and M. Imran, Artificial Intelligence and Machine Learning in 5G Network Security: Opportunities, advantages, and future research trends. arXiv preprint arXiv:2007.04490, 2020.

[32] Otoum, S., I. Al Ridhawi, and H.T. Mouftah, Preventing and controlling epidemics through blockchain-assisted ai-enabled networks. Ieee Network, 2021. 35(3): p. 34-41.

[33] Truong, V.T., L. Le, and D. Niyato, Blockchain meets metaverse and digital asset management: A comprehensive survey. Ieee Access, 2023. 11: p. 26258-26288.

[34] Gao, X., et al., An adaptive ensemble machine learning model for intrusion detection. Ieee Access, 2019. 7: p. 82512-82521.

[35] Latah, M. and L. Toker, An efficient flow-based multi-level hybrid intrusion detection system for software-defined networks. CCF Transactions on Networking, 2020. 3(3-4): p. 261-271.

[36] Wu, K., Z. Chen, and W. Li, A novel intrusion detection model for a massive network using convolutional neural networks. Ieee Access, 2018. 6: p. 50850-50859.

[37] Tavallaee, M., et al. A detailed analysis of the KDD CUP 99 data set. in 2009 IEEE symposium on computational intelligence for security and defense applications. 2009. Ieee.

[38] Ingre, B. and A. Yadav. Performance analysis of NSL-KDD dataset using ANN. in 2015 international conference on signal processing and communication engineering systems. 2015. IEEE.

[39] Aggarwal, P. and S.K. Sharma, Analysis of KDD dataset attributes-class wise for intrusion detection. Procedia Computer Science, 2015. 57: p. 842-851.

[40] mbusaidi, M.A., et al., Building an intrusion detection system using a filter-based feature selection algorithm. IEEE transactions on computers, 2016. 65(10): p. 2986-2998.

[41] Al-Qatf, M., et al., Deep learning approach combining sparse autoencoder with SVM for network intrusion detection. Ieee Access, 2018. 6: p. 52843-52856.