Blockchain Based Security and Trust Mechanisms for Vehicular Ad hoc Networks

Abdullah Alharthi

School of Computing and Communications
Lancaster University

A thesis submitted in partial fulfillment for the degree of
Doctor of Philosophy

December 2022

This thesis is dedicated to my loving parents. This couldn't have been possible to complete my doctoral studies without their endless love and encouragement. I adore you both and am grateful for everything you have done to help me".

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. The material has not been submitted, in whole or in part, either for a degree at this, or any other university. This thesis does not exceed the maximum permitted word length of 80,000 words including appendices and footnotes, but excluding the bibliography.

Abdullah Alharthi

Abstract

In the near future, intelligent vehicles (IV) will be part of the Internet of Things (IoT) and will offer valuable services and opportunities that could revolutionize human life in smart cities. The Vehicular Ad-hoc Network (VANET) is the core structure of intelligent vehicles. It ensures the accuracy and security of communication in vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) modes to enhance road safety and decrease traffic congestion. However, VANET is subject to security vulnerabilities such as denial-of-service (DoS), replay attacks and Sybil attacks that may undermine the security and privacy of the network. Such issues may lead to the transmission of incorrect information from a malicious node to other nodes in the network. Therefore, a biometrics blockchain (BBC) framework to secure data sharing among vehicles in VANET and to retain statuary data in a conventional and trusted system is designed. In the proposed framework, the biometric information is used to keep a record of the genuine identity of the message sender, thus preserving privacy and provide conditional anonymity. The suggested BBC approach provides security and trust among vehicles in VANET, as well as the capability to track identities as needed. To show the feasibility of the suggested framework utilizing the urban mobility model, simulations in OMNeT++, veins, and SUMO were performed. The framework's performance is assessed in respect to packet delivery rate, packet loss rate, and computational cost. The results demonstrate that our unique model outperforms previous techniques.

Vehicles, on the other hand, find it challenging to assess the authenticity of received messages in non-trusted environment. The primary challenges in VANET are trust, data accuracy, and dependability of data broadcast over the communication channel. To protect against these threats, the majority of researchers have proposed cryptography-based solutions to verify the sender's legitimacy but are incapable of preventing the broadcast of false or malicious messages from a legal sender. Therefore, in this thesis, we propose a formal technique to compute and classify trust for vehicles in networks. Vehicles can evaluate the received messages, calculate the vehicle’s reputation, and check the message correctness based on numerous factors. The latest reputation of the vehicle will be stored on the blockchain. A machine learning approach is employed to classify the trust. Comprehensive tests are carried out using the dataset in order to measure the efficiency of the proposed ensemble-based learning and feature selection based on random forest. The outcomes of the experimentations reveal that the suggested ensemble learning approach with attribute selection produces an accuracy of 99.98%, which is higher than the baseline models studied in this thesis.

Publications

A. Alharthi, Q. Ni and R. Jiang, A Privacy-Preservation Framework Based on Biometrics Blockchain (BBC) to Prevent Attacks in VANET, in **IEEE Access**, vol. 9, pp. 87299-87309, 2021, DOI: 10.1109/ACCESS.2021.3086225.

A. Alharthi, Q. Ni and R. Jiang, Internet of Vehicles: Design, Architecture, and Security Challenges, **Internet of Things: A Hardware Development Perspective, Taylor & Francis**, 2022, pp. 205-2022, ISBN 9780367641467

A. Alharthi, Q. Ni, R. Jiang and M. Khan. ‘’A Formal Method of Trust Computation in VANET: A Spatial and Temporal Behavioral Approach’’, International Conference on Smart Technologies in Urban Engineering (STUE-2022), June 2022.

A. Alharthi, Q. Ni and R. Jiang, M A Khan, A Formal Method for Trust Computation and Classification in Vehicular Ad Hoc Network, **IEEE Internet of Things Journal** (IOTJ), IEEE, (Submitted)

Acknowledgements

I would like to acknowledge my supervisors. To begin, I want to offer my sincere gratitude to Professor Qiang Ni for everything he has done for me during my doctoral studies, including his all-encompassing support and very genuine care, as well as his advice and inspiration all through my Ph.D. Professor Ni you are the best teacher I've ever had. Second, I'd like to express my gratitude to Dr. Richard Jiang for his assistance as a co-supervisor. It's an honor for me to be a student at Lancaster University.

Additionally, I would want to express my gratitude to all of my friends at the School of Computing and Communications (SCC), all of whom have made a significant contribution to my time here and the successful completion of my Ph.D.

I must also express my appreciation to Prof Mohammad Ayoub Khan, University of Bisha, Saudi Arabia, for his inspiration and mentorship during the thesis writing.

My deepest gratitude goes to the almighty God, to whom I owe every step of this journey from the beginning to the very end.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ITS	Intelligent Transportation System
WAVE	Wireless Access in Vehicular Environment
DSRC	Dedicated Short Range Communications
GPS	Global Positioning System
IEEE	Institute of Electrical and Electronics Engineers
VANET	Vehicular Ad Hoc Network
V2V	Vehicle to Vehicle
V2I	Vehicle to Infrastructure
V2X	Vehicle-to-Everything
RSU	Road Side Unit
OBU	On-board Unit
ECU	Digital Control Unit
AU	Application Unit
BSM	Basic Safety Message
WSN	Wireless Sensor Network
PHY	Physical Layer
MAC	Medium Access Control
V_i	Identity of vehicle *i*
KU_i	Public-key of vehicle *i*
V_p	Vehicle pseudo id
KR_i	Private-key of vehicle *i*
v_{reg}	Vehicle registration number
v_{mod}	Vehicle model
v_{chas}	Vehicle chassis number
v_{rank}	Vehicle ranking
O_i	On-board unit of vehicle *i*
D_i	User's information
d_{name}	User's name,
d_{bio}	User's biometrics data
d_{lic},	User's license
d_{rank}	User's ranking
M_i	Message sent by vehicle *i*
$\mathcal{H}$	Cryptographic hash function
C_i	Certificate of Vehicle i
M_Type	Message type: beacon, alert

$\mathcal{E}$	Encryption function
$\mathcal{D}$	Decryption function
E_{id}	Event identification number
E_{loc}	Event location
E_{type}	Event type
T	Event timestamp
t	Current time
R_i	Reputation of the sending vehicle *i*
th_f, th_R	Freshness and Reputation threshold
V_i	Identity of vehicle *i*
KU_i	Public-key of vehicle *i*
V_p	Vehicle pseudo id
KR_i	Private-key of vehicle *i*
MDCT	modified discrete transformation
EC	Euclidian distance
C_t^i	Correctness of the message at time *t*
f_c	Function of correctness
M	Message
S	Function of similarity
R_t	Reputation at time *t*
w_i	Weight of the function
f_f	Function of freshness
f_d	Function of distance between the message sender
α	The age of node/vehicle
θ	The participation degree of the node
f_T	Function of trust
i_{Reg}	Registration date of vehicle *i*
n	Number of vehicles in the network.
m	Total number of messages
V_i	Vehicle *i*
PDR	Packet Delivery Ratio
TP	Transmission Power
β	coefficient
TTL	Time to Live

Chapter 1

Introduction

With the rapid development of smart cities, the number of intelligent vehicles in mobile ad-hoc networks has raised significantly. The number of intelligent vehicles on the planet is predicted to reach 2 billion within the next ten years. [1]. With an increasing number of vehicles on the road around the world, different transportation challenges like road accidents, congestion, and air quality arise. As reported by the World Health Organization (WHO), traffic accidents claim over 1.25 million human lives each year, making them the tenth biggest cause of human deaths worldwide [2]. Consequently, there has been a surge in global interest in addressing transportation challenges. Several countries have launched numerous projects in this arena, which has resulted in the birth of Intelligent Transportation Systems (ITS). ITS makes use of advances in information and communication technology (ICT) to enhance transportation performance and traffic safety. ITS-UK is committed to improve transport in regards to traffic efficiency and safety, lowering journey times, allowing smart parking technology, and turning the environment green in the United Kingdom [3]. ERTICO is in charge of ITS improvement in Europe, providing safer, greener, and more intelligent vehicle mobility [4]. The Transportation Department (DOT) in the United States is promoting public transport innovation via its comprehensive plan 2015-2019 [5]. The design involves improved transportation mobility while reducing environmental effect. Japan is actively pursuing this field in Asia by establishing ITS regulations to enhance transportation [6]. ITSB is committed to improving the standard of transportation in Brazil. This involves controlling

the increased transportation need on current infrastructure, decreasing congestion, and ensuring overall traffic safety [7]. ITS in Australia is expanding as a result of its strategy plan 2013-2018. Through improvements in the ICT field, this scheme aims to improve effective and traffic transportation [8]. The overarching goal of all of these programs is to enhance people 's standard of life through improved and secure transportation.

Therefore, vehicular ad-hoc network (VANET) has been created to handle such large number of vehicles. In VANET, each vehicle is equipped with wireless communication devices named onboard units (OBU). These devices have hardware security chip to store sensitive information of the vehicle. Each vehicle represents a node in the network and it has the capability of sending and processing information. Improving road safety, reducing the traffic accidents, and enhancing traffic efficiency are several aims of forming VANETs [9]

The most advanced ITS technology, known as VANET, enables vehicles to interact with one another and nearby roadside units (RSUs) in order to address a number of transportation-related problems, including fewer road accidents, improved traffic monitoring, lessening traffic congestion, and offering on-board vehicle riders with entertainment.

Inadequate security could endanger the prospective advantages of Network applications. Consequentially, a variety of security objectives must be met in order to prohibit malicious attackers from manipulating exchanged sensitive information while protecting drivers' private information and establishing their responsibility in the event of an accident. To summarize, vehicle communications must not be a vulnerable link in terms of data security, instead it should offer users with at least the same standard of safety which is afforded in the absence of vehicular communications.

1.1 Research Problem

As the purpose of VANETs is to offer safety and comfort for intelligent vehicles, the information shared between vehicles plays a crucial part in vehicular communications. A relevant factor to consider is determining the context in which information might be trusted. The information transferred between vehicles is particularly important for safety-related applications; therefore, fast and precise sharing of this information could greatly reduce the frequency of fatal traffic accidents. The information could, however, become harmful if an attacker tamper with it. Applying security assessments is therefore crucial in

order to fend off such threats.

Under normal circumstances, peer vehicles in a VANET do not know each other's real identity. However, if an authenticated user becomes malicious and launches an "insider attack," rapid termination and tracing should be achievable without any additional overhead; consequently, assuring conditional privacy. Transparent reputation computation and trust establishment among the vehicles are the next most basic requirements. Due to the high variety and mobility of vehicular networks, the nearby vehicles cannot be totally trusted; this is regarded as a severe problem if the network is plagued by the presence of several malicious vehicles. Therefore, as the next step following proper authentication, we require a trust management approach [9]–[11] which not only assists the vehicles in determining the trustworthiness of the received messages, and even assists network operators in determining the sanctions or incentives for appropriate vehicles. Typically, a vehicle's trust value is based mostly on ratings of its previous behavior generated by other peer nodes, although some studies have addressed a variety of additional elements in determining a nodes' trustworthiness. According to the VANET security standards, we should assure secure communication with the least amount of computing difficulty possible, i.e., cancelling conventional encryption and decryption approaches, and limiting access control of any message to reputable and trustworthy nodes only once we have created a trusting environment between the vehicles.

Upon the introduction of the Bitcoin blockchain in 2008, industries and academia moved their attention to methods that may ensure the operation of centralized networks. From that point forward, certain VANET research projects concentrated on techniques for enhancing efficiency, ensuring privacy, putting in place trust models, and implementing security utilizing blockchain technology [12].

1.2 Research Motivation

The significant increase in the number of Intelligent vehicles across the globe have arisen numerous issues, such as traffic accidents, cyberattacks and data privacy. Therefore, it is of vital importance to design a secure, data traceable, and efficient platform yet decentralized in nature. Additionally, the growing number of IV have placed significant obstacles in the way of building an efficient and secure VANET such as:

- Centralization: Cloud service platforms are required by a typical VANET in order to secure and control data in a central database. More data storage is required because of the need for data exchange among vehicles increases. Following attacks or malicious meddling, remarkable data breaches may occur, that can be the catalyst for subsequent events which can be challenging to regulate.
- Security threat: The privacy and security of a vehicle are seriously threatened by the usage of wireless communication in VANETs since it is simple to watch and manipulate data as it is shared (e.g., they can be illegally tracked or remotely hijacked). The distribution of inaccurate messaging could result in possibly fatal traffic accidents, and false data can be transmitted by illicit vehicles, disrupting regular data communications

To overcome these obstacles, there must be urgent improvements in privacy and trust levels to guarantee that all communication processes are secure and have integrity. A secure and reliable decentralized data sharing system must be created to guarantee that VANETs can continue operating normally.

Data authentication, privacy, and trust models have been identified by researchers as the most important factors in improving security for VANETs. Therefore, blockchain technology has been investigated by numerous researchers and academics as it offers great potential in the security, credibility, tamper resistance, and traceability of VANETs [11]-[19].

1.3 Thesis Aim and Objective

This thesis aims to design a novel blockchain-based privacy perseveration and trust computation framework in VANET. First objective is to devise biometric blockchain (BBC) framework that provide a decentralized, secure, and trusted communication environment in VANET. Blockchain technology has been recognized as an excellent response to the problems of centralization, privacy and security while processing, controlling, and exchanging data within peer-to-peer networks. On one hand, vehicle to vehicle and vehicle to infrastructure communication must retain anonymity in order to keep vehicle identity private. On the other hand, this anonymity must have conditions attached to ensure that the vehicles may be tracked by authorities should a dispute arise. Hence, biometrics will be used with blockchain to achieve conditional anonymity in VANET. Furthermore, vehicle node-oriented solutions eliminate dishonest nodes from VANET

simultaneously measuring node trustworthiness. Such tactics, nevertheless, take no account of message quality and assume that "if a node is trusted, therefore the messages it delivers are likewise credible." In accordance with the message-oriented strategy paradigm, data quality is the only aspect that influences data interchange dependability. Once a large dataset is deployed, such algorithms analyse a collection of messages with transferred data provided by trusted nodes. It ensures accurate classification of nodes as trustworthy and malicious by making use of Machine Learning and thus provide

Second, objective is to investigate about trust and reputation mechanism in VANET. The trust and reputation model can ensure non-repudiation and establish a reputation for the vehicles based on the correctness of the transmitted messages and reputation credit maintained. Through the formal methods of a trust management system that is based on reputation and identity evaluation, it may be able to reward trustworthy cars while flagging untrustworthy vehicles in VANETs, therefore assuring trustworthy message broadcasting. The objective is to propose the ensemble-based machine learning model to classify the trust into trustworthy and non-trustworthy vehicles which ensures accurate classification of nodes as trustworthy or malicious by making use of machine learning and thus assure the establishment of an environment of trust between, vehicles, reduce the malicious vehicles and therefore attacks, and enhance the security of VANET.

1.4 Research Contributions

The contribution of the research work are as follows:

- Firstly, we propose biometrics blockchain (BBC) framework to make communication in VANET more secure. In this framework the biometrics features are combined with blockchain to provide secure transmission of messages, track, and identify the malicious vehicles which broadcast untrusted information. Moreover, each vehicle is assigned with biometric ID to guarantee the trust.
- Secondly, we derive formal methods of a trust based on reputation. The proposed framework is based on the spatial, temporal and behavioural paraments such as reputation, message correctness, participation degree, message similarity, message freshness, and vehicle age to computer the trust.
- Thirdly, we examined the detection of trustworthiness in the proposed system, a binary-classification problem is examined, and machine-learning methods are frequently used to address similar classification issues. The majority voting rule is

used in ensemble learning using feature selection based random forest, which selects the decision tree class with the greatest votes as the classifying result.

- Fourthly, we conduct simulation for trust computation and classification. An ensemble learning with feature selection for random forest algorithm has been used to classify trustworthiness.

1.5 Thesis Structure

The research work presented in this thesis focusses on the privacy preservation, trust and reputation computation, and classification of the trust for VANET. The thesis is organized into seven different chapters as highlighted below.

Chapter 2 presents the background, preliminaries and literature review about bout intelligent vehicles, vehicular ad-hoc network (VANET) and blockchain technology, privacy preservation, reputation and trust. The chapter focuses on the VANET (Vehicular Ad Hoc Network) that plays an important role in saving drivers' lives and possessions by disseminating critical event information with the progression of vehicular technology. The chapter presents various types of communication such as Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication and Vehicle-to-Everything (V2X), where everything represents cyclists, pedestrians, and any other entities that can communicate with the vehicle. The second focus in this chapter is on the privacy preservation and blockchain that provides distributed public digital event database to incorporates each event that has transpired and been shared between participating nodes. The third focus in this chapter is on trust and reputation. The fundamental goal of the trust and reputation models in VANET is to assure secure and trusted data dissemination by detecting dishonest vehicles and removing compromised messages from the network.

Chapter 3 introduces the proposed framework using blockchain and biometrics features. The privacy preservation includes biometrics features of the driver and storing the information in the blockchain network. The chapter proposed a biometrics blockchain (BBC) framework to secure data sharing among vehicles in VANET and to retain statuary data in a conventional and trusted system. The proposed framework takes the advantage of biometric information to keep a record of the genuine identity of the message sender, thus preserving privacy. This chapter discusses the establishment method for the security and trust between vehicles in VANET alongside the capacity to trace identities whenever required.

Chapter 4 proposes the formal methods of trust computation. This chapter provides a paradigm for trust computation in the VANET environment. The trust computation is based on knowledge gathering, processing and the generation of quantifiable value. The chapter proposed a trust metric for spatial knowledge, temporal experience, and behavior pattern based on a large number of trust attributes to represent the characteristics and experiences of the vehicle during the interaction in the network.

Chapter 5 proposes a new ensemble learning mechanism for trust computation and classification. This chapter provides a detailed discussion on the dataset, features a paradigm for trust computation in the VANET environment. The collecting, processing, and creation of measurable value are the foundation of the trust computation. This chapter suggested a set of trust metrics based on a large number of trust attributes to reflect the properties and experiences of the vehicle during network interaction which are spatial awareness, temporal experience, and behavior manner. The ensemble learning with feature selection in random forest is applied to classify the trust data into trustworthy or non-trust worthy.

Chapter 6 presents the results analysis obtained from the simulation. The result is broadly divided into two parts viz. privacy preservation and trust computation. The chapter discusses the required experimental setup along with simulation parameters. The result discussed in this chapter are packet delivery rate (PDR), security analysis, computational cost, trust computations and classification, area under curve (AUC), receiver operating characteristic (ROC), efficiency comparison, feature score and confusion matrix of different models.

Chapter 7, the last chapter, presents the conclusion of this research work by reviewing the contributions of the proposed methods and discussing the practical implications for the trust in VANET. Future research directions are focused on the implementation of the proposed framework in the hardware infrastructure like on-board unit (OBU), road side unit (RSU) and servers.

Chapter 2

Background and Related Work

In this chapter, we mainly introduced the background information about the VANETs and blockchain. In addition to, types of communication, VANET security challenges and security requirements. Related work shows the suggested solutions and recent improvements in the area.

2.1 Architecture of VANET

The implementation of Mobile Ad-hoc Network (MANET) strategies for wireless communication between vehicles in a vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication pattern is known as Vehicular Ad-hoc Networks (VANETs). The applicability of VANETs can be seen in Intelligent Transportation Systems (ITS). Due to the harsh settings in which they function, nodes in VANETs are characterized by high mobility, diversified driver activity, highly dynamic architecture, multi-hop communications, and strong security and privacy requirements. Through a range of data detectors in the form of sensor devices, the ITS framework supports the intelligent gathering, processing, and sharing of different forms of context-aware information by vehicles with one another [20]. Vehicles' capability to collect and transmit data in this way can be used for a variety of applications, including cases of emergencies warning, adaptive cruise control, traffic monitoring services, automatic tolling, and better road safety.

2.2.1 Intelligent Vehicles (IV)

Various components such as the On-board Unit (OBU), Digital Control Unit (ECU), Application Unit (AU), cameras, and a range of instruments such as sensing devices, GPS, RADAR, and safety devices are installed in most intelligent vehicles. Each one of these components are connected by high internal buses such as Controller Area Network (CAN), Local Interconnect Network (LIN), Media Oriented Systems Transport (MOST), and Ethernet [21] as shown in Figure 2.1. For example, a vehicle's distance sensor, recognizes the existence of another vehicle in its sensitivity range zone. This data from the vehicle is communicated with other vehicles via interior data buses and OBU, suggesting that the radius should be increased.

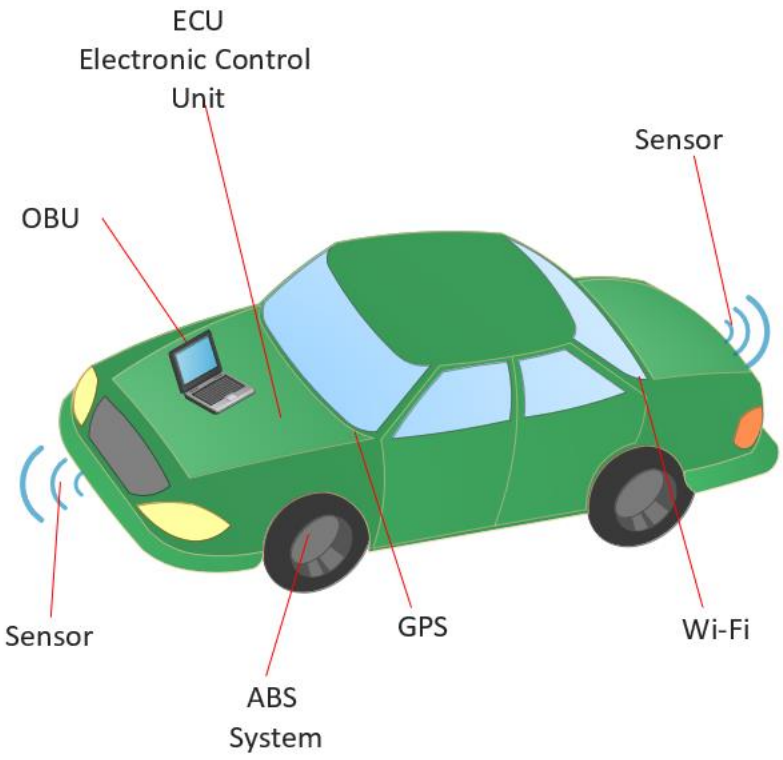


Figure 2.1 Components of Intelligent Vehicle

2.2.2 Roadside Unit

RSUs (Roadside Units) are strategically placed components that serve as a link between vehicles and network infrastructure. RSUs are stationary devices that are stationed at certain places such as traffic signals, lampposts, and mobile towers [22]. The most important components of RSU are the electronics equipment, operating system, and software running on

it. This programme communicates with nodes on one hand and connects with infrastructure on the other. RSU is frequently and advantageously used for large-scale network broadcasting of communications.

2.2.3 Communications

The WAVE/IEEE 802.11p/DSRC standards [23-24] and detailed in[25] establish VANET communication requirements. Control channels for safety applications and service channels for non-safety applications are the two types of wireless channels. Vehicles use the control channel to send Basic Safety Messages (BSMs) at a rate of 10 messages per second. Each vehicle inside the IEEE 802.11p has 300-meter radio range that receives the BSM, analyze it, and broadcast it to its neighbors. The BSM broadcasts among vehicles will aid in the development of collective consciousness, allowing for the prevention of traffic accidents. Vehicles can detect the state of other vehicles in a radius of many miles in a matter of seconds, allowing them to avoid risky conditions. For instance, if a vehicle breaks down in the center of the road following a very long and abrupt bend, incoming motorists will not be able to detect the car until it is too late.

In VANETs, however, because the broken-down vehicle will broadcast BSMs as shown in figure 2.2, all vehicles within 300 meters and beyond will receive the signals and navigate away far before the visible encounter (due to BSM flooding Depending on the objective of the presently running VANET applications, vehicles in VANETs can employ one of three communication modes. The use cases for all three-communication modes are: Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), and Vehicle-to-Pedestrian (V2P) are the three modes of transportation. Vehicle-to-Everything (V2X) refers to all three modes at the same period. In VANETs, vehicles are intended to have Internet connectivity via cellular connections (4G/LTE) and/or vehicle-to-infrastructure connectivity. The addition of Internet connectivity to VANET (V2V, V2I, V2P) communication seeks to enhance the capabilities of vehicles and the functionality of VANET applications (informing authorities in case of an emergency, road map downloads, etc.)

Vehicle-to-Vehicle (V2V) Communication - Vehicles mostly employ V2V communication to disseminate their BSMs. V2V communication is used in a variety of additional VANET applications, both safety and non-safety. Figure 2.2 depict different V2V communication use cases.

Vehicle-to-Infrastructure (V2I) Communication - Infrastructure unites (such as traffic lights, tollgates, roadside devices, and so on) will have communication capabilities in VANETs. RSUs are fixed base stations that are mounted on the sides of road sections to act as wireless LAN access points for cars within their radio range. They are primarily used to increase the overall service of VANETs, especially in the event of a dense vehicular network, because their radio range is significantly longer than an individual vehicle's radio range and all RSUs are interconnected. RSUs will flood the broadcast BSMs of vehicles, allowing them to be heard by vehicles in a considerably broader region. RSUs can also be used to notify authorities in the event of an accident and for Internet access in specific cases

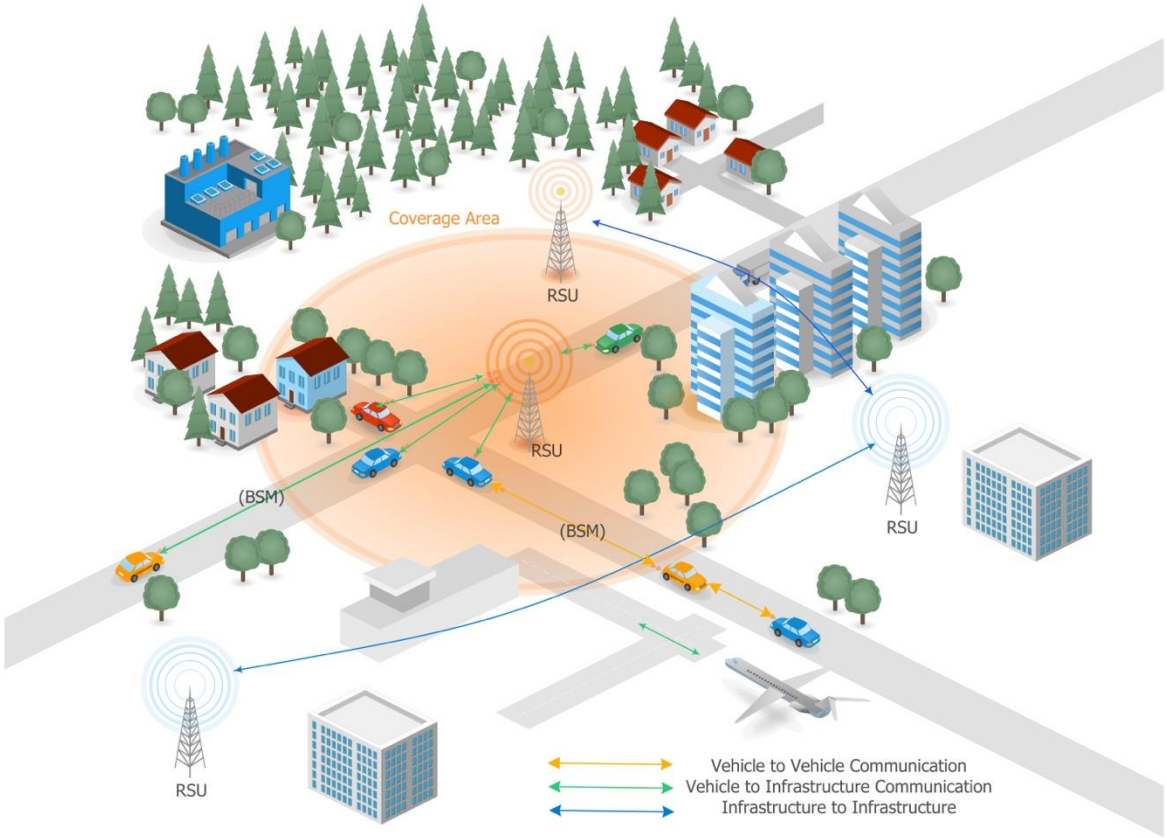


Figure 2.2 Vehicular Communications

2.2 VANET Security Challenges

Despite the various benefits for safety and driving assistance, it also comes with it a slew of security and privacy dangers, because to the vulnerabilities connected with the freely available wireless channel. The bandwidth may readily be exploited to do nefarious actions like as impersonation, eavesdropping, signal jamming, and so on. Vehicular networks are subject to wireless eavesdropping by attackers, and location data can be acquired for tracking objectives. A vehicle position sample is made up of three parts: vehicle's ID, location, and time.

Furthermore, attackers can deduce the true identity of a particular vehicle from its traffic-related basic safety messages as shown in figure 2.2 which will lead to the real vehicle's position [25]. As previously stated, the VANET structure protocol stack is based on the OSI model, and practically all of the levels are included. As a result, security risks affect the entire stack, from the application layer to the physical layer, engaging several entities, posing multiple security issues, and impeding various ways of communication (V2V, V2I, V2X). As a result, while creating a security framework for VANET, it is necessary to evaluate vulnerabilities throughout the whole stack and offer mechanisms for overcoming obstacles at each tier. The subsections that follow go through the different security criteria in depth, as well as malicious user assaults that test these requirements.

2.3 Security Requirements for VANET

a) Confidentiality

Data Confidentiality guarantees that the contents of messages/data are only given to authorized persons and are protected from unwanted and unauthorized exposure during both storage and transit. Confidentiality is a fundamental necessity in the VANET, and it is achieved by implementing particular access control policies and cryptographic techniques on the data that is stored and delivered. This criterion is considerably more critical in the military use of VANET, when information leak is not only a security violation, but also a life-threatening situation.

b) Integrity

Data integrity is verified if data is sent from source to destination without external (unknown and unauthorized) intervention or manipulation, and data correctness and dependability are guaranteed at the destination. If a malicious attacker modifies transmitted

data in the VANET, ostensibly to provide security, it might cause traffic congestion or undesirable route deviations for vehicles.

c) Availability

In order to utilize the network's resources, it is vital for them to be accessible when needed. Given the VANET's strict message delay requirements, if attackers take the majority of the available bandwidth, legitimate users will experience a denial of service (DoS). As a result, how we respond to DOS assaults has a significant impact on network availability for legitimate users.

d) Authentication

Authentication is the process of identifying network users using their unique IDs and passwords/biometrics in order to enable them access to resources. By using the user certificate and signature verification, VANET ensures that not only the message received by the receiver has originated from an authorized user.

e) Conditional Privacy Preservation

To guarantee safety and security, important and confidential user information is hidden from undesired and unauthorized actors. To avoid location monitoring or impersonation in the VANET, users must keep their identities hidden or use regularly changing pseudonyms IDs. In VANET, total privacy is hard to achieve since the user's identity must be exposed and tracked in the event of an emergency, such as an accident investigation that requires the user's location and private information. Even in the case of Pseudo IDs, cryptographic mechanisms and ID generators that are safe yet simple are required to facilitate tracing.

f) Non-Repudiation

When the recipient recognizes the sender, non-repudiation assures that the sender accepts full responsibility and cannot dispute sending the information. In unusual instances, digital signatures contained in the communication might help to avoid any disputes.

g) Trust

It is rapidly becoming the most crucial prerequisite for a successful VANET security system. Even though we can check and validate received transmissions, we cannot trust them due to the large number of entities involved and the disparity in their histories.

2.4 Blockchain Technology

A blockchain is simply a virtual record of transactions that is replicated and disseminated throughout the blockchain's complete network of computer systems. Every block on the chain includes a set of transactions, and whenever a new transaction takes place on the blockchain, a reference of this certain transaction is updated to the ledger of every participant. Blockchain is a sort of distributed ledger technology in which transactions are stored with an unbreakable cryptographic signature known as a hash. The blockchain technology is based on public distributed event ledger which incorporates each event that has transpired and been shared between participating nodes. It comprises a verifiable, definite record of every single incident that has ever taken place [26]. The majority of network nodes converges to validate each event in the blockchain database. There are two main blockchain types: public and private. The public blockchain is open access, meaning any entity can join and interact with it without needing an approval from a third party. The private blockchain is typified by controlled access [27]. The blockchain technologies can be applied to any kind of network viz. wire or wireless. Since, blockchain is applied in the VANET, so obviously it uses wireless infrastructure. Administrators can control who can view, join, and write in the blockchain. They can create consensus groups, due to which the private blockchain can become centralized. While public blockchain does not have this weakness, being completely decentralized and capable of withstanding malicious attacks [28]. Once a full node of a public blockchain is connected to other nodes in the network, the process of constructing a full blockchain will begin. The blockchain has some of the following basic features:

Distributed and trustless environment - It is possible to add any node to the blockchain to validate and synchronize the blockchain content in a distributed way without the requirement for a central control. This adds security and avoids any single point of failure. This helps to create trust in an otherwise trustless system.

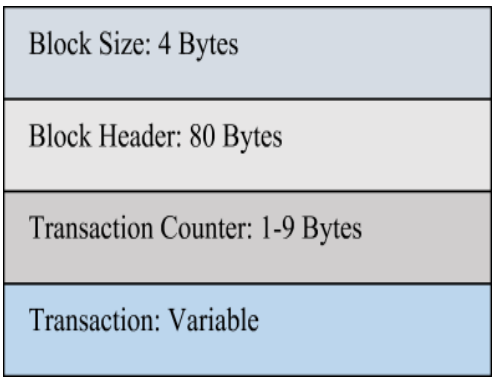


Figure 2.3. Block header Structure

Immutability - Once recorded in the blockchain then no piece of information can be modified or deleted from the network. What is more, adding information arbitrarily is not possible

Privacy and anonymity - The blockchain help users to benefit from privacy and allow users to join anonymously that means that users cannot access each other's information. This helps to ensure that the system is secure, anonymous and private [29].

Structure of Blockchain

The block has a header, metadata and a collection of transactions as shown in figure 2.3 The block header size is 80 bytes, while the transaction size is variable and varies based on the nature of application [16].

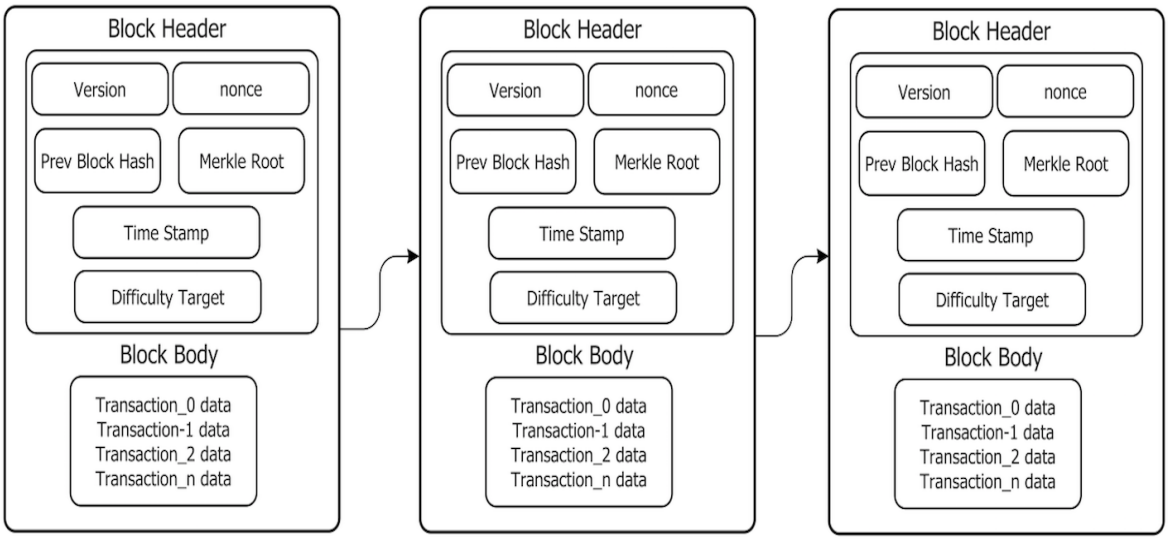


Figure 2.4 Blockchain structure

Each successive block extends the blockchain, resulting in a comprehensive log of transaction records. The network may authenticate blocks via cryptographic methods. As shown in figure 2.4 each block includes a timestamp, the hash value of the preceding block ("parent," and a nonce, that is a random integer used to validate the hash. This approach protects the integrity of the blockchain all the way back to the first block (the "genesis block"). Because modifications to a block in the chain directly affect the respective hash value, hash values are distinctive and misconduct may be successfully prevented. The block can be added to the chain if the majority of nodes in the network accept via a consensus process on the legitimacy of operations in a block and the legitimacy of transactions in this block. These nodes are then

divided into two categories: non-mining nodes and mining nodes. *Non-mining Nodes* - Since the intent of non-mining nodes is to only receive and broadcast requests for data sharing transactions, they do not need the same number of resources as a mining node. It is worth noting that all nodes maintain a complete and authenticated copy of the blockchain and the sensors in relation to the smart contracts. It is believed that all nodes have a legitimate access to the blockchain network, and in each round of transactions, the vehicle sensors upload data to the blockchain network.

Mining Nodes - These nodes are responsible for validating data-sharing transfers and compiling them into data blocks. These nodes are required to use computer computing capabilities on a regular basis in order to solve cryptographic problems and submit blocks to the blockchain network. Since each vehicle has a legitimate link to the blockchain, the vehicle will encrypt the data gathered with a private key before forwarding it to the blockchain with a signature as a request for storage.

Transaction and Consensus Mechanism - A blockchain transactions are created by acquiring a data packet from the vehicles. The gateway carries out a series of transactions that generates vast numbers of data, including data, control and results transactions. Consequently, the data are located on the blockchain by a reference pointer. The contract name, type of transaction, data relation, the sender/receiver address, block number, signature, public and private keys are typically used in the transaction format. For a quick and stable consensus algorithm, selecting safe or effective blockchain nodes is a critical factor. The safe nodes that act as a miner are chosen on the basis of several factors, including computer power, storage capacity, prestige, mining costs, production and bandwidth. We have chosen the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) to classify the credibility of the blockchain mining node [30]. Nowadays, proof-of-work (PoW) has established itself as an important consensus framework for validating transactions through mathematical challenges. The mining node gathers all transactions before a block is created in the Merkle tree and iteratively hashes the data it collects. The process of hashing terminates when the hash of transactions becomes equal to or less than a pre-determined target value (T_h) called as a threshold as expressed in equation (2.1) [31- 32]. H represents SHA-512 hash function and b_c presents the current block.

$$H(n||H(b_c)) \leq T_h \tag{2.1}$$

The probability to discover nonce of proof of H can be expressed as the equation (2.2) [33-34].

$$P(H \leq T_h) = \frac{T_h}{2^{512}} \quad (2.2)$$

After successfully computing the target hash, the miner sends the proof to each node in the blockchain network, along with data transactions and other data, in order for other miners to re-compute and thereby connect/add the new block to the network.

2.5 Related Work in Blockchain-based VANET

VANETs creates an open access environment that poses significant challenges in terms of security and privacy, rendering it unsuitable for real-world implementation. Under the Identity-based Batch Verification (IBV) scheme for V2V and V2I communication in VANETs proposed by Zhang et al.[35], a secure device is used to protect privacy. What is more, pseudo identities are generated locally as the system's master key is stored by every device. On the other hand, the unpredictable risk of system exposure to powerful attackers ensues from this type of storage of the system's master key. What is more, the resultant communication over-head and scalability issues were not considered in this scheme. Many researchers have focused on the raising issues in VANETs security such as privacy, anonymity and traceability.

Authors in [36] proposed a hybrid approach which fortifies the scheme by make use of pseudonymization with self- attestation. It's not required to govern these unless system stability compromise is not an issue. An authentication protocol without certificates was proposed by Tan et al. [37], which involves vehicle identity authentication among vehicles and Road Side Units. In order to enhance users' VANETs key security, [38] created a mechanism for secure authentication. However, a trusted third party is responsible for the execution of these two systems in the centralized system. Distributed security is not provided to identify malevolent users, Wu et al. [39] employed group signatures and one-time authentication. Unfortunately, bilinear pairing operations are costly at the tracing stage, which renders acing suspicious messages inefficient. To achieve pseudonyms control, [40] designed secure RSUs management. Regrettably, it is not possible to work efficiently in fields with RSUs which are presumed to be constrained and have low computation power in conventional VANET s. Decentralized, secure, blockchain-based, independent, and intelligent transportation systems were proposed by Yuan and Wang [41]. Rowan et al. [42] put forward an inter-vehicle session key establishment protocol and blockchain-based solution to secure communications. A privacy approach called (PPAS) was presented by Chuang et al. [43] for the purpose of achieving communication between infrastructure and vehicle in V ANET s, which fulfilled authentication between the vehicle and the RSU as well as the majority of security conditions.

However, this scheme is used to communicate with vehicles and does not provide a distributed system. An anonymous on-board network authentication protocol was proposed by Peng [44]. While this protocol ensures efficient user authentication and anonymity, the fake vehicles will not be identified. Authors in [45] suggested vehicular networks were suitable for blockchain-based anonymous authentication, and Lu et al. [14] designed an anonymous VANET reputation system based on the blockchain.

Although [45] and [14] preserve user privacy in the authentication process, they are not compatible with Bitcoins and their schemes do not involve a vehicle announcement method. A different option to bypass the restrictions of saving many anonymous certificates in advance was proposed by Lu et al. [46], which preserves conditional privacy at the same time.

In all the previous works authors have proposed pseudo numbers to protect real identities of the users which is vulnerable to guess using prior-knowledge. In our approach we have used biometric data of the user to generate a unique pseudo identity. The proposed methods utilize modified discrete cosine transformation and hash function to achieve the privacy. During the registration phase when vehicle information along with driver's detail will be sent to TA. This information will contain the fingerprint of the driver to ensure the identity. The on-board unit (OBU) will have finger print scanner. The authentication of the driver's identity will be done using modified discrete cosine transformation (MDCT). During the registration, the driver will put his fingerprint which is further processed by MDCT to generate a cancellable fingerprint template. In order to design a trusted, secured, and decentralized an autonomous framework was proposed for intelligent transportation system (ITS). To design an efficient VANET the blockchain shall be integrated in self-managed manner [47]. The blockchain technology has ability to combine with multiple applications at a system level that can enables the smart contract system within VANET. The blockchain can be easily integrated with many useful applications such as vehicle insurance, traffic regulation, vehicle tax and weather forecasts with privacy and trust. This blockchain based secure multiple channels of communication between vehicles improves the data sharing security. A blockchain-based mechanism that would protect the user's private data in the course of providing and updating vehicle technology such as remote wireless software was proposed [48]. In another work authors proposed a trust management system for intelligent vehicle which is based on blockchain technology. The vehicles can verify the messages from other vehicles using Bayesian inference models. In this method vehicles generate rating for other vehicles. The offset value of the specific vehicle's trust is determined by the RSU. The data is aggregated into blocks to enhance the traffic

efficiency and the safety is characterized by reliability factor. In another work of electric vehicle (EV), a decentralized blockchain smart grid system was proposed to reduce the overall charging cost of EV users and the grid network's power fluctuation level. A specific blockchain-based EV energy storage program for EV battery charge rate, capacity, grid dynamics, and EV user behaviour were all put forward for consideration [49]. In another work authors have proposed a technique in vehicle- to-grid (V2G) networks to enable data sharing without compromising user data safety using blockchain. This mechanism supports anonymity and audits that includes data registration and maintenance procedures based on blockchain technology [50]. In another related work authors have examined the management safety in connection with charging piles and EVs. Authors have proposed a safety model based on smart contracts and lightning networks to enhance the transaction security of charging piles and EVs [51]. The blockchain's role in enhancing IoT security was evaluated and found efficient [52]. The blockchain authentication protocol was proposed that enables cognitive radio network spectrum sharing that serves as a means to access wireless bandwidth in a competing CR to be able to use as media access control (MAC) protocol [53]. Specoins, the virtual currency they proposed, will be used as payment for the access to spectrum[53].

2.6 Trust and Reputation in VANET

In latest years, reputation and trust-based systems have become increasingly vital in ensuring wireless communications security. The reputation of an actor may be described as the information gathered from other nodes about that entity that aids in making a trust judgment. Generally, the need for trust occurs only in environments characterized by ambiguity, such as e-commerce. Due to their incapacity to acquire information about vehicles outside their sensing range, this trait is popular among network nodes in wireless communications. The employment of reputation and trust-based systems as a security solution for wireless networks is motivated by a variety of factors, including node uniqueness in MANETs, the lowest cost of creating nodes in WSNs, and cryptography failure in the face of internal assaults [54].

2.6.1 Trust in VANET

Researchers in VANET have employed trust as a strategy to improve security, with various meanings. It is defined as a trustor's conviction in the dependability of a target node with the goal of achieving a trust objective under specific circumstances [55]. In other words, trust is an

evaluator's assessment of a character based on previous experiences with the target entity and/or the views of trustworthy nodes [56]. Trust can be defined as a trustor's confidence in a trustee based on previous interactions between the two and suggestions from the trustor's neighbours about the trustee.

2.6.2 Importance of Trust

VANET is a large-scale network that primarily includes the transmission of safety warnings. The detection and revocation of illegal vehicles, as well as their data, is critical in VANET in order to ensure a safe environment. When a valid vehicle receives a safety notification, it should be checked for credibility and authenticity before accepting and transmitting it to other nodes. To accomplish so, a trust model is necessary that can assess the trustworthiness of data received in order to improve network efficiency by maintaining a safe environment for the propagation of trustworthy messages.

2.6.3 Types of Trust

a) Direct Trust

First-hand trust is based on a trustor's direct views of a specific node and the communication between the two [57]. Some academics describe knowledge as "direct information received by the trustor in order to evaluate the trustee using specified metrics based on the participating nodes and services" [58]. Direct trust is seen to be more important than indirect trust; yet, both are considered when evaluating a vehicle [59]. Figure 2.5 describe direct and indirect trust among vehicles.

b) Indirect Trust

Indirect trust expresses a trustor's neighbours'/trusted actors' views on the target node (trustee), taking into account previous interactions with the requesting node. Some scholars explain indirect observation by combining reputation and experience. Reputation is a correlation between the trustor and the trustee based on the trustor's believe in the trustee's ability to carry out a task, whereas experience is a correlation between the trustor and the trustee based on the trustor's belief in the trustee's ability to perform a specific task [55].

In ad hoc networks, trust is formed directly with one-hop neighbors and indirectly with other nodes, whereas the latter is dependent on recommendations from other trusted parties. However, unlike direct trust, indirect trust requires an initial authentication process, such as

certification authorities authorizing all the communicating nodes. Wireless networks provide for both direct and indirect trust mechanisms, although wired networks often use indirect trust. Direct trust is the trust score based on the direct interactions, both current and historical, between two vehicles; indirect trust is the highest direct trust value granted to the trustee by all of its neighbouring vehicles; and the trust computations aggregate weighted direct and indirect trust.

Practically, it is useful to combine direct and indirect observations of the target vehicle. However, it is agreed that direct trust is more important than indirect trust, yet both factors are considered while evaluating a vehicle. The proposed work uses both direct and indirect trust obtained from vehicles.

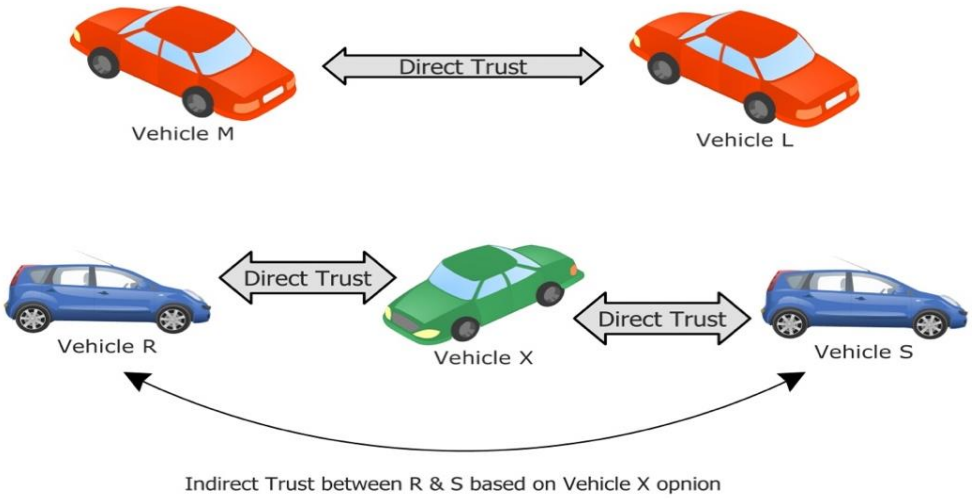


Figure 2.5 Direct and Indirect Trust between Vehicles

2.7 Categories of Trust in VANET

Vehicle trust evaluation models are divided into three main categories: data-centric, entity-centric, and hybrid trust management models:

2.7.1 Data-Centric Trust Management Models

In these trust models, data is essential, with the node computing trustworthy in the reliability and validity of received messages. These trust models are heavily reliant on prior encounters with peers, as well as feedback from nearby vehicles. Several trust models for data-oriented trust computation have been suggested in previous studies. The determination of trust in DO is based on the reliability of the received communications. [60] proposes a paradigm for data trust generation that based on specific position and timeframe. The authors' method relies on the

assessment generated by the node, which collects input from nearby cars and assigns weights to each piece of information based on two factors: position and time. Since trust is calculated throughout the entire path and data is gathered at a node, the suggested system is not suitable to dynamic and sparse situations. The author used many decision logics in his technique, including weighted voting, Bayesian inference, and Dempster–Shafer theory. In the proposed framework author, concludes that Bayesian inference outperformed Dempster–Shafer when numerous events were considered. The suggested scheme's flaw is that it is only acceptable when there is sufficient proof in favour or under a certain scheme for a specific occurrence [61]. Researchers in [62], suggested trust framework, determine the message's credibility based on numerous characteristics including semantic similarity, content conflict, and pattern familiarity. The authors concluded in their suggested trust design meet the criteria of the complex structure of VANETs. The author's study has one flaw: the model includes real-time verification of incoming packets, which is impossible under various speed and sparse situations. In their study, Shaikh and Alzahrani [63], established a trust framework relying on time frame and false position threats. The trust mechanism is decentralized and ideal for real-time use in VANETs due to its simple and linear temporal complexity. In addition, the suggested approach for the trust model identifies the fake location, time, and robustness. The computation of the message's credibility is dependent on prior knowledge of node holdings. Moreover, the event's trustworthiness determines whether the value is accepted or rejected. Marmol and Perez [64] developed a model of trust known as TRIP. In the authors' proposal, the calculation of node trust is concentrated on three parameters. Initially, straightforward interactions utilizing prior engagement with the vehicle. Then, active connection with neighbouring vehicles and any suggestions offered. Lastly, connection among both the RSU and the centralized administration, with the centralized administration sending suggestions. The reputation credits is calculated by converting all three values from criteria 1, 2, and 3 to fuzzy groups consisting of trusted if the outcome is one, non-trusted if the outcome is 2, and either trusted or non-trusted if the outcome is three. When any value is set to number two as not trusted in one of three situations of trust to accept or reject, the transmission is discarded and infrastructure is notified of the untrustworthy vehicle's presence. If the score falls into the "trust" category, the transmission is approved and transmitted to the rest of the other nodes. Finally, if the vehicle reputation value is determined as either trusted or non-trusted is treated as dependable with a probability that may be adjusted and is not transmitted to network participants. We conclude that the proposed assumption is implausible. In order to establish a historical background and credibility of the delivered message of the intelligent vehicle in this

situation, it is also critical to identify the actual identities of vehicles. The Data Intensive Reputation Management approach was developed by Patwardhan et al. [65]. Integrating reputation and agreement, the system ensures the integrity of data and enhances proactive collaboration. In addition, their approach incorporates many elements, including as the frequency of contacts, enduring identification, and a defined range of credible resources, for establishing trustworthiness connections between unidentified existent nodes. Data credibility is contingent upon consensus agreement between competitors or receipt from reliable sources. Furthermore, the authors assumed that every node should have a distinct, persistent identity, which breaches identity privacy. Researchers in [66], introduced a new method for assessing as well as propagating trust models. The authors modelled the trustworthiness of information communicated between nodes using historical record trust, trust recommendation, and task trust concepts in their trust framework. The proposed paradigm focuses on a binary deployment in which information is simply trusted or not trusted. The binary qualifier restricts this scenario depending on limited knowledge or, especially, circumstances that are ambiguous. Moreover, the most essential elements, such as privacy and resilience, are not adequately discussed.

Researchers in [67], suggested a Traffic Safety Event trust modelling paradigm. In their solution, the event-based Reputation System (ERS) is employed to prevent vehicles from broadcasting malicious, hacked, and unreliable warning signals. In addition, the technique employs collaborative assessment and repute adaption systems with two types of value level, activity trust and action reputation, to concurrently compute action rigor and dependability. The time required to convey reliable information with peers is the fundamental shortcoming of the proposed paradigm. Authors in [68], present a trust paradigm, LSOT for VANETs relying on two sets of assessment techniques: ratification-based trust and reference-based trust. Researchers, examine the traffic conditions and unpredictable dispersion patterns of VANETs in their study. In addition, the LSOT paradigm functions in a completely dispersed manner. Three weight parameters, namely number, time decay, and context, were used to the trust computation in order to properly evaluate total trust. The primary flaw of this paradigm is that the authors failed to distinguish between the message and the node's trust.

2.7.2 Entity-Centric Trust Management Models

Entity-oriented (EO) tries to eliminate dishonest nodes by evaluating the node's trustworthiness. The EO assesses the node's trustworthiness and detects the existence of a malicious vehicle on the network. Entity-oriented TMs emphasize the assessment of the message broadcasters' credibility. The message originator, as well as the trustworthiness of the trust evaluator node (EV), is crucial to the success of these trust models. The reputation of the original message sender is supported by the reputation of EV's neighbours. These trust models function well in circumstances with limited mobility and high population density, since an increasing number of neighbours may send information about a particular occurrence, allowing EV to evaluate trust based on this knowledge. However, such TMs are incapable of evaluating the reliability of data content, which is one of the primary purposes of VANET. In addition, highly mobile vehicles do not collect adequate data for trust evaluation and computation.

Several authors have produced a substantial quantity of literature on data-oriented trustworthiness. For vehicle ad hoc networks, researchers in [64] suggested a trust technique that relies on positive popularity architecture. In their research, the authors evaluate three distinct forms of data when determining the reputation score for each network node. The three estimation factors are straightforwardly communication to any preceding node, opinions and suggestions from adjacent network nodes, and recommendations from the central authority. To decide whether to accept them depending on previous requirements after the establishment of the trust value. when the created trust value indicates to not trust this sender, the communication is discarded and infrastructure is notified when a disingenuous vehicle presence the network. In all other circumstances, if the estimated trust value equals "trust," the transmission is approved. When a trust score is computed as either trusted or not trusted, then message will be acknowledged however, no broadcasting for the message in the network for other vehicles. Moreover, in their model, trust value generation is largely dependent on the node's verification of its credibility. The primary deficiency of the proposed trust model is that numerous transmitters would transmit the sender's reputation, which will cause overhead. Researchers in [69], developed a cluster-based trust paradigm for DMN in VANETs. The group head is accountable of computing and forwarding the trust to a Trusted Authority (TA). In addition, it is the responsibility of the TA to eliminate a misbehaving node from a system utilizing data obtained from the group head. Key disadvantage of such suggested method would be that it generates an excessive amount of data owing to continual reporting, which affects network performance. In addition, network connection data between groups head, TA, and cars is

absent. Author in [70], proposed a prototype system when two cars collide, each vehicle creates a profile of the other vehicle. The suggested TM is predicted on the principles of trust labelling and trustworthiness, as well as sociological trust. The estimate of dependability is determined by the interplay of node outline records. EO paradigm technique has many flaws. Firstly, because VANET is very complex, and vehicle engagement is restricted in time; this makes it difficult to collect sufficient information to assess trust. Second, if the node is reliable, the information it delivers is either correct or inaccurate. Finally, the researcher provides a method for labelling trustworthiness and expressing trust using probabilities, as well as a trust framework for vehicular purposes for trustworthiness and services. The suggested trust model by Gerlach lacks formalization of the architecture, which is a drawback. In addition, their findings did not examine a mix of the various levels of trust. In their research, Minhas et al. [71] presented task and expertise trust model were provided for assessment criteria, proposing task and expertise trust model-based trustworthiness for assessment measures. This approach also enables a moving node to thoroughly evaluate an occurrence by initiating queries to neighbouring actors, however it restricts the number of responses obtained. Researchers' comprehensive trust management technique combines task and expertise criteria, which are included into the priority-based model and utilized to choose qualified advisers. The advisers collect comments utilizing the majority-opinion technique. In addition, depending on the aggregation of input obtained from advisers, time and proximate location were also evaluated. The researchers presume that authorities have set duties and are expected to behave in a particular manner. The work's disadvantage is resilience needs to be adequately examined. Researcher in [72], developed a Reputation Management-based trust paradigm for VANETs. This study employed a pattern analysis approach in order to evaluate node's reliability. In addition, the reputations of advisors are utilized as weights for computing the message generator's total reputation. The primary shortcoming of the author's method is the contrasted produced messages similarity relying on the Euclidean distance algorithm which separates two nodes. The study in [73], presented a trust framework approach that relies on CH's reputability and voting. This study uses the facts strategy to disseminate authentic information in order to enhance their reputation. Additionally, voting is performed across the network nodes in order determine who will be the group head.

In addition, rewards are provided to nodes in the shape of weights during election. When the weight is high, it indicates more confidence in the node. Although this idea is intriguing, it is ineffective in a rural, highly mobile environment where only a small number of vehicles engage

in the elections. Haddadou et al. [74] presented a trust model based on economic incentives. The authors adopted a novel strategy when score ratings is allocated in a dispersed way. The score ratings can be either raised or lowered depending on network participant behaviour. Additionally, the score ratings diminish with every network attack happens. Zhang et al. [75] suggest a trust plan using the Chinese remainder theorem (CRT). This study focusses on protecting node confidentiality and providing authenticity. This approach relies on tamper-resistant components (TPD) identification, road side unites, and trusted authority. The suggested technique's weakness is centralization, dependent on road side unites and trusted authorities, and inapplicable in remote regions where there is no VANET services.

Researchers in [76], presented a fuzzy logic-based trust strategy for evaluating direct node trust. The author included honesty, cooperation, and accountability into their fuzzy logic-based strategy. The primary drawback of this strategy is its restricted connectivity range, as the method is not distributed.

2.7.3 Hybrid Trust Management Models

Dual trust architecture incorporated both object and data trust design characteristics. Dual trust designs assess trustworthiness utilizing the reliability of nodes and information transmitted. To put it another way, these TMs assess the reliability of data based on the reliability of vehicles, considering a trade-off between data authenticity and sender's reputation. Consequently, a vehicle's reputation and the observations of its neighbours consider a crucial factor in assessing trust. Due to of significant management transmissions must be handled in a relatively short amount of time, these rust frameworks are complex. In recent years, a number of researches published regarding trust built on dual trust frameworks. These designs assessed the credibility of vehicles and calculated the dependability and credibility of data using modelling outcomes. Sedjelmaci and Senouci [77] established a concept of trust focussed on the precision of VANET. Researchers asserts that trust designs cover the fundamental properties of VANET, such as the mobility of nodes and the rate of topology change. Using a monitoring mechanism, the authors assert that the suggested lightweight model will defend against the most hazardous threats, including black-hole, wormhole, and Sybil attacks. In addition, the suggested method is separated as two intruder monitoring system levels. The framework's first component is based on collaborative detection, while the second section deals with an RSU-processed global detection system. The primary disadvantage of the suggested system is that electing the cluster head would be a time-consuming and network delaying task. By conveying safety and security-

related messages, Dhurandher et al. [78] suggested a system based on Reputation and Plausibility Checks. The researchers build their study on vehicular protection throw credibility and reliability assessment method, that employs three concepts inside its methodology: action modifying transmission, information aggregation, and fake incident production. Detection range of the suggested method, which just 50 meters, is the primary downside. In addition, detection relies on the cars' integrated sensors. Trust Model with Delayed Verification for Message Relay was presented by researchers in [79] as a trustworthiness strategy for vehicular networks. The scientists categorized data flow into four separate classifications depending on the relative importance of safety-related communications, ranging from high to low. (1) background, (2) best-effort, (3) multimedia, and (4) audio transmission. The primary shortcoming of the suggested trust design considers their assumption where a malicious node will act consistently during the whole travel through the vehicular network; this approach is flawed in the VANET.

Authors in [80], developed a trustworthiness strategy utilizing a decentralized reputability mechanism that piggybacks on suggestions. Each node resends data contributes with new assessment of the reliability about the message. The trustworthiness algorithm is guided by a number of credibility parameters, including straightforward and passive trustworthiness, originator credibility, and location orientation. The primary flaws of the system proposed by the researchers consider the failure to offer adequate and comprehensive information regarding the technique. In addition, the article claims that originator data is controlled by the algorithm, but did not specify how TM's reputable information will be updated. The approach presented in Chen et al. [66] is relying on the data transmission and assessment paradigm. This concept is built on the distributed and collaborative dissemination of trustworthiness messages. In their approach, the authors address two fundamental properties of VANETs: network scalability and system efficacy. In addition, these two features include information assessment relying on the widespread existence of erroneous knowledge and data in vehicular network.

Researchers in [81], presented a dual trustworthiness strategy, specifically mix-mode hybrid confidence control method for vehicle networks. The writer's plan is appropriate in both populated and countryside environments. This strategy dependent on the rating method. The score rating is determined by examining the past record of the sender node and validating the received message. The primary disadvantage of the strategy is the absence of a central authority and VANET infrastructure.

2.8 Related Work in Blockchain based Trust Management

The fundamental goal of the trust and reputation models in VANET is to assure secure and trusted information dissemination by identifying malicious vehicles and eliminating messages which have been tampered with. Yang suggested a VANET architecture based on similarity-based trust and reputation in [72], which needed messages to be verified after receipt. A similarity mining technique was used to calculate the similarity between related non-linear data. Within a reputation evaluation model, agents' recommendations were combined with outcomes. When the message content is checked, the trust and reputation values are changed immediately. Another algorithm for trust management is BARS, a blockchain-based anonymous reputation system published in [15]. This approach employs two types of blockchain authentication: presence-based and absence-based. The public keys are given pseudonyms to ensure vehicle anonymity. The reputation of a vehicle is determined by the broadcast messages that are kept in a single blockchain. According to the study's findings, BARS can improve the reliability of broadcast messages while also protecting driver privacy. To secure VANET communications, the authors of [68] recommend using a Lightweight Self-Organized Trust (LSOT) framework. Self-organizing nodes collect trust certificates and recommendations in this paradigm. Li and Song [82] proposed an Attack-Resistant Trust (ART) model to assess the data and vehicle trustworthiness in VANETs. Data trust is used to verify data, whereas node trust is used to determine the trustworthiness of nodes in VANETs. Experiments were carried out to test the ART model's ability to protect against malicious attacks, with the findings demonstrating that it is capable of doing so. Three trust indicators are offered in [83] based on the factors of reputation, experience, and expertise. The vehicle's reputation indicates how well it has been able to exchange data with all pertinent units up to this point. The trustor's ability to communicate with the trustee so far is measured by the experience component. The trustee's knowledge establishes direct trust (the vehicle providing data). Primiero et al. [84] proposed a proof-theoretic model for VANET trust and reputation that included an expansion of the secured natural deduction method [85]. To verify the function, a consistency check was performed at each interaction of the vehicles that were subjected to the algorithm. As a result, this reputation system evaluated the parameterized feedback messages from the standpoint of a temporal measure and ranked the relevant service attributes. Traditional VANET architectures based on centralised systems are incapable of dealing with the increasing complexity of ITS systems. Javaid et al. [86] presented a blockchain-based approach for VANET data exchange and trust management. Each vehicle is

given its unique cryptographic fingerprint by the DrivMan system, which is then used to validate data source. The use of certificates issued by infrastructure units protects vehicle privacy. DrivMan can be used to offer data integrity and originality for the secure and dependable functioning of intelligent vehicles in VANETs.

The Internet of Vehicles (IoV) is expanding at a breakneck speed, posing substantial issues in terms of storing large volumes of data, intelligent administration, and safe data management [87]. Lu et al. [15], [14] presented the BARS system, which enables for the trusted management of VANETs via blockchain. According to their concept, the credibility of vehicles is established using a reputation score technique that evaluates previous events. A crypto trust point (cTp) based on the blockchain, according to Singh et al. [88], can be used to achieve shared data security among peer vehicles. Nisha et al. [89] presented a blockchain-based VANET authentication and revocation model. While these system designs to preserve vehicle privacy, they do not consider communications security.

Rakesh et al. [17] proposed a blockchain-based message distribution service for VANETs, which is similar to [88]. Despite the fact that both solutions provide adequate security for vehicular communication, neither solves privacy concerns. The amount of data generated by VANETs was highlighted in a study by Xiaodong et al.[89] , who demonstrated how mobile edge computing (MEC) can decrease the number of resources needed by blockchain-based VANETs. Although MEC decreases blockchains' computing overhead, it does not make them completely decentralised. Furthermore, the authors recommended Trust Bit, an incentive-based vehicle communication protocol, in [90]. They use blockchain with a distinct crypto ID which is assumed to be issued by the vehicle manufacturer to construct a secure IV communication and reward system. The authors of [91] proposed a safe platform for sharing and storing data within VANETs that utilises the consortium blockchain; nevertheless, this adds more overhead. Blockchain can be utilised to offer secure named data networking (NDN) caching within VANETs, as described by Hakima et al. [92]. Lei et al. [93] also investigate the use of blockchain-based dynamic key management in heterogeneous ITS systems. Although the aforesaid methods are reliable and secure, the privacy of the vehicles cannot be protected, putting people at risk.

2.9 Chapter Summary

This chapter offered an overview of the essential ideas of vehicle ad-hoc networks. This covered the basic structure of the VANET as well as its essential parts. in **section 2.1**, and major security challenges in VANET were provided in **section 2.2**. Moreover, Security requirements for VANET were presented in **section 2.3.** A general overview about blockchain technology introduced in **section 2.4.** Related work of Blockchain-based VANET reviewed in **section 2.5.** A brief introduction about trust and reputation and importance of trust VANET provided in **section 2.6.** Vehicular trust management models are generally categorized in three groups, namely data-centric, entity-centric, and hybrid trust management models which presented in **section 2.7** We intensively reviewed the related work in blockchain based trust management in **section 2.8**. Finally, we summarise research gaps and challenges in **section 2.9**. The next chapter will introduce a privacy preservation framework for VANET based on blockchain with the aspects of VANET security.

Chapter 3

A Privacy Preservation Framework for VANET

With the rapid advancement in smart cities, the number of intelligent vehicles in mobile ad-hoc networks has raised significantly. In the next 10 years, the number of intelligent vehicles is expected to reach 2 billion across the globe [94]. Therefore, vehicular ad-hoc network (VANET) has been created which is equipped with wireless communication devices named as on-board unit (OBU). These devices have hardware security chip to store sensitive information of the vehicle. Communications in VANET, can be categorized into vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication. Sharing valuable information about the traffic between the vehicles will be via the dedicated short-range communication radio (DSRC) [95]. Each vehicle represents a node in the network and it has the capability of sending and processing information. Improving road safety, reducing the traffic accidents, and enhancing traffic flow are several aims of forming VANET[95]. However, due to the high mobility and volatility of vehicles in VANET, various attacks could be performed during the exchange of messages between vehicles among the network leading to severe impacts. Due to the decentralized nature of VANET the process of identifying misbehaving vehicle or users has become difficult task [96].

3.1 Privacy Preservation

The privacy and authentication of the data were the biggest concerns of researchers in VANET in order to improve the security. With this intention, blockchain technology has attracted many academicians and researchers for its enormous advantages to be gained in terms of providing and playing major role for managing, controlling, and securing VANET [97]. The main characteristics of blockchain are the decentralized architecture which means that data will be stored in a peer-to-peer network which is applicable for intelligent vehicles. Moreover, security of blockchain as a distributed secure ledger that provides an essential solution to the issues in security and privacy in VANET due its cryptographic protocols [28]. Furthermore, blockchain will provide anonymity for vehicles and therefore it will be difficult to trace and discover the original identity of the vehicle due to the cryptography nature of blockchain. On the one hand, communication in VANET shall provide privacy preservation by employing anonymity of vehicles. The authorities shall be able to track the anonymity to identify the malicious vehicle, thus it should be made conditional. The objectives of this chapter are as follows:

1) Single registration: For ease of use, a VANET authentication system should support a single registration process whereby vehicles are only required to register once before being able to send messages to other road users.
2) Message authentication: To ensure that received messages are credible, the roadside units (RSUs) or vehicles should have the capacity to authenticate messages by verifying the identity of the sender and checking the message's timeliness and integrity
3) Preserving privacy: A vehicle's genuine identity must not be visible to other vehicles or RSUs and it should not be possible for a malicious actor to acquire identities through the analysis of any identity intercept.
4) Traceability: An efficient system should be in place allowing the trusted authority (TA) to trace the true identity of the vehicle if malicious behaviour takes place, e.g., false messages are transmitted to confuse other vehicles.

The contributions of this chapter are as follows:

1) A biometrics blockchain (BBC) framework is proposed to make communication in VANET more secure.
2) The biometrics features are combined with blockchain technology to provide reliable transmission of data, tracking the data exchanged and identification of the vehicle responsible in the case of falsely messages.
3) The performance of the framework is evaluated in terms of packet delivery rate, packet loss rate and computational cost.
4) Due to the requirements of the legacy system, diligence and statute, the vehicle registration data is kept by the Motor Vehicle Department. The data relating to vehicle communications in VANET is stored in blockchain to make it secure.

3.2 Preliminary VANET

VANET has started to play a major part in saving drivers' lives and possessions by broadcasting crucial event information with the progression of vehicular technology. Two main types of communication are associated with VANET: Vehicle-to-Vehicle (V2V) and Vehicle-to- Infrastructure (V2I) communication. Today, a third type of communication which is quite popular as well: Vehicle-to-Everything (V2X), when a vehicle communicate with everything acts cyclists, pedestrians, and any other entities [98]. In V2I, RSU will be located at both sides of the road as shown in figure 3.1 and vehicles driving through will communicate with RSU.

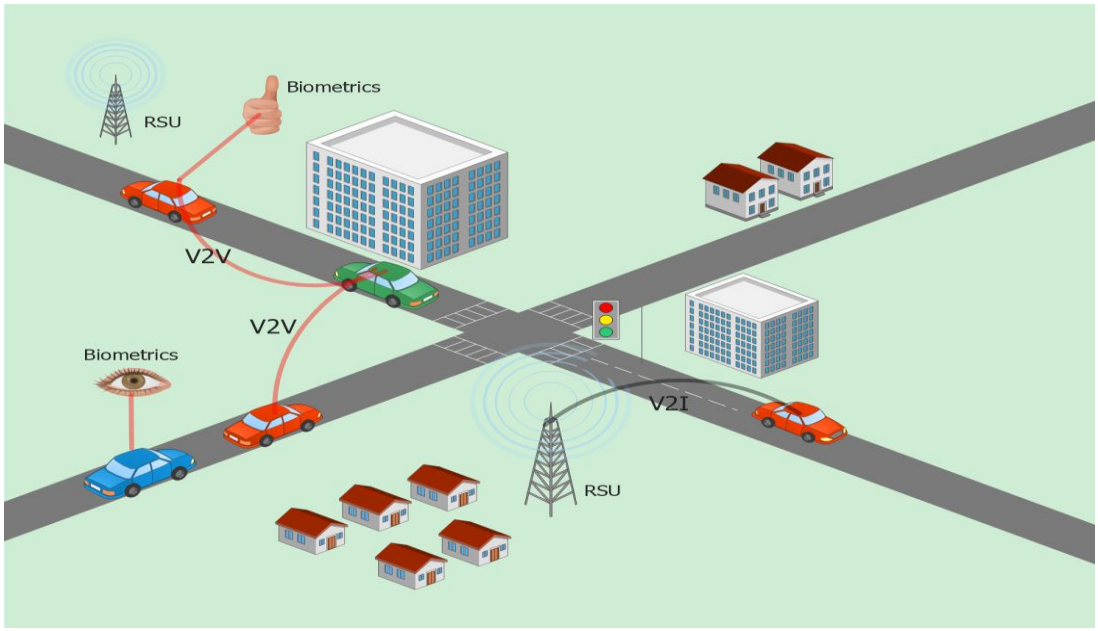


Figure 3.1: Proposed vehicular network

The protocol Wireless Access in Vehicular Environments (WAVE) provides the foundational standard for Dedicated Short-Range Communication (DSRC); this operates within the 5.9 GHz. The WAVE works on the IEEE 802.11p standard [99]. Communication with nearby vehicles is achieved using On-Board Units (OBUs), together forming ad hoc networks allowing for the distribution of communication. One of the main goals of the VANET is to use safety messages to communicate with additional vehicles when reporting events, such as warnings, accident information, weather reports, information on traffic jams, and reports of ice cover among others. One needs to distribute certain event information rapidly, accurately, and with as little delay as possible because failure to do so can cause injuries to drivers and damage to vehicles. Among the key aspects used to guarantee communication security in VANET are node and event message trust therefore it becomes important to evaluate their reliability periodically [100].

3.3 Proposed Biometrics Blockchain Framework

We have proposed a blockchain based framework that uses biometrics in VANETs to protect privacy, with vehicles employing a public-private key pair provided by the TA to communicate with other parties. By employing blockchain techniques, such a decentralized framework will be trustworthy, secure, and allow messages to be disseminated securely. Standard blockchain is associated with cryptocurrency; but this blockchain handles safety event messages with no employment of cryptocurrency. From this point safety event messages will be employed as event messages. This novel BBC design is suitable for protecting the security of safety messages within VANET in real-world scenarios. The blockchain will retain and manage event message history alongside each vehicle's trust level reliably, immutable, and with good distribution. Every country will have one unique blockchain with independent management and maintenance to record vehicle information.

3.3.1 Entities of Framework

In this section we describe the related entities for the proposed framework as follows:

a) Trusted Authority

The trusted authority (TA) is responsible for initializing the system, deploying smart contracts, registering vehicles and revoking registrations. The assumption is made that the TA has significant capacities for computation and communication and will not be working with any other party.

b) Motor Vehicle Department (MVD)

The Motor Vehicle Department has several responsibilities which include vehicles registration, maintaining vehicles records, the MVD authorizes the TA to issue the certificates and public keys to the vehicles after the verification process is completed from MVD.

c) Vehicle

The vehicle undertakes services for the driver and will carry an OBU that cannot be tampered. The assumption is made that the preloaded information carried by the OBU is protected against malicious attack. Additionally, the vehicle will employ the OBU to communicate wirelessly with other entities

d) Road Side Unit (RSU)

The roadside unite is roadside infrastructure which has the capability of communicating wirelessly with all vehicles inside a defined range. It is capable of receiving instant messages from vehicles, verifying them and passing them on either to nearby vehicles or the traffic management centre.

e) Blockchain

The blockchain represents the decentralized foundational architecture of BBC. It is responsible for secure handling of the transactions (safety messages) exchanged by vehicles across the network.

f) Messages

Messages in VANET can be classified into two forms or groups which are beacon and safety event messages. The former is disseminated at set intervals to give information to all vehicles in an area of the position such as driving information to allow all the vehicular nodes in an area to be cooperatively aware in order to manage traffic. Safety event messages are disseminated when critical events are present, e.g., hazards, traffic accidents, etc.

3.3.2 Biometrics based Authentication

During the registration phase the vehicle information along with driver's detail will be sent to TA. This information will contain the finger print of the driver to ensure the identity. The on-board unit (OBU) will have finger print scanner. The authentication of the driver's identity will be done using modified discrete transformation (MDCT). During the registration, the driver will put his fingerprint which is further processed by MDCT to generate a cancellable fingerprint template (Cf_T). The cancellable biometric system will transform the biometric identity of the

driver and store the cancellable reference template in the database in the TA and OBU. The Cf_T is used in registration of vehicle to grant access to the vehicular ad hoc network. The authentication of the driver is achieved via two processes: enrolment and authentication. During enrolment, the vehicle driver's biometrics data Cf_T is registered in the database. At the time of authentication, the driver's biometrics is captured as Cf_T^* which is compared against other reference template in the database at TA. The Euclidean distance (ED) has been deployed to match the captured and reference data based on the similarity. If both the data matches, the driver will be registered on the VANET.

Step 1: From the databases, the driver's fingerprint image is taken and subjected to image pre-processing along with thinning. In fingerprint classification, fingerprint thinning is very vital as the pre-processing of the classification. The method of attaining the image's skeleton and eliminating all redundant pixels for obtaining a new clarified image is called thinning. The skeleton could be as one pixel and it can illustrate the image's topology.

Step 2: The extraction of fingerprint features or minutiae points is done.

Step 3: From minutiae points along with their locations, a binary fingerprint feature vector (FV) is extracted.

Step 4: Utilizing the MDCT approach, the FV is subjected to transformation.

Step 5: By assessing the Euclidian distance between the probe Cf_T and Cf_T^* in the database, matching is performed. The MDCT's process for obtaining Cf_T and matching process is elucidated as follows:

3.1.2.1 MDCT

A cancelable transformation is created for securing the fingerprint feature vector F_V, of a dimension $M \times N$ where $N \leq M$ is utilizing MDCT. The proposed transformation is created upon the DCT matrix. A linear transformation function $f : R \rightarrow \hat{R}$ is the DCT in which a set of real numbers $R \in \{x_0, \ldots x_{K-1}\}$ is mapped into another set of real numbers $\hat{R} = \{X_0, \ldots X_{K-1}\}$ as per the below equation (3.1).

$$X_i = \frac{1}{2}\left(x_0 + (-1)^i x_{K-1}\right) + \sum_{n=1}^{K-2} x_n \cos\left[\frac{\pi}{K-1} ni\right] \quad (3.1)$$

Where $i = 0, \ldots K-1$. DCT is highly invertible. So here MDCT is proposed to make it non-invertible. Assume a column vector as shown in equation (3.2):

$$V_l = \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ . \\ v_p \end{bmatrix} \tag{3.2}$$

The values in the vector V_l belongs to unique positive integers. The column V_l is produced using random distribution which is deemed to be the transformation key. An important part is played by the transformation key in the cancelable biometric system. Thus V_l represents that l number of transformation values can be produced.

The V_l is employed to create sub-matrix R, donated as $\bar{R}$, i.e. i^{th} column of the $\bar{R}$ is the $v_i{}^{\text{th}}$ column of R for $(i = 1,2,\ldots.p)$. Therefore, size of $\bar{R}$ is $p \times M$.

The $\bar{R}$ is created from a random factor v_i, while partial DCT-based has been obtained on F_V using a non-invertible Cf_T by the following equation (3.3):

$$Cf_{T(p\times N)} = \bar{R}_{(p\times M)} F_{V(M\times N)} \tag{3.3}$$

The $\bar{R}$ is created from the DCT matrix, therefore, cancellable transformation in equation (3.3) tends to be a non-invertible because of $\bar{R}$ a column-reduced sub matrix of R. Finally, the M_p matching process is done by using the following equation (3.4):

$$M_p = \begin{cases} 1 & \text{ED}(Cf_T, Cf_T{}^*) < th \\ 0 & \text{Otherwise} \end{cases} \tag{3.4}$$

The distance D, between $Cf_T, Cf_T{}^*$ is computed as follows:

$$\text{ED}(Cf_T, Cf_T{}^*) = \frac{g^Y g}{Cf_T{}^Y Cf_T + Cf_T{}^{*Y} Cf_T{}^*} \tag{3.5}$$

Where $g = Cf_T - Cf_T{}^*$ and Y is transpose of a real vector/matrix.

The symbols used in the proposed model is presented in table 3.1.

Table 3.1: Notations

Notation	Description
V_i	Identity of vehicle *i*
KU_i	Public-key of vehicle *i*
V_p	Vehicle pseudo id
KR_i	Private-key of vehicle *i*
v_{reg}, v_{mod}, v_{chas}, v_{rank}	Vehicle registration number, model chassis number, ranking
O_i	On-board unit of vehicle *i*
D_i	Driver's information
d_{name}, d_{bio}, d_{lic}, d_{rank}	Driver name, biometrics data, license, ranking
M_i	Message sent by vehicle *i*
$\mathcal{H}$	Cryptographic hash function
C_i	Certificate of Vehicle i
M_Type	Message type: beacon, alert
$\mathcal{E}$	Encryption function
$\mathcal{D}$	Decryption function
E_{id}	Event identification number
E_{loc}	Event location
E_{type}	Event type
T	Event timestamp
t	Current time
R_i	Reputation of the sending vehicle *i*
th_f, th_R	Freshness and Reputation threshold

3.4 System Model

As detailed previously, it is crucial for VANET that vehicles should be registered. Verification of a vehicle's physical attributes must be undertaken before they can participate in the network. All vehicles must undergo such protocols to be awarded a valid certificate and so be allowed to join the network. It should be noted that as with every other transaction that forms part of a vehicle's blockchain, the genesis block uses a public and private key supplied by the TA. At the time of verification of vehicle's information and the vehicle's authorized user's biometric data is used by the TA. In this instance, biometric data will serve as continuous identity information.

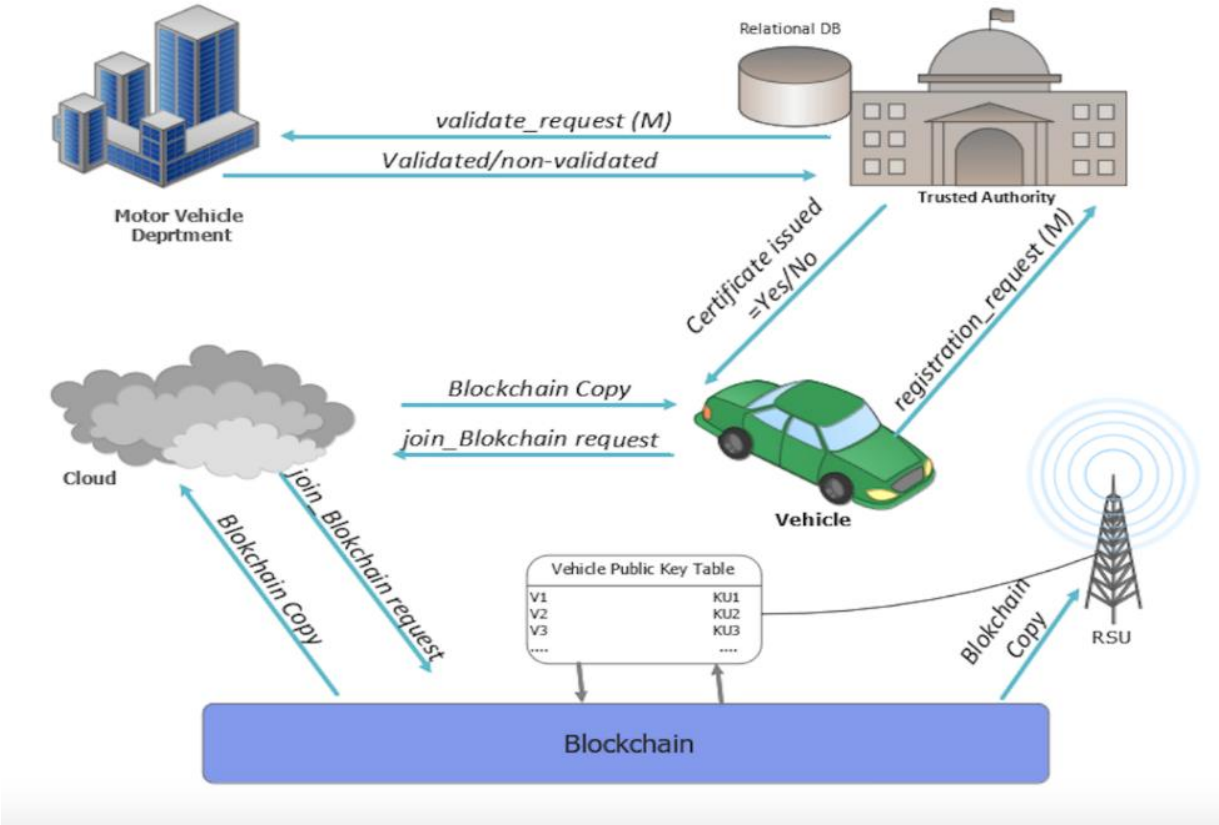


Figure 3.2. Registration Process

3.4.1 Vehicle registration on BBC

When a vehicle joins the system for the first time, it needs to register with a TA. The biometric data of the vehicle's authorized user and the vehicle data is sent to the TA in order to get a pair of keys in return. The registration process begins with obtaining a real identity of the vehicle from the MVD, and sending the biometric data of the vehicle's authorized user and vehicle ID to the TA. Thereafter, the TA verifies the existence of the real vehicle identity from MVD, if verified, the TA will verify the biometric information of authorized vehicle's user, if this too is verified, the TA then generates the certificate that includes a pair of keys according to the vehicle ID and user's biometric data as shown in figure 3.2. The TA saves the biometrics information in highly secure database that will be further used to track the real-identity of vehicle in the case of malicious activity. The complete process of registration is shown as follows:

Vehicle Registration Process
Input: V_i, D_i,
Output: Success/Failure
Begin

fetch: V_i: $v_{reg} \leftarrow O_i, v_{mod} \leftarrow O_i, v_{chas} \leftarrow O_i, v_{rank} \leftarrow O_i\}$

Compute $d_{bio} \leftarrow Cf_{T_{(p\times N)}}$ the IMAGE matching process is done by using

Compute $\mathrm{ED}(Cf_T, Cf_T^*)$ *from equation (3.5)*

if $(\mathrm{ED}(Cf_T, Cf_T^*) < th)$ **then.**

fetch: D_i: $\{d_{name} \leftarrow O_i, d_{lic} \leftarrow O_i, d_{rank} \leftarrow O_i\}$

Join the network

generate KU_i, KR_i

$M_i = Sign_{KR_i}(KU_i, V_i, D_i \| \mathcal{H}(KU_i \| V_i \| D_i \| d_{bio}))$

Certificate=registration_request (M_i)

if *Certificate=TRUE*

Join blockchain $(\boldsymbol{C_i})$

else

error: unauthorize driver

endif

endif

End

3.4.2 Vehicle Joining

Once the vehicle is successfully registered, it can join the blockchain. The vehicle can join the chain by getting update copy of the blockchain. Thus, the vehicles in VANET can download and append to the blockchain.

In the proposed framework, the blockchain performs the function of a distributed ledger that saves the important historical data of vehicles along with safety messages. Any vehicle that experiences a critical event, such as an accident, will broadcast the safety message to neighboring vehicles in the network.

Vehicle Joining Process

Input: C*i*, M_i

Output: Success/Failure

Begin

status=validate_ certificate (Ci)

if *status=valid*

$V_p = \mathcal{H}(d_{bio})$

Update vehicle pseudo id

GetBlockchainCopy ()

while *(Event)*

Broadcast (M_i)

endwhile

else

error: invalid certificate

end if

end

3.4.3 Message Reception

Every vehicle in VANET will continuously receive message from the network. Based on the message type and priority level, appropriate action will be taken by the vehicle. The process of message retrieval is shown as follows:

```
Get Message Process
Begin
    M=GetMessage ()
    if M_Type=safety
        status=Validate_Message ()
        if (status=validated)
                Broadcast(M)
                If node=miner
                    AppendToBlockchain(message)
                endif
        else
                discard(M)
        endif
    endif
End
```

3.4.4 Message Broadcast

All the vehicles in VANET broadcast the information about their positions, status and other details through a message called beacon messages. Whenever a vehicle wants to transmit these messages to any proximate RSUs and vehicles, then message will be encrypted and signed the sending vehicle's private key. The message may consist of details such as the event ID, event type, event timestamp, and event location. receiving the message by the vehicles, the messages are first checked the freshness of the received message using the verification process. After receiving the message by the vehicles, the messages are first checked the freshness of the received message using the verification process. The receiving vehicle compares the timestamp, reputation rate of the sender, verifies the message, and decrypts it using the sender's public key. If the verification process is successful and the message is valid, then the receiving vehicle can rebroadcast the event message to other vehicles. The process for broadcast and verification is shown as follows:

Broadcast Message Process

Begin
    $M=Sign_{KR_i}(\mathcal{E}(E_{id}||\ E_{loc}||\ E_{type}||T))$
    SendToAll(M)
End

Verify Message
Begin
    if $(t - T) > th_f$
        if $R_i > th_R$
            Broadcast $(\mathcal{D}_{KU_i}(M)$
        endif
    endif
End

3.4.5 Blockchain Update

Another important task in the proposed framework is to add the message to the blockchain, which is called as blockchain update with new blocks. The complete process for adding a new message in the blockchain is shown as follows:

Input: Message
Output: Success/Failure
Begin
    Block: structure {H_1, KU_i, Vi, Di, T}
    $Block=UnSign_{KU_i}(H(M))$
    $H_2= \mathcal{H}(Block)$
    If $H_1= H_2$ *then*
        Validate Transaction
        Append Transaction to Block
        Distribute Update of blockchain
        Finish ()
    else
        error: malicious message

        discard (M)
    endif
End

When the message is generated from the vehicle then it should be added to the blockchain in the form of a transaction. However, adding every message to the blockchain will lead to the communication and computational overhead. Therefore, we have proposed adding only safety messages that are validated.

A block in the proposed framework consists of cryptographic function H, public-key of vehicle i as KU_i, vehicle identity as V_i, driver's information as D_i, and timestamp as T. This block is signed by the vehicle who wants to join to the VANET. In this case, tuple (H_1, KU_i, V_i D_i) constitute message M. Once the message is received then it is unsigned and re-hashed resulting into new hash called H_2. Now, both the hashes ($H_1$and H_2) will be compared and if they are found equal then transaction will move next step which is called validation. The decision to add a transaction to the blockchain is made by consensus algorithm. This means that the majority of nodes in the network must agree that the transaction is valid. The transaction validation procedure is one of the key elements enabling blockchain functionality. The two primary varieties of blockchain, Proof of Work (PoW) and Proof of Stake (PoS), have independent block validation processes.

In this work we have used PoW mechanism. The mining node gathers all transactions before a block is created in the Merkle tree and iteratively hashes the data it collects. The process of hashing terminates when the hash of transactions becomes equal to or less than a pre-determined target value (T_h) called as a threshold $H(n||H(b_c)) \leq T_h$. The H represents SHA-512 hash function and b_c presents the current block. The probability to discover nonce of proof of H can be expressed as $P(H \leq T_h) = \frac{T_h}{2^{512}}$.

After successfully computing the target hash, the miner sends the proof to each node in the blockchain network, along with data transactions and other data, in order for other miners to re-compute and thereby add the new block to the network to form updated blockchain.

De-registration Process

The de-registration process may be invoked by the registering authority or it may be requested by the vehicle depending on the situation. The complete process for de-registering vehicle is shown below.

Vehicle De-registration Process

Input: Certificate C_i, request R_i
Output: Success/Failure
Begin
    If *(**C_i**=valid and **R_i**=TRUE) then*
        De-register vehicle
        Update Vehicle Data from TA/CA
    endif
End

3.5 Chapter Summary

In this chapter, a novel biometrics blockchain based framework has been proposed to improve the security and privacy of VANET. The privacy and authentication of the data were the biggest concerns of researchers in VANET in order to improve the security has been discussed in **section 3.1.** The **section 3.2** the basics of blockchain and VANET. The **section 3.3** discussed the use of biometrics in VANETs to protect privacy, with vehicles employing a public-private key pair provided by the TA to communicate with other parties. The focus of this section is on the storing the data based on blockchain mechanism. Trusted Authority, Motor Vehicle Department, Vehicle, Road Side Unit, Blockchain, Messages are the entities of the proposed framework which has been presented in section **3.3.1.** At the time of registration phase when vehicle information along with driver's detail will be sent to TA. This information will contain the finger print of the driver to ensure the identity. The on-board unit (OBU) will have finger print scanner. The authentication of the driver's identity will be done using modified discrete transformation MDCT. This process of biometrics-based authentication has been discussed in **section 3.2.2.** The system modelling of the proposed system has been discussed in section **3.4.** The sections present the complete process staring from when vehicle joins the network till the vehicle leave the network. This section includes discussion on vehicle registration on BBC, joining the blockchain, receiving the message, message broadcast, update to blockchain, and finally the de-registration when the vehicle leaves the network.

Chapter 4

Formal Methods of Trust Computation

New issues have arisen as a result of the evolution of VANETs, and reputation must be considered because it is critical to know whether vehicles can be trusted on such a network. Through a distributed trust management system, a vehicle might determine its direct relationship with its peers; however, future communication with the same vehicles cannot be ensured. It's difficult to keep track of a vehicle's reputation history when it's anonymous and changes pseudonyms frequently. Storing data on a large number of vehicles may also cause scalability issues. Trust management could be an efficient solution to address VANET security and privacy challenges. It may be possible to reward trustworthy vehicles and flag bad vehicles in VANETs by implementing a trust management framework based on reputation and identity evaluation, therefore ensuring trustworthy message broadcasting. The centralised and distributed models of trust management can be divided into two categories [101]. In the centralised architecture, trust management is handled by a centralised server [102]–[104]. The administration of a centralised server, on the other hand, necessitates a vast number of resources and is vulnerable to attacks by people with malevolent intent who can cause severe problems as a result of single points of failure. Researchers [60], [105], [106] have sought to address these issues by employing a distributed architecture in which trust is assessed by Roadside Units (RSUs) rather than a centralised model. The distributed model has solved the problem of single points of failure by implementing a system in which each RSU's communication ranges are primarily responsible for trust management. VANETs encounter critical and difficult security concerns as they strive to secure the dependability of transmitted information [72]. There have been numerous studies on improving VANET security [15], [68],

[82]; however, there has been a lack of investigations and research on the methods for detecting fake data. A valid node in a VANET can deliver false data to its neighbours with reasonable ease. Worse, this type of information can be distributed at incredible speeds [84]. This information could be utilised to predict driving behaviour. As a result, an evaluation system is needed to determine the relative credibility of the information given in messages in order to avoid erroneous data impacting driving decisions. In an untrusted environment, vehicles have difficulty determining the legitimacy of incoming messages. Trust, data accuracy, and dependability of data being broadcasted via the communication channel are the primary challenges in VANET. Depending on a variety of characteristics, vehicles can determine how trustworthy a given vehicle is based on how well it processes the received message. Therefore, a formal method of trust computation of vehicles is needed. The proposed framework is based on the spatial, temporal and behavioral attributes such as reputation, message correctness, participation degree, message similarity, message freshness, and vehicle age to compute the trust.

4.1 Terms, Definition and Symbols

The trust T_{ij} of a vehicle V_i can be defined as $T \in [0, 1]$ which is evaluated by the vehicle j based on existing knowledge, interactions and behavior in a specific context and time. Depending on the interaction between the vehicles in the network, the trust score may be different. If two vehicles have different experience after interaction, their trust in the vehicle and its interactions might be different. A trust score is therefore a mix of the vehicle specific attributes and the interaction factors. The quality of the interaction and the trust score are both affected by the score T, which expresses the distinctive characteristic score value of the vehicle.

The average trust of $T(V_i)$ of a vehicle i can be defined as by the equation provided (4.1).

$$T(V_i) = \frac{\sum_{j=1}^{N} T_{ij}}{N}, j = 1,2,3, \dots, n \tag{4.1}$$

Here, $j = 1,2,3, \dots, n$ represent peer vehicle in the network. The trust T_{ij} of vehicle i computed by vehicle j is based on existing reputation R_{ij}^{t} of vehicle defined in equation (4.2) and (4.3).

$$T_{ij} = f_T(R_{ij}^{(t)}), 0 < R_{ij}^{(t)} \leq 1 \tag{4.2}$$

$$T_{ij} = \begin{cases} 0, & R_{ij}^{(t)} < t_h \\ 1, & R_{ij}^{(t)} > t_h \end{cases} \qquad T_{ij} \in [0,1] \tag{4.3}$$

The f_T represent the function of trust which is based on the current reputation $R_{ij}^{(t)}$. The vehicle will be trusted if value of the reputation is greater than the pre-defined threshold t_h as shown in equation (4.3).

Table 4.1: Symbol used in the model

C_t^i	Correctness of the message at time *t*
f_c	Function of correctness
M	Message
S	Function of similarity
$R_{ij}^{(t)}$	Reputation at time t or current time, $(t-1)$ represent existing or past reputation.
f_f	Function of freshness
f_d	Function of distance between the message sender
α	The age of node/vehicle
θ	The participation degree of the node
f_T	Function of trust
i_{Reg}	Registration date of vehicle i
n	Number of vehicles in the network
m	Total number of messages
V_i	Vehicle i

4.2 Trust Attributes Exploration

The trust attributes exploration can be divided into three phases. The first phase is to explore set of attributes that defined the characteristics of the vehicle. The second phase is to explore the ways or methods for discovering those attributes. The third phase is about the exploration of assessment processes that vehicle will employ to assess the attributes.

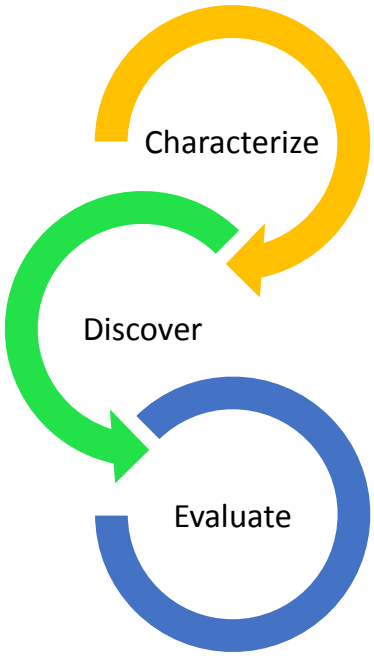


Figure 4.1. The trust attributes exploration

To determine whether a vehicle can be trusted to do the desired mission, the characteristics should be examined. The first of these is integrity, followed by purpose, reliability, and efficiency. The degree of trust of the neighbor node is defined by its integrity, while motivation indicates how motivated a node is to complete the assigned mission. The predictability of peer vehicle activities specifies how consistent they are. Lastly, competency indicates the vehicle's capacity to complete the mission at hand. The integrity of a node defines how its overall ethical status. The integrity of a vehicle defines its overall ethical status. This can indicate a node's honest, or moral standing. This is also known as the vehicle's ability to respond and perform effectively. Prediction and trust are tools for reducing unpredictability. Trust should reach beyond predictability to be effective. Differently, the predictability of the node is incomplete to induce a competitor node to take a mistake and place itself in a vulnerable position. It can, however, assist a node estimate the predictability of another node. The node's qualities can be divided into three classifications: spatial knowledge, temporal experience, and behavioral pattern. The detail of each qualities has been discussed in the following section.

4.2.1 Spatial knowledge

In the trust computation the spatial knowledge features before the interaction, performs the duty of perception for any vehicle. For example, if two vehicles are positioned within the same region, it is logical to expect that they will be trustworthy. Vehicle position, vehicle kind,

vehicle age, identification, and GPS location, technology are a few of the factors that contribute to geographical knowledge about vehicles.

4.2.2 Temporal Experience

Temporal features refer to the vehicle that generates substantial information at a certain time. The value of time-based expertise for the trust evaluation can be extracted using parameters which include network duration, number of packets sent, and number of packets deleted. It is acceptable to believe that the high network duration, frequency, and packet transfer ratio reflect greater trust. However, the outcomes of these encounters may differ from the explanation; it is critical that each individual behavior for future communications be preserved in the form of temporal experiences. This encounter could be Boolean response showing the success of a specific transaction. Furthermore, using machine learning algorithms, the aggregated experience related to a certain connection, time, and context can be utilized to intelligently generate a trust score.

4.2.3 Behavioral Pattern

The behavioral pattern is described by the vehicle's activity as evaluated by different vehicle. Every vehicle in VANET must achieve the pledge to enhance the outcomes of the overall vehicular network in a cooperative environment. A pattern of interactivities is formed by behavioral traits such as feedback.

4.3 Attack Model

The following is a list of the different forms of VANET attacks carried out by malicious vehicles.

a) Transmission Interruption or dropping misbehaviour: This is a common type of attack that exist in most of the network. In this attack, the malicious vehicles purposefully can disrupt transmission of data or repeatedly drop data packets, forcing sender vehicles to rebroadcast these packets or possibly halt data transmission process. This misbehaviour can result in excessive energy consumption and a significant increase in transmission time.

b) Spoofing attacks: The second type of attack in VANET is spoofing attack. In this vehicle that are malicious pretend to have the best data transfer capability or access to a wide range of resources. Malicious vehicle that deceives normal vehicles, after

that malicious vehicle provides a poor delivery expertise and bogus contents to disrupt the entire network.

c) Compromised road side units: The RSU are compromised because they are dispersed throughout the highway and not always protected by network providers. As a result, these entities are believed to be semi-trustworthy and vulnerable to attacks. The attacker can add, delete, and tamper with the data contained in an RSU after breaking into it. Large-scale intrusion attacks, on the other hand, are exceedingly unlikely due to attackers' restricted capabilities. Furthermore, the hacked RSU can't be exploited by intruders for an extended period of time due to the network operators' periodic security checks. On the basis of these factors, therefore, attackers are expected to compromise just a small percentage of RSUs for a relatively short period of time.

4.4 Problem Formulation

Reputation system is a key requirement since it indicates how trusted, secure, or credible a peer vehicle would be in any encounter with a different vehicle. As a result of this unique requirement for reputation assessment, a high-standard, effective reputation characterization and monitoring method is needed. Hence, our model offers a different approach which is based on participation degree, vehicle age, and computation of vehicle reputation.

The current reputation $R_{i,j}^{(t)}$ of the vehicle i computed by vehicle j for any message at time t is derived by the correctness of the message C_t^i, age α, participation degree θ where, the smoothing coefficient β, and existing reputation $R_{i,j}^{(t-1)}$ which can be computed by equations (4.5) and (4.6).

$$R_{i,j}^{(t)} = \begin{cases} \beta R_{i,j}^{(t-1)} + (1-\beta)\left((\alpha + \theta + C_t^i)/3\right), & \alpha > 0 \\ (\beta)\left((\alpha + \theta + C_t^i)/3\right) & \alpha = 0 \end{cases} \tag{4.5}$$

$$\text{Where} \begin{cases} \alpha = \left((t - t_0)/((t - t_{Reg}) + (t - t_0)\right) \\ \theta = \frac{N^+}{N}, 0 \le \theta \le 1 \\ N = N^+ + N^- \\ N^+ = \sum_{k=1}^{n} M_k, N^- = \sum_{l=1}^{n} M_l, N > 0 \\ C_t^i = f_c(M), C_t^i \in [0,1] \end{cases} \quad N = \text{total number of messages}$$

$$f_c(M) = (S\left(M_i, M_j\right).w_1 + f_f\ (M_i).w_2 + f_d\left(M_i, M_j\right)).w_3)/3 \tag{4.6}$$

The coefficient β, a value which can be assigned based on previous experience is between 0 and 1 ($0 \leq \beta \leq 1$). The is a β weight that is assigned to existing trust. For example, if existing trust has been given an initial weight as 0.4, then (1- β) will be 0.6.

The participation degree θ is an important parameter to compute the trust of vehicle. The participation degree indicates the amount of participation in the network. More participation will indicate that vehicle is active which can be achieved using the number of messages classified as correct which were sent with respect to the total number of messages sent. The N^+ represents the total number of messages which were classified as correct or positive. The N^- represents the total number of messages which were classified as incorrect or negative. The N represents the total number of messages which includes correct and incorrect message. The vehicle age α represents the real age of the vehicle and age in the network. The real age refers to the date when vehicle was registered with MVD while the network age refers to time elapse between the first participation time $t_0$and current participation time t. In the equation (4.5), the vehicle age α is a normalized value between (0 and 1). The value $\alpha = 0$, represent that vehicle has just joined the network, therefore, it will have no exiting reputation. The correctness of the message C_t^i as defined using equation (4.5). The function of correctness f_c is defined over many parameters such message similarity, freshness of the message, closeness of senders, and authenticity of the message. The $f_c(M)$ has a value between 0 and 1. If $f_c(M) < t_h$then C_t^i will be classified as 0 otherwise 1. The trust of all the vehicles can be stored into $n\ x\ n$ matrix as follows:

Table 4.2: Trust matrix

	V_1	V_2	V_3	…	V_n
V_1	T_{11}	T_{12}	T_{13}	…	T_{1n}
V_2	T_{21}	T_{22}	T_{23}	…	T_{2n}
V_3	T_{31}	T_{32}	T_{33}	…	T_{3n}
…	…	…	…	…	…
V_n	T_{n1}	T_{n2}	T_{n3}		T_{nn}

Finally, the trust values of all the vehicles are stored into $n\ x\ n$ matrix as shown in table 4.2. For example, if we have $n = 25$ vehicles, so we need to define a matrix (25 x 25) to store all the trust values. In the matrix, the trust of a vehicle i computed by vehicle j is represented by the T_{ij}. This matrix is further used by the RSU and vehicles to know whether a vehicle is trustworthy or not.

4.4.1 Message Similarity Measurement

The same message can be received from multiple vehicles at the similar time for the same event at similar location. With each vehicle broadcasting many messages per second, the arrival rate of safety messages can quickly exceed the rate of digital signature verification. Since not all communications can be confirmed, mechanisms for selecting which messages to examine are required. The proposed method uses sender position, direction, trust and time, to reduce the number of irrelevant messages verified. The message which has minimum distance are classified as similar and anyone will be chosen to verify the signature. There are many algorithms to measure the similarity between the messages such as Euclidian distance [107], Manhattan distance [108], Jaccard similarity [109] and cosine-similarity [110]. Every similarity algorithm has its own advantages and disadvantages. The Euclidean distance is easy to compute and implement, however variables which has the largest value greatly influence the result. The Cosine similarity can be used for both categorical and continuous variable. However, it doesn't work efficiently with nominal data. The Manhattan Distance is easy to generalize into higher dimensions but can't be used to compute the nominal values. We found that Jaccard Similarity among all has less complexity and easy to implement. Moreover, it can be used for all kind of variables such as continuous, discrete, categorical and nominal. The table 4.3 shows that all the messages are produced by multiple vehicles and describe the same event, which they have observed. The location which consists of Latitude and Longitude where the message was produced by the vehicles at specific time. Every vehicle has reputation value, speed and direction which will be considered in computing the similarity.

Table 4.3. Scenario of multiple event messages

Message	Sender	Time	Location	Speed	Direction	Reputation
M_1	V_1	00:01:23	(21° 32' 49.4772", 39° 13' 33.5496")	50	180	0.8
M_2	V_2	00:01:24	(21° 32' 49.4772", 35° 13' 33.5486")	49	181	0.6
M_3	V_3	00:01:23	(21° 33' 49.4772", 39° 13' 33.5496")	50	175	0.7
M_4	V_4	00:01:25	(21° 32' 49.4772", 39° 13' 33.5496")	48	200	0.5
M_n	V_n	00:01:24	(21° 33' 49.4772", 39° 13' 33.5496")	50	180	0.8

The table 4.3 present a set of objects such as Message, Sender, *Time*, *Location*, *Speed*, *Direction and Reputation* that can be used to find out the similarity, The column message contains all the messages received from different vehicles such as message $M_1$is received b vehicle V_1at time 00:01:23. The location of the sender vehicle was represented in tuple of latitude and longitude as (21° 32' 49.4772",39° 13' 33.5496"). The speed of vehicle $V_1$was recorded as 50 Km/h while the direction towards vehicle was heading was 180^0. The vehicle has exiting reputation of 0.8. In order to determine more likely authentic message a similarity computation is performed on every pair of messages M_i, M_j based Jaccard distance formula using below formula in equation (4.7).

$$D(A,B) = 1 - J(A,B); J(A,B) = |A \cap B| / |A \cup B| \tag{4.7}$$

4.4.2 Message Freshness

The freshness f_f of the message can be obtained from the time difference between the message transmission time t_t and message receiving t_r as shown in equation (4.8).

$$f_f = t_r - t_t \quad s.t \begin{cases} \text{TRUE}, f_f \geq t_h \\ \text{FALSE}, f_f \leq t_h \end{cases} \tag{4.8}$$

If f_f is greater than or equal to the pre-determined threshold then message is classified as fresh, otherwise not fresh. The threshold $t_h$is determined based on the estimated packet delivery time in traffic congestion situation. If message is delayed beyond the threshold, then it indicates less reputation about the vehicle. A higher value for the f_f indicates higher reputation about the sender vehicle.

4.4.3 Sender Proximity and Event Location

The distance plays an important role to find out the proximity between the event location and the vehicle which has observed the event. The message which sent by two or more vehicles must be obtained in order to find out the accuracy of messages. The distance from the event can be obtained by finding the latitude and longitude easily using below equations (4.9) – (4.10).

$$d = R * C \quad \text{where,} \begin{cases} \mathrm{C} = 2.\operatorname{atan} 2(\sqrt{a}, \sqrt{(1-a)}) \\ \mathrm{a} = \mathrm{Sin}^2\left(\frac{\Delta\emptyset}{2}\right) + \cos\emptyset_1 \,.\, \emptyset_2 . \mathrm{Sin}^2\left(\frac{\Delta\lambda}{2}\right) \end{cases} \tag{4.9}$$

$$D = Min\,\{d_1, d_2, d_3, \ldots, d_n\},\ n > 0 \tag{4.10}$$

The d represents the distance between the pair of latitude and longitude. The radius of the earth is denoted by R(6,371km). The C represents a well-known *haversine* formula to find out shortest distance over the earth's surface. Here, $\emptyset$ denotes latitude ($\emptyset_1$- latitude 1, $\emptyset_1$- latitude 2), and λ denotes longitude (λ_1- longitude 1, λ_2- longitude 2). The lower value of d_i between the vehicles indicates that message or event being reported will be correct.

4.5 Chapter Summary

The **section 4.1** presented a formal definition of the trust. The trust T_{ij} of a vehicle V_i can be defined as $T \in$ [0, 1] that vehicle i can assess relying on the current details, interactions and behavior in a specific context and time. After definition the important symbol used in the mathematical formulation is discussed in the **section 4.2**. The trust attributes must be addressed to determine whether a vehicle can be trusted to fulfil the needed duty. The vehicle's features can be divided into three classifications: spatial knowledge, temporal expertise, and behavioral type. which is discussed in **section 4.3**. A brief introduction to the possible attack model in VANET is presented in **section 4.4**. The mathematical formulation of the trust model is presented in **section 4.5**. The reputation $R_{i,j}^{(t)}$ of the vehicle i computed by vehicle j at time t is the correctness of the message C_t^i, age α, participation degree θ, the smoothing coefficient β, and existing reputation $R_{i,j}^{(t-1)}$ are defined in this section. This section also presents the trust of all the vehicles that can be stored into $n\,x\,n$ matrix. The mathematical formulation on message similarity, message freshness, and sender proximity and event location are the part of this section.

Chapter 5

Ensemble Learning for Trust Classification

As VANET has developed, new problems have emerged, and reputation must be taken into account since it is crucial to understand if vehicles can be trusted on a network. A vehicle may establish a direct contact with its peers via a distributed trust management system, but future communication with the same vehicles cannot be guaranteed. When a vehicle is anonymous and regularly switches pseudonyms, it might be challenging to monitor its reputation history. Scalability problems may also result from storing data on a lot of different vehicles. Regarding VANET security and privacy issues, trust management may be an effective approach. By establishing a trust management system based on reputation and identity evaluation, it may be feasible to reward reliable vehicles and identify unreliable ones on VANETs, and assuring reliable message broadcasting. There are two categories that may be used to both the distributed and centralised trust management methods [101]. A centralised server manages trust management in the centralised architecture [101], [102], [104]. On the other hand, managing a central server requires a tremendous amount of resources and is susceptible to attacks from those with bad intentions who may cause serious issues as a result of single points of failure. Instead of using a centralised paradigm, [60], [105], [106] have used a distributed design in which trust is evaluated by Roadside Units (RSUs). The distributed model has solved the problem of single points of failure by implementing a system in which each RSU's

communication ranges are primarily responsible for trust management. VANETs encounter critical and difficult security concerns as they strive to secure the dependability of transmitted information. [111], [112]. There have been numerous studies on improving VANET security [113]–[115]; however, there has been a lack of investigations and research on the methods for detecting fake data. A valid node in a VANET can deliver false data to its neighbours with reasonable ease. Worse, this type of information can be distributed at incredible speeds [116]. This information could be utilised to predict driving behaviour. As a result, an evaluation system is needed to determine the relative credibility of the information given in messages in order to avoid erroneous data impacting driving decisions. In an untrusted environment, vehicles have difficulty determining the legitimacy of incoming messages.

Trust, data accuracy, and dependability of data being broadcasted via the communication channel are the primary challenges in VANET. Depending on a variety of characteristics, vehicles can determine how trustworthy a given vehicle is based on how well it processes the received message. Therefore, a formal method of trust computation of vehicles is needed. The proposed framework is based on the spatial, temporal and behavioral paraments such as reputation, message correctness, participation degree, message similarity, message freshness, and vehicle age to compute the trust.

5.1 System Model

Figure 5.1 highlights the suggested model general scheme architecture, which essentially involves four elements: vehicles, RSU, trust management mechanism, and blockchain. Each component purpose is outlined below.

5.1.1 ***Vehicles:*** Each vehicle in VANETs is equipped with an on-board unit (OBU) capable of wireless communication, allowing it to interact peer vehicles and RSUs and sharing of data [34]. In addition, we believe that each OBU contains a tamper-proof device (TPD) for storing sensitive data like secret keys. TPD stores critical data in a physically secure environment. The cryptographic computations and system parameters should be kept in the TPD to prevent manipulation.

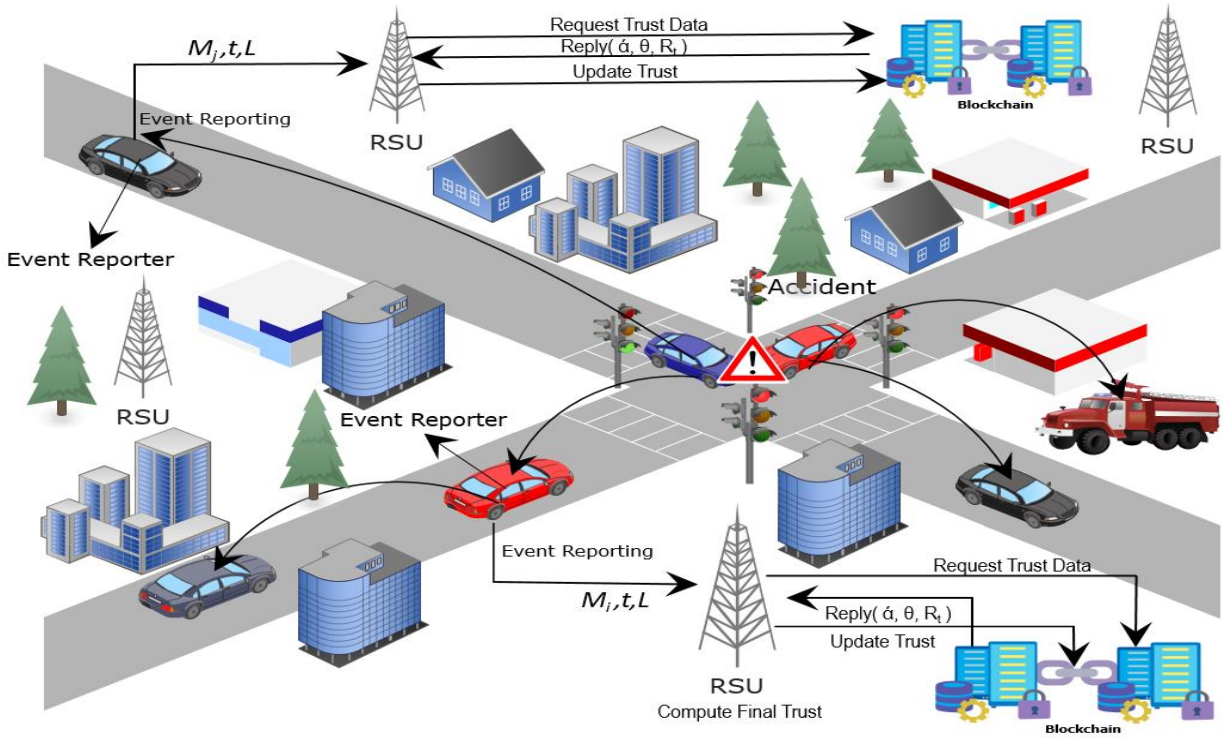


Figure 5.1. Proposed Trust Model

5.1.2 ***Road Side Unit (RSU):*** Roadside-unites have a greater calculation capability and communication range than vehicles. Furthermore, most RSUs are often positioned on both sides of the street, giving high-speed vehicles easy access to surrounding RSUs. In addition, the chance of the RSU being hijacked is reduced, and restoration is faster. As a result, rather of maintaining vehicles, we have the RSU to maintain the blockchain.

5.1.3 ***Trust Management Model:*** In the VANET, a vehicle's credibility is extremely important. As a result, the trust model calculates the event message initiator's trust based on the previous reputation credit and the message correctness based on a number of characteristics.

5.1.4 ***Blockchain:*** Blockchain's Proof-of-Work (POW) and duplicate decentralized copies produce extremely high security and reliability, and as a result, it has been widely investigated and applied [38]. To strengthen the security of VANETs, we use blockchain in our proposed model. Figure 5.2 depicts the architecture of the blockchain employed in this thesis. It's an ordered list of blocks with a block head and a block body for each block. The head contains details such as the previous block address and timestamp, among other things. The VehicleList, which is located in the body as shown in figure 5.3, is used to record comprehensive transaction information. Each vehicle's ID, trust value, and type are stored in a VehicleList. It's important to note that vehicles might be harmful or non-malicious.

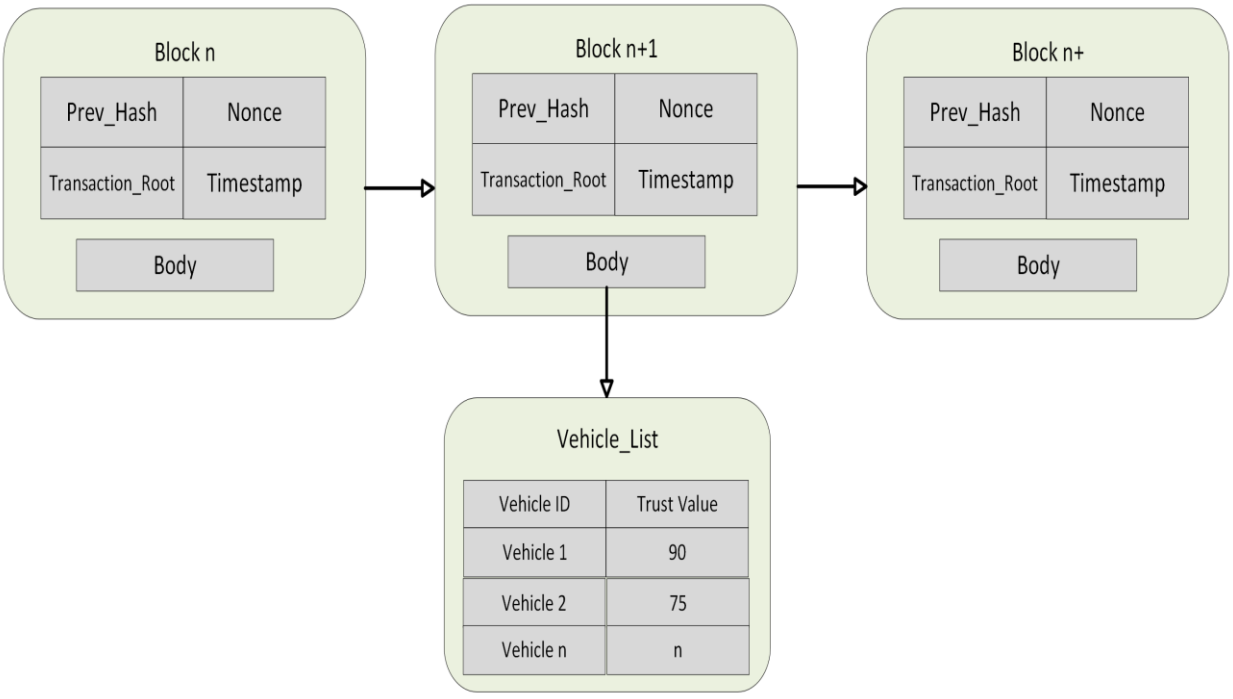


Figure 5.2. Blockchain data model for VANET

5.2 Design Goals

The objectives of this work are to evaluate, record, and disseminate vehicle trust in vehicular communications. As a result, the major goals are each vehicle activity may be fairly assessed, and that all vehicles can, if necessary, obtain dependable neighbor trust values. The following objectives should be met by the design of a trust management system.

1) *Decentralization:* Centralized trust management solutions may become obsolete as the number of smart vehicles grows. As a result, the trust management system must fully utilize distributed nodes, such as RSUs and vehicles. Ratings from message receive are used to calculate trust values, which would then be stored in the RSU, ensuring the system's reliability and scalability.
2) *Tamper-Proofing:* road side unites are typically dispersed both sides of the street and are subject to being hacked. The dependability of trust management will be harmed if data contained in the exploited RSU is tampered with. Due to the attackers' limited capacity, a large-scale compromise of RSUs is improbable. As a result, the trust models must be able to resist the compromising of a small number of RSUs.
3) *Consistency*: Vehicles must frequently traverse across many RSUs due to their high mobility. In this circumstance, maintaining a consistent database and exchanging trusted data among RSUs becomes a difficult challenge for distributed trust models in VANETs.
4) *Timeliness:* The total rating of a vehicle relying on the previous activities is represented by its trust value. This number may change over time depending on the trustworthiness of recent communications provided by this vehicle. As a result, the trust values held in RSUs must be updated on a regular basis.

5.3 Design Overview

In the proposed model, biometrics is integrated with blockchain as a decentralized trust management and reputation platform with VANET. The exchange of data between vehicles is the basis of VANET. Each vehicle will have biometric information. When a vehicle (referred to as a reporter) detect an event with the equipped sensors, it immediately generates and sends a message about the event to the nearby vehicles and RSUs. The event message contains of several fields such as event type and vehicle biometric ID which have the biometric information of the authorized user. When a vehicle/RSU receives the message, in one dimension it

calculates the reputation of the event reporter based on vehicle age, participation degree, and the existing reputation credit of the event reporter. In the second dimension, receiver will calculate the credibility of the received message based on several factors as explained in the later sections. Once both dimension, reporter vehicle reputation and the message correctness calculated. We combine both reputation of the vehicle and correctness of the message to get the final trust value of the vehicle. If the final trust result above the predefined threshold after the combination of both dimensions vehicle reputation and message correctness, then, the vehicle will be rewarded an extra credit associated with biometric ID of the vehicle authorized user. Moreover, the reputation credit of a vehicle will be stored in the biometric blockchain. The proposed trust model and reputation will provide higher trustworthiness among vehicles which will lead to higher security and accuracy of messages being transmitted in VANET.

5.4 Miner Election and Block Generation

Due to the decentralized nature of the network, there is no single point of control for the blockchain. Therefore, a miner from all RSUs is selected on a regular basis to build new offset blocks, as a result of this process. The proof-of-work (PoW) based miner election mechanism is commonly employed in blockchain-based systems. The nodes in the network modify the nonce on a regular basis and then calculate the block's hash values, including the nonce. The miner is the one who gets the hash value below a certain threshold and is able to publish their block. Figure 5.3 depicts the format of the reputation data update block. It has a block header as well as a block body. The header contains the following information: RSU ID, block ID, block generation timestamp, nonce, hash of preceding block, hash threshold, for confirming the legitimacy of this block. The body of the block is used to store the vehicle's or trust value. The nodes with more powerful computing capacity will have an easier time obtaining the proper nonce and hence winning the election because the threshold for all nodes is the same based on the challenge shown in the equation (5.1):

$$Hash(\rho\ , t\ , prev_{Hash}\ , RSU_{id}) \leq Hash\ threshold \tag{5.1}$$

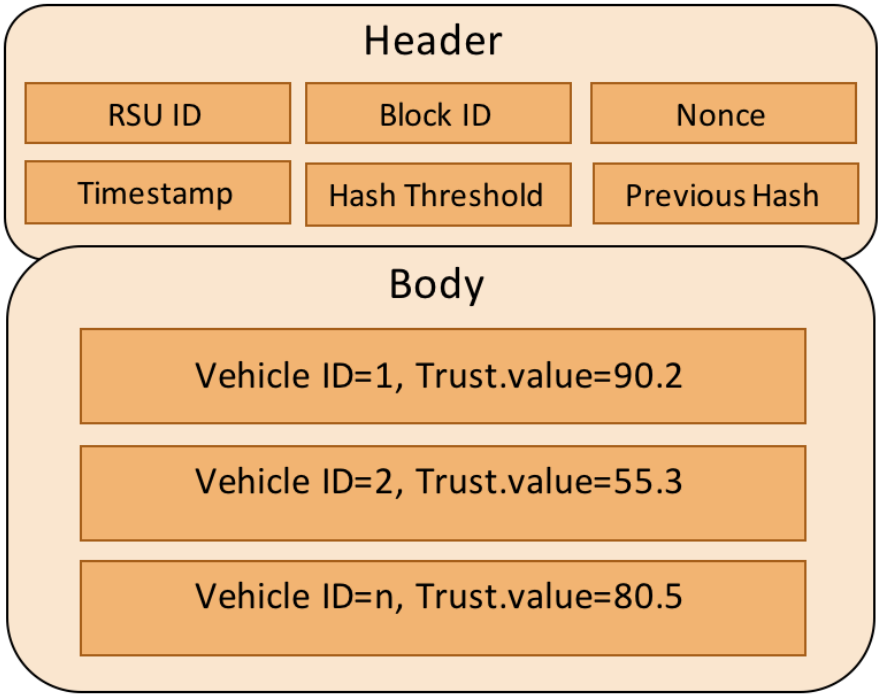


Figure 5.3. Block Format for VANET

The ρ is the nonce and t represents time. When the RSU receives a block from the miner, it must verify the nonce's legitimacy before adding it to its blockchain. This is possible for the RSU to receive a large number of blocks at once. In such a scenario, the blockchain begins to fork to make the process faster. To deal with this problem, a distributed consensus technique is implemented. Each RSU picks a fork and adds additional blocks to it.

The branch with the most RSUs grows quicker than the others over time. Finally, the network's distributed consensus is formed by discarding everything but the longest branch. It also needs to keep the blocks that were made by each RSU and attempt to append them to blockchain network at a later time. As a result, all RSUs employ the exact blockchain version, therefore aids in network stability.

5.5 Machine Learning Model

As a first stage, a substantial volume of network traffic data is collected under both normal and abnormal conditions induced by various scenarios. In order to create classifiers, data must be acquired using packet sniffers, however they must contain the appropriate network properties or specified network characteristics. As a result of the lack of the VANET dataset's that contains features related to reputation, we have developed our simulated data. Therefore many

common network parameters, such as packet length, have been considered such as packet length, TTL, total forwarded packets, total backward packets, failure type, option type, road condition, speed, weather, time scenario, lane type, traffic scenario, packet type, and location etc as shown in Figure 5.4.

	Source	Destinatio	tection Tar	Time Stamp	TTTL	Trust	OT	Failure Type	Road Condition	Speed Scenario	Time Scenario	ather Scen	Lan Type	Traffice Scenario	acket Ty	Latitude	Longitude
2	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	191	0.796961	3	1	Dry	constant	Dawan	Windy	Curve	Car on 3 sides	General	21.55947	39.20696
3	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	202	0.184845	1	1	wet	Accelerating	Night	Foggy	Straight	Car on 2 sides	Traffic	21.56394	39.20991
4	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	198	0.71733	3	2	Icy	constant	Dusk	Raining	Intersectic	Car on 4 sides	General	21.55821	39.21849
5	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	70	0.246915	3	1	wet	constant	Day	Windy	DownHill	No cars	General	21.55516	39.17695
6	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	72	0.649204	2	3	Icy	Deaccelerating	Night	Snowing	Winding	No cars	Safety	21.58123	39.23123
7	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	238	0.522136	3	3	Icy	constant	Day	Clear	Winding	No cars	Traffic	21.54673	39.19348
8	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	245	0.970512	1	0	Icy	Accelerating	Night	Windy	Intersectic	Çar on 1 side	General	21.57964	39.18276
9	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	129	0.476865	3	2	Icy	constant	Dawan	Raining	Winding	Car on 4 sides	Traffic	21.54989	39.23862
10	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	92	0.283233	0	3	Dry	Deaccelerating	Day	Foggy	Intersectic	No cars	Traffic	21.5506	39.19483
11	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	180	0.069922	1	2	Icy	Deaccelerating	Day	Snowing	Intersectic	Çar on 1 side	General	21.58186	39.20773
12	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	150	0.962196	1	1	Dry	constant	Night	Foggy	Straight	Car on 3 sides	General	21.54847	39.24068
13	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	81	0.358866	1	0	wet	constant	Dawan	Raining	UpHill	Car on 3 sides	Safety	21.54859	39.24526
14	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	229	0.640388	1	2	Icy	constant	Dusk	Clear	UpHill	Car on 3 sides	Traffic	21.55105	39.23865
15	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	183	0.134814	2	3	wet	Deaccelerating	Dusk	Foggy	UpHill	No cars	General	21.54653	39.2193
16	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	198	0.204285	2	0	Icy	Deaccelerating	Night	Clear	Intersectic	Car on 3 sides	Traffic	21.55325	39.19321
17	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	245	0.134519	0	2	Dry	Deaccelerating	Day	Raining	Intersectic	Çar on 1 side	Traffic	21.5552	39.21152
18	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	128	0.852354	2	0	wet	constant	Night	Windy	DownHill	Car on 4 sides	Safety	21.55334	39.25871
19	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	80	0.826901	3	3	wet	Accelerating	Dusk	Snowing	Curve	Car on 3 sides	Safety	21.57553	39.2352
20	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	76	0.724673	3	2	Dry	constant	Day	Raining	Winding	Çar on 1 side	Safety	21.55805	39.23326
21	20.100.16	20.100.16	20.100.16	Sun Dec 5 12:06:34	103	0.373956	0	0	Dry	Accelerating	Day	Foggy	DownHill	Car on 4 sides	Traffic	21.55021	39.20944

Figure 5.4. Snapshot of dataset attributes

However, due to the large data dimensionality, the computational complexity of the proposed classifier may significantly rise. The external network data should thus be analysed for additional features. Classifier development will be made easier by pre-processing the acquired network data. To begin, normalized data makes the machine learning training process more efficient since it can be encoded using one-hot-vector because it contains a threshold to help identify trustworthy packets from non-trustworthy packets. If we want that the data should not be influenced by how we measure things, then we should normalize the data.

5.5.1 Data normalization process

This means changing the data so that it fits into a smaller range, like [0.0, 1.0]. The A_{min} and A_{max} is the minimum and maximum values of an attribute A. The mapped new attribute is A' is computed as shown in equation (5.2):

$$A' = \frac{A - A_{min}}{A_{max} - A_{min}} \tag{5.2}$$

To identify trustworthiness in the proposed system, a binary-classification problem is

examined, and machine learning methods are frequently employed to solve such classification issues. The machine learning techniques chosen are tree-based, including decision trees, XGBoost, and random forest. The tree-based machine learning algorithms are the subset of supervised learning algorithm. The task of classification and regression are performed by constructing a tree-like structure for determining the target variable class or value based on the properties of the input data. In many applications, tree-based algorithms are one of the most widely used machine learning methods. The tree-based algorithms separate every root into two branches at every depth level, beginning with the top node and progressing below as shown in Figure 5.5. The choices, also known as the leaves, are the terminal end branches where they do not split further. At every depth, there are conditions that questions the feature values. According to the binary response, the next branch will be chosen. It will continue to divide until we reach one of the leaves, at which point it will stop splitting. By looking at the final leaf, we can figure out what is going to be predicted. In the field of supervised learning, tree-based algorithms are widely regarded as one of the most effective and widely used approaches. In addition to great accuracy and stability, tree-based algorithms also provide models with simplicity of comprehension. The fact that they map non-linear connections successfully, as opposed to linear models, is a significant advantage. They are capable of dealing with any type of difficulty that may arise in classification or regression. All types of data science challenges are being solved using methods such as decision trees, random forests, and gradient boosting.

5.5.2 Decision Tree

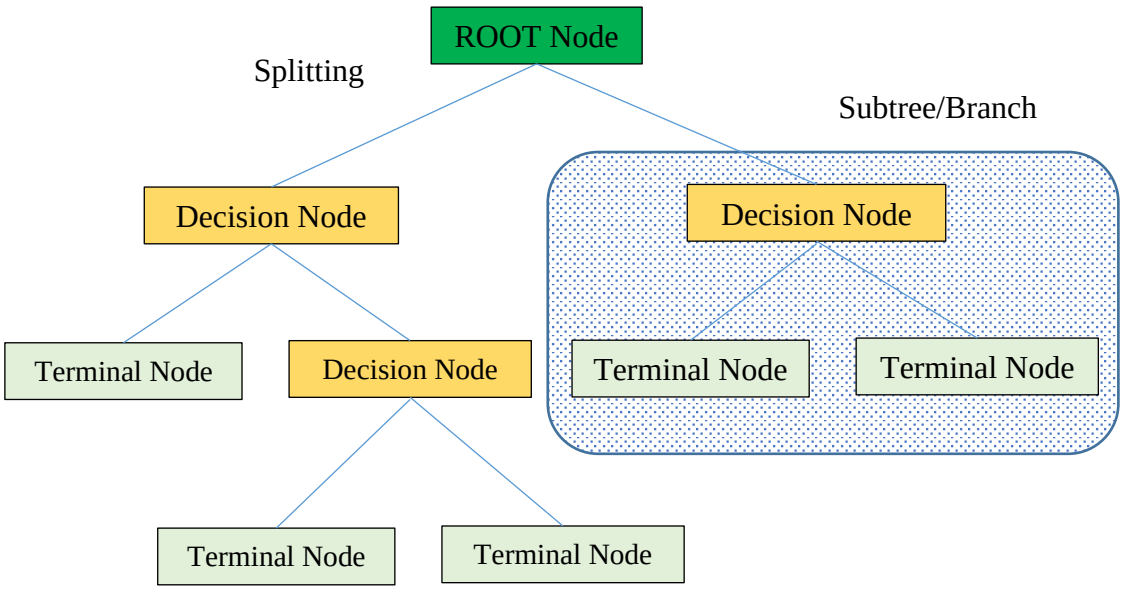


Figure 5.5. Decision Tree

The decision tree contains root node, decision node, and terminal node as shown in figure 5.5. The detailed explanation of these terms are as follows:

Root Node: This depicts the overall community, which is gradually subdivided into two or more identical sets. *Splitting process:* This is the process of dividing a node into two or more sub-nodes/branches as shown in figure 5.5. *Decision Node:* It is referred to as a decision node when a sub-node divides into other sub-nodes after it has formed. *Leaf/ Terminal Node:* Leaf and Terminal nodes are nodes that do not split. *Pruning:* Pruning is the term used to describe the process of removing sub-nodes from a decision node. You might say that the process of splitting is the inverse of the process of splitting. *Branch / Sub-Tree:* A branch or a sub-tree is a portion of a tree that is smaller than the total tree.

Parent and Child Node: A node that is subdivided into sub-nodes is referred to as the parent node of sub-nodes, whilst sub-nodes are the children of the parent node, and vice versa. Algorithms for creating decision trees are typically top-down, with each step selecting a variable that best divides the collection of objects. Different algorithms employ different measures to determine which is the best decision. These are often used to assess the homogeneity of the target variable across the subsets of data. In the below we list some of metrics that help in evaluation of the decision. Estimate of the positive correctness can be defined by equation (5.3):

$$E_p = TP - FP \tag{5.3}$$

Here, number of false positives (FP) are deducted from number of true positives (TP). The Positive or true are those in which the feature was able to find and categorize appropriately. A higher number indicates that more positive instances were found by the feature.

Gini Impurity

The CART (classification and regression tree) algorithm uses the Gini to determine the likelihood that a randomly selected element from the set will be incorrectly labelled if it is randomly labelled according to the distribution of labels in the subset. The Gini impurity is computed by add all the probability p_m of an item with label m being selected times the probability $\sum_{l \neq m} p_l = 1 - p_m$ of a mistake in categorizing the particular item. The Gini impurity can be computed for a set of items which has K classes $m \in \{1,2,\ldots,K\}$. Let the

p_l be the fraction of items which is labelled with class l. Then Gini impurity can be compute as shown in equation (5.4):

$$I_G(p) = \sum_{m=1}^{K} \left(p_m \sum_{l \neq m} p_l \right) = \sum_{m=1}^{K} p_m \, (1 - p_m \,) = \sum_{m=1}^{K} (p_m \, - p_m^2 \,)$$
$$= \sum_{m=1}^{K} (p_m) - \sum_{m=1}^{K} (p_m^2)$$
$$= 1 - \sum_{m=1}^{K} (p_m^2) \tag{5.4}$$

The entropy H can define by equation (5.5):

$$H(T) = I_E(p_1, p_2, \dots, p_K) = - \sum_{m=1}^{K} p_m \, log_2 \, p_m \tag{5.5}$$

5.5.3 XGBoost

The Gradient Boosted Trees (GBT) technique is classified as supervised learning since it relies on the approximation method, which is optimized via the use of specific loss functions, as well as the use of numerous regularization procedures, to achieve its results. For the sake of our investigation, we are seeking for a function that can improve on the performance of the given model.

As a result, the loss function serves as a useful indicator of the accuracy of our model's predictions. In general, if the predicted outcomes $\hat{y}_i$, are very similar to the real values y_i, then the loss is the least significant; conversely, if the predictions are substantially different from the original values, then the loss is the most significant. The loss may be calculated with the help of equations (5.6).

$$Loss = \, |y_i \, - \hat{y}_i| \tag{5.6}$$

In accordance with the value of Loss function, the model is iterated for updating until the best result is obtained. The binary cross entropy (Log loss) method has been used to get the classification result. In XGBoost, we have a large number of trees to choose from [117-118]. If we assume G trees in the XGBoost then the prediction model can be expressed by equation (5.7)

$$\sum_{g=1}^{G} f_g \tag{5.7}$$

Where f_g is the prediction from the decision. Therefore, the predictions using all of the decision trees can be shown by (5.8):

$$\hat{y}_i = \sum_{g=1}^{G} f_g(x_i) \tag{5.8}$$

The $x_i$represents the feature vector for the data point i. The N is the total number of rows, and M is a list of classes. Loss function and regularization at iteration t must be minimized as part of the objective function. The below equation (5.9) may be used to define XGBoost's goal function [117 – 118].

$$\mathcal{L}^t = \sum_{i=1}^{n} l.\,(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t)\ ;\ \text{s.t}\ x = \hat{y}_i^{(t-1)} \tag{5.9}$$

5.5.4 Random Forest

A random forest, as the title suggests, is comprised of a large amount of independent decision trees that collaborate to overcome issues. Using the random forest, every independent tree generates a class assumption, and the class with the most votes becomes the class predicted by our model. Random forest is based on the wisdom of crowds, which is a basic yet efficient notion. The reason why the random forest model performs so well is known as the random forest effect.

The decision tree (DT) is a widely used categorization technique that employs the divide and conquer strategy. A decision tree is composed of decision nodes and leaf nodes, which indicate a judgment test on one of the characteristics and accordingly, the outcome class. By comparison, XGBoost is an ensemble learning technique optimized for speed and efficiency by combining many decision trees through the gradient descent approach. Apart from these algorithm, additional models such as Random Forest (RF) is another ensemble learning techniques depending on the largest group election voting, in which the decision tree class with the most votes is chosen as the classification result.

5.5.5 Ensemble Learning with Feature selection in Random Forest

Most tree structure machine learning models employ ensemble learning, which can lead to better results than single models like linear regression or KNN. In order to make feature selection easier, feature importance estimations are made throughout the model construction

process. In random forest we assume that we have N instances and feature vector is $\{f_i\}_{i=1}^{N}$and O_i then data D can be defined as $D = \{(f_1, O_1),\ldots, f_N, O_N)\}$ and very feature vector can be defined as $f_k = (f_{k1}, \ldots, f_{xd})$. At each node, the features f_i and threshold t are chosen in order to reduce the amount of diversity that results in the offspring nodes. The Gini criteria, is frequently used to assess diversity.

Gini criterion - The class C_1=non-Trustworthy and class C_2=Trustworthy, which is the most accurate way to determine the number of instances in a node with regard to two classes.

If we assume the set S as an example in the present node the $S = S_1 \cup S_2$. Where $|S|$= Number of items in the S then we can define $\hat{P}$ as in equation (5.10):

$$\hat{P}(\dot{S_j})=\frac{|S_j|}{|S|} \; and \; \hat{P}(\dot{C_i}|CS_j)=\frac{|S_j \cap C_i|}{|S_j|} \tag{5.10}$$

The variaton and Ginnin index can be defined as showing in equation (5.11) - (5.12):

$$g(S_j) = \sum_{i=1}^{2} \hat{P}(C_i|CS_j)\left(1 - \hat{P}(C_i|CS_j)\right) \tag{5.11}$$

$$G = \hat{P}(S_1)g(S_1) + \hat{P}(S_2)g(S_2) \tag{5.12}$$

When numerous designs (low learners) are educated on the same dataset and then combined to get better results. and then integrated to produce superior results, this is termed ensemble learning. If weak models are coupled appropriately, they can yield more accurate predictions. Training a meta-model, which considers many weak models, allows stacking to provide a prediction that takes into account the results of all of the models that were trained. The input data to the random forest can be defined as shown in equation (5.13):

$$D = \{(f_1, O_1),\ldots, f_N, O_N)\} \tag{5.13}$$

The output of the algorithm is an ensemble or weak learner can be defined as shown in equation (5.14) – (5.15):

$$h = \{h_1(f), ...,h_{\mathrm{k}}(f)) \tag{5.14}$$

$$h_k(f) = h(f|\Theta_k) \tag{5.15}$$

The margin function in ensemble learning can be defined as shown in equation (5.16):

$$\widehat{M}(\dot{F}, o) = \hat{P}_k(h_k(f) = o) - \underbrace{max}_{j \neq o} \hat{P}_k(h_k(f) = j \quad (5.16)$$

The margin function is the percentage of votes for the proper class that surpasses the percentage of votes for the second-best class. The strength of the random forest as ensemble can be defined as shown in equation (5.17):

$$s = \mathbb{E} \in_{x,y} M(\dot{F}, o) \quad (5.17)$$

The error using Chebyshev inequality can be defined as follows in equation (5.18):

$$e = P_{x,y}(M(F, o) < 0 \leq P_{x,y}(|m(F, o) - s| \geq s) \leq \frac{V(M)}{s^2} \quad (5.18)$$

5.5.6 Complexity of Machine Learning Algorithms

In a dataset if the number of cases is N and the number of attributes is f, and the number of trees generated is T, accordingly temporal difficulty of a decision tree can be computed as $O(N^2 f)$. The complexity of XGBoost will be $O(NfT)$ On the other hand complexity of random forest can be computed as $O(N^2 \sqrt{f} T)$. To reduce the computational time a multi-processing capability can be applied. If the number of the processor available for processing is P then the time complexity of random forest will be $(O \frac{(N^2 \sqrt{f} T)}{P})$.

5.6 Chapter Summary

The **section 5.1** highlights the suggested general system design, which consists primarily of four elements: vehicles, RSU, trust management model, and blockchain. Each element's purpose has been thoroughly described. In the next section **5.2**, the design goals are presented. The design goal of the proposed system are decentralization, tamper-proofing, consistency, and timeliness. The **section 5.3** present design overview of the proposed system where biometrics is integrated with blockchain as a decentralized trust management. The proposed trust model and reputation will provide higher trustworthiness among vehicles which will lead to higher security and accuracy of messages being transmitted in VANET. The **section 5.4** presents discussion on miner election and block generation. The **section 5.5** is dedicated for machine learning model used in the proposed model. In this section discussion starts with the data normalization process and ends with the computational complexity of the model. For data normalization min-max method is used. To identify trustworthiness in the proposed system, a

binary-classification problem is examined, and machine learning methods are frequently employed to resolve various categorization tasks. The machine learning techniques chosen are tree-based, including decision trees, XGBoost, and random forest. Apart from these algorithm, additional models such as Random Forest (RF) is another ensemble learning technique used for categorization relying on the overall election law, in which the decision tree class with the most votes is chosen as the classification result. At the end a discussion on the complexity of the machine learning model is discussed. The section concludes the complexity of random forest to be computed as $O(N^2\sqrt{f}T)$. It is also emphasized that to reduce the computational time a multi-processing capability can be applied. If the number of the processor available for processing is P then the time complexity of random forest will be $(O\frac{(N^2\sqrt{f}T)}{P})$.

Chapter 6

Simulation and Result Analysis

The simulation and findings are presented in this chapter are intended to validate the validity and efficacy of the proposed model. First, this chapter presents results and analysis of privacy-preservation framework to prevent attacks in VANET. Second, this chapter presents the results obtained from trust computations and classification. Additionally, this chapter discusses the results of the simulated implementation of the system models presented in Chapter 3, 4, 5 and the proposed solutions to the identified objectives at the onset of this thesis. Section 6.1 present obtained results and findings on the privacy preservation and blockchain framework, section 6.2, discusses that of the formal methods and results obtained, section 6.3, presents the result obtained from the ensemble-based machine learning algorithm.

6.1 Privacy Preservation

Vehicle anonymity must be utilized in VANET communications in order to protect users' privacy. It should be made conditional since the authorities would be able to trace the anonymity to locate the malicious vehicle. Other vehicles or RSUs should not be allowed to know the true identification of a vehicle, and a malicious actor must not be able to acquire identities by examining any identity intercept. Without detection by the RSU or other drivers, the messages should be completely invisible. We achieved this by creating a unique id using biometrics and taking the hash (SHA-512) as shown by equation (6.1):

$$Vp = \mathcal{H}\ (dbio \leftarrow\ Cf_{T_{(p\times N)}}) \tag{6.1}$$

Now, every vehicle has a unique 512-bit pseudo identity. The MVD and TA can only find the real identity of a vehicle by looking at the database stored in the blockchain. The traceability will be needed whenever there is a malicious activity of any fake message is generated by a vehicle.

6.1.1 Experimental Setup

The simulation has been performed using OMNeT++, Veins and SUMO to demonstrate the correctness of the proposed model using IEEE 802.11p/1609.4 protocols. The trace control interface (TraCI) which is a middle interface between OMNeT++ and SUMO++ that provides a TCP based communication between these two simulators. To evaluate the model, we have chosen the parameters as shown in table 6.1 and 6.2 which consist of 100 vehicles with a maximum speed of 50 m/s. The length and width of the vehicle is 4m and 2.5m respectively. The number of RSU in the experiment are 15 while the coverage of a single RSU is 2 km.

Table 6.1: SUMO simulation parameters

Parameters	Values
Simulation time	3000 s
Queue length of the MAC	10
Bit rate of MAC	15 Mbps
Maximum Transmission Attempts	20
Transmission Power	100 mW
Contention Window of MAC	10
PHY. Sensitivity	-80 dBm
Interval to update	0.01s

Table 6.2: OMNET simulation parameters

Parameters	Values
Number of vehicles	100
Max. speed of Vehicle	50 m/s
Maximum Acc.	3 m/ s^2
Maximum Dec.	5 m/s^2
The length of the vehicle	4 m
The width of the vehicle	2.5 m
RSUs No	15
Coverage of RSU	2 km.
Sigma	0.5

6.1.2 PDR without Denial-of-Service attack

The packet delivery ratio (PDR) is a critical metric for evaluating the effectiveness of the proposed mechanism in VANET. The effectiveness of the technique is dependent on the simulation's numerous parameter settings. The most important factors are packet size, node count, transmission range, and network topology.

The packet delivery ratio can be computed by dividing the overall number of data packets have reached to recipients by the overall number of packets delivered from origins. Specifically, the packet delivery ratio the percentage of packets transmitted from the originator to those obtained at the endpoint. Effectiveness of the communication improves whenever the packet delivery proportion is high. It is arithmetically defined as an equation (6.2).

$$PDR = \frac{S_p}{R_p} \tag{6.2}$$

Where S_p the total packets sent by the source vehicle and R_p is the total number of packets gathered by the target vehicle. Figure 6.1 shows the impacts of PDR in the absence of an attacker in the network. The proposed model has high PDR of 0.99, however, BC-VANET [96], ASC [119] and LAKAP [120] have PDRs of 0.98, 0.94, and 90 respectively. The average PDR of the proposed model is 0.97 with 20 vehicles in the network. We can observe that the highest delivery rate is 0.98 which has dropped slightly when the number of vehicles is increasing in the network. We can observe that PDR of the proposed BC-VANET method, on the other hand, was greater, with a much steadier decline as the number of nodes increased, because the proposed solution required a less intensive computing activity to execute the algorithm as compare to others. This resulted in a decrease in packet transmission delay, resulting in a higher PDR.

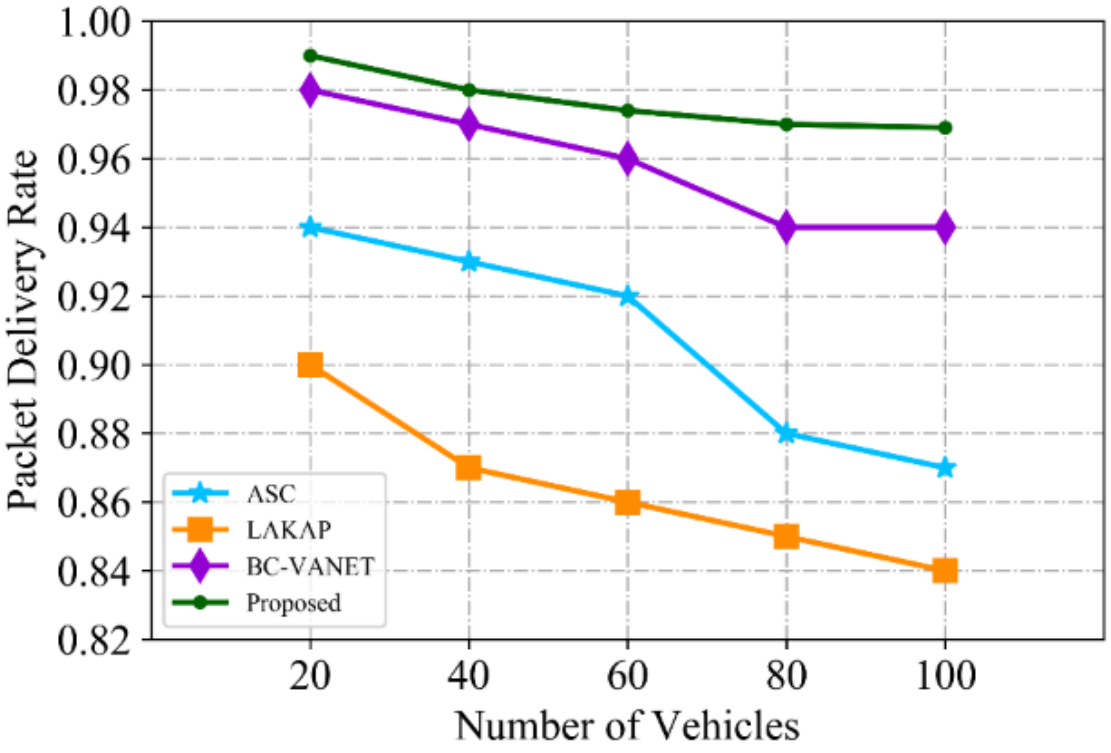


Figure 6.1: Packet Delivery Rate without DoS Attack

6.1.3 Packet Loss Without Denial-of-Service attack

The ratio of packets that never made it to their target vehicle to all of the packets that were sent from the source vehicle is known as packet loss. It can be represented mathematically as a formula (6.3) .

$$PLR = \frac{R_p}{S_p} \tag{6.3}$$

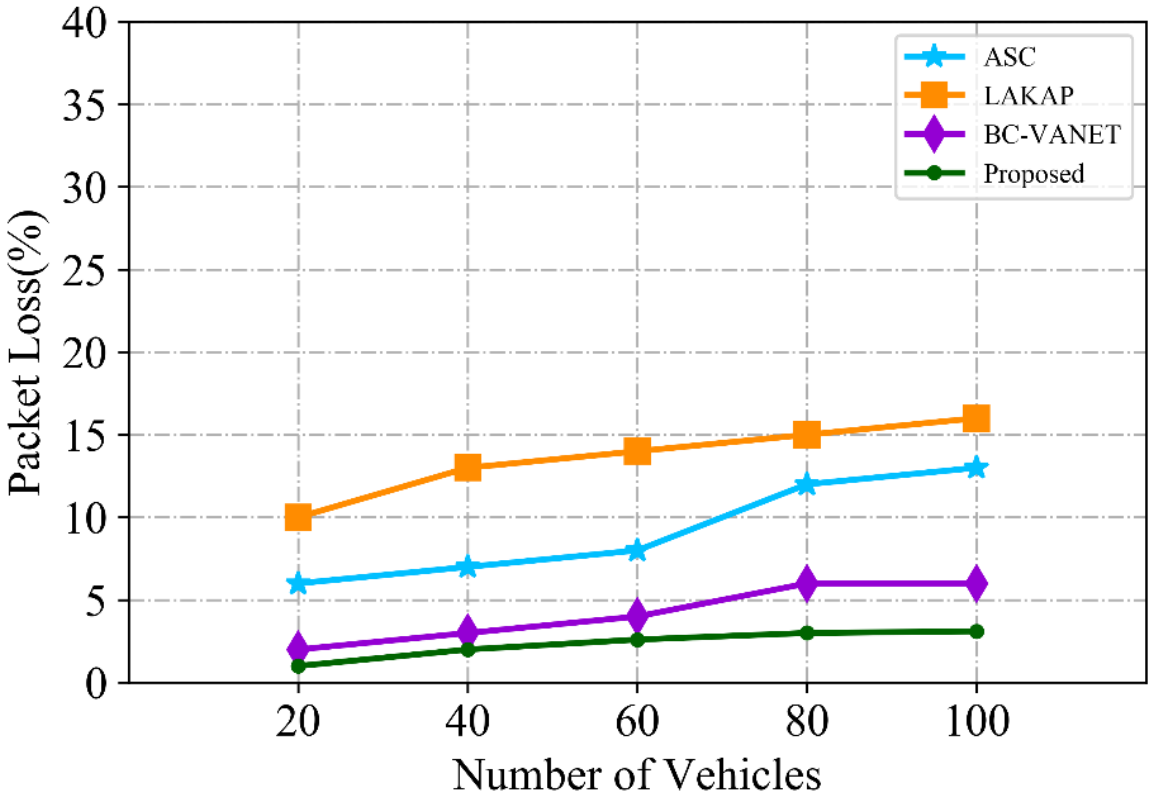


Figure 6.2: Packet loss without DoS Attack

Where S_p the total packets send by the source vehicle and R_p the is the complete number of packets gathered by the target vehicle. Figure 6.2 shows the packet loss is lower when there are fewer vehicles in the network and increases slightly when the number of vehicles increased. When the simulation conducted and number of vehicles is 20 [119] packet loss is about 6% and rise to 13% when the number of vehicles reached 100. Moreover, packet loss ratio is approximately 10% when there is 20 vehicles in the simulation in relation to [120]. In contrast, the proposed model has the lowest packet loss rate about 2% at the beginning of the simulation with 20 vehicles in the network and rise to 4% with 100 vehicles in the network.

6.1.4 PDR with Denial-of-Service attack

In the simulation, a denial-of-service attack was conducted to analyse the packet delivery ratio. Figure 6.3 shows the effect on PDR when there is an attacker in the network. The proposed model has high PDR of 0.96, however, BC-VANET [96], ASC [119] and LAKAP [120] have PDRs of 0.94, 0.75, and 0.70 respectively. We can observe that the highest PDR is 0.96 with an attacker present in the network; however, this drops to 0.92 as the number of vehicles in the network increase. Due to a less, intensive computing operation being required to run the

algorithms in the proposed technique, the PDR was higher and decreased more steadily as the number of nodes rose. PDR increased as a result of a reduction in packet transmission delay.

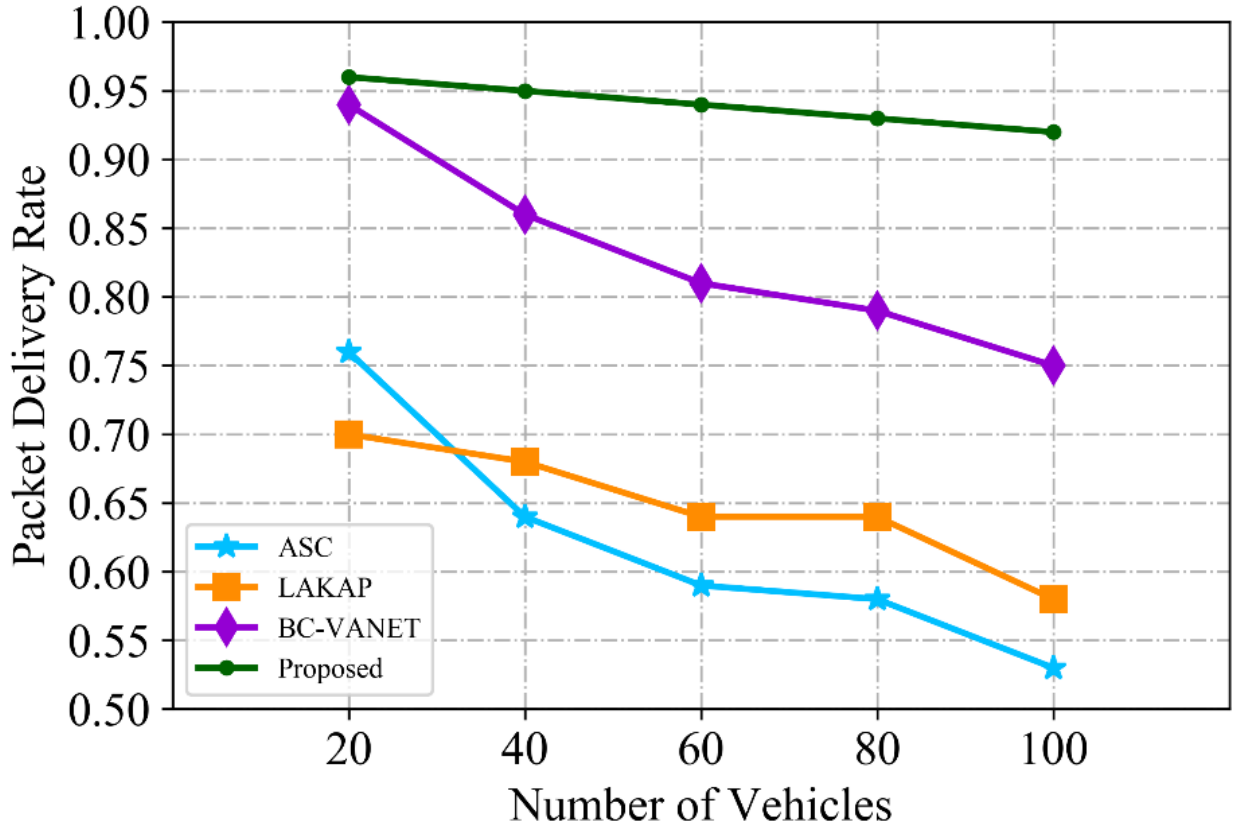


Figure 6.3: Packet Delivery Rate with DoS Attack

6.1.5 Packet Loss with Denial-of-Service attack

In the simulation, a denial-of-service attack was conducted to examine the packet loss ratio. The impact on PLR when there is an attacker in the network is depicted in Figure 6.4. The proposed solution kept packet loss at a very low level, in comparison to other solutions. When the number of vehicles were less, the packet loss was low and slightly increased when the number of vehicles increased. However, the proposed biometric based blockchain model has the lowest packet loss ratio due to the effective algorithms which protect against DoS attack compared to BC-VANET [96], ASC [119] and LAKAP [120].

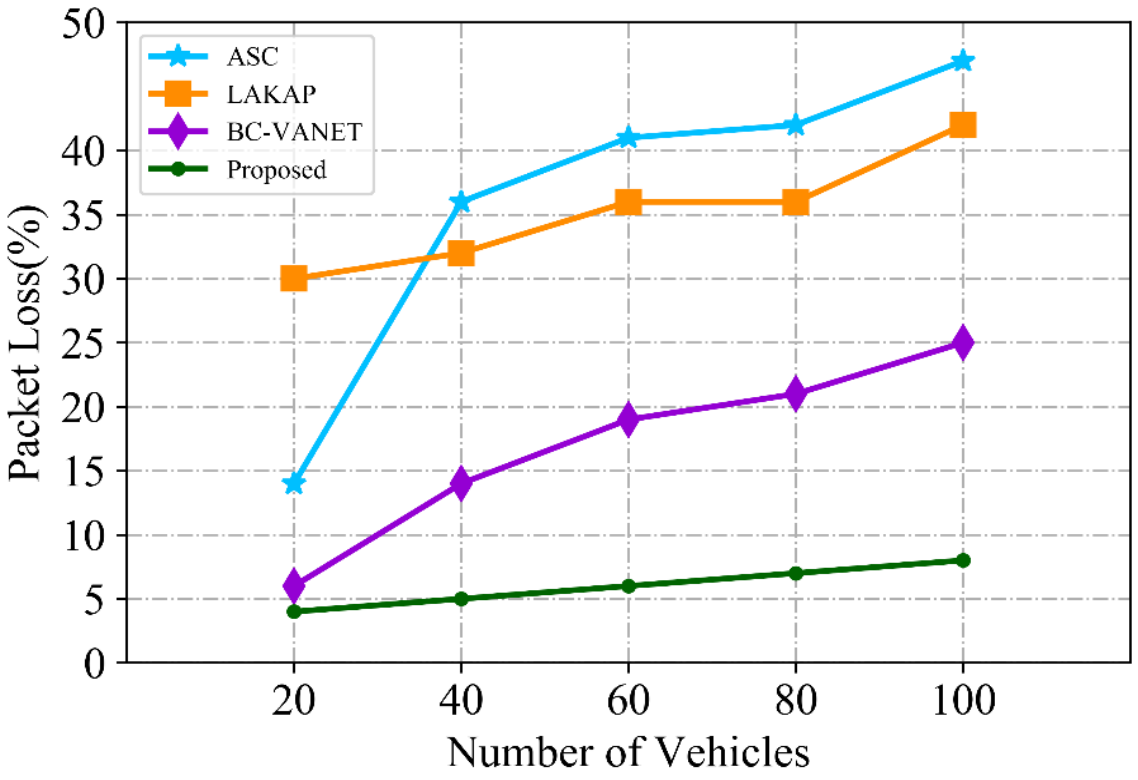


Figure 6.4: Packet Loss Rate with DoS Attack

6.1.6 Computational Cost

The proposed model has less computational cost as compared with BC-VANET [96], ASC [119] and LAKAP [120]. In the beginning with 20 vehicles the proposed model has a time cost of 0.1ms while with BC-VANET [96], ASC [119] and LAKAP [120] the cost is 0.13, 2.8 and 4.0 respectively as shown in figure 6.5 . As the number of vehicles in the network increases from 20 to 100, the computational cost also increases from 0.1ms to 0.3ms, however, this is still very low compared with existing approaches. All approaches displayed a rise in computational cost during execution phase as the number of nodes rose from 40 to 90. This resulted from network's adoption of new technologies such as big data technology with a feature of low latency. Considering 5G and 6G technologies can support extremely high concurrent connection and offer dependable data transfer to support large data collecting services, the computational cost of the proposed model decreased as the number of vehicles increased. Consequently, the complexity of computing processes was reduced.

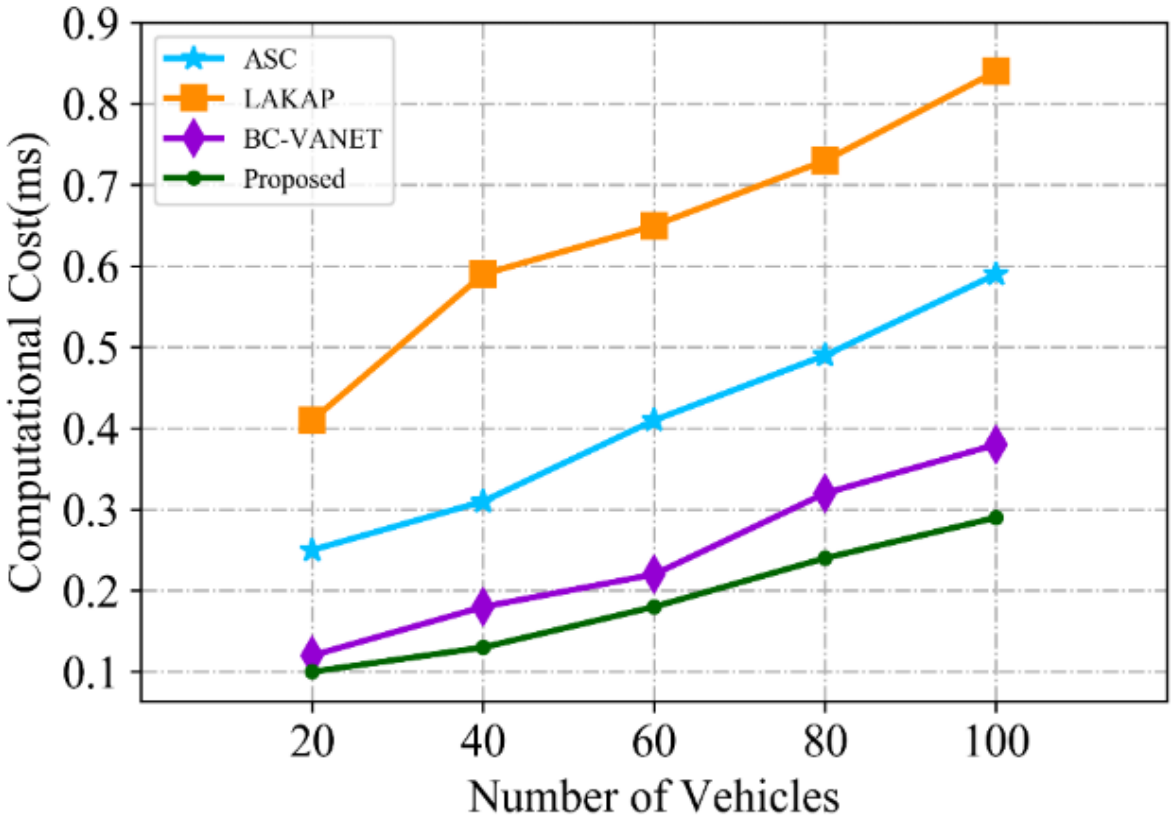


Figure 6.5: Computational cost of the proposed model

6.1.7 Security Analysis

We analyze the security of our proposed model which meets the security criteria.

a) Secure Registration

The registration of the vehicle ensures security using public and private keys. The keys are stored in the OBU which is a tamper proof memory. The vehicle is registered only when the registration data is verified with MVD. The data is sent to network by signing with private key of the vehicle as shown in equation (6.3).

$$Mi=Sign_{KR_i}(KUi, Vi, Di||\mathcal{H}(KUi \parallel Vi \parallel Di)) \quad (6.3)$$

The length of the RSA signature key is 1024 bit long. The signature is a bit of string of $\lceil log_2 N \rceil$-1 bits that provides strong security.

b) Data Integrity

We can denote the cost of authentication as T_a which has H as hash function, signature and match functions. Additionally, E_c represents the encryption cost. The T_{hash} can be evaluated using the hash function. The vehicle authentication process takes 2 * T_{hash} and

further time t_δ to compare hash functions $H_1$and H_2. Since, t_δ is small and therefore it can be neglected. The cost involved in vehicle authentication for hashing will be $2 * H$ while the encryption cost can be derived by adding the hash cost as $1 * H + E$. The T_a may be derived by combining the encryption, decryption and verification (Hashing) cost as expressed in equation (6.4):

$$T_a = 2 * H + E + D \tag{6.4}$$

c) Privacy preservation and Traceability

The identify of a vehicle should not be disclosed to anyone. The messages must be absolutely undetectable to the RSU or other drivers in order to function properly. We accomplished this by employing biometrics to create a unique ID and then taking the hash (SHA-512), every vehicle has a unique 512-bit pseudo identity. The MVD and TA can only determine the real identify of a vehicle by examining the blockchain-based database. Traceability will be required if a vehicle generates a fraudulent message or engages in any malicious action.

6.1.8 Summary of Results

The presented results demonstrate that the framework can protect the security of messages within VANET in real-world scenarios. The proposed framework not only provide security and trustworthiness of the communication between vehicles, but also keep anonymity without exposing the original identity of authorized users. Additionally, biometric is combined with blockchain technology to provide reliable transmission of data, keep track of data being exchanged and identify the responsible vehicle in the case of false messages. The simulation under the OMNET++, veins and SUMO is carried out to demonstrate the viability of the proposed framework. The performance of the framework is evaluated in terms of packet delivery rate, packet loss rate and computational cost. Therefore, the obtained results reveal that the proposed model is superior in comparison with existing approaches. As a part of future work, we will extend the model for computing ranking and reputation of vehicles and drivers using machine learning techniques.

6.2 Reputation Computations

The computation of trust of any vehicle is solely dependent on the previous reputation and some temporal and behaviour parameters. The reputation is a key requirement as it is indicating the level to which a peer vehicle would be trusted, secure, or reliable in any contact with other vehicle. As a result of this unique requirement for reputation assessment, a high-standard and effective reputation characterization and monitoring system is needed. Hence, our model offers a different approach which is based on participation degree, vehicle age, and computation of vehicle reputation. The reputation $R_{i,j}^{(n)}$ of the vehicle i computed by vehicle j for the n^{th} message at time t is the correctness of the message C_t^i, age α, participation degree θ, the smoothing coefficient β, and existing reputation $R_{i,j}^{(n-1)}$ which can be computed by equation (6.5).

$$R_{i,j}^{(t)} = \begin{cases} \beta R_{i,j}^{(t-1)} + (1-\beta)\big((\alpha + \theta + C_t^i)/3\big), & \alpha > 0 \\ (\beta)\big((\alpha + \theta + C_t^i)/3\big) & \alpha = 0 \end{cases} \tag{6.5}$$

$$\text{Where} \begin{cases} \alpha = \big((t - t_0)/((t - t_{Reg}) + (t - t_0)\big) \\ \theta = \frac{N^+}{N}, 0 \le \theta \le 1 \\ N = N^+ + N^- \\ N^+ = \sum_{k=1}^{n} M_k, N^- = \sum_{l=1}^{n} M_l, N > 0 \\ C_t^i = f_c(M), C_t^i \in [0,1] \end{cases} \quad N = \text{total number of messages}$$

$$f_c(M) = (S(M_i, M_j).w_1 + f_f(M_i).w_2 + f_d(M_i, M_j)).w_3)/3 \tag{6.6}$$

Data which is generated for VANET simulation contains most of the typical network parameters such as packet length, TTL, total forwarded packets, total backward packets, failure type, option type, road condition, speed, weather, time scenario, lane type, traffic scenario, packet type, and location etc as shown in Tab. 6.3. All the 24 features describe the scenario of a vehicle that constitutes spatial $\mathbb{S}$, temporal $\mathbb{T}$ and behavior $\mathbb{B}$ parameters. The dataset defines the important numerical attributes which participate in reputation computation are number of packets forwarded, the average size of packet, the time when message has originated, message communication type and status, time to live in the network and the port number.

Table 6.3. Dataset description

Feature Name	Description
source	IP address of the source vehicle
destination	IP address of the destination vehicle
detection_target	IP address of the detection node
destination_port	Destination port address
Total_Fwd_Packets	Total number of forwarded packets
Total_Bkwd_Packets	Total number of backward packets
Total_Length_of_Fwd_Packets	Size of the forwarded packets
Total_length_of_Bkwd_Packets	Size of backward packets
Flow_Packet_Per_Second	Flow per second of packet
Average_packet_Size	Average size of the packet
Time_Stamp	Time when packet originated
TTL	Time to live
Reputation	Reputation of the sender
OT(Type of communication)	0: default, 1 request, 2 reply, 3 transmission result
Failure	0: no malicious behavior, 1: malicious behavior), 2: Failure caused by non-malicious behaviours. 3: transmission result
Road_Condition	'Dry', 'wet', 'Icy'
Speed Scenario	'Accelerating', 'constant', 'Deaccelerating'
Time_Scenario	'Dawan', 'Day','Dusk', 'Night'
Weather Scenario	'Clear', 'Foggy', 'Raining', 'Snowing', 'Windy'
Lane_Type	'Winding', 'UpHill', 'Straight', 'Intersection', 'Curve', 'DownHill'
Traffic_Scenario	'Car on 1 side', 'Car on 2 sides', 'Car on 3 sides', 'Car on 4 sides', 'No cars'
Packet_Type	'General', 'Safety', 'Traffic'
Latitude, Longitude	Position of vehicle

The source in table 6.3 contain the IP address of the vehicle which is generating the data or requesting for any service. Similarly, the destination field contains the IP address of the destination vehicle. The detection_target contain the IP address for the node to be detected for the computing the trust. The destination port at which the data or service is available. The Total_Fwd_Packets field will contain the number of packets that has been forwarded by the vehicle in particular session. Similarly, Total_Bkwd_Packets will not contain number of backward packets. The Total_Length_of_Fwd_Packets field contains the size of the forwarded packets while Total_length_of_Bkwd_Packets will contain the size of backward packet. The field Flow_Packet_Per_Second the rate of packet arrival or departure in the network. The Average_packet_Size refers to the average length of the packet in the network. The Time_Stamp refers when the packet was sent. The time to live or the life of the packet in the network is represented by TTL. The computed value of the reputation is represented by the field 'Reputation'. The option type such as (0: default, 1 request, 2 reply, 3 transmission result) is represented by OT. There is different type of failure such as (0: normal behaviour, malicious behaviour, 2: Failure caused by a legitimate activity. 3: transmission result is measure by the field called 'failure'. In the VANET, the road condition is also an important parameter that affect the quality of service. The road condition could be 'Dry', 'wet', 'Icy'. The Speed Scenario at any time could be 'Accelerating', 'constant', 'Deaccelerating' while the timing scenario is recorded as 'Dawan', 'Day','Dusk', 'Night'. The 'Weather Scenario' can attain different values such as 'Clear', 'Foggy', 'Raining', 'Snowing', 'Windy'. The Lane_Type at different location could have different types such as 'Winding', 'UpHill', 'Straight', 'Intersection', 'Curve', 'DownHill'. The Traffic_Scenario is captured among one type from the 'Winding', 'UpHill', 'Straight', 'Intersection', 'Curve', 'DownHill'. The packet in network could be of three types such as 'General', 'Safety', 'Traffic'. The position of the vehicle is indicated by capturing Latitude, Longitude.

All the describe 24 features are used in the proposed model to compute the reputation and trust of a vehicle. The table 6.4 present sample output generated after computation of the reputation on the dataset. This table contains vehicle IP address, port number, time stamp, Time-To-Live (TTL) and computation value of reputation. From the table 6.4 it can be seen that reputation has continuous value between 0 and 1.

Table 6.4. Sample output of the reputation computation

IP address	Port	Time Stamp	TTL	Reputation
20.100.168.125	4244	Thu Jan 27 02:32:40 2022	147	0.947461244
20.100.168.34	2624	Thu Jan 27 02:32:40 2022	86	0.61587248
20.100.168.177	1064	Thu Jan 27 02:32:40 2022	116	0.656448565
20.100.168.187	1892	Thu Jan 27 02:32:40 2022	221	0.593758417
20.100.168.210	3417	Thu Jan 27 02:32:40 2022	213	0.995342545
20.100.168.95	1260	Thu Jan 27 02:32:40 2022	218	0.615054778
20.100.168.228	2221	Thu Jan 27 02:32:40 2022	162	0.834318201
20.100.168.171	4892	Thu Jan 27 02:32:40 2022	171	0.443103195
20.100.168.4	1329	Thu Jan 27 02:32:40 2022	242	0.921932154
20.100.168.11	3739	Thu Jan 27 02:32:40 2022	202	0.897358896
20.100.168.105	3951	Thu Jan 27 02:32:40 2022	255	0.703708723
20.100.168.236	2959	Thu Jan 27 02:32:40 2022	138	0.221071998
20.100.168.90	1316	Thu Jan 27 02:32:40 2022	161	0.913946538
20.100.168.53	4766	Thu Jan 27 02:32:40 2022	111	0.724440567
20.100.168.225	4510	Thu Jan 27 02:32:40 2022	112	0.324995418
20.100.168.230	2240	Thu Jan 27 02:32:40 2022	171	0.444987762
20.100.168.192	2768	Thu Jan 27 02:32:40 2022	104	0.679105299
20.100.168.52	1064	Thu Jan 27 02:32:40 2022	218	0.361281827
20.100.168.167	3071	Thu Jan 27 02:32:40 2022	173	0.297529623
20.100.168.114	3443	Thu Jan 27 02:32:40 2022	220	0.438722483
20.100.168.115	2145	Thu Jan 27 02:32:40 2022	70	0.780719787
20.100.168.223	2311	Thu Jan 27 02:32:40 2022	184	0.519106395
20.100.168.33	3110	Thu Jan 27 02:32:40 2022	243	0.097114853

6.3 Ensemble Machine Learning for Trust Classification

Trust, data accuracy, and dependability of data being broadcasted via the communication channel are the primary challenges in VANET. In this section, we will show the results obtained from the distributed trust model and reputation system for vehicle networks. The results evaluate the received messages, calculate the vehicle reputation and the message correctness based in numerous factors. The new reputation credit will be stored in the blockchain. Comprehensive tests are carried out using the dataset in order to measure the performance of proposed ensemble learning and feature selection based random forest.

6.3.1 Experimental Setup

As the experimental dataset for demonstrating the viability of our model, we have generated simulated data for the VANET which has 24 attributes, a typical trust detection dataset, which we consider to be a good fit. A detailed description of the dataset is shown in Table 3 The total number of generated 1000000. The non-trustworthy rate is 2.08661%, according to the data available. For the purpose of normalizing all the features, we employ **the min-max normalization technique which is one of the most often used methods of data normalization.** Specifically, the lowest value of each characteristic is turned into a 0, the highest value is changed into a 1, and all other values are transformed into a decimal between 0 and 1. The simulation has been carried out on an Intel Pentium processor running the Python 3.6 programming language (Windows 10 operating system, 2.6GHz Intel Core i7 processor, 16.0GB RAM). Minor operations were performed on the datasets, such as data combining, empty variable elimination, removing irrelevant attributes, and new attribute mappings, to make them more suitable for classification. A description of sampled dataset is shown in Table 6.4 and Figure 6.6 that contains 39567 trustworthy instance and 6260 as non-trustworthy instances.

Table 6.5. Sampled Dataset

Class Label	Number of Instances	
	Original	Sampled
Trustworthy	791339	39567
Non-Trustworthy	208661	6260

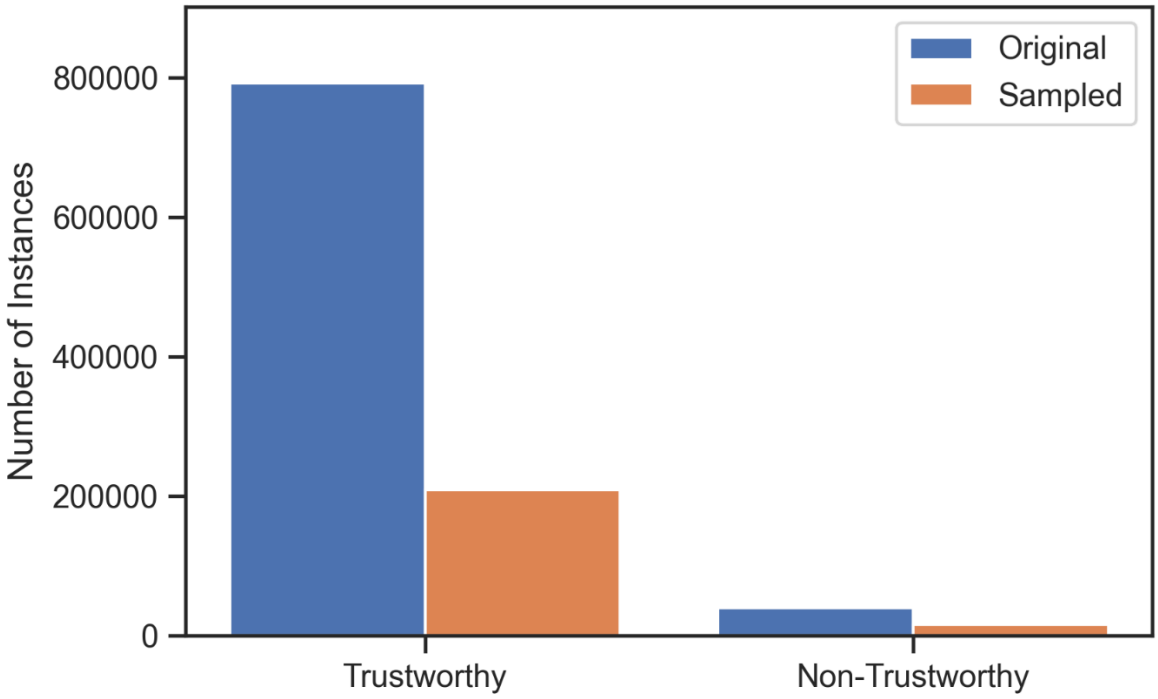


Figure 6.6. Sampled data from Dataset

6.3.2 Evaluation Metrics

In this section, we compare our proposed feature selection-based random forest approach to the baseline decision tree and Xgboost algorithms, as well as to other existing methods. Feature selection based random forest is evaluated using two measures, which are as follows:

(1) AUC (Area Under Curve): AUC is an abbreviation for Area Under Curve. It is used **to represent the accuracy degree** of feature selection-based random forest selection. In this case, the area between the ROC curve and the coordinate axes is equal to the value. Greater correctness of our suggestion is indicated by a higher value of the metric.

(2) Confusion metrics: A confusion metrics is used to present the predicted Vs actual classification

(3) Time measurement: It is used to assess the effectiveness of feature selection based random forest model for the proposed trust classifications. For purposes of this article, the processing time of a single data instance is defined as the sum of the time spent on detection of non-trustworthiness. The lower value indicates the efficiency of the model as compared to others.

6.3.3 Baselines

For the purpose of evaluating the performance of our approach, we selected three classification methods for detecting non-trustworthiness as baselines.

Decision Tree algorithm: The decision tree is an essential classification tool that is focused on the split and conquer strategy. It is composed of deciding nodes and leaf nodes, which represent a choice check one of the attributes and a classification of the outcome class, respectively.

Random Forest*:* Random Forest is an ensemble learning classification technique that uses the consensus voting law to select the class with the highest polls from decision trees as the categorization outcome.

XGBoost algorithm*:* is an ensemble learning technique that uses the gradient descent approach to integrate numerous decision trees in order to increase speed and performance.

Ensemble learning and feature selection*:* An ensemble feature selection (FS) approach is used to boost the confidence in the selected features by computing the average of the characteristic priority sets provided through the four chosen tree-based machine learning techniques. Ensemble learning is a strategy that combines two or more ML algorithms to achieve better results than when the algorithms are employed individually. Rather than depending on a single model, the predictions from several models are merged utilizing a combination strategy to get a single more accurate prediction. In this work we have applied FS on DT, XGBoost and random forest. However, we found the random forest to be the best among these.

6.3.4 Experimental Results and Analyses

Random forest performance is evaluated in terms of accuracy and efficiency, and the results are confirmed using ensemble learning and feature selection. The following are the experimental findings and analyses.

a) *Area Under Curve Analysis:* The AUC values for decision tree, XGBoost, and FS-RF ensemble learning and comparison approaches are plotted in Figure 6.7 with respect to the subset size.

The subset is randomly chosen from the dataset, with subset sizes of 4000, 8000, 12000, and 16000. These numbers are derived from the values of parameters described in previously

published papers as well as from our own experimental experiments. While the parameter accepts values within the specified ranges, trial findings indicate that FS-RF performance is superior and considerably varies. This can see from the Figure 6.7, the FS-RF algorithm produces a higher AUC value when the subset size is varied. XGBoost has a somewhat lower AUC but higher than decision tree. Also, from the Figure 6.8 the true positive rate and false positive rate. The AUC obtained is 99.98%.

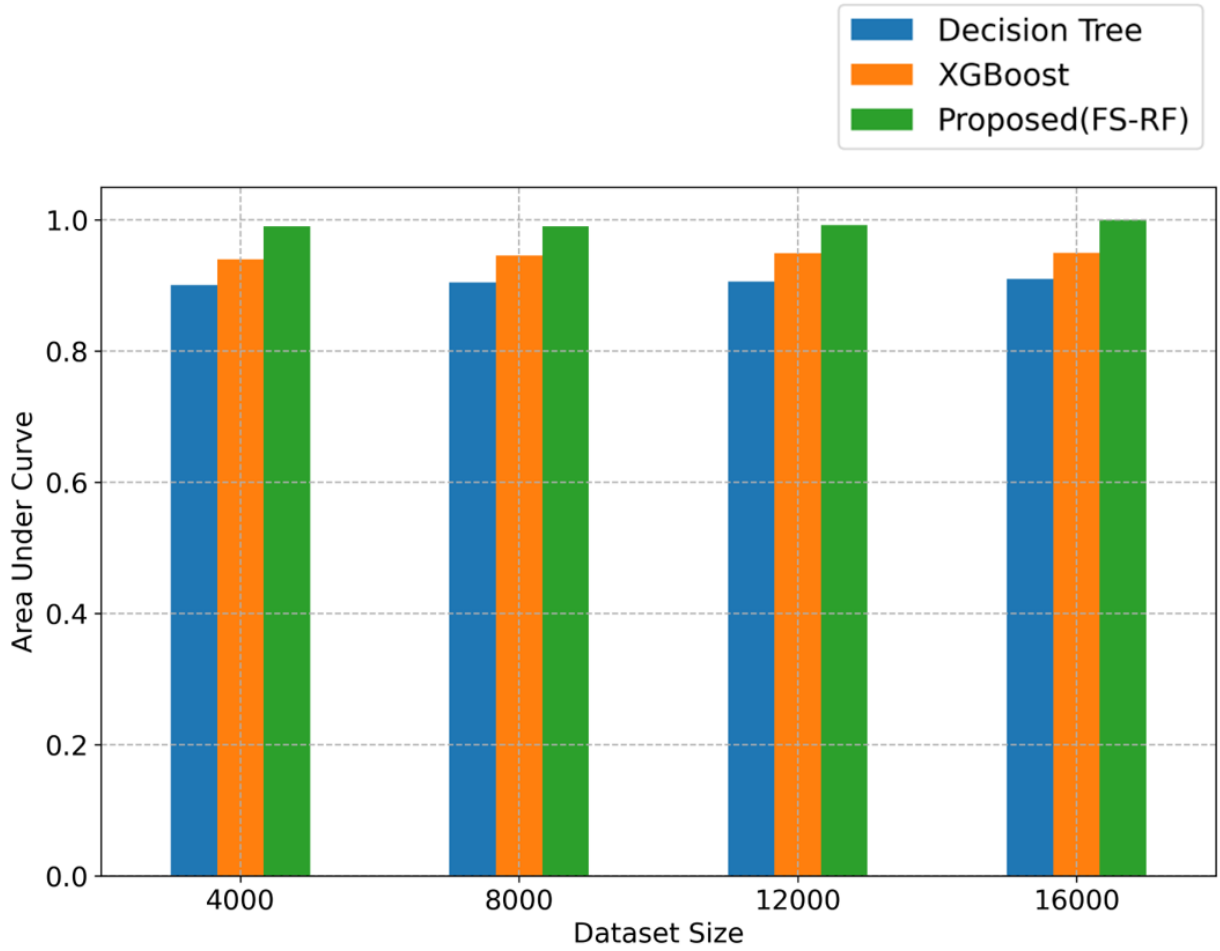


Figure 6.7: AUC comparision of different approaches with varying size dataset

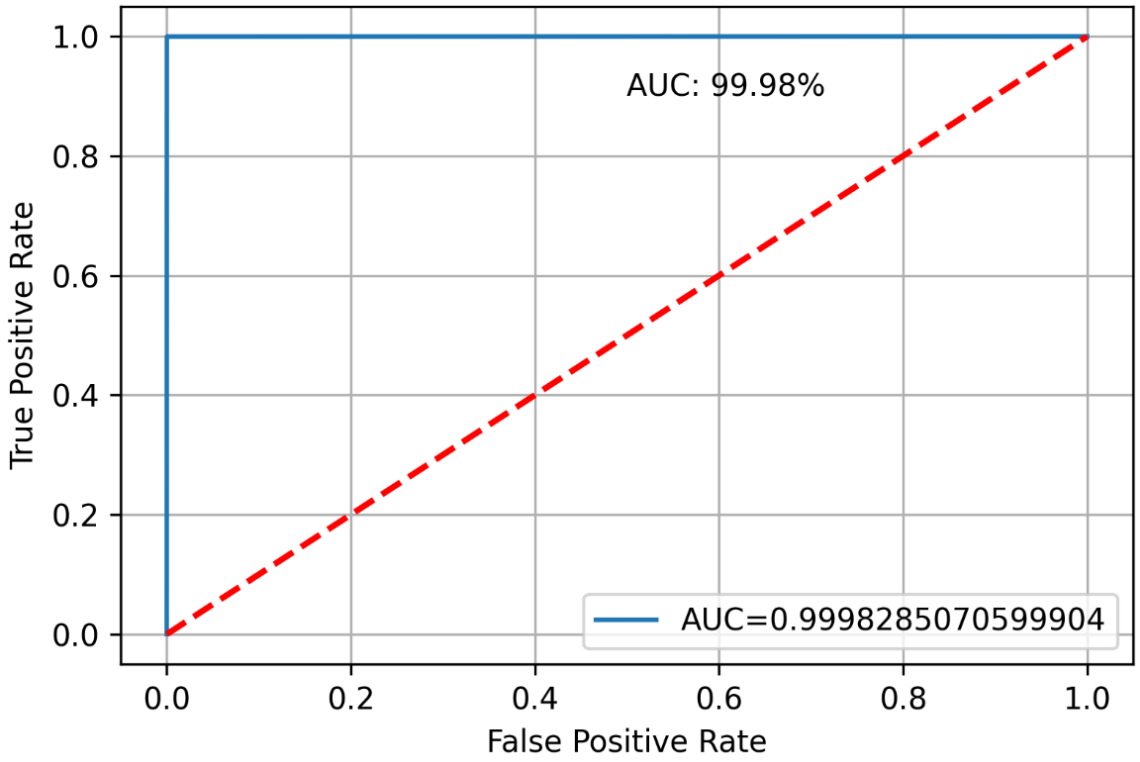


Figure 6.8: ROC for Radom Forest with feature selection

b) *Time Analysis:* The time cost of each technique is presented in Figure 6.9 in relation to the subset size. The subset size ranges from 4000 to 16000. The decsion tree time vaires from 1.5 seconds to 3.8 seconds while as the XGBoost varies from 1.1 seconds to 3.8 second. From the Figure 6.9, we can see, proposed FS-RF is more efficient than the baselines in this comparison which has a time cost 1.1 second to 1.7 seconds. **The time varies due to the fact that our solution integrates the ensembling learning and feature seelction mechanisms, both of which have the potential to significantly lower the time cost.**

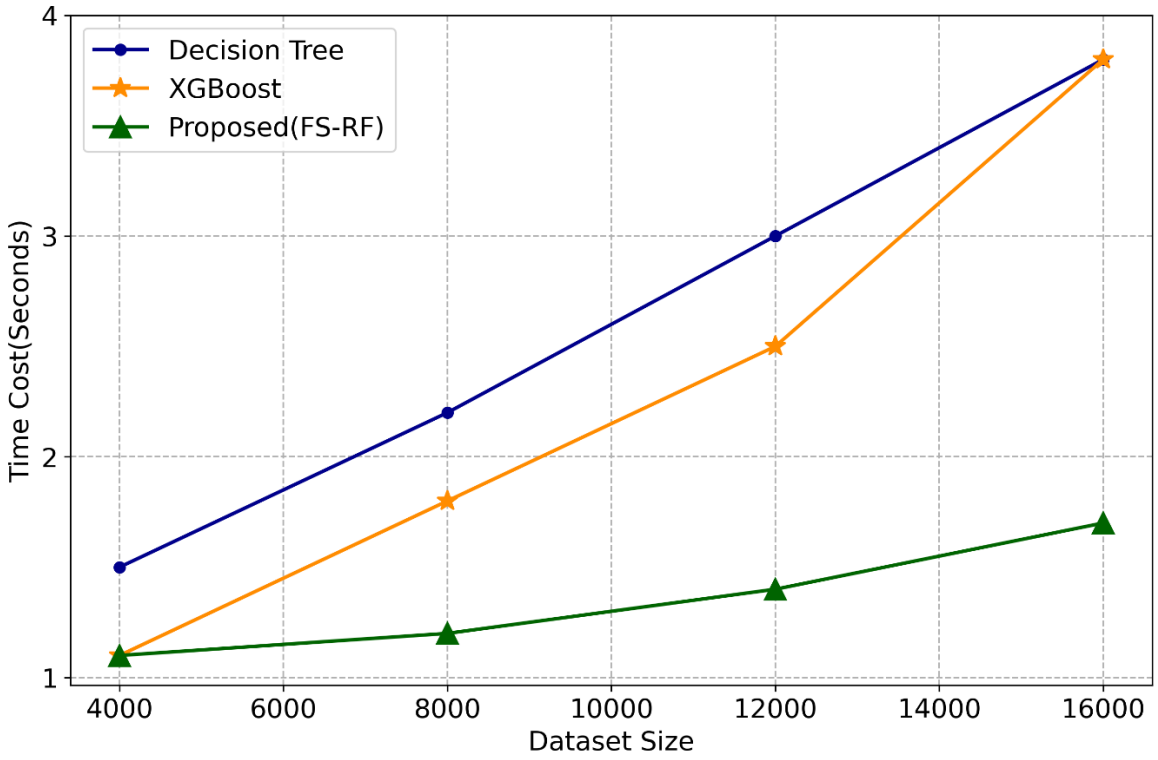


Figure 6.9: Efficiency comparison(time) of different approaches with varying size dataset

c) *Feature Selection Analysis:* The suggested feature selection approach was evaluated on subgroups to determine its effectiveness in examining the attributes. Table 6.6 contains a list of the top four most essential attributes, as well as the appropriate weights assigned to each feature. As shown in Table 6.6, the reputation has the value of 0.468 which is the highest weight that is used to indicate a trustworthy and non-trustworthy. The lowest value in feature selection is 0.0106 for TTL. Additionally, the average length of the packets is key characteristic. For instance, the larger packet size, shows untrustworthiness vehicle. The value of option type and failure type are other parameters to indicate the trust.

Table 6.6. Feature score

Reputation	0.468
Average_packet_Size'	0.3733
OT	0.0201
Failure_Type'	0.0184
Total_Bkwd_Packets'	0.0151
Total_length_of_Bkwd_Packets'	0.015
Destination port'	0.0148
Latitude	0.0142
Longitude	0.0116
Total_Fwd_Packets'	0.0133
Flow_Packet_Per_Second'	0.013
Total_Length_of_Fwd_Packets'	0.0126
TTL	0.0106

As per the Table 6.7, when the evaluation is performed on the dataset, the ensemble learning with **feature selection based random forest** produces better results in comparison to baseline method such as decision tree, XGBoost and random forest. The proposed model (FS-RF) has achieved the accuracy of 99.981 while as decision tree, XGBoost, RF, FS-DT, FS-XGBoost has achieved 98.128, 99.035, 99.835, 99.398, and 99.796 respectively. Since decision tree has the weakest correctness and the longest processing duration machine learning models, therefore, XGBoost and random forest, were chosen for inclusion in the stacking ensemble model, and the single model with the greatest efficiency. FS-RF, was chosen to be the meta-classifier in the second layer. Because of the use of stacking to merge these models, the accuracy, F1, and precision reach 99.98 %, indicating that all of the trained untrustworthy vehicles can be identified.

Table 6.7. Performance Evaluation metrics

Model	Accuracy (%)	Precision (%)	Recall (%)	F1(%)
Decision Tree	98.128	98.350	98.128	98.177
XGBoost	99.035	99.096	99.035	99.045
RF	99.835	99.837	99.835	99.835
FS DT	99.398	99.424	99.398	99.404
FS XGBoost	99.796	99.799	99.796	99.796
FS RF	99.981	99.980	99.97	99.98

Accuracy: Accuracy is a metric for classification models that measures the number of correct predictions.

Precision*:* Proportion of positive(trustworthy) identifications that was actually correct.

Recall*:* proportion of all actual positives was identified correctly

F1-score: is a metric which takes into account both *precision* and *recall*

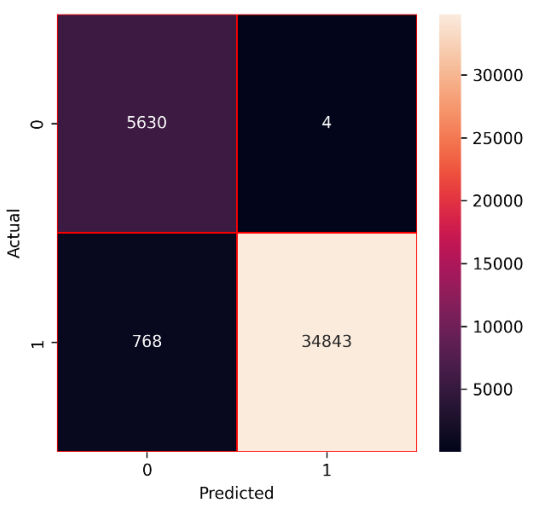
Figure 6.8 represents confusion matrix for the baseline model (a) – (c) and ensemble learning with feature selection (d) – (f). From the Figure 6.8(f) this can be seen that 35603(This is true positive) has been correctly classified as trustworthy and 5634 (This is true negative) has been classified as non-trustworthy while 8 (This is false negative) instances have been misclassified out of total 90% testing instances (41245).

TP $(True\ Positive)$ – The data was actually trustworthy and it has been classified as trustworthy by the algorithm

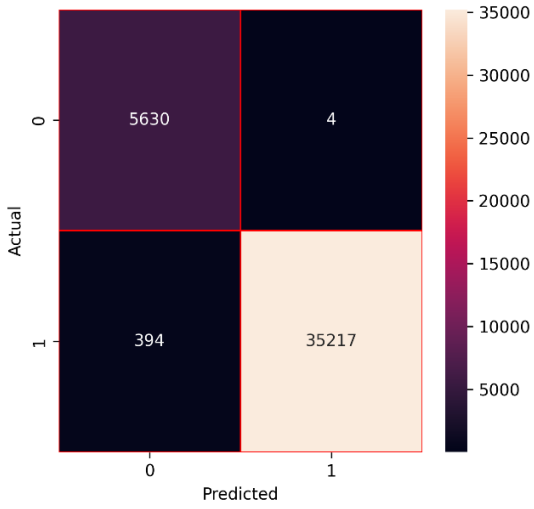
TN (True Negative) – The data was actually non-trustworthy and it has been predicted as non-trustworthy.

FN (False Negative) – The data was actually non-trustworthy and it has been classified as trustworthy

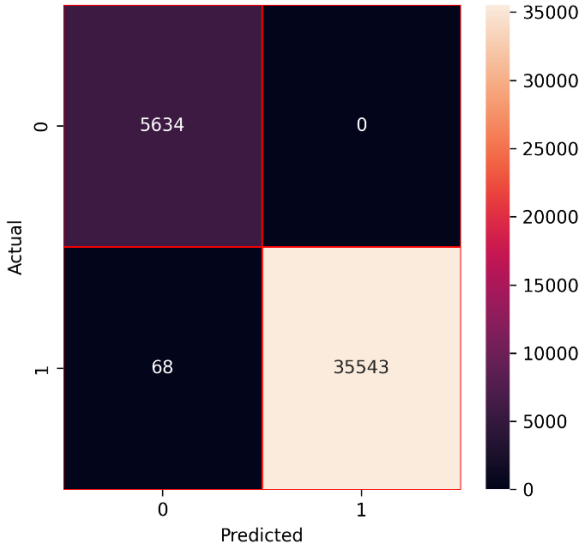
FP $(False\ positive)$ – The data was actually trustworthy and it has been classified as non-trustworthy by the algorithm



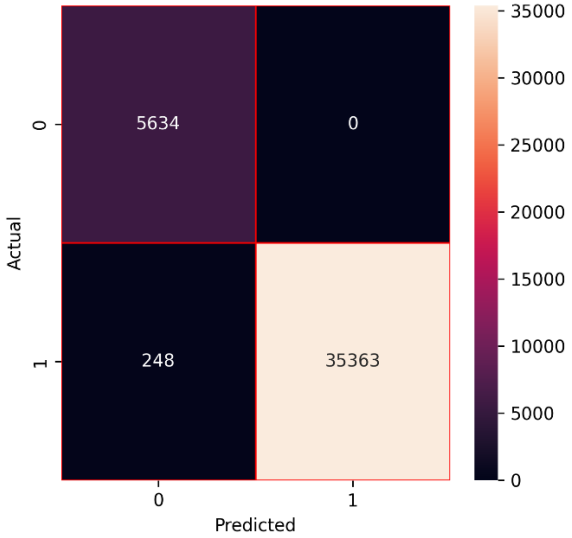
(a) Decision Tree



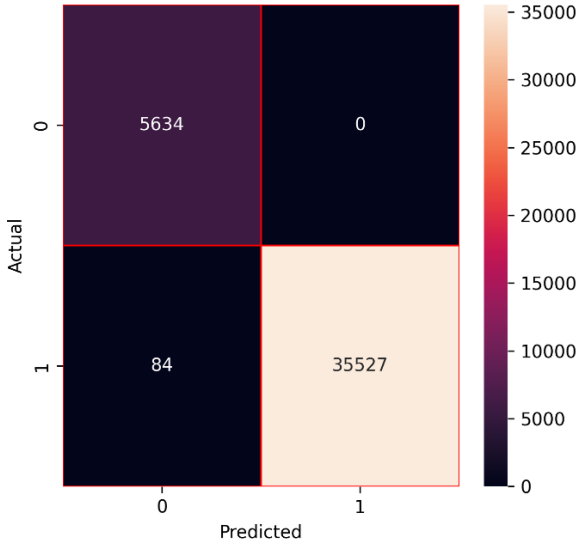
(b) XGBoost



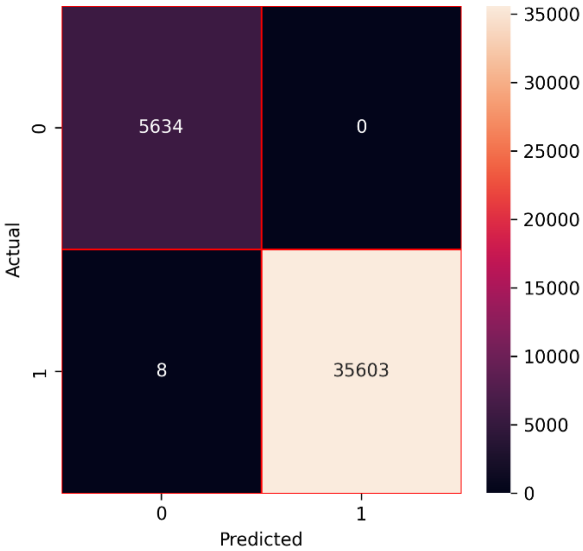
(c)Random Forest



d) Decesion Tree with feature selection



(e) XGBoost with feature selection



(f) Radom Forest with feature selection

Figure 6.8: Confusion matrix of different models

6.4 Chapter Summary

The simulation and findings were presented in this chapter are intended to validate the validity and efficacy of the proposed model. The **section 6.1** presented the results obtained from the privacy preservation by employing anonymity of vehicles. This section discussed that the authorities shall be able to track the anonymity to identify the malicious vehicle, thus it should be made conditional. The vehicle's genuine identity must not be visible to other vehicles or RSUs and it should not be possible for a malicious actor to acquire identities through the analysis of any identity intercept. The **section 6.1.1** presents discussion and simulation on the VANET tools such a OMNeT++, Veins and SUMO to demonstrate the correctness of the proposed model using IEEE 802.11p/1609.4 protocols. **Section 6.1.2** present the effect of packet delivery rate and packet loss with or without Denial-of-Service attack. This has been observed that the proposed model has the lowest packet loss and highest PDR as compared to BC-VANET [3], ASC [42] and LAKAP [43]. The section **6.1.6** presents the computational cost of the proposed model which is less than the existing models. Thereafter, section **6.1.7** presents the security analysis by employing the secure registration, data integrity, privacy preservation and traceability. Finally, the summary of the section is presented in section **6.1.8.**

The computation of reputation is presented in **section 6.2**. This section presents the 24 features of the dataset in the table 6.3. The table 6.4 illustrate the sample output generated after computation of the reputation on the dataset.

The **section 6.3** presents the result obtained from the ensemble machine learning for the trust classification. The evaluation is based on the AUC, confusion metrics, and time measurement for the baseline method *Ensemble learning and feature selection:* An ensemble feature selection (FS) approach is used to boost the confidence in the selected features by computing the average of the characteristic priority sets provided through the four chosen tree-based machine learning techniques. In this work we have applied FS on DT, XGBoost and random forest. However, we found the random forest to be the best among these while using feature selection ensemble technique.

Chapter 7

Conclusion and Future Work

VANET is the state-of-the-art technology in the domain of ITS, where vehicles communicate with each other and adjacent roadside unit (RSU) to partially solve transportation issues such as reducing traffic accidents, better traffic management, minimizing traffic congestions, and providing infotainment to on-board vehicular users. In this chapter we present the summary of each chapter and future work.

7.1 Summary of the Thesis

Chapter 1, introduces this thesis by providing a brief overview and background of the thesis. This has been highlighted that over a period 10 years the number of intelligent vehicles in mobile ad-hoc networks has raised will reach 2 billion. In order to deal with this enormous number of automobiles, a VANET was established. Moreover, it presents the aim of the research problem in vehicular communications. To overcome the security and safety obstacles, there must be urgent improvements in privacy and trust levels to guarantee that all communication processes are secure and have integrity. A secure and reliable decentralized data sharing system must be created to guarantee that VANETs can continue operating normally. This is the motivation if the thesis which has been discussed. The aim and objective of the thesis addresses the privacy perseveration and trust computation framework in VANET. First objective is to devise biometric blockchain (BBC) framework that provide a decentralized, secure, and trusted communication environment in VANET. To accomplish this ambition, the single registration, message authentication, privacy preservation, and traceability has been considered.

The trust and reputation system in VANET is the second objective of this study. The trust and reputation model ensures non-repudiation and establishes a reputation for the vehicles based on the accuracy of the sent messages and the reputation credit that is preserved. Using a trust management system based on reputation and identity evaluation, it may be possible to reward trustworthy automobiles while reporting untrustworthy vehicles on VANETs, ensuring trustworthy message broadcasting. Classifying trust data into trustworthy and non-trustworthy vehicles is the primary goal of this research.

The **chapter 2,** presents background a systematic review of the existing work on VANET's security and trust. Moreover, it presents the architecture of the VANET that are characterized by mobility, diverse driver behavior, highly dynamic topology, multi-hop communications, as well as strict security and privacy requirements due to the hostile environments in which they operate. Additionally, it presents a discussion about intelligent vehicles and network components such as RSU, ECU, CAN, and communication model (V2V, V2I, V2X). The chapter presents security challenges of VANET such as data confidentially, integrity, availability, authentication, non-repudiation, trust, and privacy preservation. The fundamental of the blockchain technology has been discussed in this chapter, the two main blockchain types: public and private has been the part of this discussion. Furthermore, transaction, consensus algorithm and mining process has been elaborated. The related work in blockchain based VANET has been presented systematically. Finally, the fundamentals, importance, and type of trust in addition to different categories of trust such data centric, entity centric, and hybrid have been presented in this chapter.

In **chapter 3**, the blockchain based privacy preservation framework for VANET has been proposed. Privacy and authentication of the data concerns were discussed to improve the security. The framework is proposed to make communication in VANET more secure. The biometrics features are combined with blockchain technology to provide reliable transmission of data, tracking the data exchanged and identification of the vehicle responsible in the case of falsely messages. The performance of the framework is evaluated in terms of packet delivery rate, packet loss rate and computational cost. Due to the requirements of the legacy system, diligence and statute, the vehicle registration data is kept by the Motor Vehicle Department. The data relating to vehicle communications in VANET is stored in blockchain to make it secure. Moreover, preliminary nomenclature definition of VANET and blockchain and different communication model such as vehicle-to-Vehicle, vehicle-to-Infrastructure

communication and Vehicle-to-Everything is presented. The most widely used protocol in VANET is Wireless Access in Vehicular Environments (WAVE) that provides the foundational standard for Dedicated Short-Range Communication (DSRC); which operates within the 5.9 GHz is discussed in this chapter. Moreover, a blockchain based framework that uses biometrics in VANETs to protect privacy, with vehicles employing a public-private key pair provided by the TA to communicate with other parties is proposed. By employing blockchain techniques, such a decentralized framework will be trustworthy, secure, and allow messages to be disseminated securely. This also presents various entities and authentication mechanism which is based on biometrics features where driver's identity is extracted using the modified discrete transformation. Thereafter, a system model and step-by-step description of vehicle registration process, joining process, message reception, message broadcast, append to blockchain, and finally de-registration process has been presented.

In **chapter 4**, mathematical model and derivation have been presented. The formal definition of the trust has been presented that is evaluated by a particular vehicle based on existing knowledge, interactions and behavior in a specific context and time is presented in this chapter. Depending on the interaction between the vehicles in the network, the trust score may be different. The trust score is therefore a mix of the vehicle specific attributes and the interaction factors. The quality of the interaction and the trust score are both affected by the score T, which expresses the distinctive characteristic score value of the vehicle. Trust attributes has been discussed in addition to the spatial knowledge, temporal experience, and behavioral pattern have been discussed. The chapter presents attack model such as transmission, interruption dropping misbehavior, spoofing attacks, and compromised RSU. The mathematical derivation of the trust and reputation model and the message similarity has been proposed. In order to determine more likely authentic message a similarity computation is performed on every pair of messages M_i, M_j based Jaccard distance formula has been presented in this chapter. The distance plays an important role to find out the proximity between the event location and the vehicle which has observed the event. The message which sent by two or more vehicles must be obtained in order to find out the accuracy of messages.

The **chapter 5,** presents ensemble learning for trust classification. The new issues have arisen as a result of the evolution of VANET, and reputation must be considered because it is critical to know whether vehicles can be trusted on a network. Trust management could be an efficient solution to address VANET security and privacy challenges. The goal of this chapter is to

evaluate, record, and disseminate vehicle trustworthiness in vehicular networks. Moreover, machine learning for trust classification is presented. The substantial amount of network traffic data is collected under both normal and abnormal conditions induced by various scenarios. This chapter also discussed about data normalization and to identify trustworthiness in the proposed system, a binary-classification problem is examined, and machine learning methods are frequently employed to solve such classification issues. The machine learning techniques chosen are tree-based, including decision trees, XGBoost, and random forest. Most tree structure machine learning models employ ensemble learning, which can lead to better results than single models like linear regression or KNN. In order to make feature selection easier, feature importance estimations are made throughout the model construction process. Finally, the complexity of the algorithm is discussed.

The **chapter 6**, presents the discussion on the results obtain from the proposed privacy framework and formal method of trust computation. The simulation and findings are reported in this chapter are meant to validate the validity and efficacy of the suggested model. First, it gives findings and analysis of privacy-preservation architecture to avoid attacks in VANET. second, it presents the simulation setup and parameters used in the experiment. The ssimulation has been performed using OMNeT++, Veins and SUMO to demonstrate the correctness of the proposed model using IEEE 802.11p/1609.4 protocols. The trace control interface (TraCI) which is a middle interface between OMNeT++ and SUMO++ that provides a TCP based communication between these two simulators. Moreover, the first part of this chapter presented the results and findings in terms of packet delivery rate without DoS attack, packet loss without DoS attack and computational cost.

The second part of this chapter shows the findings derived from trust calculations and categorization. This chapter analyses the results of the simulated implementation of the system models described in Chapter 3, 4, 5 and the recommended solutions to the defined objectives at the commencement of this thesis. Moreover, this chapter examines the formal techniques and results achieved and describes 24 features are used in the proposed model to compute the reputation and trust of a vehicle. The feature contains IP address, port number, time stamp, TTL and computation value of reputation. Additionally, it investigates the ensemble machine learning for trust classification, and presents experimental dataset for demonstrating the viability of our model which is generated simulated data for the VANET. Furthermore, it presents various metrics of evaluation such as AUC, confusion metrics and time measurement.

In this research we have applied FS on DT, XGBoost and random forest. However, we found the random forest to be the best among these. Finally, this chapter discusses and compare the obtained results with other existing work.

7.2 Future Work

The results show that the framework can protect the security of messages in real-world situations with VANET. The proposed framework not only protects and trusts the communication between vehicles, but it also protects the identities of people who are allowed to use it. Additionally, biometric is combined with blockchain technology to provide reliable transmission of data, keep track of data being exchanged and identify the responsible vehicle in the case of false messages. The performance of the framework is evaluated in terms of packet delivery rate, packet loss rate and computational cost. Therefore, the obtained results reveal that the proposed model is superior in comparison with existing approaches. For future research direction of this work, it is worthwhile to evaluate efficiency of the network communication among vehicles under more severe security attacks and analyze its behavior accordingly, such as a Sybil attack, eavesdropping attack, message reply and jamming attacks. Applying attacker's behavior patterns is difficult because it can be as sophisticated as an attack intended to avoid observation by adjusting its behavior in accordance to protocol changes. Moreover, we will extend the model for computing ranking and reputation of vehicles and drivers using machine learning techniques. Furthermore, the authenticity of the recommended messages requires further assessment based on similarity, especially when there are insufficient recommending neighbors. The RSU aggregates trust values based on ratings provided by message receivers. All RSUs collaborate to create a trustworthy and consistent database using blockchain principles. It has been found that a decentralized trust management system, can substantially assist vehicles in evaluating the credibility of their neighbors and establishing an efficient and secure intelligent transportation network. However, greater performance and outcomes can be obtained by implementing and evaluating the proposed model in a real-world setting. Lastly, trust and reputation management schemes for a broader area might be investigated, with intercommunication among several regional blockchains. The proposed model can potentially be customized to use different data aggregation and dissemination applications.

Bibliography

[1] I. Altaf and A. Kaul, “A Survey on Autonomous Vehicles in the Field of Intelligent Transport System,” pp. 11–31, 2022.

[2] “Global status report on road safety 2018.” [Online]. Available: https://www.who.int/publications/i/item/9789241565684. [Accessed: 21-Apr-2022].

[3] “ITS United Kingdom : Better Transport Through Technology – The Intelligent Transport Society for the United Kingdom (ITS (UK)) is a not-for-profit public/private sector association financed by members’ subscriptions. It promotes the benefits of Intelligent Transport Solutions (ITS).” [Online]. Available: https://its-uk.org.uk/. [Accessed: 21-Apr-2022].

[4] “Homepage » ERTICO.” [Online]. Available: https://ertico.com/. [Accessed: 21-Apr-2022].

[5] I. Transportation Systems Joint Program Office, “USDOT’s Intelligent Transportation Systems (ITS) Strategic Plan 2015-2019,” 2015.

[6] “ITS Japan（English）.” [Online]. Available: https://www.its-jp.org/english/. [21-Apr-2022].

[7] “ITS Brasil - Intelligent Transportation System.” [Online]. Available: https://www.itsb.org.br/. [Accessed: 21-Apr-2022].

[8] R. Brendha and V. S. J. Prakash, “A survey on routing protocols for vehicular Ad Hoc networks,” *2017 4th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2017*, Aug. 2017.

[9] T. N. D. Pham and C. K. Yeo, “Adaptive trust and privacy management framework for vehicular networks,” *Veh. Commun.*, vol. 13, pp. 1–12, Jul. 2018.

[10] C. Kalaiarasy and N. Sreenath, “An incentive-based co-operation motivating pseudonym changing strategy for privacy preservation in mixed zones in vehicular networks,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 1, pp. 1510–1520, Jan. 2022.

[11] C. Wang, X. Cheng, J. Li, Y. He, and K. Xiao, “A survey: applications of blockchain in the Internet of Vehicles,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2021, no. 1, pp. 1–16, Dec. 2021.

[12] A. ANDERBERG *et al.*, “Blockchain Now And Tomorrow,” 2019.

[13] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, “Blockchain-based decentralized trust management in vehicular networks,” *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1495–1505, 2019.

[14] Z. Lu, W. Liu, Q. Wang, G. Qu, and Z. Liu, “A privacy-preserving trust model based on blockchain for VANETs,” *IEEE Access*, vol. 6, pp. 45655–45664, Aug. 2018.

[15] Z. Lu, Q. Wang, G. Qu, and Z. Liu, “BARS: A Blockchain-Based Anonymous Reputation System for Trust Management in VANETs,” *Proc. - 17th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. 12th IEEE Int. Conf. Big Data Sci. Eng. Trust. 2018*, pp. 98–103, Sep. 2018.

[16] R. Shrestha, R. Bajracharya, A. P. Shrestha, and S. Y. Nam, “A new type of blockchain for secure message exchange in VANET,” *Digit. Commun. Networks*, vol. 6, no. 2, pp. 177–186, May 2020.

[17] R. Shrestha, R. Bajracharya, and S. Y. Nam, “Blockchain-based Message Dissemination in

VANET," *Proc. 2018 IEEE 3rd Int. Conf. Comput. Commun. Secur. ICCCS 2018*, pp. 161–166, Dec. 2018.

[18] J. Kang *et al.*, "Blockchain for secure and efficient data sharing in vehicular edge computing and networks," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4660–4670, Jun. 2019.

[19] N. Malik, P. Nanda, X. He, and R. Liu, "Trust and reputation in vehicular networks: A smart contract-based approach," *Proc. - 2019 18th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. IEEE Int. Conf. Big Data Sci. Eng. Trust. 2019*, pp. 34–41, Aug. 2019.

[20] S. Oubabas, R. Aoudjit, J. J. P. C. Rodrigues, and S. Talbi, "Secure and stable Vehicular Ad Hoc Network clustering algorithm based on hybrid mobility similarities and trust management scheme," *Veh. Commun.*, vol. 13, pp. 128–138, Jul. 2018.

[21] S. Baras, I. Saeed, H. A. Tabaza, and M. Elhadef, "VANETs-based intelligent transportation systems: An overview," *Lect. Notes Electr. Eng.*, vol. 474, pp. 265–273, 2018.

[22] M. Milton Joe and B. Ramakrishnan, "Review of vehicular ad hoc network communication models including WVANET (Web VANET) model and WVANET future research directions," *Wirel. Networks*, vol. 22.

[23] H. Tan, M. Ma, H. Labiod, A. Boudguiga, J. Zhang, and P. H. J. Chong, "A Secure and Authenticated Key Management Protocol (SA-KMP) for Vehicular Networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9570–9584, Dec. 2016.

[24] I. Standards, *IEEE Guide for Wireless Access in Vehicular Environments (WAVE) Architecture IEEE Vehicular Technology Society Sponsored by the Intelligent Transportation Systems Committee*. 2019.

[25] J. B. Kenney, "Dedicated short-range communications (DSRC) standards in the United States," *Proc. IEEE*, vol. 99, no. 7, pp. 1162–1182, 2011.

[26] M. S. Sheikh, J. Liang, G. E. : Manzoor, and A. Khan, "A Comprehensive Survey on VANET Security Services in Traffic Management System," 2019.

[27] A. Reyna, C. Martín, J. Chen, E. Soler, and M. Díaz, "On blockchain and its integration with IoT. Challenges and opportunities," *Futur. Gener. Comput. Syst.*, vol. 88, pp. 173–190, Nov. 2018.

[28] S. Roy, M. Ashaduzzaman, M. Hassan, and A. Rahman Chowdhury, *BlockChain for IoT Security and Management: Current Prospects, Challenges and Future Directions; BlockChain for IoT Security and Management: Current Prospects, Challenges and Future Directions*. 2018.

[29] P. Raj, K. Saini, and C. Surianarayanan, "Blockchain Technology and Applications; Chapter Chapter 926 Pages Identification of Blockchain-Enabled Opportunities and Their Business Values: Interoperability of Blockchain WithN.S. Gowri Ganesh ," 2020.

[30] A. Khan, M.A., Algarni, F., Quasim, M.T, Alharthi, *Decentralised Internet of Things, Studies in Big Data,*. Springer International Publishing, 2020.

[31] V. Balioti, C. Tzimopoulos, and C. Evangelides, "Multi-Criteria Decision Making Using TOPSIS Method Under Fuzzy Environment. Application in Spillway Selection," *Proc. 2018, Vol. 2, Page*

637, vol. 2, no. 11, p. 637, Jul. 2018.

[32] R. L. Rivest, A. Shamir, and D. A. W agner, “Time-lock puzzles and timed-release Crypto,” 1996.

[33] M. Mahmoody, T. Moran, S. Vadhan, and Ø. Øøôô»»××××ºººöúúöö ººº ù»»øøðñ¸××ðð», “Time-Lock Puzzles in the Random Oracle Model.”

[34] L. Zhu, C. Chen, X. Wang, and A. O. Lim, “SMSS: Symmetric-Masquerade Security Scheme for VANETs,” 2011.

[35] C. Zhang, R. Lu, X. Lin, P.-H. Ho, and X. Shen, *An Efficient Identity-based Batch Verification Scheme for Vehicular Sensor Networks*. 2008.

[36] G. Calandriello, P. Papadimitratos, J.-P. Hubaux, and A. Lioy, *Efficient and Robust Pseudonymous Authentication in VANET*. 2007.

[37] H. Tan, D. Choi, P. Kim, S. Pan, and I. Chung, “Secure Certificateless Authentication and Road Message Dissemination Protocol in VANETs,” *Wirel. Commun. Mob. Comput.*, vol. 2018, 2018.

[38] P. Vijayakumar, M. Azees, A. Kannan, and L. J. Deborah, “Dual Authentication and Key Management Techniques for Secure Data Transmission in Vehicular Ad Hoc Networks; Dual Authentication and Key Management Techniques for Secure Data Transmission in Vehicular Ad Hoc Networks,” *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 4, p. 1015, 2016.

[39] Q. Wu, J. Domingo-Ferrer, and Ú. González-Nicolás, “Balanced Trustworthiness, Safety, and Privacy in Vehicle-to-Vehicle Communications; Balanced Trustworthiness, Safety, and Privacy in Vehicle-to-Vehicle Communications,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, p. 559, 2010.

[40] B. Qin, Q. Wu, J. Domingo-Ferrer, and W. Susilo, “Distributed Privacy-Preserving Secure Aggregation in Vehicular Communication,” *2011 Third Int. Conf. Intell. Netw. Collab. Syst.*, 2011.

[41] Y. Yuan and F.-Y. Wang, *Towards blockchain-based intelligent transportation systems; Towards blockchain-based intelligent transportation systems*. 2016.

[42] S. Rowan, M. Clear, M. Gerla, M. Huggard, and C. M. Goldrick, “Securing Vehicle to Vehicle Communications using Blockchain through Visible Light and Acoustic Side-Channels,” Apr. 2017.

[43] M. C. Chuang and J. F. Lee, “PPAS: A privacy preservation authentication scheme for vehicle-to- infrastructure communication networks,” *2011 Int. Conf. Consum. Electron. Commun. Networks, CECNet 2011 - Proc.*, pp. 1509–1512, 2011.

[44] X. Peng, “A novel authentication protocol for vehicle network,” *2016 3rd Int. Conf. Syst. Informatics, ICSAI 2016*, pp. 664–668, Jan. 2017.

[45] N. Malik, P. Nanda, A. Arora, X. He, and D. Puthal, “Blockchain Based Secured Identity Authentication and Expeditious Revocation Framework for Vehicular Networks,” *Proc. - 17th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. 12th IEEE Int. Conf. Big Data Sci. Eng. Trust. 2018*, pp. 674–679, Sep. 2018.

[46] R. Lu, X. Lin, H. Zhu, P.-H. Ho, and X. Shen, “ECPP: Efficient Conditional Privacy Preservation Protocol for Secure Vehicular Communications,” pp. 1229–1237, Jun. 2008.

[47] B. Leiding, P. Memarmoshrefi, and D. Hogrefe, “Self-managed and Blockchain-based Vehicular Ad-hoc Networks.”

[48] A. Dorri, M. Steger, S. S. Kanhere, and R. Jurdak, “BlockChain: A Distributed Solution to Automotive Security and Privacy,” *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 119–125, Dec. 2017.

[49] C. Liu, K. K. Chai, X. Zhang, E. T. Lau, and Y. Chen, “Adaptive Blockchain-Based Electric Vehicle Participation Scheme in Smart Grid Platform,” *IEEE Access*, vol. 6, pp. 25657–25665, May 2018.

[50] F. Gao, L. Zhu, M. Shen, K. Sharif, Z. Wan, and K. Ren, “A Blockchain-Based Privacy-Preserving Payment Mechanism for Vehicle-to-Grid Networks,” *IEEE Netw.*, vol. 32, no. 6, pp. 184–192, Nov. 2018.

[51] X. Huang, C. Xu, P. Wang, and H. Liu, “LNSC: A Security Model for Electric Vehicle and Charging Pile Management Based on Blockchain Ecosystem,” *IEEE Access*, vol. 6, pp. 13565–13574, Mar. 2018.

[52] N. Kshetri, “Can Blockchain Strengthen the Internet of Things?,” *IT Prof.*, vol. 19, no. 4, pp. 68–72, 2017.

[53] K. Kotobi and S. G. Bilen, “Secure Blockchains for Dynamic Spectrum Access: A Decentralized Database in Moving Cognitive Radio Networks Enhances Security and User Access,” *IEEE Veh. Technol. Mag.*, vol. 13, no. 1, pp. 32–39, Mar. 2018.

[54] A. Srinivasan, J. Teitelbaum, J. Wu, M. Cardei, and H. Liang, “Reputation-and-Trust-Based Systems for Ad Hoc Networks,” *Algorithms Protoc. Wirel. Mob. Ad Hoc Networks*, pp. 375–403, Mar. 2008.

[55] N. B. Truong, T. W. Um, B. Zhou, and G. M. Lee, “From Personal Experience to Global Reputation for Trust Evaluation in the Social Internet of Things,” *2017 IEEE Glob. Commun. Conf. GLOBECOM 2017 - Proc.*, vol. 2018-January, pp. 1–7, Jul. 2017.

[56] S. A. Soleymani *et al.*, “Trust management in vehicular ad hoc network: a systematic review,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2015, no. 1, Dec. 2015.

[57] Y.-H. Wang, “A Trust Management Model for Internet of Vehicles.”

[58] U. Jayasinghe, A. Otebolaku, T. W. Um, and G. M. Lee, “Data centric trust evaluation and prediction framework for IOT,” *Proc. 2017 ITU Kaleidosc. Acad. Conf. Challenges a Data-Driven Soc. ITU K 2017*, vol. 2018-January, pp. 1–7, Jun. 2017.

[59] A. Mahmood, W. E. Zhang, Q. Z. Sheng, S. A. Siddiqui, and A. Aljubairy, “for Software-Defined Heterogeneous,” *Secur. Priv. Trust IoT Environ.*, pp. 203–226, 2019.

[60] M. Raya, P. Papadimitratos, V. D. Gligor, and J.-P. Hubaux, “On Data-Centric Trust Establishment in Ephemeral Ad Hoc Networks,” pp. 1238–1246, Jun. 2008.

[61] J. Zhang, “A survey on trust management for VANETs,” *Proc. - Int. Conf. Adv. Inf. Netw. Appl. AINA*, pp. 105–112, 2011.

[62] S. Gurung, D. Lin, A. Squicciarini, and E. Bertino, “Information-oriented trustworthiness evaluation in vehicular ad-hoc networks,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7873 LNCS, pp. 94–108, 2013.

[63] R. A. Shaikh and A. S. Alzahrani, “Intrusion-aware trust model for vehicular ad hoc networks,” *Secur. Commun. Networks*, vol. 7, no. 11, pp. 1652–1669, Nov. 2014.

[64] F. Gómez Mármol and G. Martínez Pérez, “TRIP, a trust and reputation infrastructure-based proposal for vehicular ad hoc networks,” *J. Netw. Comput. Appl.*, vol. 35, no. 3, pp. 934–941, May 2012.

[65] A. Patwardhan, A. Joshi, T. Finin, and Y. Yesha, “A data intensive reputation management scheme for vehicular ad hoc networks,” *2006 3rd Annu. Int. Conf. Mob. Ubiquitous Syst. MobiQuitous - Work.*, 2006.

[66] C. Chen, J. Zhang, R. Cohen, and P. H. Ho, “A trust modeling framework for message propagation and evaluation in VANETs,” *2010 2nd Int. Conf. Inf. Technol. Converg. Serv. ITCS 2010*, 2010.

[67] N.-W. Lo and H.-C. Tsai, “A Reputation System for Traffic Safety Event on Vehicular Ad Hoc Networks,” *EURASIP J. Wirel. Commun. Netw.*, 2009.

[68] Z. Liu, J. Ma, Z. Jiang, H. Zhu, and Y. Miao, “LSOT: A Lightweight Self-Organized Trust Model in VANETs,” 2016.

[69] U. Khan, S. Agrawal, and S. Silakari, “Detection of Malicious Nodes (DMN) in vehicular ad-hoc networks,” *Procedia Comput. Sci.*, vol. 46, pp. 965–972, 2015.

[70] M. Gerlach, “Trust for vehicular applications,” *Proc. - Eighth Int. Symp. Auton. Decentralized Syst. ISADS 2007*, pp. 295–302, 2007.

[71] U. Farooq Minhas, J. Zhang, T. Tran, and R. Cohen, “A Multifaceted Approach to Modeling Agent Trust for Effective Communication in the Application of Mobile Ad Hoc Vehicular Networks,” *Appl. Rev.*, vol. 41, no. 3, p. 407, 2011.

[72] N. Yang, “A Similarity based Trust and Reputation Management Framework for VANETs,” *Int. J. Futur. Gener. Commun. Netw.*, vol. 6, no. 2, 2013.

[73] A. Jesudoss, S. V. Kasmir Raja, and A. Sulaiman, “Stimulating truth-telling and cooperation among nodes in VANETs through payment and punishment scheme,” *Ad Hoc Networks*, vol. 24, no. PA, pp. 250–263, Jan. 2015.

[74] N. Haddadou, A. Rachedi, and Y. Ghamri-Doudane, “Trust and exclusion in Vehicular Ad Hoc Networks: An economic incentive model based approach,” *2013 Comput. Commun. IT Appl. Conf. ComComAp 2013*, pp. 13–18, 2013.

[75] J. Zhang, J. Cui, H. Zhong, Z. Chen, and L. Liu, “PA-CRT: Chinese Remainder Theorem Based Conditional Privacy-Preserving Authentication Scheme in Vehicular Ad-Hoc Networks,” *IEEE*

Trans. Dependable Secur. Comput., vol. 18, no. 2, pp. 722–735, Mar. 2021.

[76] S. Guleng, C. Wu, X. Chen, X. Wang, T. Yoshinaga, and Y. Ji, "Decentralized Trust Evaluation in Vehicular Internet of Things," *IEEE Access*, vol. 7, pp. 15980–15988, 2019.

[77] H. Sedjelmaci and S. M. Senouci, "An accurate and efficient collaborative intrusion detection framework to secure vehicular networks," *Comput. Electr. Eng.*, vol. 43, pp. 33–47, Apr. 2015.

[78] S. K. Dhurandher, M. S. Obaidat, A. Jaiswal, A. Tiwari, and A. Tyagi, *Securing Vehicular Networks: A Reputation and Plausibility Checks-based Approach*. 2010.

[79] K. Chaker Abdelaziz, N. Lagraa, and A. Lakas, *Trust Model with Delayed Verification for Message Relay in VANETs*. 2014.

[80] F. Dötzer, L. Fischer, and P. Magiera, "VARS: A vehicle ad-hoc network reputation system," *Proc. - 6th IEEE Int. Symp. a World Wirel. Mob. Multimed. Networks, WoWMoM 2005*, pp. 454–456, 2005.

[81] I. A. Rai, R. A. Shaikh, and S. R. Hassan, "A hybrid dual-mode trust management scheme for vehicular networks," *Int. J. Distrib. Sens. Networks*, vol. 16, no. 7, Jul. 2020.

[82] W. Li, H. Song, and S. Member, "ART: An Attack-Resistant Trust Management Scheme for Securing Vehicular Ad Hoc Networks; ART: An Attack-Resistant Trust Management Scheme for Securing Vehicular Ad Hoc Networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 4, 2016.

[83] N. B. Truong, G. M. Lee, and G. Myoung, "Poster Abstract: Trust Evaluation for Data Exchange in Vehicular Networks," *2017 IEEE/ACM Second Int. Conf. Internet-of-Things Des. Implement.*, vol. 2, p. pages, 2017.

[84] G. Primiero, F. Raimondi, T. Chen, and R. Nagarajan, "A proof-theoretic trust and reputation model for VANET," *Proc. - 2nd IEEE Eur. Symp. Secur. Priv. Work. EuroS PW 2017*, pp. 146–152, Jun. 2017.

[85] G. Primiero, "A calculus for distrust and mistrust," *IFIP Adv. Inf. Commun. Technol.*, vol. 473, pp. 183–190, 2016.

[86] U. Javaid, M. N. Aman, and B. Sikdar, "DrivMan: Driving trust management and data sharing in VANETs with blockchain and smart contracts," *IEEE Veh. Technol. Conf.*, vol. 2019-April, Apr. 2019.

[87] T. Jiang, H. Fang, and H. Wang, "Blockchain-Based Internet of Vehicles: Distributed Network Architecture and Performance Analysis; Blockchain-Based Internet of Vehicles: Distributed Network Architecture and Performance Analysis," *IEEE INTERNET THINGS J.*, vol. 6, no. 3, 2019.

[88] M. Singh and S. Kim, "Crypto trust point (cTp) for secure data sharing among intelligent vehicles," *Int. Conf. Electron. Inf. Commun. ICEIC 2018*, vol. 2018-January, pp. 1–4, Apr. 2018.

[89] X. Zhang, R. Li, and B. Cui, "A security architecture of VANET based on blockchain and mobile edge computing," *Proc. 2018 1st IEEE Int. Conf. Hot Information-Centric Networking, HotICN 2018*, pp. 258–259, Jan. 2019.

[90] M. Singh and S. Kim, “Trust Bit: Reward-based intelligent vehicle commination using blockchain paper,” *IEEE World Forum Internet Things, WF-IoT 2018 - Proc.*, vol. 2018-January, pp. 62–67, May 2018.

[91] X. Zhang and X. Chen, “Data Security Sharing and Storage Based on a Consortium Blockchain in a Vehicular Ad-hoc Network,” *IEEE Access*, vol. 7, pp. 58241–58254, 2019.

[92] H. Khelifi, S. Luo, B. Nour, H. Moungla, and S. Hassan Ahmed, “Reputation-based blockchain for secure NDN caching in vehicular networks,” *2018 IEEE Conf. Stand. Commun. Networking, CSCN 2018*, pp. 1–6, Oct. 2018.

[93] A. Lei, H. Cruickshank, Y. Cao, P. Asuquo, C. P. A. Ogah, and Z. Sun, “Blockchain-Based Dynamic Key Management for Heterogeneous Intelligent Transportation Systems,” *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1832–1843, Dec. 2017.

[94] L. Yadav, S. Kumar, A. Kumarsagar, and S. Sahana, “Architechture, Applications and Security for IOV: A Survey,” *Proc. - IEEE 2018 Int. Conf. Adv. Comput. Commun. Control Networking, ICACCCN 2018*, pp. 383–390, Oct. 2018.

[95] M. A. Al-Shareeda, M. A. Alazzawi, M. Anbar, S. Manickam, and A. K. Al-Ani, “A Comprehensive Survey on Vehicular Ad Hoc Networks (VANETs),” *2021 Int. Conf. Adv. Comput. Appl. ACA 2021*, pp. 156–160, 2021.

[96] A. Shahid Khan, K. Balan, Y. Javed, S. Tarmizi, and J. Abdullah, “Secure Trust-Based Blockchain Architecture to Prevent Attacks in VANET.”

[97] M. A. Khan and K. Salah, “IoT security: Review, blockchain solutions, and open challenges,” *Futur. Gener. Comput. Syst.*, vol. 82, pp. 395–411, May 2018.

[98] O. Kaiwartya *et al.*, “Internet of Vehicles: Motivation, Layered Architecture, Network Model, Challenges, and Future Aspects,” *IEEE Access*, vol. 4, pp. 5356–5373, 2016.

[99] Y. L. Morgan, “Notes on DSRC & WAVE standards suite: Its architecture, design, and characteristics,” *IEEE Commun. Surv. Tutorials*, vol. 12, no. 4, pp. 504–518, Dec. 2010.

[100] K. N. Qureshi, S. Din, G. Jeon, and F. Piccialli, “Internet of Vehicles: Key Technologies, Network Model, Solutions and Challenges with Future Aspects,” *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1777–1786, Mar. 2021.

[101] Z. Lu, G. Qu, and Z. Liu, “A Survey on Recent Advances in Vehicular Network Security, Trust, and Privacy,” *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 760–776, Feb. 2019.

[102] X. Li, J. Liu, X. Li, and W. Sun, “RGTE: A reputation-based global trust establishment in VANETs,” *Proc. - 5th Int. Conf. Intell. Netw. Collab. Syst. INCoS 2013*, pp. 210–214, 2013.

[103] Q. Li, A. Malip, K. M. Martin, S. L. Ng, and J. Zhang, “A reputation-based announcement scheme for VANETs,” *IEEE Trans. Veh. Technol.*, vol. 61, no. 9, pp. 4095–4108, 2012.

[104] N. Bißmeyer, J. Njeukam, J. Petit, and K. M. Bayarou, “Central misbehavior evaluation for VANETs based on mobility data plausibility,” *VANET’12 - Proc. 9th ACM Int. Work. Veh. Inter-NETworking, Syst. Appl.*, pp. 73–82, 2012.

[105] X. Huang, R. Yu, J. Kang, and Y. Zhang, “Distributed reputation management for secure and efficient vehicular edge computing and networks,” *IEEE Access*, vol. 5, pp. 25408–25420, Nov. 2017.

[106] A. I. Awad *et al.*, “A Categorized Trust-Based Message Reporting Scheme for VANETs,” *Commun. Comput. Inf. Sci.*, vol. 381 CCIS, pp. 65–83, 2013.

[107] I. Dokmani´c, R. Parhizkar, J. Ranieri, and M. Vetterli, “Euclidean Distance Matrices Essential Theory, Algorithms and Applications.”

[108] S. Craw, “Manhattan Distance,” in *Encyclopedia of Machine Learning and Data Mining*, C. Sammut and G. I. Webb, Eds. Boston, MA: Springer US, 2017, pp. 790–791.

[109] “View of Comparing sets of patterns with the Jaccard index.” [Online]. Available: https://journal.acs.org.au/index.php/ajis/article/view/1538/817. [Accessed: 27-Jul-2022].

[110] J. Han, M. Kamber, and J. Pei, “Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems),” 2011.

[111] F. Qu, Z. Wu, F. Wang, and W. Cho, “A Security and Privacy Review of VANETs,” *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 2985–2996, Dec. 2015.

[112] R. G. Engoulou, M. Bellaïche, S. Pierre, and A. Quintero, “VANET security surveys,” *Comput. Commun.*, vol. 44, pp. 1–13, May 2014.

[113] J. S. Weng, J. Weng, Y. Zhang, W. Luo, and W. Lan, “BENBI: Scalable and dynamic access control on the northbound interface of SDN-Based VANET,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 822–831, Jan. 2019.

[114] J. Lin, W. Yu, X. Yang, Q. Yang, X. Fu, and W. Zhao, “A Real-Time En-Route Route Guidance Decision Scheme for Transportation-Based Cyberphysical Systems,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2551–2566, Mar. 2017.

[115] Y. Zhang, J. Weng, J. Weng, M. Li, and W. Luo, “Onionchain: Towards Balancing Privacy and Traceability of Blockchain-Based Applications,” Sep. 2019.

[116] J. Cheng, P. Qin, M. Zhou, Z. Huang, and S. Gao, “Key properties of connectivity in vehicle ad-hoc network,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9864 LNCS, pp. 328–339, 2016.

[117] J. Hatwell, M. M. Gaber, R. Muhammad, and A. Azad, “gbt-HIPS: Explaining the Classifications of Gradient Boosted Tree Ensembles,” 2021.

[118] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, “A comparative analysis of gradient boosting algorithms,” *Artif. Intell. Rev.*, vol. 54, no. 3, pp. 1937–1967, 2021.

[119] B. Ying and A. Nayak, “Anonymous and lightweight authentication for secure vehicular networks,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 10626–10636, Aug. 2017.

[120] M. Wazid *et al.*, “Design of Lightweight Authentication and Key Agreement Protocol for Vehicular Ad Hoc Networks.” in IEEE Access, vol. 5, pp. 14966-14980, 2017, doi: 10.1109/ACCESS.2017.2723265.