



# Politecnico di Bari

Repository Istituzionale dei Prodotti della Ricerca del Politecnico di Bari

Digital twin and security solutions for intelligent transportation systems

This is a PhD Thesis

*Original Citation:*

Digital twin and security solutions for intelligent transportation systems / Ali, Wasim Ahmed Mohamed. - ELETTRONICO. - (2024).

*Availability:*

This version is available at <http://hdl.handle.net/11589/271220> since: 2024-06-27

*Published version*

DOI:

Publisher: Politecnico di Bari

*Terms of use:*

(Article begins on next page)

03 July 2024



Politecnico  
di Bari

Department of Electrical and Information Engineering

ELECTRICAL AND INFORMATION ENGINEERING

PH.D. PROGRAM

SSD: ING-INF/04-AUTOMATICA

**Final Dissertation**

**DIGITAL TWIN AND SECURITY SOLUTIONS FOR  
INTELLIGENT TRANSPORTATION SYSTEMS**

by

Wasim Ahmed Mohammed Ali

*A thesis submitted for the degree of  
Doctor of Philosophy*

**Supervisor:**

Prof. Maria Pia Fanti

**Co-supervisor:**

Prof. Gennaro Boggia

**Coordinator of Ph.D. Program:**

Prof. Mario Carpentieri, Ph.D.

*Cycle XXXVI - November, 1<sup>st</sup> 2020 - January 31<sup>st</sup> 2024*



Politecnico  
di Bari

Department of Electrical and Information Engineering

ELECTRICAL AND INFORMATION ENGINEERING

PH.D. PROGRAM

SSD: ING-INF/04-AUTOMATICA

**Final Dissertation**

DIGITAL TWIN AND SECURITY SOLUTIONS FOR  
INTELLIGENT TRANSPORTATION SYSTEMS

by

Wasim Ahmed Mohammed Ali

---

*A thesis submitted for the degree of  
Doctor of Philosophy*

**Referees:**

Prof. Karinne Ramirez

Prof. Gregory Faraut

**Supervisor:**

Prof. Maria Pia Fanti

---

**Co-supervisor:**

Prof. Gennaro Boggia

---

**Coordinator of Ph.D. Program:**

Prof. Mario Carpentieri, Ph.D.

---

*Cycle XXXVI - November, 1<sup>st</sup> 2020 - January 31<sup>st</sup> 2024*

## *Dedication*

*This work is dedicated to my wonderful family:  
My beloved wife **NOORA**, my greatest supporter and  
source of strength throughout this challenging journey.  
My lovely Kids **AHMED and AL-NASSER**,  
the joy and inspiration in my life.*

## Acknowledgements

First and foremost, my heartfelt gratitude goes to God and my parents, whose prayers and support have been my anchor throughout this journey, driving me toward success.

I extend my sincere appreciation to those who contributed to the completion of this research. My deepest thanks to my supervisor, **Prof. Maria Pia Fanti**, for her invaluable guidance and support from the inception of my PhD program. I also appreciate the collaboration of my co-supervisors, professors, and lab colleagues, enriching my research experience.

Special recognition is reserved for my wife, **Noora** and my two sons, **Ahmed and Naser**, who were always with me through the journey of struggles and the moments of success, and understanding and patience sustained me during the challenges of this academic endeavour.

I owe a debt of gratitude to my brothers and sisters, with special mention to my sister **Olfat** and my brother **Mohammed**, for their constant unconditional support.

I am grateful to two influential persons who inspired and supported me throughout my educational journey: **Dr. Naser Ibraheem**, a steadfast friend who never left me midway, and my uncle **Dr. Hasan Aljaradi**, who supports, offering advice and encouragement without hesitation.

A profound acknowledgement goes to my former teachers in previous studies, **Prof. Abdulbasit Darem, Prof. Sandhya P., and Dr. Manasa K. N.**; their continuous support and belief in my potential were pivotal in motivating me to embark on this transformative PhD journey.

Finally, I take pride in the culmination of this work, acknowledging with gratitude the support and contributions of everyone who has been part of this transformative journey. Thanks to everyone who played a role in my life and supported me, contributing to this achievement.

## Abstract

Intelligent Transportation Systems (ITS) are poised to transform the transportation landscape by enabling seamless integration of technologies, enhancing road safety, and optimizing traffic flow. To fully realize the potential of ITS, it is crucial to address the challenges of cybersecurity and sustainability. This thesis explores innovative approaches to integrate security measures and sustainability strategies into vehicular networks.

First, a comprehensive review of digital twin (DT) technologies in the context of ITS is presented. This review highlights the potential of digital twins to enhance cybersecurity by providing a holistic view of vehicular networks and enabling proactive mitigation of threats.

Next, a novel Intrusion Detection System (IDS) is proposed for Vehicular Ad Hoc Networks (VANETs). The IDS leverages decision tree-based machine learning techniques to detect anomalies and identify potential intrusions with high accuracy.

To address sustainability challenges, an optimization framework for electric vehicle (EV) routing in logistics operations is presented. The framework minimizes charging/discharging costs while considering the shortest path for each EV, optimizing route planning, and contributing to reduced environmental impact. Real-world case studies validate the effectiveness of the proposed optimization method.

Finally, a simulation-based study on traffic networks and communication protocols is conducted. The study employs a hybrid methodology that integrates SUMO, OMNeT++, and VEINS frameworks to model and simulate interactions within the dynamic urban setting of Bologna, Italy. Focusing on attacks against VANET networks through IEEE 802.11p protocol / WAVE standard messages, the simulation-based approach enhances vehicular network security and contributes to sustainability by ensuring the reliability and efficiency of communication protocols.

In conclusion, the contributions of this thesis provide a strong foundation for future research in ITS. We can create a more secure, efficient, and sustainable transportation ecosystem by applying DT framework on ITS and integrating cybersecurity measures and sustainability strategies.

# Contents

<b>List of Publications</b>	<b>ix</b>
<b>List of Figures</b>	<b>xi</b>
<b>List of Tables</b>	<b>xiii</b>
<b>Abbreviations</b>	<b>xiv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	2
1.1.1 Digital Twin . . . . .	2
1.1.2 Intelligent Transportation System . . . . .	4
1.1.3 DT and ITS . . . . .	6
1.2 Motivation . . . . .	7
1.3 Problem statement . . . . .	8
1.4 Thesis Contributions . . . . .	9
1.5 Thesis Structure . . . . .	12
<b>2 Background about Digital Twin in Intelligent Transportation Systems</b>	<b>14</b>
2.1 Introduction . . . . .	14
2.2 Digital Twin in Transportation Systems . . . . .	19

---

2.2.1	Digital Twin in Electromobility . . . . .	22
2.2.2	Digital Twin Networks in Smart Transportation . . . . .	23
2.3	Technologies in Digital Twin for Electromobility . . . . .	24
2.3.1	Internet of Things . . . . .	25
2.3.2	Virtual Sensors . . . . .	26
2.3.3	Internet of Vehicles . . . . .	27
2.3.4	5G networks . . . . .	28
2.3.5	Artificial Intelligence for Autonomous Vehicles . . . . .	30
2.3.6	Discussion . . . . .	32
2.4	Current Issues in EVs . . . . .	32
2.4.1	Tracking and Monitoring . . . . .	33
2.4.2	Battery Management Systems . . . . .	35
2.4.3	Connectivity . . . . .	37
2.4.4	Security and privacy . . . . .	38
2.4.5	Data Analytics . . . . .	39
2.4.6	Intrusion Detection Systems . . . . .	41
2.4.7	Discussion . . . . .	43
2.5	DT Solutions and Future Directions . . . . .	43
2.5.1	Cost-Effective and Reliability . . . . .	44
2.5.2	Visualization and Charging Time . . . . .	44
2.5.3	Intrusion Detection Systems for EV and AV . . . . .	44
2.5.4	Case studies . . . . .	45
2.6	Chapter Summary . . . . .	46
<b>3</b>	<b>Communication Security for Intelligent Transportation Systems</b>	<b>49</b>
3.1	Security and Intelligent Transportation Systems . . . . .	49



---

3.1.1	Introduction . . . . .	49
3.1.2	Intrusion Detection Systems . . . . .	50
3.1.3	Network Attacks . . . . .	52
3.1.4	Machine Learning Techniques . . . . .	53
3.1.5	Current Datasets for Intrusion Detection Systems . . . . .	58
3.2	Vehicular Ad-hoc Networks . . . . .	64
3.2.1	Security in Vehicular Ad-hoc Networks . . . . .	66
3.2.2	Simulation for Vehicular Ad-hoc Networks . . . . .	70
3.3	Electric Vehicles Routing Simulation and Optimization . . . . .	75
<b>4</b>	<b>Routing Optimization for Electric Vehicle Using SUMO Simulation</b>	<b>78</b>
4.1	Introduction . . . . .	78
4.2	Background: Routing Optimization Review . . . . .	79
4.2.1	Electric Vehicle Routing Problem . . . . .	79
4.2.2	Simulation of Urban Mobility . . . . .	80
4.2.3	Traffic Control Interface . . . . .	81
4.2.4	Optimization Techniques for EVRP . . . . .	81
4.3	Modeling and Methodology . . . . .	82
4.3.1	Model Initialization and Input Data Preparation . . . . .	82
4.3.2	Simulation-Based Rerouting Using TraCI . . . . .	83
4.3.3	Output Analysis . . . . .	84
4.4	Case Study: The Southern Italian Region of Puglia . . . . .	85
4.4.1	Description of the simulation model . . . . .	85
4.5	Results and Discussion . . . . .	89
<b>5</b>	<b>Intrusion Detection System for Vehicular Ad-Hoc Network</b>	<b>93</b>

---

5.1	Introduction . . . . .	93
5.2	Intrusion Detection System for VANET . . . . .	97
5.2.1	VANET Security & IDS . . . . .	97
5.2.2	Attacks on Vehicular Networks . . . . .	99
5.3	The Proposed IDS Model . . . . .	103
5.3.1	System architecture . . . . .	104
5.3.2	Benchmark description . . . . .	105
5.3.3	Data Pre-processing . . . . .	106
5.3.4	Feature Engineering . . . . .	107
5.3.5	The ML model . . . . .	110
5.3.6	Evaluation methods . . . . .	113
5.4	Experimental Setup . . . . .	115
5.5	Results and Discussion . . . . .	120
<b>6</b>	<b>Security and Intelligent Transportation Systems</b>	<b>122</b>
6.1	Introduction . . . . .	122
6.2	Simulation Approach for VANET Security . . . . .	125
6.3	The IEEE 802.11p - WAVE Protocol . . . . .	128
6.3.1	Dedicated Short-Range Communications . . . . .	128
6.3.2	Wireless Access in Vehicular Environments . . . . .	129
6.4	Simulation Tools for VANET . . . . .	130
6.4.1	SUMO, OMNeT++ and VEINS . . . . .	131
6.5	Case Study: Bologna City - Pasubio Region . . . . .	133
6.5.1	Simulation Description . . . . .	133
6.5.2	Attacks Simulation . . . . .	136
6.6	Result and Discussion . . . . .	138

Contents	viii
<b>7 Conclusion</b>	<b>140</b>
<b>References</b>	<b>143</b>

# List of Publications

## International Journals

- Ali, W. A., Fanti, M. P., Roccotelli, M., & Ranieri, L. (2023). A Review of Digital Twin Technology for Electric and Autonomous Vehicles. *Applied Sciences*, 13(10), 5871. Doi: 10.3390/app13105871.
- Ali, W. A., Sandhya, P., Roccotelli, M., & Fanti, M. P. (2022). A Comparative Study of Current Dataset Used to Evaluate Intrusion Detection System. *International Journal on Engineering Applications*, 10(5). Doi: 10.15866/irea.v10i5.21030
- Ali, W. A., Roccotelli, M., Gennaro Boggia, & Fanti, M. P. (2024). Intrusion Detection System for Vehicular ad Hoc Network Attacks Based on Machine Learning Techniques. *Information Security Journal: A Global Perspective*. Doi: 10.1080/19393555.2024.2307638.

## International Conferences

- Ali, W. A., Roccotelli, M., & Fanti, M. P. (2022, May). Digital Twin in Intelligent Transportation Systems: A Review. In *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)* (Vol. 1, pp. 576-581). IEEE. Doi: 10.1109/CoDIT55151.2022.9804017.
- Ali, W., Cacho Estil-Les, M. D., Mangini, A. M., Roccotelli, M. and Fanti, M. P. (2023). Electric Vehicles Routing Simulation and Optimization under

Smart Charging Strategies. Proceedings of the 35rd European Modeling & Simulation Symposium (EMSS 2023). DOI: 10.46354/i3m.2023.emss.021.

- Ali, W. A., Mangini, A. M., Júlvez, J, Mahulea, C., & Fanti, M. P. (2024). Toward Enhancing Security in Intelligent Transportation: A Simulation-Based Approach. The 12th IFAC SafeProcess 2024 Symposium is to be held in Ferrara, Italy, in June 2024. (Accepted)

# List of Figures

2.1	Vehicular Ad Hoc Networks (VANETs) . . . . .	18
2.2	Visualization of the geographical distribution of the analyzed articles. . . . .	19
2.3	Digital Twin model in transportation. . . . .	19
2.4	DTN architecture for electromobility. . . . .	24
2.5	Key Technologies in DT for Electromobility. . . . .	25
2.6	Current Issues in Electric Vehicles. . . . .	33
3.1	Classification of an intrusion detection system. . . . .	51
3.2	Machine learning techniques used in IDS. . . . .	54
3.3	Basic architecture of VANETs. . . . .	65
3.4	Security in VANET. . . . .	67
3.5	Taxonomy of VANET Simulation Software. . . . .	70
3.6	Map and Vehicles in SUMO simulator. . . . .	72
3.7	High-level architecture of Veins. . . . .	74
4.1	SUMO Simulation view of Region of Puglia . . . . .	85
4.2	Traffic on connected nodes and charging points . . . . .	88
4.3	Blocked EVs in the waiting process . . . . .	90

---

5.1	Cyber-attacks goals on VANET based on CIA triad. . . . .	101
5.2	Flowchart of ID. . . . .	103
5.3	Multi-level IDS model architecture. . . . .	104
5.4	Random Forest structure. . . . .	108
5.5	Structure of XGBoost Algorithm. . . . .	112
5.6	Feature selected performance with evaluation matrices. . . . .	116
5.7	Performance of model on CICIDS2017 with different evaluation matrices. . . . .	117
6.1	Architecture of DSRC communication. . . . .	124
6.2	Vehicular Ad Hoc Networks (VANETs) . . . . .	126
6.3	VEINS Architecture with SUMO and OMNeT++ . . . . .	132
6.4	TraCiDemo11p.h implementation in OMNeT++ . . . . .	133
6.5	Sample of parameters setup in omnet.ini file . . . . .	134
6.6	Realistic traffic environment in Bologna city. . . . .	135
6.7	RSU and attackers position in OMNeT++ . . . . .	136
6.8	First attack: broadcast fake accident advertisement . . . . .	137
6.9	Second attack: DDoS attacks on the entire network . . . . .	138

# List of Tables

2.1	Classification based on the technologies and services applied for DT. . . . .	47
2.2	DT reviews comparison based on the analyzed technologies. . . . .	47
3.1	A Comparison of The Various Datasets. . . . .	62
4.1	Distance Between the Nodes . . . . .	86
4.2	Distance Between the Nodes . . . . .	88
4.3	SCENARIOS RESULTS . . . . .	91
5.1	Common attacks on availability. . . . .	102
5.2	Selected features by RF. . . . .	109
5.3	Selected features by FCBF. . . . .	110
5.4	Confusion Matrix (CM). . . . .	113
5.5	Selected features by FCBF. . . . .	114
5.6	Model evaluation with a different set of features in CICIDS2017 dataset . . . . .	115
5.7	Model evaluation results in three different cases. . . . .	118
5.8	Comparison between our work and literature. . . . .	119
6.1	Simulation Results . . . . .	139



# Abbreviations

ADAS	Advanced Driver Assistance System
AI	Artificial Intelligence
AR	Augmented Reality
AVs	Autonomous Vehicles
BMS	Battery Management Systems
BSM	Basic Safety Message
CAN	Controller Area Network
CAVs	Connected and Autonomous Vehicles
CPS	Cyber-Physical System
DdoS	Distributed Denial-of-Service
DL	Deep Learning
DoS	Denial-of-Service
DSRC	Dedicated short-range communications
DT	Digital Twin
DTN	Digital Twin Networks
DTree	Decision Tree
ECU	Electronic Control Unit
EVRP	Electric Vehicle Routing Problem
EVs	Electric Vehicles
EXT	Extra-trees classifier
FCBF	Fast Correlation Based Filter
HPO	Hyperparameter Optimization
ICT	Information Communications Technology
IDS	Intrusion Detection Systems
IoT	Internet of Things

IoV	Internet of Vehicles
IPS	Intrusion Prevention System
ITS	Intelligent Transportation Systems
IVN	In-Vehicle Networks
KNN	k-nearest neighbors
LSTM	Long short-term memory
MANET	Mobile Ad-hoc Networks
ML	Machine Learning
OMNeT++	Objective Modular Network Testbed in C++
RF	Random forest
RL	Reinforcement Learning
RRQ	Route Requests
RSU	Road Side Units
SMOTE	Synthetic Minority Oversampling Techniques
SoC	State of Charge
SoH	Battery State of Health
SUMO	Simulation of Urban MObility
TraCi	Traffic Control Interface
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything
VANETs	Vehicular Ad-Hoc Networks
VEINS	Vehicles in Network Simulation
VR	Virtual Reality
VSs	Virtual Sensors
VSA	Vendor Specific Action
WAVE	Wireless Access in Vehicular Environments
WSA	WAVE Service Advertisement
WSM	WAVE Short Message
XGBoost	Extreme Gradient Boosting

# Chapter 1

## Introduction

Within the context of intelligent transportation systems (ITS), this work completes an intensive research journey that has spanned the delicate intersections of digital twin (DT) technology and security solutions. An investigation begins with a fundamental study into the more significant uses of DT technology across numerous sectors. After navigating the complexities of this digital replication phenomenon, the research gradually evolved to focus on specific applications within the automotive industry, delving into the implications of DT in electric vehicles (EVs), self-driving vehicles, and the broader landscape of ITS.

The first steps of this journey started with reviewing DT applications in many areas and industries, resulting in a comprehensive review study on DT and ITS. This comprehensive review served as a vital point of reference for later, more focused inquiries into the function of DT in improving traffic systems and bolstering security standards within ITS.

As the research progressed, the emphasis switched to the intersection of DT concepts and network security, focusing on intelligent transport networks. During this phase, the effectiveness of applying security solutions based on artificial intelligence (AI) and machine learning (ML) algorithms gained center stage. A thorough assessment of datasets used for evaluating intrusion detection systems (IDS) resulted in a comprehensive review laying the foundation for this research's subsequent stages. In alignment with the DT principle, the study seamlessly flowed from theoretical inquiry to practical implementations. The utilization of

freely available simulation tools for traffic systems, particularly in the Puglia region of Italy, became a focus point. Simulation-driven optimization studies emphasized the influence of traffic on travel times, underlining the important need for validation and potential revisions for best outcomes. In this part of the research, the practical simulation proves the significance of simulation tools in addressing routine optimization challenges.

Simultaneously, the network security-focused component of the research journey emerged by building a multi-level IDS based on ML techniques. This proposed system was evaluated on CICIDS2017 dataset, showcasing its excellence in detecting a wide range of attacks on Vehicular Ad-Hoc Networks (VANETs). This achievement became a cornerstone for tackling security issues within ITS.

The research progressed with an in-depth examination of network simulation technologies that could be seamlessly connected with traffic simulation systems. The integrated simulation of part of the Italian city of Bologna proved important, yielding a realistic network dataset that will serve as a core resource for future research and system evaluations. The primary purpose of this simulation-driven was to study the network communication protocols within VANETs to improve vehicle network security. By understanding wireless access in vehicular environments protocols and messages, we can employ advanced detection algorithms and undertake careful analysis to detect and block network attacks.

This PhD thesis is a complete trip through the confluence of DT technology and security solutions inside ITS. From foundational explorations to specialized inquiries and theoretical frameworks to practical applications, this research contributes to the evolving discourse on the subject. It highlights the depth and breadth of knowledge already gained in the dynamic and evolving field of ITS.

## **1.1 Background**

### **1.1.1 Digital Twin**

The advent of the DT concept has heralded a transformative era in modeling complicated physical systems. DT is defined as a virtual representation of a

physical object, system, or process, constantly updated with real-time data from its tangible counterpart [1]. This dynamic synchronization enables the DT to serve as a substitute for the physical system, opening opportunities for simulation, prediction, and optimization. Adopting DT becomes crucial in literature and research as these entities become integral to different industries.

DT demonstrates the capacity to effect substantial transformations across various sectors, including manufacturing, healthcare, and infrastructure management. The utilization of such technology in the manufacturing industry is crucial due to its ability to enhance product design, optimize production processes, and minimize downtime [2]. This technology is essential in determining the trajectory of the industrial sector due to its significant effects on the efficiency and productivity of manufacturing processes. In the healthcare sector, DT is being employed to facilitate the creation of novel therapies, ensure ongoing health surveillance, and deliver individualized attention to patients; this signifies a fundamental change in the methodology of healthcare provision [3]. In infrastructure management, DT offers solutions to improve traffic flow, predict asset failures, and increase resource utilization. This phenomenon underscores DT's critical role in forming more intelligent and sustainable municipalities, which has substantial implications for urban development and planning [4]. The multidisciplinary nature of DT raises significant scholarly interest in the subject. Academically, researchers conduct various research on the complexities of their implementations and the far-reaching impacts they apply to multiple industries.

The fundamental components become apparent upon closer inspection of the technical complexities of DT. These elements comprise the digital representation, which is a communication channel that facilitates real-time data exchange between the physical system and its DT model. The physical object represents the tangible system mirrored in the DT, and the communication channel integrates data from sensors, actuators, and other supplementary sources.

Many modeling techniques can be employed to construct DT, including physics-based, data-driven, and hybrid models. The availability of big data from these applications determines the modeling technique option and creates opportunities for a thorough investigation into the efficacy of various modeling methodologies using DT techniques.

Nevertheless, the extensive implementation of DT is full of obstacles. The collecting, storing, and processing of large volumes of real-time data presents a notable challenge in the realm of data management [5]. This is considered a current challenge in the research community concerning infrastructure, technology, and the creation of data administration systems. Model complexity represents an additional obstacle, as the development and upkeep of precise and current models of intricate systems necessitate substantial labor and proficiency. In addition, data security and privacy are of the utmost importance in DT, necessitating scrupulous consideration to safeguard the confidentiality and integrity of the information undergoing processing.

However, despite these obstacles and challenges, the prospective direction of DT is promising. Technological developments such as the Internet of Things (IoT), AI, and ML are expected to contribute to the increased sophistication and broader utilization of DT.

The application of DT in ITS seamlessly aligns with the broader vision of smart cities, where adaptive and responsive systems are at the heart of urban development [6]. By employing real-time data generated by vehicles, infrastructure, and an intricate network of sensors, DT fosters a comprehensive understanding of traffic scenarios. This, in turn, enables predictive modeling and scenario simulations, paving the way for transportation systems to adapt dynamically to ever-changing conditions. While DT has already proven its transformative potential across various sectors, its impact on ITS unveils an exciting new frontier ripe for exploration. This thesis delves explicitly into the realm of DT and its intricate applications within the ITS. Throughout the subsequent chapters, we will explore this transformative technology comprehensively and its potential to provide solutions to transportation infrastructure in smart cities.

### **1.1.2 Intelligent Transportation System**

ITS has unquestionably become a powerful influence, completely transforming the transportation industry to improve efficiency, safety, and sustainability in our constantly growing mobility networks [7]. This paradigm change is demonstrated by a wide range of technologies, including real-time traffic management systems,

intelligent traffic signal control, the rapid growth of EVs and autonomous vehicles (AVs) technologies, and the widespread of vehicular networks, including vehicle-to-vehicle (V2V) and vehicle to infrastructure (V2I) communication [8].

EVs operate on electricity rather than traditional fossil fuels and are crucial in addressing air pollution and decreasing greenhouse gas emissions [9]. This makes EVS a vital component of a cleaner and more sustainable transportation system [10]. AVs, which are equipped with autonomous driving capabilities, can enhance safety and improve traffic flow simultaneously. This envisions a future where accidents and congestion on roadways are minimized [11].

The integration of EVs and AVs into the current transportation system is a challenging but crucial task. This requires the deployment of ITS to navigate this era of transformation successfully. The role of ITS is vital as it enables the efficient management of EV charging infrastructure, enhances AV routing techniques, and delivers up-to-date information to drivers and cars. This coordinated collaboration aims to promote a balanced and effective transport system that accommodates and actively encourages the widespread use of these innovative technologies [12].

Beyond this integration, ITS offers extensive benefits, including optimizing traffic flow, reducing congestion, and improving public transportation efficiency. As a result, travel times are decreased, and overall productivity is increased [13]. In addition, ITS substantially contributes to safety by providing up-to-date traffic information and alerts, promoting safe driving behaviors, and facilitating automated emergency response systems [14]. However, this transforming journey has its significant challenges. Integrating ITS systems with existing infrastructure provides complicated logistical and technological issues, requiring careful consideration of compatibility and interoperability. Ensuring adequate data privacy and security methods becomes crucial, mainly as these systems rely significantly on real-time data transmission and communication [15]. Moreover, the high expenses associated with establishing and maintaining ITS infrastructure demand careful planning and efficient resource allocation to assure long-term sustainability [16].

### 1.1.3 DT and ITS

In these advancements in technology and revolutionary changes, the idea of DT is emerging as a crucial factor in enhancing and optimizing the capabilities of ITS. When it comes to transportation, combining DT with ITS has the potential to improve efficiency, safety, and sustainability objectives. DT can function as dynamic representations for specific vehicles and the broader transportation infrastructure, such as network components and charging infrastructure. Moreover, DT is essential in forecasting energy usage trends, optimizing charging timetables, and evaluating the real-time condition of EV components. DT can replicate various situations within the AV industry, optimize routing algorithms, and improve decision-making processes by thoroughly comprehending the vehicle's virtual model. Furthermore, this integration can significantly enhance predictive maintenance techniques and total system reliability by continually monitoring the status and performance of vehicles and infrastructure components. This anticipates possible difficulties and enables proactive maintenance, thereby minimizing downtime.

In the context of ITS security, this real-time capability becomes a strategic advantage, where DT enables ongoing monitoring of vehicular networks, analyzing real-time data streams to promptly detect anomalies in data traffic, communication patterns, and overall system behaviors. By leveraging this real-time data, DT contributes to the identification of potential security threats, cyberattacks, and unauthorized access attempts. Moreover, DT facilitates the implementation of predictive security measures by simulating various real-time scenarios, attacks, and vulnerabilities. This simulation empowers system operators to anticipate and mitigate real-time risks, enhancing vehicular networks' overall resilience against evolving cybersecurity challenges. In summary, the availability of real-time data and employing it in designing the DT significantly reinforces the security framework of ITS, ensuring a dynamic and responsive defense mechanism in the face of emerging threats.



## 1.2 Motivation

The evolution of EVs and AV or self-driving into Intelligent ITS represents a significant leap forward in refining traffic management, ensuring drivers' and passengers' safety, and fostering sustainability within smart mobility networks. As self-driving vehicles strive to diminish road accidents and alleviate traffic congestion caused by human error [17], EVs contribute to a cleaner and more sustainable transportation paradigm by relying on electricity instead of traditional fossil fuels [18]. Notably, recent research underscores challenges in the security and privacy of communication systems within self-driving and EVs [19], [20]. Addressing these challenges is the core focus of this thesis, which uniquely explores and introduces route optimization strategies, IDS for VANETs, and simulation tools of traffic networks and their communication dynamics to fortify the security of VANETs.

The thesis draws motivation from four crucial factors applicable to electrified and self-driving automobiles. The Security Factor emphasizes the need for a comprehensive security system that is both lightweight and durable, capable of online detection, and specifically designed to boost protection against potential threats. Furthermore, the Economic Factor emphasizes the significant economic consequences of optimizing routes for driverless and EVs, resulting in cost reduction and increased road capacity. The attacks Type Factor acknowledges the limitations of conventional security measures in identifying different forms of assaults on the external communication of autonomous and EVs. This underscores the significance of IDS in enhancing security. Finally, the Safety Factor prioritizes the creation of an intelligent, responsive response system for both types of vehicles, guaranteeing a prompt switch to safe modes to protect passengers and drivers.

To address these motivations, the thesis introduced a tree-based ML model using techniques such as decision trees, random forests, XGBoost, and extra trees. These techniques play a pivotal role in enhancing detection rates and minimizing false alarms within the framework of both autonomous and EVs. These technologies have the advantage of real-time operation, fault tolerance, and self-organization. Furthermore, the unique contribution of this thesis lies in the simulation of vehicular networks that serve in the application of route optimization strategies. Moreover,

the simulation of communication in traffic networks enhances the ability of IDS to identify different types of attacks on the network. This simulation not only refines the security of VANETs but also ensures a more resilient and responsive defense mechanism against emerging cybersecurity threats in electric and self-driving vehicles.

## 1.3 Problem statement

The development and operation of EV and AV systems in ITS highlight the importance of growth in communication systems. Recent research highlights the significant difficulties associated with ensuring the security of communication systems in ITS [21]. Understanding the need to safeguard these networks and implementing strong security measures is essential for promoting progress and universal acceptance of the revolution in vehicular networks in ITS [22].

Nevertheless, differentiating between normal and abnormal/malicious behavior in vehicular communication, particularly VANETs in ITS, is a complex challenge. The complexity stems from the constantly developing infrastructure and the volatile physical environment inherent in vehicular networks. The security problems in VANETs, a vital component of ITS, can be classified as follows [23]:

- **Online Detection:** Real-time identification of online threats is crucial when designing security systems. In ITS, online detection is vital for addressing security threats in dynamic vehicular networks. Rapid response is crucial for ensuring passenger safety and maintaining the integrity of real-time communication, which is essential for applications like collision avoidance. This ensures that packets can be exchanged between the source and target without any delay, thereby ensuring the safety of passengers and drivers.

- **High Mobility:** The detection algorithm depends on network behavior to collect essential characteristics. However, the high mobility in VANETs complicates this process, making collecting information about the type and number of features challenging. Technologies like fuzzification are necessary to bridge this gap caused by the rapid mobility inherent in VANETs.

- **Security Systems:** Conventional security systems are inadequate for protecting the communication systems of vehicles within VANETs. Therefore, developing new security systems or changing existing protection mechanisms is necessary.

- **Internal Attacks:** Although successful in countering external attacks, encryption methods are insufficient in thwarting internal attacks within VANETs. The absence of adequate security measures motivates researchers to create sophisticated IDS that can identify and stop internal attacks, therefore enhancing the security of VANETs within the larger context of ITS.

DT plays a pivotal role in digitally mirroring the environment, providing a foundation for developing software and models that effectively simulate real-world networks. It goes beyond simulation and representation, incorporating elements of construction, design, operations, results analysis, model experimentation, and generating outcomes that closely parallel reality. This comprehensive functionality empowers researchers and specialists to thoroughly assess the digital model, understanding its strengths and limitations before constructing the physical counterpart. The main focus of this thesis is to address VANETs security challenges by building the security framework in DT technology. This digitalization aims to strengthen and enhance protection mechanisms, ensuring the resilience of communication systems and contributing to the secure development of VANETs functionalities within ITS.

## 1.4 Thesis Contributions

The thesis contributes to addressing several challenges related to sustainability and security in ITS. The primary objective is improving vehicle network security effectiveness within ITS. In addition to the focus on security, the contributions made by this research extend to sustainability, including cost reduction and the implementation of smart charging strategies. The major contributions are outlined below:

- **Comprehensive Review of Digital Twin Technologies:** Review of DT in ITS: This work is the basic building block on which this thesis is based, as all previous studies related to DT were reviewed, starting from their inception

until their use in various fields, especially in ITS and smart cities. The exploration encompasses technologies integrated with DT, emphasizing their impact on the development of vehicular networks. Notably, the study delves into the capabilities of DT to enhance security measures and contribute to the sustainability of vehicle networks through real-time monitoring and response capabilities.

- **Intrusion Detection Systems for VANETs:**

The thesis proposes a new IDS designed explicitly for VANETs. The proposed IDS uses decision tree-based ML techniques. The proposed model combines two feature selection techniques, Random Forest (RF) and Fast Correlation-Based Filter (FCBF), and four tree-based algorithms, extreme gradient boosting decision tree (XGBoost), RF, Decision Tree (DTree), and ExtraTrees classifier (EXT), for classification performance. Moreover, the model is evaluated using CICIDS2017 as a benchmark dataset and Python ML libraries in Jupyter Notebook, such as scikit-learn and Pandas. Experiment results show that the proposed model using the stacking method achieves 99.86% attack detection accuracy, 99.85% precision with Hyperparameter Optimization (HPO), and 99.83% attack detection accuracy without using HPO.

- **Routing Optimization under Smart Charging Strategies:**

The Electric Vehicle Routing Problem (EVRP) is a crucial topic that has been addressed in this thesis to solve the problem of optimizing the routing of EVs for logistics operations. The optimization approach solves the EVRP and is formulated as an integer linear programming problem. The goal is to minimize the charging/discharging cost, considering the shortest path for each EV that must deliver a charge to a group of customers. The methodology integrates smart charging strategies, contributing to the sustainability aspect of ITS by minimizing charging costs. Moreover, to validate the performance of the proposed optimization method, Simulation of Urban MObility (SUMO) software was used to model and simulate the solution of the EVRP. To demonstrate the effectiveness of this method, a real case study in the Puglia region (Italy) was considered. Additionally, different

traffic scenarios were simulated in the SUMO environment, emphasizing the impact of traffic on travel times, thereby addressing sustainability through efficient routing and reduced travel times.

- Simulation of traffic and communication protocols:

**Comprehensive Network Modeling:** The thesis provides a comprehensive modeling framework for understanding complex interactions within ITS by simulating traffic networks and communications protocols. The model has been built within the dynamic framework of an urban setting, employing the SUMO simulation testbed for the city of Bologna, Italy, as a realistic traffic simulation model. This study uses a hybrid methodology integrating the SUMO, Objective Modular Network Testbed in C++(OMNeT++), and Vehicles in Network Simulation (VEINS) framework to simulate communication within vehicle networks. The main emphasis is examining the different types of attacks that could happen on the VANETs network through various messages facilitated by the IEEE 802.11p protocol / WAVE (Wireless Access in vehicular Environments) standard. The primary objective of this simulation-based study is to improve vehicular network security by leveraging the inherent benefits of using WAVE standard messages and applying and analyzing techniques to detect intrusions in the network.

In summary, the contributions of this thesis transcend the realms of security and extend into sustainability. Leveraging DT capabilities enhances vehicular network cybersecurity, while the proposed IDS achieves remarkable attack detection accuracy. Additionally, smart charging strategies optimize electric vehicle routing, minimizing charging costs. The simulation-based study on traffic networks and communication dynamics contributes to improving vehicular network security. These contributions collectively enhance ITS security and sustainability, addressing crucial challenges and fostering a more secure and efficient transportation ecosystem. These diverse contributions will be detailed and discussed in the subsequent chapters, highlighting their profound impact on advancing the field of ITS.

## 1.5 Thesis Structure

This thesis is presented in three main parts, comprising six chapters in total. Part I deals with Optimizing Electric Vehicle Routing in Large-scale Traffic Networks using simulation tools. Part II presents an IDS for Vehicular ad Hoc Network Attacks Based on ML Techniques. Finally, Part III introduced a simulation-based study for enhancing security in intelligent transportation.

Starting from Chapter One, the introductory background of the domain, motivation towards the study, Problem Statement, and Thesis Contributions have been inscribed. The start-of-art was divided into two chapters to study the thesis's main components separately and then link them in subsequent chapters. Chapter two presents the background of DT in ITS and all emerging techniques that apply to DT technology. Chapter three discussed the background of communication security for ITS, particularly in VANETs and EVs routing problems and optimization.

Part I starts by illustrating chapter four, which spans the study of the applications of Routing Optimization, followed by the technological challenges towards the ITS and EVRP. The chapter also discusses the simulation tools widely applied to solve EVRP and explains the methodology used in modeling and addressing the problem. A Case Study was conducted in this chapter to solve EVRP in the Southern Italian region of Puglia. At the end of this part, the result was presented with a comprehensive discussion and future work.

Part II embraces the Intrusion Detection System for VANETs. Chapter Five presents a background on the Intrusion Detection System for VANETs, as well as techniques and tools. The chapter comprehensively explains the proposed IDS model and the methods, followed by the experimental setup for the model and the results achieved with discussion. The chapter sheds light on IDS significance in ITS security and potential future work.

Part III illustrates the thesis contribution of security and ITS. Chapter Six contains a comprehensive examination of VANETs security employing simulation tools, both Traffic and network simulation software, starting from a study on the simulation approach for VANETs security and the possible attacks that could be performed on VANETs simulation. The chapter discussed the network protocols

---

and standards for vehicular networks, such as IEEE 802.11p protocol / WAVE, and simulation tools like SUMO, OMNeT++, and VEINS. This chapter conducted a case study by simulating the entire vehicular network and injecting different types of attacks. The study aimed to highlight the importance of practical and realistic simulations and the importance of network protocols to secure VANETs. The study concluded by presenting the results and discussing future work.

As a conclusion of this dissertation, Chapter 7 provides final reflections, emphasizes vital discoveries, and presents a future outlook, offering valuable insights for the broader research community.

## **Chapter 2**

# **Background about Digital Twin in Intelligent Transportation Systems**

### **2.1 Introduction**

In the era of the fourth industrial revolution, EV industry is transforming the manufacturing field of transportation systems. Today, most of the features and services of EVs are realized through smart technology. Conventional vehicles with internal combustion engines significantly contribute to the consumption of fossil fuels and the emission of greenhouse gases, such as carbon oxides and hydrocarbons. To overcome this issue, EVs have been designed and improved over the past few years [24]. Many industrial fields adopted IoT technologies to enhance the electromobility industry and make this transformation smarter. Vehicles are becoming smart objects using sensors that form the basis for IoT networks. In turn, the amount of data these sensors provide will constitute a qualitative shift in the concept of EV management systems. These sensors will cover all vehicle parts to monitor all movements and changes during the movement and charging, as well as monitor engines and the internal components. The generated data from the sensors need to be collected and synchronized, then analyzed and processed to improve EV service quality and assist EV management system in decision-making. Despite the significant development of these technologies, there are still difficulties in using physical sensors; this led many researchers to introduce the concept of



virtual sensors (VSs) for electromobility [25]. The VS system can analyze, predict, and estimate the vehicle behavior, Battery State of Charge (SoC), and availability of charge points.

Improving the EV user experience relies on three essential components: the physical entities in the real world, the virtual models, and the data-driven by these models. Integrating these components requires a simulation framework to simulate large-scale traffic scenarios [26]. The distribution of charging points, EV volume, and all dynamic operations in the EV network should be managed effectively and safely. For this goal, simulation platforms can be introduced to simulate the EV's network components and their interaction with each other. The simulation of operations could help us understand the nature of the physical product in each stage and collect information about the product characteristics, which can be helpful in the development process. Most simulation platforms support the concept of DT, which provides an excellent capability to simulate real-world entities in the industrial environment. DT concept is known as a virtual replica of a real-world object that can give the ability to study the development of physical objects in a digital situation/environment. DT was considered one of the world's ten latest strategic innovations in 2019, providing autonomous objects (e.g., self-driving cars), immersive technologies such as virtual reality (VR), augmented reality (AR), and quantum computing [27]. The main idea of this technology is to replicate the physical object's behavior in a virtual environment that can produce the same output as the actual physical object.

Adopting DT can reinforce the development of the industry and the academic section. Digital data can improve an engineering system's intelligence concerning analytical evaluation, extrapolative diagnosis, and performance optimization. Then, the results of the analyses can be used to make the product or process run better in the physical environment [28]. Academically, DT concept was introduced in 2002 by Grieves et al. [29] in a special summit on product life-cycle management at the University of Michigan Lurie Engineering Center. The first adoption of this technology is by Tuegelet al. [30], who presents a digital framework to reproduce the structural behavior of an aircraft. Indeed, DT technology is widely used in multiple industrial sectors to facilitate maintenance operations and predict failures, allowing machines and humans to interact with each other. In particular, DTs

are used in a wide range of applications, including transportation, manufacturing, medicine, business, education, and more. Integrated machine-driven, electrical, and computer software systems can be simulated in the virtual workspace through DT technology [31]. One important example of such systems is EV, for which new technology is needed to optimize vehicle performance continuously. DT technology can be an innovative solution for EV optimization. Many features of an electric EV can be computerized to increase its efficiency, performance, and smartness.

Integrating big data analytics, IoT, and AI technologies with DT leads to new significance, prospects, and challenges. Furthermore, an intelligent DT model can only be created using advanced AI technologies applied to the data [32]. It can tune the significant challenges for manufacturing, such as improving process stability, fault diagnosis, reducing downtime, and optimizing logistics processes [33]. AI can further enhance DT technology by using analytical models to process raw data into valuable digital forms. In this context, ML algorithms and technologies are currently used in EVs. In particular, ML algorithms can be used effectively if combined with predictive testing tools and DT technology. The importance of DT is also proved in the security and monitoring systems. Lu et al. [34] presented a DT-enabled anomaly detection system based on industry foundation classes for asset monitoring solutions. The proposed framework was evaluated using a case study to control the Heating, Ventilation, and Air Conditioning system, and the system efficiently contributes to monitoring building assets.

There is a lack of literature reviews that explore the use of DT in ITS, especially for EVs and AVs. Due to the rapid development of the smart mobility industry, there is an urgent need to study and analyze the challenges and issues that the new generation of transportation will produce and how these can be addressed [35]. In response to the growing adoption of DT technology in ITS, this survey investigates recent literature, particularly emphasizing the application of DT in electromobility and self-driving systems. By exploring the integration of emerging technologies, communication tools, and DT concepts, our findings reveal the potential of DT technology to address challenges such as cost-effectiveness, reliability, visualization, charging time, and intrusion detection in electric and AV networks. Furthermore, this survey highlights the role of data analytics and ML techniques

in securing ITS networks and improving overall efficiency. The research was conducted in multiple databases, including Scopus, Google Scholar, and Web of Science, to identify relevant studies published in the lasquicky years. In addition, ResearchGate was considered as an additional database and information source.

The provided review is carried out systematically, as shown in Figure 2.1, considering specific domains within smart manufacturing and transportation in which DT technology is applied in combination with IoT, ML, AI, and 5G technologies. First, a set of papers was selected from the publication databases considering the following keywords: DT, ITS, EVs, Big data, 5G, IoT, and AVs. From such a paper set, a subset of articles is selected on the basis of their relevance and quality. Then, the current issues in EV services, such as tracking, monitoring, Battery Management Systems (BMS), connectivity, privacy, and security, are discussed, as well as how they can be addressed effectively through DT technology. The next EV revolution has been highlighted, i.e., electric AV, and the importance of data analytics roles in applying DT in such a context has been discussed.

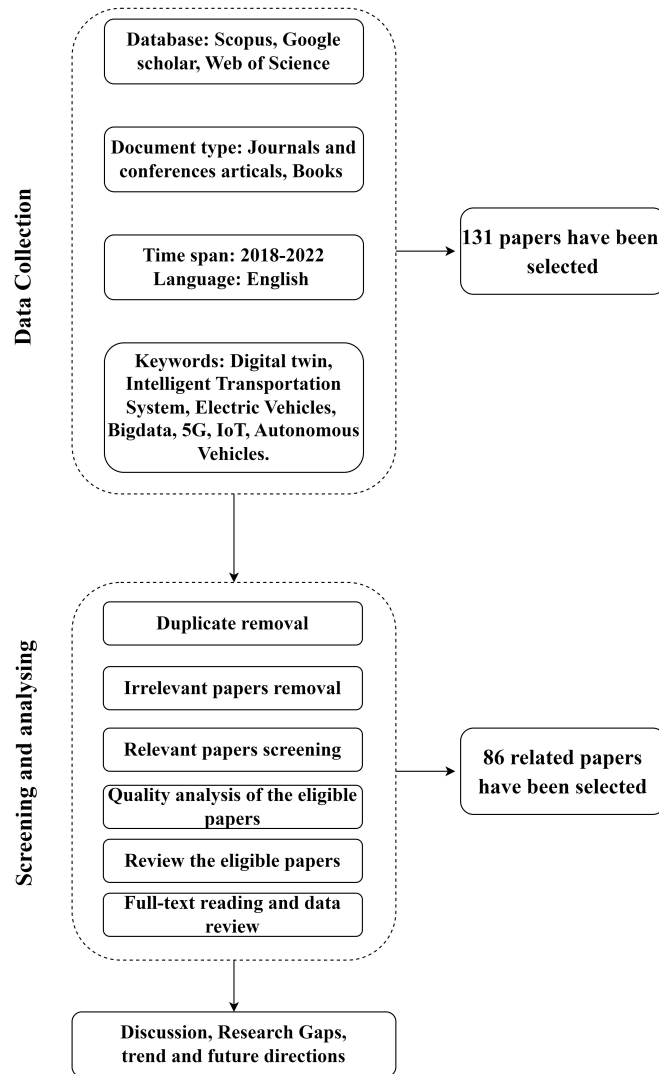


Figure 2.1: Vehicular Ad Hoc Networks (VANETs)

In figure 2.2, a geographical distribution map has been created to display the dispersion of the research articles examined in this chapter. The map visually represents the number of papers originating from each continent, providing a clear overview of the global research landscape in this field. This illustration allows readers to easily grasp the contributions on DT and ITS made by researchers from different parts of the world, showing the regions with the most research output.



Figure 2.2: Visualization of the geographical distribution of the analyzed articles.

## 2.2 Digital Twin in Transportation Systems

With the advancement of Big Data, IoT, and AI, a new generation of information technology, geographic and global positioning data, is to be handled. Combining these technologies with DT technology is a critical element of the digital wave trends and takes the lead in transportation applications for planning, maintenance, security, and other aspects.

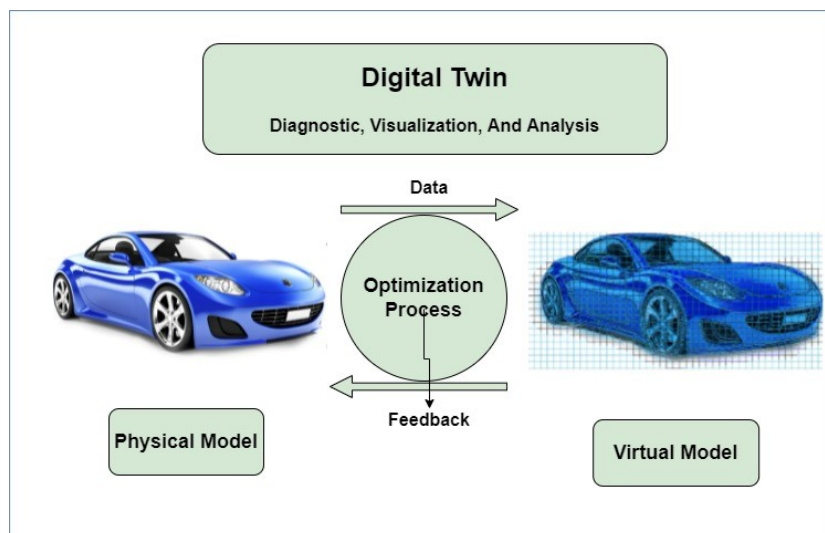


Figure 2.3: Digital Twin model in transportation.

The DT has the potential to improve the transportation sector by providing a digital identity, synchronized visualization, and virtual and real interaction (see Figure. 2.3). The DT technology utilizes intelligent technical advantages such as controlling traffic perception, road warning, and emergency response. Furthermore, it can provide transportation solutions and new paths, such as intelligent driving, which will increase efficiency and safety and allow convenient traffic management. In this context, Wang et al. [36] introduce a DT framework for connected vehicles using an Advanced Driver Assistance System (ADAS). They used vehicle-to-cloud communication to calculate the advisory speed based on the information that could be collected from the sensors on the vehicles. The proposed model helps the driver to control the speed intelligently. Another example of using DT in the cloud is presented by Alam and Saddik [37], which developed a DT model for the cloud-based Cyber-physical system (C2PS). They described the key properties of the C2PS and introduced a telematics-based prototype driving assistance application for the vehicular C2PS.

In this context, a review paper on DT technology with smart EVs has been done by Bhatti et al. [38]. The review has divided the smart vehicle systems into different categories: autonomous navigation control, advanced driver assistance systems, vehicle power electronics, vehicle health monitoring, BMS, and electric power drive systems. As a result of this research, smart EVs and DT technology are investigated theoretically to see what impacts their integration can have in the near future.

Due to the amount of data generated by transportation systems, ML and Deep Learning (DL) technologies are employed to create an ITS. The application of intelligence in the transportation field is increasing rapidly, and improving the performance of these systems has become the focus of research. The existing transportation systems can be controlled, analyzed, and operated using DT. Using ML and DL in DT can collect real-time data and provide adequate services to the service provider and the end-user [32]. Moreover, the ITS can effectively optimize and coordinate traffic conditions based on DL and DT technologies by monitoring the flow of people, traffic, and roads. It can also maximize the duration of traffic lights and find the signal light scheme with the shortest transit time. For example, Zhihan et al. [39] proposed a DL algorithm to solve the security problems of

the ITS. The proposed model assured a response time to emergency alerts and increased the prediction accuracy. Moreover, vehicles will travel faster because they can better adapt to the road environment, transmit data more quickly, and develop routes considering traffic patterns. Another example of how DT and AI technologies are utilized in transportation is traffic management, prediction, and congestion avoidance. Kumar et al. [40] introduced an ITS that uses ML, fog/edge analytics, data lakes, DT, and blockchain. The authors used cameras to collect environmental information and then run edge analytics on the collected data. The DT generated the virtual car model to simulate the real-world scenario. This work uses ML and DL algorithms to predict drivers' intentions. By creating a virtual vehicle model, non-autonomous drivers were able to make better decisions depending on the current traffic scenario and the intents of other drivers.

The traveler's driving experience is also important, as the DT can reduce and redistribute waiting time at intersections. Dasgupta et al. [41] worked on a DT approach for adaptive traffic signal control to improve the user driving experience. They developed DTs to emulate vehicles close to the intersection and vehicles' waiting time at the immediate upstream intersection. The proposed model can balance waiting time across a signalized network to improve the travel driving experience in congested areas, and it can be scalable on the city-wide network. While the Data analytics in the DT concept is still developing, Aslani et al. [42] developed a real-time DT simulation model that can provide a performance measure. The study also demonstrated the data scarcity required for real-time applications that rely on high real-time frequency connected corridor data streams. In summary, the DT uses all the gathered data and accurately captured city signs to achieve new insights into urban traffic on different sides, such as the road supply and traffic demand, optimizing the road network structure through traffic simulation, and improving overall traffic efficiency in the city. In addition, using DT in transportation can improve the decision-making execution, safety, and stability of vehicle driving and accelerate intelligent and safe driving.

### 2.2.1 Digital Twin in Electromobility

Many applications of electromobility have been involved in research activities to give rise to the phenomenon known as smart electromobility. The rapid growth of smart control systems led to several developments in this industry. The data generated through this growth are the key factor in improving the smart mobility sector. The DT's strength lies in collecting and visualizing data and conducting statistics in which advanced analysis tools are used to improve manufacturing processes and help decision-making.

For charging an EV, it is commonly vital to physically attach/connect the EV plug to a charger located in a household or a public place through a charging cable. Considering that EVs and self-driving accomplished by automated driving will derive into general use, physical charging is not manageable, and automatic charging should come in place. Generally, there are two options to charge automatically: park a vehicle in an accurate position so that the vehicle's charging connector of the vehicle automatically fits the charging cable of an available charger or wireless charging. Shikata et al. [43] introduced a DT's vehicle simulation technique, focusing on two factors: 1) power consumption and 2) ride comfort. The simulated environment contains a vehicle model for interpreting the physical performance of the vehicle. An electronic control unit (ECU) has also been simulated as a prototype for regulating the simulated environment. They also developed an automatic charging system for EVs to charge the vehicle automatically after parking in an accurate position.

BMS in electromobility is also essential concerning battery life, safety, and reliability. It relies on different types of sensors and actuators on the EV to provide real-time battery performance. Using an IoT platform to build a DT for BMS in the cloud boosts the robustness of the BMS. Wang et al. [44] reviewed the solutions for BMS issues based on DT, such as the problems related to real-time estimation, dynamic charging control, and dynamic equalization control in a smart BMS. Another important application of the EV is the ADAS, built to enhance driver experience and passengers' and pedestrians' safety by decreasing vehicle accidents and alerting drivers of possible dangers. Liu et al. [45] introduced a new vision-cloud data fusion approach to enhance the performance of visual guidance



systems by leveraging DT technology and cloud servers. This work is one of the effective studies to visualize the cloud DT data and support the ADAS or driver's decision-making.

### **2.2.2 Digital Twin Networks in Smart Transportation**

Digital Twin Networks (DTN) is the natural evolution of the development of DT technologies in the modern era. DT of any physical object is the first cell of the DTN; thus, DTN can be defined as a set of virtual digital representations of different groups of physical objects connected by a high-speed communication medium to configure an integrated virtual system. The data exchange between the virtual model and physical object in the DT is done in a one-to-one unidirectional way. The operational changes in the physical object will directly affect the virtual model but not the opposite.

On the other hand, DTN allows comprehensive data exchange between DTs and physical assets in a multidirectional manner [46]. Recently, transportation has encountered issues that increase with the development of urban cities, such as traffic congestion and accidents. In this context, DTN can provide a better solution for such a complex environment and help to optimize the entire transportation system. DTN also offers innovative transportation services such as traffic information reporting, vehicle security, and data sharing. To keep pace with the rapid progress in the electric mobility sector, we need to use and integrate DTN technology with EV networks in smart cities, whether autonomous or non-autonomous vehicles, which will provide high possibilities for managing and improving transportation network systems, not only at the city level but also at a broader level.

The DTN architecture shown in Figure 2.4 is composed of three layers: physical, network, and virtual. The physical layer consists of EVs, charging stations, roads, and facilities. These entities are connected to the network layer through sensors that transmit data about the vehicle positions and velocities, the road traffic, and the charging station status. The network layer receives information from the physical layer by communication services provided through 5G or WIFI technologies. Moreover, this layer sends information and data to the virtual layer, which is composed of a network of DTs and servers. At the virtual level, the DTs

are connected to collaborate in executing the simulation and computation tasks based on new enabling technologies devoted to decision-making, analysis, and maintenance issues (such as AI, AR, and ML). In this context, Dai et al. [47] proposed a new DTN model to build network topology and integrate it with the IoT network. The adopted system significantly solved many problems, such as computation offloading and resource allocation problems.

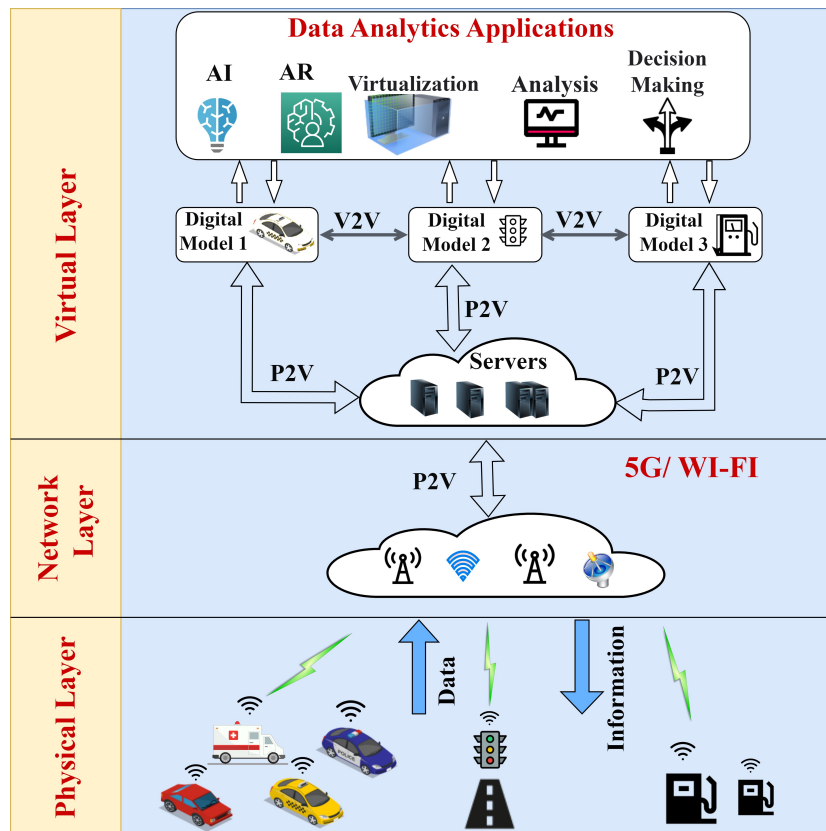


Figure 2.4: DTN architecture for electromobility.

## 2.3 Technologies in Digital Twin for Electromobility

DT applications are increasingly used in electromobility and combined with enabling technologies like IoT, 5G, big data, ML, virtual systems, and advanced communication interfaces. Critical functionalities such as real-time monitoring,

predictive analysis, or cloud computing may be impacted. However, the DT's main concept and basic architecture were always common with all these technologies. Figure 2.5 shows the key technologies to be used with DT for electromobility, such as the IoT, VSs, 5G, Data analytics, and AVs. We will discuss the main contributions of DT in detail in this section and how DT technology will provide services to heterogeneous fields in different communication networks.

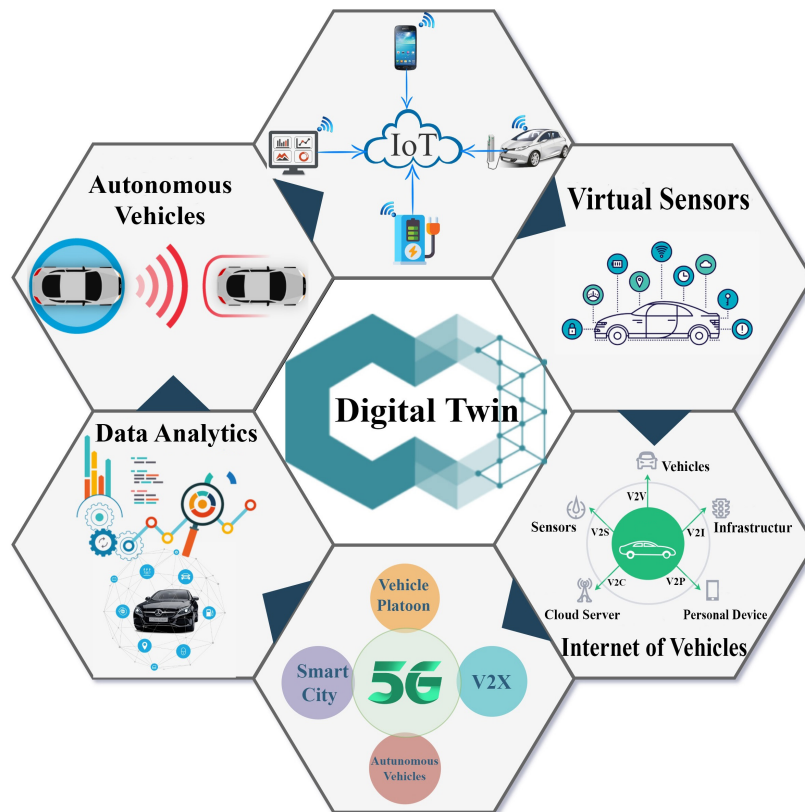


Figure 2.5: Key Technologies in DT for Electromobility.

### 2.3.1 Internet of Things

Recently, IoT technology has been used in smart and electric mobility. In the electric mobility revolution, the application of DT technology is facilitated by advanced data analytics and IoT. Digital and physical interactions are changed by integrating DT and IoT platforms. IoT enables connection and intelligent access

to physical devices, and the DT can handle challenges related to integrating IoT and data analytics, which facilitates rapid real-time analysis and decisions. In electromobility, IoT establishes a wide platform with connected vehicles that can send the data from physical devices to the cloud or local servers. Then, the role of the DT lies in dealing with this information, simulating resources by creating DT models, establishing virtual connections, and integrating with AI. Therefore, the DT technology is a substantial technology to improve performance innovatively in electromobility and advances monitoring, analytics, and predictive capabilities [48].

Zhao et al. [49] introduced an IoT-based and DT-based model that enabled tracking solutions for safety management. The proposed framework allows a real-time safety tracking mechanism for detecting stationary behavior and self-learning genetic position to recognize abnormal conditions and obtain an accurate location. The combination of IoT and DT technologies can help operators and stakeholders make the necessary technological improvements for making electromobility smarter by connecting DTs of smart vehicles, simulating and managing the EV fleet data network for value-added services that can not only improve the driving and charging experience of users but generating benefits for the entire sector. Being able to simulate the connection of EVs through IoT can, for instance, allow for optimally scheduling the recharge of such vehicles by prioritizing the charging operations based on real-time states of charge, available infrastructure, and user preferences. Finally, DT technology facilitates smart functionalities by leveraging data gathered consistently from IoT devices. Moreover, DT enables features such as predictive maintenance, network analysis, energy efficiency optimization, streamlined resource distribution, and real-time network supervision. Additionally, seamless interaction between various networking devices is possible, as their digital counterparts are platform-agnostic and can be managed using standardized methods without concern for the specific technical details of individual devices [50].

### **2.3.2 Virtual Sensors**

The EVs use many environmental sensors to perceive and act according to their perceptions. This section illustrates how logical entities called VSs can be used

in a DT framework as the virtual twins of physical sensors. VSs can support new services for smart transportation, addressing issues related to battery charging and route planning for EV drivers based on parameter estimation/prediction. In particular, VS derives new data from existing information generated by the physical sensors and utilizes a data processing algorithm to process the data input and produce the required output [51]. For example, Roccotelli et al. [25] introduced and designed new VSs to enhance the EV charging experience. The proposed model provides a smart charging service that allows drivers to find the best charging point for their vehicles. Another example is given by Gruosso et al. [52], who proposed a methodology for estimating the state of charge in EVs. The method relies on the VS and other measurements available in the vehicle, such as speed, acceleration pedal position, and battery voltage. VSs also play an essential role in enhancing user experience and optimizing EV services, which could support the growth of the EV market.

In [53], Fanti et al. developed a new EV service to improve user experience when preparing for the trip. They designed three VSs that help the driver predict the cost and required energy for the journey. The proposed VSs can estimate the energy demand based on historical data from past trips. Some relationships between the sensors can be determined over time by the virtualization platform, for example, by utilizing and exploiting ML technologies to improve the functioning of the sensors. Therefore, combining the technology of VSs with the EVs simulation model can provide a way to solve complicated issues such as battery management, vehicle energy management, and vehicle control.

### **2.3.3 Internet of Vehicles**

By guaranteeing greenhouse gas reduction and fuel efficiency, EVs have attracted a progressively greater share in the private automobile market [54]. Nevertheless, the charging problem is still a barrier to the EV business's growth with the present battery technology. Thus, it is essential to have a wide charging infrastructure area that holds fast charging poles, battery swapping stations, and individual charging points for faster EV battery charging. In this scenario, the EV model is simulated with the DT to replicate the EV in the real world accurately. Introducing a DT

model makes it easy to simulate mobility behaviors and interactions to study the efficiency of the charging pole and EVs from the demand side and supply side.

DT for IoV creates virtual representations of vehicles and traffic systems, connecting the virtual world and physical world spaces. This allows for real-time vehicle and traffic system performance monitoring, improved situational awareness, feasibility forecasting, and decision-making through comprehensive, multidimensional modeling. Continuous, real-time interactions between the virtual and physical realms are necessary for accurate simulations [55].

One example of the potentiality of DT for IoV is in [56]. In this work, a simulation platform based on DT models is proposed to replicate and simulate the charging and discharging processes of large-scale EV fleets in different kinds of dynamic scenarios. To replicate the realistic mobility of EVs, the operational motion parameters such as steering angle, orientation angle, and acceleration are integrated into the mobility model to simulate the realistic trajectories. A tracking method of the dynamic position and orientation is proposed to synthesize the reasonable trajectory of EVs in day-to-day traffic conditions. In this framework, in which the key elements of the real world are considered, the DT of EVs and charging points behave and interact with each other as real entities. Then, the simulation result can be used to evaluate schemes for the deployment and management of EVs and charging infrastructures. In addition, the DT simulation platform can be used to verify the design of deployment of charging infrastructure and related impact on smart grid [57]. Advancements in this area can solve the problem of managing and exploiting dependent traffic data. The huge traffic data available could help constitute a DT that creates a virtual representation of the physical vehicles via various communication means.

#### **2.3.4 5G networks**

The 5G network supports a wide range of applications in different industries. The enhancement in 5G network communication impacts Industry 4.0. Such industries are smart cities, military applications, health care systems, and transportation using IoT. The 5G network makes rapid changes in wireless communications and improves wireless network performance by increasing capacity, improving

reliability, lowering latency, and increasing network speed [58]. The previous cellular technologies depend on fixed infrastructure while coming to 5G networks significantly enhanced the use of small cells and mobile cell sites to increase network access in congested areas. The 5G network has various applications and dynamically changes latency, bandwidth, and reliability requirements. These requirements have a high impact on the deployment of the 5G network in EVs.

The 5G network is primarily used in EVs for communication between the EV components. The rapid transformation and fusion between industry and communications systems have resulted in significant highway renovations, especially in self-driving. This development affected many applications, such as the roll-out of 5G networks, the Internet of Vehicles, and the adoption of Cellular Vehicle-to-Everything (C-V2X) connectivity. As a result, when the 5G is connected, vehicles exchange traffic data, highways, traffic signals information, roundabouts, etc., without human interference [59]. The 5G network-connected vehicles will generate a massive amount of data and have more autonomous functions. The 5G network integrated with DT can address key variables such as capacity, reliability, mobility, latency, and security. Recently, a prediction method for 5G-enabled IoV in Real-time traffic using DT concept was introduced in [60]. Hu et al. worked in IoV solutions and introduced a DT-assisted real-time traffic data prediction model using 5G communication. As a result of this work, the proposed model proved to optimize the scheduling of traffic resources and mitigate possible traffic jams at peak times. The authors believe the proposed method can be more accurate than others by analyzing the traffic flow and velocity data measured by IoV sensors and transmitted over 5G communications.

Deng et al. [61] proposed a combined approach of DT, reinforcement learning, and expert knowledge for the self-optimization of current 5G networks performance, and they described potential application scenarios for 6G to solve the problem of end-to-end delay and reduce the processing time at the local servers in many emerging critical applications. However, Dong et al. [62] adopted a DT framework of the current network. The proposed framework based on DL algorithm achieved lower energy consumption with minimal computing complexity. Jagannath et al. [63] proposed an innovative DT framework as an expandable approach for data-oriented modeling and real-time simulation of extensive systems on

5G-supported IoT networks. The DT framework employs a tiered architecture for decentralized deployment on cloud computing platforms. Additionally, it facilitates the application of AI/ML engines for event identification and prediction models. It can be concluded that the use of 5G in combination with the DT technology is essential in real-time scenarios, like the IoV framework, in which the safety and efficiency of the vehicle fleet traveling depend on the velocity and quality of the exchanged data.

### **2.3.5 Artificial Intelligence for Autonomous Vehicles**

The recent research in EVs focuses on AV, also known as self-driving or driverless cars, i.e., vehicles driven without human intervention. Such vehicles are electric since electric propulsion is more effortless to be autonomously governed. With advanced technology, the vehicle will sense the surrounding environment, plan the route, and drive safely, thanks to AI and ML technology [64]. The AV is still under testing and has not yet become popular globally, but in the coming years, the AV will occupy the global market and vehicle industry due to its great benefits. Although AV has been an active research and development area in the last decades, it still faces many challenges in developing an entirely safe automated vehicle system. Road conditions, traffic conditions, weather conditions, and communication expansion have helped the growth of AV systems.

Car navigation systems assist in controlling and making decisions based on the prior knowledge (sensors or road map) that feeds into the system. Lopes et al. [65] proposed an efficient approach for vehicle navigation systems based on the velocity optimization paradigm. The maximum speed is adjusted to the curve of the road, and the car follows a smooth path to the lane's center. The approach is integrated into car navigation architecture and evaluated in two separate simulators before being tested in AV prototypes. The local route and road geometry are required for autonomous driving. Therefore, Jo et al. [66] proposed a hybrid local route generation method. According to the history of performance and map availability, the algorithm can precisely choose the best route between the available options. The proposed method is validated and verified in real traffic conditions in an urban area in Korea. In fact, verification and validation are significant challenges in AV for



safety assessment. The authors in [67] introduced a systematic review to investigate the current verification and validation software used in AV. They discussed the simulation environments and more specific approaches such as mutation testing, fault injection, techniques for cyber-physical systems, adversarial examples, corner cases, and formal methods.

Yang et al. [68] developed a framework that integrated the intelligent driving model with human factors such as driving mode and their reactions and expectations on the road to enhance autonomous driving performance. The proposed model helps to reinforce the efficiency and safety of AV. Reinforcement Learning (RL) is a widely used ML technique to train the agent on rewarding and punishment approaches. This approach effectively works with the AV industry as the RL algorithm learns from the driver's actions to increase a certain reward or take a decision. Masmoudi et al. [69] designed a framework for car-following based on video frame processing using RL algorithms. The framework is based on navigation decisions and automated object detection. The proposed model achieved promising results and acceptable car-following behavior in AVs.

Employing RL in the AV industry is innovative and will lead to some benefits. The more information the algorithm processes, the more efficient the algorithm becomes, and the better the results can be obtained. Software providers that support DT concept have begun integrating reinforcement learning into their tools. For example, Flexsim software [70] recently introduced the RL model and the possibility of connecting external systems with the Flexsim models. Such additions will make the software a 3D design tool and a data analysis tool, enhancing the concept of Rassolkin et al. [26] specify tasks required for a specialized unsupervised prognosis and control platform for energy system performance estimation for AV. They develop several test platforms using DT and ML algorithms to optimize self-driving EV' electric propulsion drive systems and monitor sensors autonomously. In addition, Venkatesan et al. [71] propose a pre-estimation of the service requirement of EV motors for AV using intelligent DT that employ Artificial Neural Networks and fuzzy logic in MATLAB/Simulink for monitoring and prognosis of permanent magnet synchronous motor distance.

### **2.3.6 Discussion**

The performed analysis highlights DT as a promising technology for ITS, and the current literature does not exhaustively present the deployment of DT, in combination with the other discussed technologies, in EVs or AVs. Although some articles generally talk about DT technology and EVs, there is no comprehensive coverage of all the technical aspects of applying this technology to electromobility. There is a lack of coverage of the critical technologies in DT for electromobility, especially for the next generation of self-driving systems. For instance, the use of data analytics in this area is promising as data is the most crucial key in the era of AI, which leads to optimizing performance and enhancing security. During our research, there were insufficient reviews that considered the importance of data generated by ITS and how the data could be utilized in DT to generate value-added electromobility services. The aim is to cover this gap by giving a wider spectrum of analysis. It can be concluded that having DTs of EVs and AVs, in combination with other technological solutions like IoV, VSs, IoT, and 5G, can facilitate technological advancements and give the ability to optimize both the single vehicle technology and the traveling and charging operations of EV fleet.

## **2.4 Current Issues in EVs**

EV development is considered a successful and promising solution to electrifying the transportation sector, and the use of EVs can lead to several environmental benefits, reducing gas emissions. However, EVs are integrated with all emerging technologies, such as smart grids. This integration will result in several technical and logistic issues affecting EV diffusion. Figure 2.6 shows the current and significant issues affecting the electromobility sector regarding EV tracking and monitoring, the BMS, connectivity, security, and privacy issues.

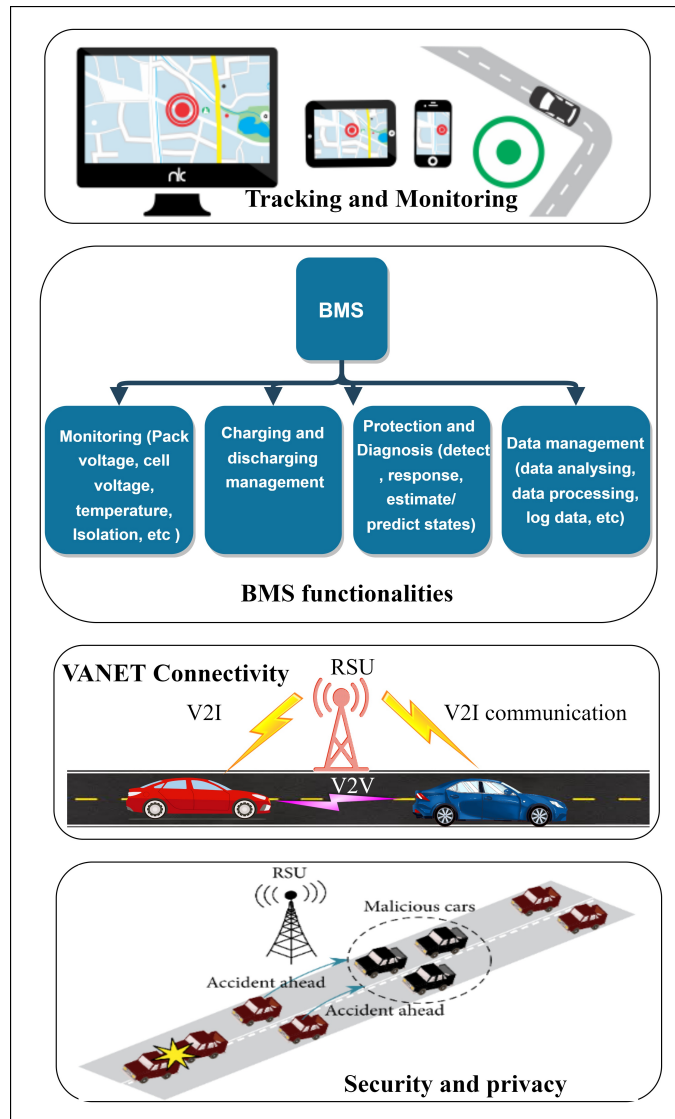


Figure 2.6: Current Issues in Electric Vehicles.

### 2.4.1 Tracking and Monitoring

The use of EVs leads to important changes in mobility. Numerous advantages are linked to using EVs, including zero emissions, improved fuel economy, and lower running costs, making them perfect cars for everyday use. Security concerns, route planning, and battery charging are all challenges that vehicle monitoring systems

may help to address. When planning long trips, optimal tracks, battery life, state of health, and charging time have to be planned and monitored.

An increasing number of charging stations are appearing. Since they are not as widely spread and prominent as gas stations, regular route planning and daily charging station check-ups are necessary. The issue of range anxiety is still critical in EVs and needs to be considered in the upcoming research. Sarrafan et al. [72] proposed a novel framework to solve this issue by introducing a real-time mixed SoC estimation algorithm. The proposed system was implemented in an advanced driver assistance system. The proposed modified method includes the recalibration technique for estimating the battery state of health (SoH), the initial battery SoC, and the adequate battery capacity, considering external parameters, the environmental factors such as the traveling factors, vehicle factors, traffic congestion factor, and the driver's behavior factor to make the model more accurate than the traditional models available in the literature where the environmental conditions and driver's behavior have been not considered. Both laboratory and field tests have been conducted using a Nissan Leaf. As a result, the SoC estimation of lithium-ion batteries improved with great accuracy and high real-time capability. In [73], a discrete scheduling process was formulated to track arbitrary power profiles and control charging, limiting it to the maximum rated power. The coordinated charging of PEVs is formalized considering the realistic case of PEVs with different charging rates. In addition, a novel algorithm has been developed to ensure plug-in EV charging and eliminate the need for a central aggregator. It guarantees tracking and not exceeding the power profile the power utility imposes while maximizing the user's comfort. The algorithm's effectiveness has been demonstrated in realistic scenarios with a heterogeneous PEV population, but no performance comparison with other methods has been provided.

Thus, the vehicle monitoring systems significantly affect the research and development and improve the operations of industrialization of EVs. In addition, Wang et al. [74] designed a remote monitoring system for EVs, a combination of remote monitoring platforms and onboard terminals installed on the vehicle. On the contrary, with respect to [73], the idea was to aggregate the data from the onboard terminal, such as battery status, temperature, SoC, running status, etc., and send them to the server cluster model for processing and monitoring. This model

can improve communication reliability and enhance the real-time performance of the system. The system structure of the EV remote monitoring system is based on CAN and GPRS technology. It aims to provide a specific guiding significance for the design of remote real-time monitoring systems for new energy vehicles. It results in the necessity to monitor the parameters of EVs to avoid electrical malfunctions because problems of this type are costly to solve. In this context, adopting the DT concept can provide solutions to the discussed issues about EV monitoring and tracking.

## 2.4.2 Battery Management Systems

Worldwide, serious challenges such as global warming and greenhouse emissions due to the use of petrol and diesel in vehicles, as well as excessive levels of toxic gas (CO<sub>2</sub>). EVs are now being promoted and widely regarded as eco-friendly and substitutes for other vehicles based on combustion engines. Rechargeable batteries are widely used as a power supply for EVs. There are many different types of batteries, including lithium-ion, lead-acid, nickel-cadmium, and nickel-metal hydride [75]. The lithium-ion battery is the most used battery, which provides a high density of electric energy, is eco-sustainable and has a long life cycle. To keep the EV battery's state of health at a good quality level, we must take proper care of the battery and adhere to appropriate charging operations (over-charging, current, voltage, or discharging). These operations may cause problems and may damage the battery and the EV.

Several studies have been done on applying IoT and cloud computing to solve the BMS issues. Recently, Kim et al. [76] developed a cloud-based battery system to monitor stationary batteries and help in fault diagnosis platforms for large-scale. As the authors claimed, it is a new model, with no comparison in the previous literature, made of a cloud-based battery condition monitoring and fault diagnosis platform, incorporating IoT-enabled wireless battery module management systems and the proposed cloud battery management platform to support onboard battery health monitoring and to provide intelligent and cost-effective maintenance of the large-scale BESSs. The hybrid filter (HF)-based condition monitoring algorithm of cells and the proposed outlier detection-based fault diagnosis algorithm are

implemented in the platform. Compared to the Kalman Filter-family and Sliding Mode Observer, the HF leads to less computational cost and chattering issues, respectively.

Tanizawa et al. [77] introduced a cloud framework for EVs to manage the battery information related to the battery replacement system. This cloud-connected battery management system will maximize the value of the shared batteries by using a location data cloud to continuously connect to the batteries, manage the SoC, and monitor changes in their characteristics. Friansa et al. [78] presented a solution for battery monitoring in a microgrid system based on IoT, but differently from [76], the authors present a smart microgrid integrating a battery pack, PV system, Intelligent Electronic Device (IED) hybrid inverter, grid connection, and electricity load where no fault diagnosis is considered. In this framework, the IoT is realized through a communication channel with the IED, data acquisition algorithm, cloud system, and Human Machine Interface (HMI). The data stored in the cloud system database are processed and analyzed to produce information that users can access through the desktop and mobile devices using ExtJS/HTML5 framework. The analytical results show good performances for the average execution and connection times within the architecture for overall BMS-IoT-based data acquisition to the cloud server. Also, the availability of monitored data shows satisfactory results for the BMS-IoT system data acquisition reliability.

The above studies have some drawbacks as the technical details were not introduced in the cloud, and the battery diagnostic algorithms that can help improve accuracy and data storage in the cloud are not exhaustively analyzed. Therefore, Li et al. [79] tried to overcome these drawbacks by introducing a cloud BMS for battery systems to improve the computational power and data storage capability of cloud computing. The proposed cloud computing system provides computational power and data storage capacity with greater speed and utility. In EVs, charging operation is done by a graded control structure controlled by an aggregator, which controls all EV charging rates. For example, Nour et al. [80] proposed a new approach for smart charging in EVs. A fuzzy logic controller controls and manages the EV charging process to maximize electric utility and EV owner benefits. In particular, the controller regulates the EV charging power based on the electricity price signal provided by the electric utility and EV battery SoC. The proposed

technique was evaluated using simulation with MATLAB/SIMULINK, and the impact of EV charging on the distribution network decreased compared with uncontrolled charging. Many EV drivers still face a limited driving range. However, in only a few years, the range of most EVs has considerably improved by increasing the battery size and improving lithium-ion battery technology. Anyway, the current EV range is still not convenient for users. The ranges of an EV and its higher purchase cost are two main deterrents to the widespread use of EVs [81]. Therefore, we need to work around this limitation of driving range in the future, and DT can support further advancements in battery technology.

### 2.4.3 Connectivity

Connected cars will inevitably improve user experiences and the ability to manage the network of interconnected vehicles. Two types of connectivity can be distinguished: Intra-vehicle connectivity and Inter-vehicle connectivity. The bandwidth requirements in vehicles have increased dramatically because of recent innovations and changes in automotive technology. However, modern vehicles have gradually developed in entertainment and networking with advanced capabilities [82]. Intra-vehicle or internal networks are designed to share data across the various subsystems, ECUs, sensors, and actuators so that a single vehicle can operate easily. Although the sensors and the networks are exclusive to original vehicle companies, these technologies typically meet standards that permit vehicle diagnostics and future applications to communicate with the vehicle using external technologies or devices.

Inter-vehicle connectivity is a networking approach that allows data to be transferred from the vehicle to other vehicles, remote computers, and other cloud infrastructures. Remote applications can combine the vehicle's data with external sources (such as traffic or weather data) [83]. Gupta and Sandhu [84] developed an authorization framework describing data exchange scenarios on the Internet of Vehicles (IoV). The development of IoV has produced a considerable amount of real-time traffic data. These traffic data are used to generate a kind of DT that connects the physical vehicles and their virtual representation via 5G. Moreover,

Khan et al. [85] presented an effective communication framework to mitigate cyber-attacks on Connected and Autonomous Vehicles (CAVs).

In this context, data will be collected from different networks and heterogeneous resources, including charge management software, weather sites, live EV modeling, battery modeling, telemetric devices, and route maps like Google Maps. Data integration in such platforms is complex and needs to support data interoperability in connected vehicles to communicate effectively and efficiently. Paper [86] deals with AV collaboration as a service and proposes a DT-based scheme to facilitate collaborative and distributed autonomous driving. Specifically, a DT is designed for each AV, and a DT-enabled architecture is developed to help AVs make collaborative driving decisions in the virtual networks. This architecture uses an auction game-based cooperative driving mechanism to decide each group's head DT and tail DT. After that, by considering each group's computation cost and transmission cost, a coalition game-based distributed driving mechanism is developed to decide the optimal group distribution for minimizing the driving cost of each DT.

#### **2.4.4 Security and privacy**

The security issues seriously influence the CAV sector due to the comprehensive connectivity over the Internet and cloud. The essential threat in CAV derives from exchanging information with other vehicles and servers. CAV can share the ID details, battery information, or vehicle location. This data exchange between CAVs and the server might be subject to cyber-attacks such as (DoS, spoofing, privacy attacks, modification, etc). Several studies in the literature surveyed the issue of security and privacy in electromobility. Many approaches and technologies have been innovated to increase the protection level of EVs. Guo et al. [87] proposed an anomaly detection system to enhance the cyber security of the steering stability control system in the CAV. The proposed approach can identify the threats to control inputs and sensors by combining two approaches, physics-based and learning-based approaches. The results have shown an improvement in the cyber-physical security of EVs. Babu et al. [88] introduced a robust authentication



protocol for charging electric cars while driving. Secure and lightweight primitives like elliptic curves and hash functions are used in the proposed protocol.

The protocol's security is being investigated to show it can handle different attacks. Kavousi-Fard et al. [89] developed a cyber-resistive model for detecting vehicle cyber-attacks. A hybrid smart structure built of wavelet decomposition technology and a modified support vector machine detects any malicious behavior in the controller area network bus. The proposed approach performs excellently in detecting cyber-attack communications while also detecting regular data. Thus, it is essential to find a way to validate the data exchange and transformation between the vehicles and infrastructures before the system implementation. DT will give the ability to validate the security and privacy over the IoV by simulating the network scheme and using AI and ML technologies; it will allow predicting threats that can harm the vehicular network [90]. Safety and security functions are of basic importance in AV. Paper [91] aims to identify a standard vehicular DT framework that facilitates the data collection, processing, and analytics phases. DT is explored to automate the decision-making process inside an AV using radar sensor data collected from initial, analytics, and reporting phases and generate reports to be sent to AVs. The recommended model aims to identify, analyze, and assess the threats and allow the user to take appropriate countermeasures to ensure safety and security using DTs in driverless vehicles. Some of the advantages presented using the car follower model include the benefit of reducing the risk of cyber-attacks and accidents.

### **2.4.5 Data Analytics**

Big Data Analytics influences ITSs by enabling EV internet connections to optimize their performance. The plug-in EVs should be connected and able to revolutionize energy use, creation, and redirection. Smart grids and EVs generate equal amounts of significant Big Data among the connected devices. As consumers and big data producers, EVs produce data from various sources, such as sensors, logs, etc. By utilizing big data technologies, the data can be used to develop smart charging algorithms, solve energy efficiency issues, develop policies and strategies for the location of electric charging stations, and turn smart cities into green cities [92].

Technology integration transforms the transport and the automotive industry by tracking, analyzing, and evaluating the demographics of EVs. The demographic data of the EV include statistics on charging stations, battery features, analysis of energy usage, and route profiles. In this framework, the goal is to overcome obstacles such as battery capacity, battery prices, charge time, and the availability of charging stations to take full advantage of EV's potential. Using data science and AI technology could improve operations and overcome the main obstacles [93]. Nowadays, every industry is affected by the role of data. We are facing an increase in the volume of data produced, particularly in transportation. The transportation industry promised to develop better information systems to optimize energy consumption in highly complex environments through the electrification of vehicles. Car manufacturers, governments, and charging infrastructure providers utilize data analysis and data science tools to use and analyze the available data to provide services for optimizing EV use.

Big data analytics is often used to assess the driving range and effectively reduce user anxiety. In [94], the authors propose an approach for classifying the range estimate. The data is classified into standard, historical, and real-time data. In addition to the range estimation, data produced from EVs can be used to determine the position of public charging stations. Different types of information, including traffic density, gas station distribution, and vehicle ownership, are used in this respect. In [95], the authors proposed a solution to site public EV charging stations using a large-scale trajectory dataset. They examined the model based on data from 11,880 taxis in Beijing. Security is one of the most important features that data analysis offers for ITS. In this direction, various ML and DL technologies have been developed to secure transportation systems and improve prediction accuracy. Lv et al. [39] developed a DT framework based on the DL algorithm, combined using the Convolutional Neural Network and Support Vector Regression. The proposed model can reduce the system data transmission delay, improve prediction accuracy, and reasonably modify pathways to reduce traffic congestion. By combining the power of predictive analytics and data intelligence, the predictive maintenance analysis of batteries can be improved. The integration of these technologies aims to achieve high battery efficiency and reliability. Sreedhar et al. [96] introduced a simulation design of common BMSs. The proposed system is designed to handle

and control battery parameter values such as battery voltage, consumption of power, current, and SoC.

Data analytics improved EV efficiency rapidly, and DT technology provides smart modeling capabilities in this field. Tang et al. [97] proposed a novel model utilizing an ML algorithm to increase computational efficiency and precision. In smart EVs, it is also important to focus on other essential aspects, such as the safety of drivers and passengers. A safety-based intelligent approach in EV, presented in [98], uses a fuzzy adaptive control method for vehicle following, decreasing accident rates. Another safety application proposed by Guo et al. [99] is direct yaw control, which enables EV to keep the vehicle within the allowed path and increase the stability of the steering system. Behrendt [90] developed a hybrid architecture using the DT technology to perform analytics operations and increase real-time driver safety. The information that comes from the physical smart vehicle through sensors is analyzed to detect privacy anomalies in the transportation ecosystem and minimize other privacy risks. As a result, Data science, AI, and big data tools are emerging use cases in this context. They can significantly impact the current EV market to improve the end-product performance.

### **2.4.6 Intrusion Detection Systems**

Network security has developed as a critical research subject due to the advanced development of internet and communication technologies over the previous decade. It uses firewalls, antivirus software, and IDS to keep the network and its assets safe in cyberspace. Therefore, researchers have conducted many studies in network security, with various approaches and technologies employed by researchers to develop algorithms for detecting unusual activities on different network platforms [100].

To meet network security needs in electromobility, various IDS have been developed. IDS is a security framework that continuously monitors network traffic to detect any abnormal behavior that may violate the network policy and threaten its confidentiality, integrity, or availability. The vehicular IDS architecture proposed by Loukas et al. [101] ensures onboard IDS, collaborative detection, and offloaded intrusion detection: onboard IDS where the vehicle identifies illegal behavior

on networks on its own; collaborative detection is where vehicles collaborate to determine whether they are under attack or not; offloaded intrusion detection is where the detection process is done in the cloud. In this context, Zeng et al. [102] developed an end-to-end intrusion detection using a DL algorithm to detect malware intrusions for onboard units. Different from previous intrusion detection methods, the proposed method only requires raw traffic instead of private information features extracted by humans. The performance is compared to prior methods on a public dataset and a simulated real-life VANET dataset. The experimentations show that the model can achieve higher performance with a minimum resource requirement.

Shams et al. [103] introduced a trust-aware Based IDS; they used the combination of modified promiscuous mode and support vector machine to ensure the safety of vehicles and detect malicious behavior in any network node. Liang et al. [104] proposed an intelligent IDS model based on hidden Markov methodology to filter out malicious messages and reduce detection time without affecting the detection rate. The authors claim it is the first work in the literature to model the state pattern of each vehicle in VANETs as a Hidden Markov Model (HMM) to quickly filter the messages from the vehicles instead of detecting these messages. It consists of three modules: schedule, filter, and update. In the schedule module, the Baum–Welch algorithm is used to produce an HMM and its parameters for each neighbor vehicle. In the filter module, multiple HMMs are used with their parameters to forecast the future states of neighbor vehicles with which the messages from them are filtered. In the update module, a timeliness method is used to update HMMs and their parameters. The experiments show that the IDS with FM-HMM has a better performance in terms of detection rate, detection time, and overhead. Sedjelmaci et al. [105] designed an efficient, lightweight IDS simulator for the vehicular network. The model can protect VANETs against denial-of-service attacks, false alert generation, and integrity target attacks. They used the NS-3 simulator (a discrete event network simulator for Internet systems) to present the detection mechanism's performance analysis. The simulation framework shows high-level security and an accurate detection rate.

### **2.4.7 Discussion**

In this section, the main challenges for electromobility were discussed, and how DT technology is used and can be used in future applications to address them. The complexity of integrating ITS with emerging technologies and communication tools using the DT concept emerges. In this part, several technical and logistic issues under different topics were presented that could represent a risk for the future of smart and green mobility. Researchers have done several works so far to find effective solutions, especially for monitoring and tracking the vehicles, replicating the battery and vehicle technology to improve energy consumption, charging time, driving ranges, etc.

In some cases, DTs of batteries and vehicles have helped to perform simulations and validate experiments, achieving better performances and pushing technological advancements. Future works must further develop technologies and tools able to answer the needs of connected and smart EV fleets in which a large amount of data needs to be safely exchanged, collected, and elaborated in real time, respecting security and privacy issues. In the next section, taking into account the previous challenges and issues, the benefits and significant challenges in EV networks, and how DT could provide solutions that help integrate modern technologies and provide the best services at lower costs will be discussed.

## **2.5 DT Solutions and Future Directions**

DT concept involves feeding data from the real world back into the virtual environment to improve model accuracy. This approach reduces the gap between the real and virtual worlds, allowing real-world simulation. As a result of this survey, the important challenges in EV networks were presented, as well as how DT can provide solutions.

### **2.5.1 Cost-Effective and Reliability**

The implementation of EV networks is one of the most relevant challenges due to the shortage of infrastructure and safety measurements. Providing infrastructure services to implement EV networks is very cost-effective. The maintenance of EV services is also cost-effective. The DT model of the EV can be evaluated before deployment, thus lowering maintenance costs and making DT a cost-effective option. The present state of the EV system doesn't have reliability in data transmission. Reliability is the main challenge in EV transportation systems to operate EVs safely under various conditions. In the future, EV transportation systems should maintain data reliability and scalability.

### **2.5.2 Visualization and Charging Time**

The data visualization gives complete scope for EV consumers to plan long-distance transportation. However, the EV system has visualization limitations that lead to the problem of testing EV functions and efficiency. DT integrates 3D graphics and audio with real-world objects to solve this issue using IoT and AI. Using such technologies, the operator can monitor and control the EVs and allow them to communicate and interact with DT model to improve efficiency during and after the design process. Whether the charging system is standard, fast, or quick, the charging time is still quite long. This is one of the main reasons holding back the growth of the EV industry. There is also the need to research wireless charging. The DT simulation can improve the charging time by analyzing the data from the virtual model, and the result can be used to evaluate charging infrastructures and charging efficiency.

### **2.5.3 Intrusion Detection Systems for EV and AV**

Data attacks are significantly mitigated with DT technology, which provides greater security for the mechanism against attacks and protects the privacy of EV users. The DT enables the development of high-accuracy models for real-time systems using massive quantities of operational data with expert observation. DT archi-

ture could be built to effectively detect intrusions or anomalies that behave abnormally inside the vehicular (EV and AV) network. Available studies have explored this technology for industrial applications using data analytics and ML technologies such as One-Class Support Vector Machine, Local Outlier Factor, and deep unsupervised learning. For example, in [106], Fraser et al. proposed DT architectural enhancements to improve the security of Unmanned Aerial Systems. Gao et al. [107] introduced an anomaly detection system to monitor abnormal behaviors in DT-based Cyber-Physical Systems.

Gehrmann et al. [108] investigated how the DT model and security architecture can share data and control security-critical processes. They introduced a new security framework that provides the foundation for future research work in automation and control systems. Xu et al. [109] presented a two-phase digital-twin-assisted fault approach based on deep transfer learning, detecting faults during development and maintenance in a vehicle body-side production line. Snijders et al. [110] used the Convolutional Neural Network model to improve the predictive power of DT for Cyber-Physical Energy Systems. The behavior of ten batteries was predicted using real-world data. They conclude that ML for DT can aid in maintaining a heterogeneous energy ecosystem. Castellani et al. [111] presented cluster centers, a clustering-based method, and Siamese Auto-encoders, which is a neural architecture designed for weakly supervised environments with few labeled data samples. The methods utilize a DT model to create a training dataset that replicates the machinery's usual operation. The above research employed DT to detect intrusions and abnormal behaviors in Unmanned Aerial Systems, manufacturing, and energy systems utilizing ML and DL. In conclusion, the same technologies could be applied to EVs and AVs to obtain benefits.

#### **2.5.4 Case studies**

The paper [112] establishes and compares a traffic infrastructure efficiency assessment Data Envelopment Analysis model based on DT and a traffic flow prediction model based on Long Short-term Memory. The traffic flow data of a certain road section in Zhenjiang City has been simulated and predicted. Taking the transportation infrastructure of 12 cities in J province as the research object, the two models

are verified to make the intelligent transportation facilities have a greater potential. The results show that the established DEA model based on DT can estimate the efficiency of transportation infrastructure more reasonably and accurately. Compared with other models, the traffic flow prediction model based on Long Short-term Memory is more accurate in traffic flow prediction, which can provide a reference for intelligent transportation system infrastructure investment planning.

A case study about safety and security functions in AVs is tested in [86] to demonstrate the effectiveness of the proposed approach: a vehicle follower model is analyzed when radar sensor measurements are manipulated to cause a collision. Almeaibed et al. [91] proposed a standard framework for DT that facilitates data collection from Vehicular network, processing, and analytics. A vehicle follower case study was presented to prove the model's efficiency. The vehicular model was analyzed during the manipulation of measurements of the radar sensors in an attempt to make a collision. This research can light the way for future research using the DT concept in the EV and AV industry. Another case study proposed by [84] introduced a real-time case study on smart cities to improve the defense mechanism for connected AV attacks. To reduce security and privacy attacks such as DOS, hijacking, man-in-the-middle, GPS spoofing, privacy attacks, and replay attacks, they proposed a blockchain-based architecture providing a secure and decentralized connected AV.

## 2.6 Chapter Summary

Table 2.1 shows interesting works in DT from different perspectives and applications. The table classifies the existing contributions in this field based on the use of DT in combination with other technologies and services.





This chapter aims to provide a complete survey of the existing literature in the last five years on adopting DT technology in ITS. The awareness of DT has recently been growing exponentially due to the number of applications that demonstrate their capabilities for connecting the physical and digital worlds. EVs are physical objects for which DT can overcome some important limitations and enhance their functionalities. There is a lack of coverage of the analysis of critical technologies in DT for electromobility, especially for the next generation of self-driving systems. We found insufficient reviews that consider the importance of data generated by ITS and how the data could be utilized in DT to allow value-added services to be created.

The chapter addresses the complexity of integrating ITS with emerging technologies and communication tools using the DT concept to accelerate the technological advancement goals and achievements of EVs and AVs. It highlighted the potential of DT technology based on several aspects, such as tracking and monitoring, security and privacy, data analytics, and intrusion detection, that could enhance efficient electric and AV network management.

Finally, several significant issues and major challenges that continue to influence this field of research were discussed, as well as solutions that DT can offer.

# Chapter 3

## Communication Security for Intelligent Transportation Systems

### 3.1 Security and Intelligent Transportation Systems

#### 3.1.1 Introduction

ITS represents a technological leap forward in optimizing the utilization of transportation infrastructure. By harnessing real-time data, predictive analytics, and smart communication, these systems aim to alleviate traffic congestion, enhance mobility, and reduce environmental impact. From smart traffic management to the advent of autonomous vehicles, ITS is reshaping the way people and goods move, promising a future of safer, more efficient, and sustainable transportation.

The fusion of Security and ITS has emerged as a critical factor in the advancement and sustainability of modern transportation ecosystems during this age of accelerated technological advancement and urbanization [113]. Beyond physical barriers and surveillance, the realm of transportation security has evolved to include measures that ensure the resilience of digitally interconnected infrastructures and protect against cyber-attacks [114]. ITS optimizes transportation systems in terms of environmental sustainability, safety, and efficiency through the utilization of cutting-edge technologies.

The essence of security within transportation systems encompasses a broad spectrum, ranging from protecting critical physical infrastructure such as bridges and tunnels to securing data transmitted across interconnected networks [15]. With the digitization of transportation, there is a growing need for robust cybersecurity measures to thwart potential threats and vulnerabilities [115]. The integration of sensors, communication networks, and AI in transportation networks brings about a paradigm shift, introducing new dimensions of complexity to security challenges.

As ITS grows in sophistication, the implementation of strong security practices becomes critical. It is essential to maintain a balance between technological progress and strict safety requirements in order to protect the integrity, reliability, and public trust in these technologies. This introduction establishes the context for an examination of the complex association between security and ITS, encompassing innovative resolutions, challenges, and prospective directions that will shape the intelligent and secure transportation systems of the future. The technological requirements, policy, and methodologies that are necessary to realize this objective will be elucidated in subsequent discussions.

### **3.1.2 Intrusion Detection Systems**

Intrusion Detection Systems (IDS) constitute pivotal components in the realm of cybersecurity, designed as software or tools with the explicit purpose of identifying and flagging malicious activities and unauthorized transactions within a network. These systems serve as vigilant guardians, constantly monitoring system activities, scrutinizing file integrity, and conducting pattern analyses to pinpoint potential threats. The overarching goal of an IDS is to fortify computer systems against various forms of cyber threats [116].

The functionality of an IDS is multifaceted, involving the continuous scanning of systems for vulnerabilities that might serve as entry points for malicious attacks. This proactive approach ensures a heightened state of readiness against potential security breaches. An IDS achieves its objectives by collecting data from diverse sources within networks and systems, subsequently cross-referencing this information with pre-established discriminatory patterns. The outcome of this comparison determines the existence of potential threats or vulnerabilities [117].

Classified based on their detection locations, IDS can be broadly categorized into two main types: Host-based IDS (HIDS) and Network-based IDS (NIDS). HIDS is tailored to detect suspicious user activities on the local host machine, while NIDS focuses on detecting potential attacks primarily through the analysis of network traffic. Both HIDS and NIDS leverage log files and databases to cross-verify and validate results, creating a comprehensive security net for the entire system [118].

The approach taken by an IDS in identifying threats can further be classified into two main categories: signature-based detection and anomaly-based detection. The former, also known as misuse-based detection, involves the continuous collection of malware signatures, which are then used to create a comprehensive database. When the system identifies a network traffic flow or activity that matches a signature in the malware database, it is flagged as abnormal, signaling a potential security breach. Conversely, anomaly-based detection relies on a comparison with a predefined baseline, involving an analysis of network packet patterns to identify deviations from the norm [119].

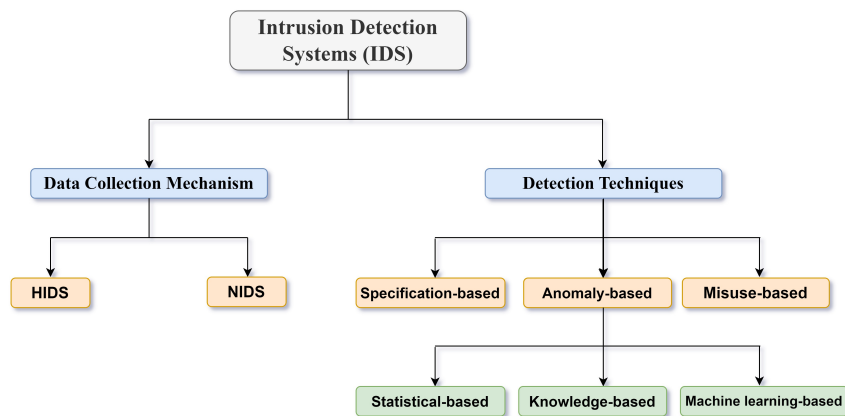


Figure 3.1: Classification of an intrusion detection system.

Figure 3.1 illustrates the classification of IDS based on their data collection methods and detection techniques. This intricate categorization underscores the diverse strategies employed by IDS to fortify cybersecurity, emphasizing the

importance of a multi-faceted approach to threat detection and prevention in modern computing environments.

### 3.1.3 Network Attacks

A network attack is defined as unauthorized actions targeted at network devices or exploiting vulnerabilities within the network infrastructure. These malicious endeavors are aimed at circumventing system protections with the intent to alter, compromise, destroy, or steal confidential data. The evolution of network attacks has witnessed a transformation in how they are classified and understood.

In the nascent stages of cybersecurity discourse, network attacks were often dichotomized into two broad categories: external and internal [120]. External attacks typically originate from sources outside the secure network perimeter, attempting to breach defenses from the external environment. Conversely, internal attacks emanated from within the network itself, posing a threat to the integrity and confidentiality of sensitive information from within the organization.

However, as the complexity and sophistication of network attacks have grown, the need for a more nuanced classification system has become evident. Notably, researchers and cybersecurity experts [121] have responded to this challenge by delineating network attacks into a more comprehensive taxonomy. This categorization, as depicted in Table I, identifies seven distinct categories based on the strategies employed in executing these attacks.

The expanded classification not only acknowledges the external and internal dimensions but also delves deeper into the specific methodologies and tactics employed by attackers. These categories include but are not limited to:

**Denial of Service (DoS) Attacks:** Overwhelming a network or system with excessive traffic to render it unavailable to legitimate users.

**Malware Attacks:** Introducing malicious software, such as viruses, worms, or trojans, to compromise the functionality or integrity of a network.

**Phishing Attacks:** Deceptive attempts to trick individuals into revealing sensitive information, often through disguised communication.

**Man-in-the-Middle (MitM) Attacks:** Intercepting and potentially altering communication between two parties without their knowledge.

**SQL Injection Attacks:** Exploiting vulnerabilities in a database by injecting malicious SQL code to gain unauthorized access or manipulate data.

**Cross-Site Scripting (XSS) Attacks:** Injecting malicious scripts into websites viewed by other users, compromising their browsing experience.

**Password Attacks:** Employing various techniques to gain unauthorized access to a network or system by exploiting weak or compromised passwords.

This refined classification not only enriches our understanding of network attacks but also aids in developing targeted and effective cybersecurity strategies to mitigate the evolving threats in today's interconnected digital landscape.

### 3.1.4 Machine Learning Techniques

An extensive survey was conducted in this section focusing on ML techniques commonly employed in IDS applications. We intend to build a robust IDS that can effectively address the unique challenges posed by the dynamic and real-time nature of VANET environments.

The survey delves into the landscape of ML algorithms utilized in IDS, aiming to identify techniques that are particularly well-suited for VANETs. By understanding the strengths and applications of various ML methods used in IDS, we aim to select algorithms that align with the specific requirements of VANETs.

ML contributes to systems acquiring and enhancing automated capabilities without direct manual intervention or explicit programming. In the context of Intrusion Detection Systems (IDS), a range of ML methods is employed. Diverse ML algorithms have been incorporated into the development of IDS, as depicted in Figure 3.2. A comprehensive review conducted by the authors in [100] elucidated the distinctions between unsupervised and supervised learning methodologies applied in the context of IDS.

Among the several ML methods applied in IDS, prominent ones include Support Vector Machines (SVM), K-means clustering, Logistic Regression, the Naive

Bayes method, Artificial Neural Networks (ANN), and Principal Component Analysis (PCA). These methods collectively contribute to the efficiency and effectiveness of IDS by enabling automated recognition and response to potential intrusion events, showcasing the versatility of ML in enhancing cybersecurity measures.

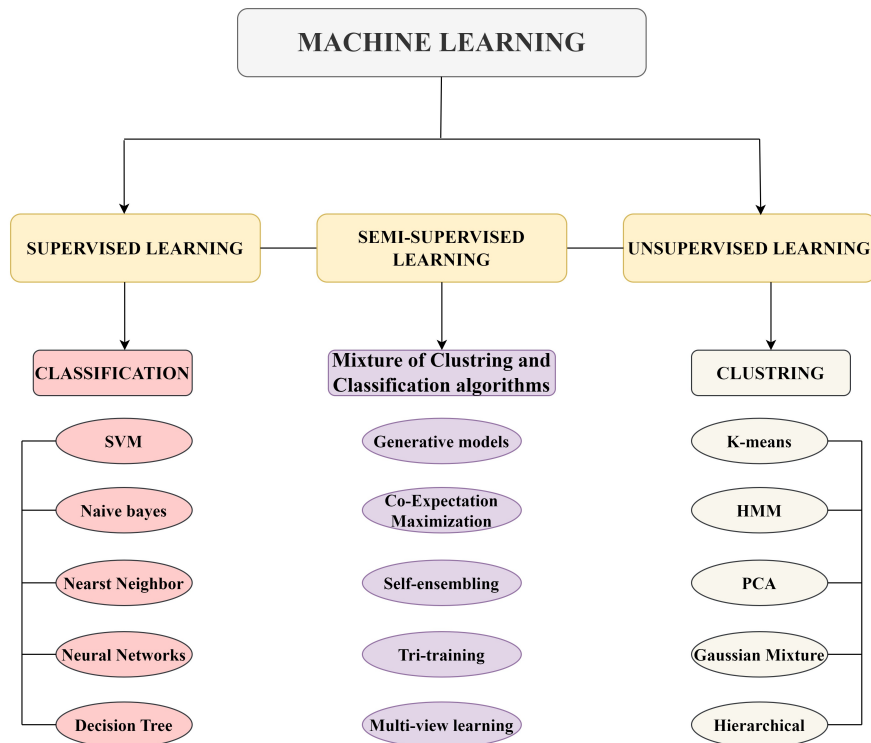


Figure 3.2: Machine learning techniques used in IDS.

These methods effectively and adeptly identify irregularities and intrusions within the network. The subsequent sections provide a detailed description of several ML methods employed to recognize and address such threats.

### Support Vector Machines

SVM algorithm is a supervised ML method commonly applied to both regression and classification tasks. Its core concept involves constructing an optimal hyper-plane capable of effectively classifying two distinct classes. A recent study by authors in [122] introduced an IDS model using the Particle Swarm Optimization (PSO) algorithm. This model integrates an information gain feature selection



method with an SVM classifier. The research demonstrates that the amalgamation of feature selection and parameter optimization for SVM results in reduced training time and enhanced classifier performance. Moreover, the proposed FS PSO-SVM model exhibits a high detection rate and minimal false-positive instances. The efficiency of this model was evaluated using the NSL-KDD Dataset, which encompasses four diverse network attack types: DoS, R2L, U2R, and Prob.

### **K-means algorithm**

K-means algorithm, a distance-based unsupervised technique, stands out as one of the most prevalent clustering methods. Its fundamental objective is to minimize the distance between points within a cluster and their respective centroid. Clustering is achieved by assessing the distances between objects for classification, consolidating observations into homogeneous groups [123]. In the context of time series data, the K-means algorithm is frequently employed to uncover patterns for matching and clustering connections within datasets. The algorithm employs  $k$  items as cluster centers, calculates the distance between each center and an item, and assigns the item to the nearest center. Subsequent updates and repetitions refine the clusters, a process particularly relevant in the context of stream mining for large datasets. Despite its efficiency, the K-means algorithm faces challenges such as anomaly detection, outlier detection, and fraud detection [124].

While the K-means algorithm is fast and efficient, its application in unsupervised methods is not as widespread. Addressing this, researchers in [19] proposed a model enhancing the K-means approach by utilizing the Calinski-Harabasz indicator to determine the optimal number of clusters for effective clustering. The proposed model underwent evaluation on two datasets: NSL-KDD and CIDS2017.

In contrast to the supervised SVM classifier, experimental findings reveal that the model proposed in [125] exhibited remarkable efficiency in both detection capabilities and time consumption. Another innovative approach presented in [126] involved the integration of two clustering and optimization methods, namely the K-means algorithm and Simulated Annealing, to achieve optimal classification. Although evaluated on the NSL-KDD dataset, the proposed method demonstrated satisfactory accuracy, showcasing improvements in IDS efficiency and misdetection rates. These studies underscore the evolving landscape of clustering algorithms,

emphasizing their potential to enhance the accuracy and efficiency of intrusion detection systems.

### **Logistic Regression**

Logistic Regression, a supervised ML approach, is adept at identifying discrete class sets. Employing sigmoid functions in its algorithms, logistic regression transforms various predictions into probabilities [127].

In IDS development, a strategy known as Sparse Regularized Optimization (SPLR) is proposed for feature selection and intrusion classification [128]. SPLR involves selecting a subset of features from the original set and utilizing them to model data for classification. Unlike traditional feature selection methods that treat feature selection and classification as separate entities, SPLR integrates both into a unified framework. In a linear SPLR model, characteristics from a dataset repository are used to construct a linear classifier, with subsequent discovery of classifier coefficients.

Binary classification methods like Logistic Regression find applications either independently or in conjunction with various data mining techniques to discern abnormal traffic from normal patterns within a network. An exemplary study by authors in [129] showcases the development of a two-stage anomaly-based Network Intrusion Detection System (N-IDS) using the UNSW-NB15 dataset. The research utilizes techniques such as Recursive Feature Elimination and Random Forests for optimal feature selection in ML. The Logistic Regression method, in combination with other data mining techniques, is then employed to effectively distinguish intrusive from normal network traffic, exemplifying the versatility of logistic regression in enhancing intrusion detection capabilities.

### **Naïve Bayes**

Naïve Bayes algorithm, grounded in the Bayes theorem, deals with probability updating based on provided evidence. This algorithm operates under the assumption of attribute independence, where the probability of one entity does not influence the probability of another [130]. It calculates the total output cases and then determines the conditional probability for each data class. It is characterized by a short implementation time, compatibility with both continuous and discrete attributes, and rapid learning capabilities.

Recent studies, such as the one presented in [131], have proposed innovative paradigms for IDS built upon the Naïve Bayes algorithm. Specifically designed to enhance the security of IoT infrastructures, the model aims to combat prevalent attacks, including DDoS attacks. DDoS attacks are particularly widespread in IoT networks due to the diverse and vulnerable resources available to attackers. The authors introduced an IDS called NB-MAIDS, comprising multiple agents, where the MA (Multi-Agent) agent is strategically deployed over the network for Naïve Bayes classifier implementation. The IoT system is equipped with sensors that gather data on network nodes exhibiting abnormal behavior. The proposed multi-agent Naïve Bayes classifier methodology demonstrates its efficacy in promptly thwarting attacks with a lower execution cost. The NSL-KDD dataset serves as the evaluation benchmark for assessing the performance of the suggested classifier.

#### **Artificial Neural Networks**

ANNs draw inspiration from the intricate connectivity of the human brain, where numerous neurons form parallel pathways, engaging in complex interactions. Nodes within ANNs are interconnected through multiple links, creating a network that possesses the ability to learn and adapt. In a study outlined by authors in [132], Convolutional Neural Network (CNN), a specific type of neural network model, is utilized. The authors emphasize the advantages of employing CNNs for intrusion detection, highlighting their ability to incorporate traffic characteristics in raw data. The CNN model demonstrates enhanced accuracy and a reduced false alarm rate, addressing challenges associated with unbalanced datasets. The study also introduces a method for transforming raw traffic data into image format, mitigating computational costs.

In a separate research endeavor [133], a model for detecting network attacks is proposed, leveraging Neural Networks and DL. Evaluated on the CICIDS2017 dataset, this neural network algorithm proves highly accurate in identifying various types of network attacks. Additionally, an innovative technique for NIDS is presented in [134], utilizing a genetic algorithm (GA) to identify an improved feature subset referred to as Fuzzy C-Means (FCM). This approach significantly enhances overall detection performance, as demonstrated through evaluations on the NSL-KDD dataset. These advancements underscore the versatility and

effectiveness of Artificial Neural Networks in bolstering the accuracy and efficiency of intrusion detection systems across diverse datasets and scenarios.

### **Principal Component Analysis**

PCA stands as a statistical technique employed for dimensionality reduction in datasets characterized by numerous variables with interrelationships. An intriguing feature of PCA is its applicability to unlabeled data in training sets, making it a valuable unsupervised learning approach for dimensionality reduction.

A pioneering anomaly detection technique, Robust PCA (RPCA), was introduced in [135]. This technique is showcased through its application to network packet capture data, revealing its impact on various network attack detection systems. The proposed RPCA technique is rigorously tested on the DARPA dataset, encompassing diverse attack scenarios such as DDoS, assaults, IP sweeps, probing, and breaking. Impressively, the model achieved the lowest false positives while maintaining a high positive rate. Furthermore, the RPCA approach demonstrated its capacity to identify network threats, even in the case of an assault that had not been previously encountered or trained, emphasizing its robustness and effectiveness in anomaly detection.

### **3.1.5 Current Datasets for Intrusion Detection Systems**

In instances where an abundance of datasets is available online, models undergo training and testing using these datasets. While training datasets serve as a crucial method to validate proposed approaches, the challenge lies in evaluating these techniques and understanding their impacts across diverse datasets [136].

In the context of our research objectives, we conducted a comprehensive survey to assess the current datasets available for IDS. This exploration is motivated by our intention to develop an IDS tailored for VANETs. The unique characteristics of VANETs, including their dynamic nature and real-time data requirements, necessitate the careful selection of datasets that align with these specific parameters.

Our goal is to choose datasets that are suitable for VANETs and capable of providing real-time data to evaluate our forthcoming IDS model effectively.

The landscape of network threats has evolved, necessitating a continuous update of datasets used for IDS testing. Although datasets typically consist of both normal and abnormal network traffic, enabling the model to classify data based on labeled examples, they may not perform optimally when faced with new attacks, such as zero-day exploits. Various techniques have proven effective for IDS; however, each method employs distinct training and testing methodologies on datasets like NSL KDD, KDD CUP 99 [137], or DARPA [138]. Evaluation and comparison of the produced models hinge on criteria such as dataset variability, feature, and parameter optimization [139].

Selecting an appropriate performance metric is critical in IDS techniques. While accuracy and the confusion matrix are commonly used metrics, precision and recall derived from the confusion matrix, offer insights into algorithm efficiency and classification quality [140]. Despite their ability to measure algorithm efficiency, these metrics may not provide sufficient granularity in estimating IDS algorithms' performance, especially considering variations in accuracy across different types of attacks and the persistence of false-positive and false-negative issues.

This section outlines the existing datasets that show an apparent absence of real-world threats, attack representation, and outdated threats, limiting the efficiency of ML IDS techniques. However, limited attention has been given to the datasets employed for evaluating and scrutinizing proposed IDS models. IDS datasets are curated by amalgamating data from diverse sources, including network traffic flows [141], encompassing details about hosts, user activities, and system configurations. This information is indispensable for analyzing network attack patterns and identifying unusual activities. Therefore, the ongoing update and classification of IDS datasets based on their advantages and limitations are imperative.

Additionally, the study presents recent advancements in IDS datasets, offering research communities valuable resources for developing efficient ML-based IDS. The paper emphasizes the necessity for real-time datasets to yield results more pertinent to real-world scenarios and to substantiate the efficiency of proposed IDS in future applications.

### The Current IDS Datasets

In the contemporary landscape, organizations utilize diverse data formats for their specific objectives, contributing to public repositories. This section elucidates relevant pre-existing datasets:

1. **DARP** The Defense Advanced Research Project Agency (DARPA) dataset, among the earliest IDS datasets, comprises real-time and offline evaluations. Modeled on a US Air Force set up with restricted personnel, data was collected through computers connected to the internet, capturing host log files or network data packets. The dataset includes 41 attributes distinguishing normal and abnormal data, featuring a variety of TCP sessions simulating different attacks [142].

2. **KDD CUP 99** Derived from the DARPA 98 dataset, the public KDD 99 dataset is a crucial resource for IDS researchers. With 41 attributes assigned to each connection instance, it designates incidents as "normal" or "abnormal," categorizing attacks into four types: DoS, Probe, R2L, and U2R. KDD99 remains a standard in the IDS research community, given the limited availability of freely accessible public datasets [143].

3. **NSL-KDD** Developed to address KDD CUP 99's limitations, the NSL-KDD dataset in [137] stands out for its reasonable record count, making it cost-effective for running experiments on the entire set. Notably, it excludes redundant data records, ensuring classification is not biased towards frequently occurring instances. With 125,973 records in the training dataset and 22,544 in the test dataset, its size allows for full utilization without random sampling, yielding repeatable and comparable results across studies.

4. **UNSW-NB15** Introduced in [144], the UNSW-NB15 dataset comprises 42 features, including both categorical and numeric attributes. Split into training (UNSW-NB15-TRAIN) and testing (UNSW-NB15-TEST) sets, this dataset covers various network attack categories such as Shell code, worms, DoS, and Backdoor. It employs innovative techniques to generate features and is available online for research use.

5. **ISCX2012** Addressing the need for dynamic datasets due to evolving intrusion patterns, the ISCX2012 dataset captures traffic like SMTP, SSH, FTP, and

HTML. With features like realistic network and traffic, labeled dataset, complete interaction capture, and diverse intrusion scenarios, ISCX2012 provides entire packet payloads and profiles in pcap format, freely accessible for researchers [145].

6. **CIC-DDoS2019** CIC-DDoS2019 [146], a recent dataset, overcomes existing limitations and is designed for evaluating IDS algorithms on DDoS attacks. It includes 11 DDoS attacks captured over a real-time network, covering protocols like MSSQL, UDP, NTP, DNS, and SNMP [147].

7. **CIC-IDS-2017 & CSE-CIC-IDS-2018** Created with crucial criteria in mind, including attack variety, access protocols, labeled data samples, and network traffic capture, CIC-IDS-2017 [148], and CSE-CIC-IDS-2018 [149] datasets offer in-depth information on attacks and a conceptual understanding of network elements. The CIC-2017 dataset, widely used in research, has a signature similar to PCAP, and the latest CIC-2018 dataset introduces updated malware detection methods, featuring seven attack types, including brute-force, DoS, and Botnet, along with extensive network traffic and system logs [150].

Table 3.1 provides a comprehensive comparison of various datasets used in IDS; each dataset serves distinct purposes, such as DARPA and KDD CUP 99 for historical context, NSL-KDD for reduced redundancy, UNSW-NB15 for diverse attack categories, ISCX2012 for dynamic scenarios, CIC-DDoS2019 for DDoS evaluations, and CIC-IDS-2017 & CSE-CIC-IDS-2018 for detailed attack information.

Table 3.1: A Comparison of The Various Datasets.

Data Set	Developed By	Features	Attack Type	Volume of data	Description
DARPA	MIT Laboratory	41	Dos,U2R, R2L, Probe R2L, Probe	-	Absence of false-positive instances, No represent of real network traffic, Attack data patterns with differing consistency.
KDD CUP 99	University of California	41	Dos, U2R, R2L, Probe	5M points	Contains redundant and duplicate data samples
NSL-KDD	University of California California	41	Dos, U2R, R2L, Probe	150,000 points	An optimized dataset of KDD CUP 99 but has a restricted number of attack types
UNSW-NB15	Lab of UNSW Canberra	49	DoS, Backdoors, worm, Fuzzers	2M points	The training set is 175,341 records, and the testing set is 82,332 records, some issues like over-fitting, dimensionality, and imbalance in the dataset
ISCX2012	University of New Brunswick	IP flows	DoS, DDoS, Brute-force, Infiltration	2M flows	consist of network scenarios with intrusive activities and labelled data instances.
CIC-DDoS2019	University of New Brunswick	80	DDoS attacks LDAP, PortMap, UDP-Lag, NetBIOS,etc	More than 10GB	containing 12 different DDoS attacks, and a completely labelled dataset can be executed using TCP/UDP.
CIC-IDS -2017 & 2018	University of New Brunswick New Brunswick	80	Brute Force FTP, Botnet Heartbleed, Brute Force SSH, Infiltration, DoS, Web Attack, and DDOS	4.6 GB for 2017 and more than 50GB for 2018	solve the issue of high-class imbalance, have a wide range of attack categories, have limitations in data samples, and analyze files created by network flow

### Limitations and Challenges

Defining or substantiating the completeness and correctness of any proposed IDS without rigorous evaluation of current datasets poses a formidable challenge. Study outcomes reveal the formidable task of creating an accurate, scalable, robust, and protective IDS. Future research is anticipated to confront several limitations and challenges in dataset usage, including:

**-Availability of Known Attacks:** Openly available datasets primarily encompass a small fraction of known attacks, presenting a significant hurdle to real-world IDS implementation.

**-Zero-Day Attacks:** Despite escalating attack speeds and data collection lags, alternative dataset production methods are necessitated to address the detection gap for zero-day attacks. Timely public availability of datasets is crucial to keeping pace with evolving IDS models.

**-Specialized Datasets:** A limited selection of datasets is available for specialized IDS applications, such as SCADA, IoT, and Tor networks, posing a constraint on



tailored model development.

**-Fresh Malware Dataset:** The continual evolution of malicious activity necessitates a fresh malware dataset, as existing approaches often rely on outdated assumptions about the absence of new threats in current datasets.

**-Documentation Challenges:** Inadequate documentation accompanying newly available datasets hinders researchers, requiring custom techniques for tasks related to data collection, documentation, anonymization, and publication due to the absence of standardized tools.

**-Privacy Concerns:** Privacy emerges as a significant challenge in collecting new attack data, compelling researchers to employ customized methods for data-related tasks, given the absence of standard tools.

**-Data Mapping and Integration:** The lack of comprehensive data documentation complicates the mapping of datasets, rendering it impractical to introduce new attacks into existing datasets. Moreover, despite the existence of standard methods for exporting realistic traffic, insufficient information exists on integrating freshly gathered data into older datasets.

## Discussion

In the dynamic realm of cybersecurity, the behavioral patterns of network cyber attacks undergo constant evolution, necessitating the frequent update of available datasets. This iterative process ensures the development of diverse network traffic scenarios and adaptable attack patterns that are easily understandable and modifiable [151]. The careful selection of an appropriate dataset is paramount, considering that some datasets are collected by individual organizations and labs for specific research objectives, often remaining proprietary. Moreover, publicly available datasets may include records incompatible with current technological standards, and their statistical insufficiency poses challenges in attaining an ideal dataset [152].

This section delves into the examination of current datasets used to evaluate developed models in network IDS. It underscores the significance of an optimal dataset encompassing all communications involving various protocols, including both normal and attack scenarios. The emphasis is on the necessity for a diverse

range of up-to-date malware and attack categories to empower new IDS models to safeguard systems in contemporary contexts effectively. Furthermore, the datasets should come with comprehensive documentation detailing the testing environment, attack system infrastructure, victim system infrastructure, and various attack scenarios.

Notably, the CIC-IDS-2017 and CSE-CIC-IDS-2018 datasets emerge as meeting these criteria, featuring seven attack categories that vividly describe contemporary attack scenarios. These datasets have garnered significant attention from developers and researchers for building IDS using benchmark datasets [153] [154] [155] [156]. Given the robust characteristics of the CIC-IDS-2017 and CSE-CIC-IDS-2018 datasets, we have chosen to utilize them in our forthcoming IDS model, which will be presented in the upcoming chapter, recognizing them as the best-suited datasets for our specific IDS application.

## 3.2 Vehicular Ad-hoc Networks

Advancements in mobile communications and current trends in ad hoc networks allow for diverse deployment architectures of vehicular networks in highways, urban, and rural environments to support various applications with distinct Quality of Service (QoS) requirements [157]. The primary objective of VANET architecture, depicted in Figure 3.3 [158], is to facilitate communication among vehicles and between vehicles and fixed roadside equipment.

Communication in VANETs involves three primary possibilities [159], with additional variations based on these fundamentals.

**V2V Communication:** Utilizing OnBoard Units (OBUs) installed on vehicles, direct vehicular communication is established without relying on a fixed infrastructure.

**V2I Communication:** Vehicles communicate with the roadside infrastructure through equipped Road Side Units (RSUs). This communication is primarily used for information and data-gathering applications.

**Hybrid Communication:** Combines both V2V and V2I. In this scenario, a vehicle can communicate with the roadside infrastructure either in a single-hop or

multi-hop fashion, enabling long-distance connections to the Internet or to vehicles that are far apart.

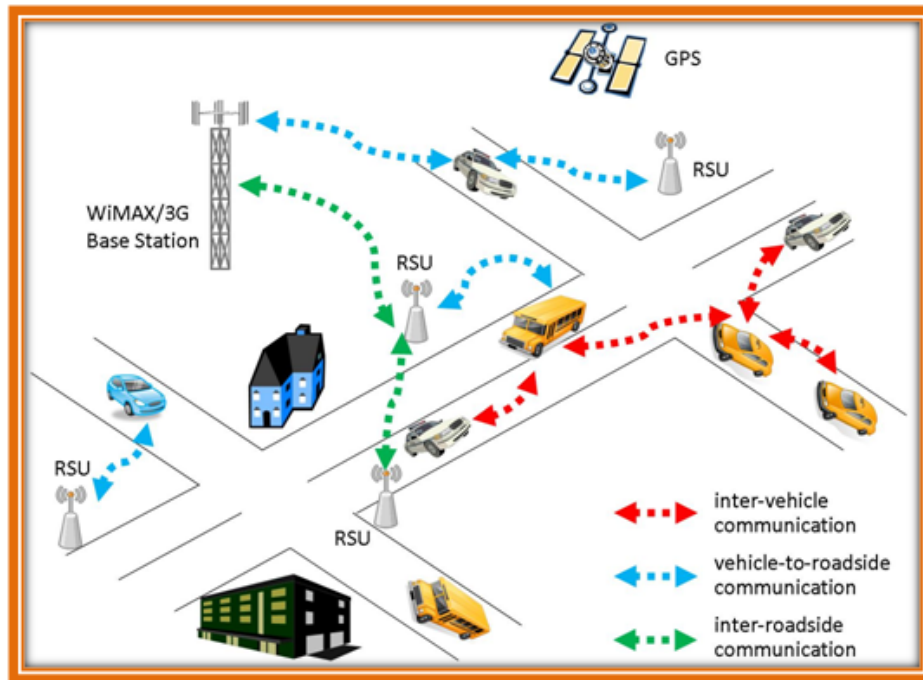


Figure 3.3: Basic architecture of VANETs.

As a specific class of Mobile Ad Hoc Networks (MANETs), VANETs inherit several characteristics, such as self-organization and mobile nodes. However, VANETs has features that make it more powerful and promising network [160], such as the following:

**Sufficient Power:** Power scarcity is less severe in VANETs compared to MANETs due to communication devices in vehicles being supported by stronger, rechargeable batteries.

**Fruitful Capabilities:** Vehicles have enough space, allowing the installation of devices with significant computing, communication, and sensing capacities, enabling powerful functions and high computational abilities.

**Predictable Mobility:** Unlike MANETs with random node movements, vehicle movements in VANETs are controlled by street topologies, traffic lights, and regulations. Future vehicle positions can be predicted based on roadway information.

**Large-Scale Application Scenarios:** VANETs are deployed in highway/urban

environments, constituting large networks with a high number of mobile nodes. In contrast, MANETs are typically studied in limited-size environments.

**Rapid Network Topology Changes:** Vehicles in VANETs move at varying speeds and change directions constantly, leading to dynamic network topologies and unstable communication links, potentially causing network partitions.

These characteristics of VANETs foster the development of diverse applications, categorized into three main groups: road safety applications, traffic efficiency and management applications, and infotainment applications.

### 3.2.1 Security in Vehicular Ad-hoc Networks

VANETs, functioning as distributed systems, present significant security and privacy challenges. Specifically, to fully harness the application potential of VANETs, particularly in traffic safety applications, security measures are essential to ensure their proper operation. Without adequate security, VANETs could potentially be exploited, posing risks to both traffic safety and management [161]. For example, a misbehaving vehicle broadcasting false warnings about an emergency braking event could lead to collisions with vehicles trailing behind. Moreover, the unique characteristics of VANETs, including limited infrastructure access, high node mobility, and the presence of valuable assets at risk, create additional complexities in security provisioning compared to general ad hoc networks [162].

Security in VANETs is a critical consideration, given the dynamic and interconnected nature of vehicular communication. Addressing the multifaceted challenges inherent in these networks involves a comprehensive understanding of various aspects. In particular, the security landscape in VANETs can be delineated into three primary aspects, each presenting unique concerns and requiring tailored solutions. See figure 3.4.

**Attack and Threats:** One of the foremost aspects of VANET security involves safeguarding against potential attacks and threats. Threat vectors encompass the entire VANET ecosystem, from the vulnerability of RSUs to the integrity of communication among connected vehicles. Security measures must contend with potential

attacks on the wireless interface, ensuring the reliability and trustworthiness of data exchanged in the network.

**Challenges:** VANETs operate in dynamic, real-time environments, presenting specific challenges for ensuring robust security. Key distribution, a fundamental challenge, involves securely managing cryptographic keys across a multitude of vehicles and RSUs. Real-time constraints dictate the need for security protocols to respond swiftly to dynamic vehicular conditions. Additionally, maintaining verification and consistency in data exchanges amid the network's dynamic nature poses a significant challenge.

**Requirements:** To establish a secure and resilient VANET environment, specific security requirements must be met. Access control mechanisms are essential for regulating interactions between vehicles and infrastructure based on predefined policies. Availability of communication services is paramount, even in the face of potential attacks. Authentication protocols play a crucial role in verifying the identity of vehicles and RSUs, preventing unauthorized entities from compromising the integrity of VANET communication.

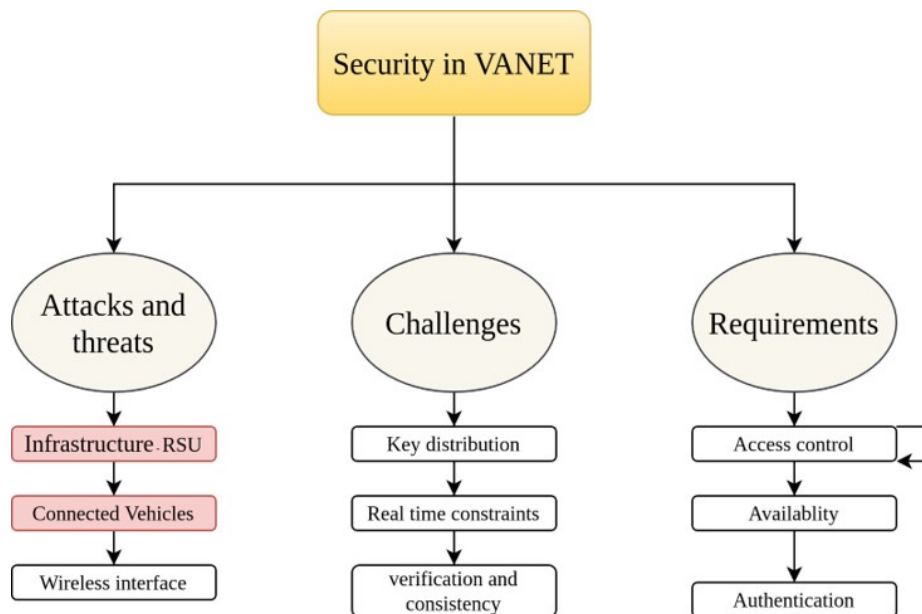


Figure 3.4: Security in VANET.

In addition, other security aspects have been considered for VANET in the following section:

**Identity management** is crucial in VANETs to ensure message integrity, authenticity, node authentication, and non-repudiation. Schemes for identity management are necessary to represent the identities of both vehicular nodes and RSUs securely. In essence, to establish its identity and eligibility for participating in VANET communications, an entity requires a certificate issued by the VANET Authority. In addition to common public key cryptography techniques, various specialized digital signature techniques, such as group signature [163], ring signature [164], and identity-based signature, are employed in VANET identity management. To ensure privacy-preserving node authentication, several anonymous authentication protocols [165] have been proposed for VANETs. Addressing the issue of nodes assuming the identity of others through Sybil attacks, various existing schemes [166] can be adopted.

**Message verification** in VANETs, crucial for ensuring message integrity and authenticity, requires each node to verify the content [167] and the digital signature of each received message. Due to the potential influx of numerous messages each node can receive, existing schemes aim to make message verification efficient and scalable. Some studies, for example, [168], attempt to reduce the resources needed for verifying each message. In contrast, others [169] aim to minimize the number of messages to be verified by each vehicle. However, these proposals currently face challenges in simultaneously ensuring efficiency, security, and application-friendliness.

For **Cooperation Enhancement and Misbehavior Detection** in VANETs, it is crucial to identify node misbehaviors to ensure the reliability of traffic safety applications and robust communications [170]. Various existing studies focus on verifying the content, such as location information, broadcasted by vehicular nodes. Additionally, some others leverage watchdog mechanisms designed for ad hoc networks to detect routing and packet relaying behaviors, which can be adapted for VANETs. To incentivize cooperation, schemes utilize mechanisms such as credit and reputation.

**Trust and Reputation** play a crucial role in VANETs and general networks, serving as influential tools to incentivize node cooperation and deter node misbehaviors [171]. Diverse trust models and metrics have been proposed, drawing from information theory [172], Bayesian theory, graph theory, and abstract algebra, applicable to both general ad hoc networks and VANETs.

However, a notable challenge remains in efficiently and reliably maintaining the reputation history of vehicular nodes, especially considering the high node mobility and potential pseudonym changes for privacy protection [173].

**Privacy protection** arises from the widespread use of wireless vehicular communications and periodic beacons in VANETs, posing significant challenges [174]. Eavesdropping on VANET communications allows a determined adversary to uncover the real identities of targeted nodes and create profiles based on application usage and personal information. To address this, various solutions have been proposed to enable privacy-preserving node authentication in VANETs [175], [176], [177]. However, the interception of periodic beacons by adversaries raises the risk of collecting the location history of any node. To safeguard the identity privacy of each node, pseudonyms are often employed instead of real identities in VANET communications. A pseudonym serves as a temporary identifier with no obvious link to the real identity, typically comprising a certificate, MAC address, and IP address. Despite this measure, adversaries can potentially deduce the real identity by aggregating personal information and movement trajectories associated with a specific pseudonym.

In summary, this discussion of security aspects highlights several critical research issues in VANETs. Addressing these open challenges is vital for the successful implementation of VANETs, making them pivotal concerns within this dissertation. Furthermore, existing security and privacy protection applications often lack meticulous consideration of the practical constraints inherent in VANETs. Therefore, as part of this dissertation, we aim to study VANET networks and build a realistic simulation that enables us to apply ML techniques to solve VANET security issues and be readily applicable in real-world VANET scenarios.

### 3.2.2 Simulation for Vehicular Ad-hoc Networks

The evaluation process of existing trust models designed for external communication systems within autonomous systems primarily relies on simulation systems [178]. Researchers emphasize the pivotal role of simulation systems, defined by Shannon as "the process of designing a model of a real system and conducting experiments with this model for the purpose of understanding the behaviour of the system and evaluating various strategies for the operation of the system" [179]. Within the domain of VANETs, simulation systems play a crucial role, necessitating the application of any new protocol or security method within these systems due to the substantial costs associated with real-world implementations. The field of VANETs employs various simulation systems classified into three categories, as depicted in Figure 3.5 [180]:

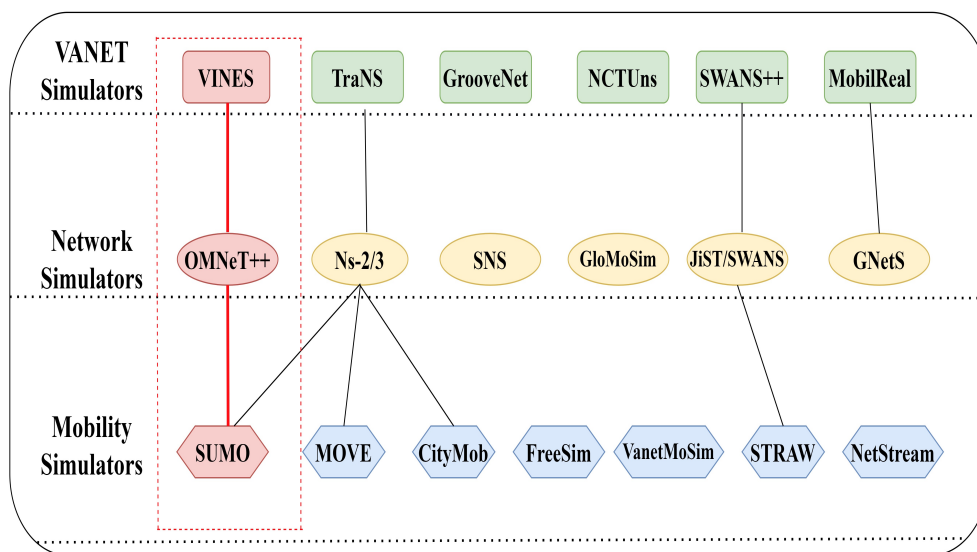


Figure 3.5: Taxonomy of VANET Simulation Software.

**Vehicular Mobility Generators:** These programs are designed to generate realistic vehicular mobility patterns within simulation environments, enabling the assessment of communication systems' performance under diverse vehicular movement scenarios.



**Network Simulators:** Focused on broader network functionalities, these simulators provide a platform to evaluate communication protocols and strategies in the context of vehicular networks, including aspects beyond vehicular mobility.

**VANET Simulators:** Specifically tailored for simulating VANETs, these programs offer a comprehensive environment to assess the performance of communication protocols and security measures within the unique dynamics of VANETs.

The utilization of simulation systems, categorized into these three classes, facilitates a cost-effective and efficient means of evaluating and refining protocols and security methods within the challenging and dynamic domain of VANETs.

In VANET simulation, researchers have harnessed the power of various tools renowned for their effectiveness and extensive capabilities. In addition to the three key simulation software tools discussed below, three more standout tools have played a pivotal role in shaping the research landscape of VANET simulation:

#### **1-Mobility simulator: SUMO**

SUMO is a prominent mobility simulator widely recognized for its capabilities. Developed by the German Aerospace Center, it has been freely available since 2001 and became part of the Eclipse Project foundation in 2017. SUMO is structured into distinct modules, each addressing specific aspects of the network to be simulated. As described in [181], SUMO facilitates traffic simulation and analysis, enabling the implementation and analysis of new traffic strategies before real-world deployment. Additionally, SUMO has been suggested as a tool for developing and validating automated driving functions through various X-in-the-loop and DT approaches. To effectively use SUMO, a network topology file is crucial, serving as the starting point for defining routes in a separate file. A typical SUMO project includes files like `file.net.xml` for network topology, `file.rou.xml` for routes within the network, and `file.add.xml` containing additional network-related information. SUMO operates as a purely microscopic traffic simulation platform, where each vehicle is explicitly defined by essential parameters such as an identifier name, departure time, and the specified route through the network, see figure 3.6. For more detailed representation, additional properties like departure and arrival characteristics, including lane choice, velocity, and position, can be defined. Vehicles are categorized by type, encapsulating their physical attributes and movement model

variables. SUMO also incorporates pollutant and noise emission classes for each vehicle, with supplementary variables enabling the definition of their appearance in the simulation's graphical user interface.

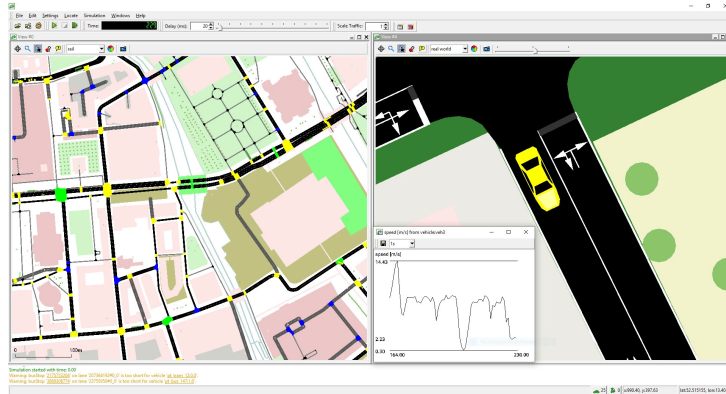


Figure 3.6: Map and Vehicles in SUMO simulator.

In terms of outputs, SUMO offers a versatile array of results for each simulation run. These outputs range from simulated induction loops to detailed information about single vehicle positions recorded at each time step for all vehicles. Additionally, SUMO provides complex data, such as details about each vehicle's trip or aggregated measures along streets or lanes. Beyond conventional traffic metrics, SUMO includes models for noise emission and pollutant emission/fuel consumption. Furthermore, the simulation allows interaction with an external application through a socket connection, enhancing its capabilities.

In SUMO project, Traffic Control Interface (TraCI) serves as a method to establish a connection between road traffic and network simulators [182]. This integration allows access to a live road traffic simulation, enabling the extraction of simulated object values and the real-time manipulation of their behavior. This dynamic interaction facilitates the control of vehicle behavior during the simulation runtime, offering valuable insights into the impact of VANET applications on traffic patterns.

TraCI employs a client/server architecture based on TCP to facilitate access to Sumo, with Sumo operating as the server in this setup. This approach enhances the ability to retrieve and manipulate information seamlessly within the simulation environment.

**2- Network simulator: OMNeT++**

On the other hand, OMNeT++ stands for Objective Modular Network Testbed in C++, and it is the most renowned network simulator. It's a component-based simulation library written in C++ developed to simulate various communication networks. OMNeT++ is not a network simulator but a framework that allows you to design and construct your network simulations [183].

Model frameworks, developed as independent projects, provide domain-specific functionalities like support for sensor networks, wireless ad-hoc networks, Internet protocols, performance modeling, and photonic networks. OMNeT++ boasts an Eclipse-based Integrated Development Environment (IDE), a graphical runtime environment, and various accompanying tools.

The framework consists of different components, including a simulation kernel library (C++), the Network Description (NED) topology description language, an Eclipse-based IDE, a graphical runtime environment (QtEnv), a command-line interface for simulation execution (CmdEnv), utilities like a makefile creation tool, and comprehensive documentation.

**3- VANET simulator: VEINS**

Delving deeper, VANET simulators represent a combination of mobility and network simulators. The network simulator focuses on modeling communication protocols and the transmission of messages among nodes, while the mobility simulator governs the movement of individual nodes. To achieve our objective of this thesis, the chosen framework for VANET simulation is VEINS [184].

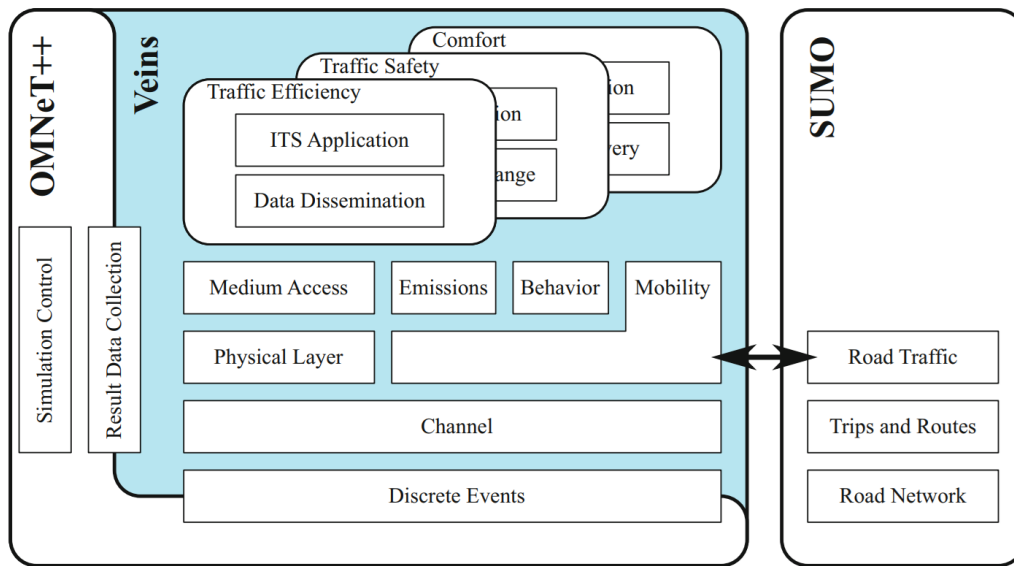


Figure 3.7: High-level architecture of Veins.

Veins is an open-source framework dedicated to simulating vehicular networks and operates based on the SUMO and OMNeT++ platforms. Figure 3.7 illustrates all the modules comprising the Veins architecture [185]. Initially, an OMNeT++ node is instantiated for each CAV within the simulation and is then synchronized with its movement in the road traffic simulator. This parallel simulation approach, encompassing both network and mobility simulations, occurs simultaneously due to a bidirectional coupling facilitated by the standardized TraCI connection protocol. This protocol enables message exchange between OMNeT++ and SUMO via TCP connections during the ongoing simulation. Veins incorporates different extensions to support various protocol stacks like IEEE 802.11p. In essence, Veins serves as an execution environment for user-programmed applications, facilitating the modeling of new environments and applications. However, it is essential to note that Veins relies on both SUMO and OMNeT++ for accurate results, and any discrepancies or bugs in either can compromise the reliability of Veins' output. Veins is compatible with Linux, Windows, and macOS, offering a versatile simulation environment with interconnected modules for seamless communication, as depicted in the previous figure.

### **3.3 Electric Vehicles Routing Simulation and Optimization**

In contemporary logistics, the integration of EVs into corporate fleets has become a prevalent practice [186]. This shift towards adopting EVs underscores the growing importance of addressing the EVRP within the realm of logistics, primarily driven by the environmental imperative to reduce carbon emissions [187]. The EVRP model entails assigning a single EV to each client node, with multiple EVs potentially utilizing a single charging station [[188]; [189]]. A fundamental element of this vision is anticipated to be the next generation of vehicles, integrating new sensing, communication, and social capabilities. By providing mobile wireless sensing and communications, vehicles can facilitate data access, which is essential for realizing smart cities.

Given the extensive adoption of EVs in logistics transportation, there has been a growing focus on EVRPs in recent times. These problems pose significant complexity, involving the joint optimization of routes and recharging strategies. The integrated recharging strategies encompass various factors, including the selection of recharging stations, the choice between fast or slow recharging modes, and decisions related to partial or full charging. The lack of supporting charging infrastructure is a pivotal factor, and deploying such infrastructure poses a challenging problem [190]. It inadvertently requires changes to existing civil infrastructure, incurring significant costs and time for implementation. While the car industry experiments with larger and more powerful batteries, suggesting coverage up to 400km without intermediate charge, it is argued that, for environmental reasons, future batteries should have reduced capacity. This creates a need for new approaches to charging EVs that overcome the lack of supporting infrastructure, adapt to the existing civil infrastructure (i.e., road network), and avoid the need for new, space-consuming, and environmentally unfriendly batteries in cars.

Crucially, SoC of EV batteries within an acceptable range is imperative to ensure optimal performance. The introduction of smart charging methods, as a novel approach, plays a pivotal role in maintaining balance by facilitating both charging

and discharging activities to prevent exceeding energy peaks [191]. This innovative technique also enables EV batteries to supply energy back to the grid during discharge. Using the on-board unit, vehicles can communicate with each other and with RSUs, enabling smart application solutions and enhancing road safety and traffic management. To effectively address the complexities inherent in the EVRP, a multi-objective optimization strategy has been employed, as detailed in [192]. Furthermore, [193] proposes a two-stage simulation-based heuristic for the EVRP. In the initial stage, EV routes are determined by considering factors such as expected waiting times at charging stations, while the subsequent stage corrects any infeasible solutions by penalizing time-window violations and late returns to the depot.

Simulation tools, such as SUMO, play a crucial role in addressing the challenges posed by EVRPs and optimizing both conventional and EV journeys [194], [195]. There are huge benefits of using such software specifically aids in solving EVRP: **Route Optimization:** Simulation software allows for the modeling and simulation of various routes for EVs. It considers factors such as traffic conditions, road layouts, and charging station locations to optimize routes. This helps in determining the most efficient paths for EVs to minimize travel time and energy consumption. **Recharging Strategy Analysis:** EVRPs involve optimizing not only the routes but also the recharging strategies. Simulation software like SUMO enables the analysis of different recharging strategies, including the selection of charging stations, the timing and duration of recharging, and the choice between fast or slow recharging. This allows for a comprehensive evaluation of the impact of different strategies on the overall efficiency of the EV fleet.

**Performance Evaluation:** Through simulation, researchers and planners can assess the performance of EV fleets under various scenarios. This includes studying the effects of different traffic conditions, vehicle types, and charging infrastructure layouts on the overall efficiency of EV operations.

**Scenario Testing:** Simulation software facilitates scenario testing, allowing stakeholders to assess the feasibility and effectiveness of different EVRPs under diverse conditions. This helps identify potential challenges and refine routing and recharging strategies accordingly.

**Risk Mitigation:** By simulating EVRPs before actual implementation, stakehold-

ers can identify and mitigate potential risks. This includes evaluating the impact of uncertainties such as fluctuating energy prices, changes in traffic patterns, and variations in charging station availability.

**Decision Support:** Simulation results provide valuable insights into the performance and feasibility of different EVRPs. This information serves as a foundation for decision-making, allowing planners and decision-makers to choose optimal routing and recharging strategies based on the simulation outcomes.

In summary, simulation software acts as a powerful tool for modeling, analyzing, and optimizing EVRP. It provides a virtual environment for testing different scenarios, refining strategies, and making informed decisions to enhance the efficiency of EV logistics in various settings. In the upcoming chapter, we will introduce an EV Routing Simulation and Optimization framework designed to address an EVRP through a detailed examination of a real-world case study.

# Chapter 4

## Routing Optimization for Electric Vehicle Using SUMO Simulation

### 4.1 Introduction

EVRP has gained significant attention in recent years, driven by the global push to adopt sustainable transportation methods and reduce environmental impact. However, the widespread adoption of EVs is hindered by factors such as limited driving range and the availability of charging infrastructure [196]. Numerous studies have addressed these challenges to develop algorithms and optimization techniques for EVRP [197] [198]. These approaches aim to optimize the routing of EVs by considering factors such as energy consumption, travel time, and charging station availability. These solutions are essential to ensure that EVs can efficiently travel between origins and destinations while considering the constraints imposed by their limited battery capacity [199]. In recent years, microscopic traffic simulators have become an essential tool for understanding the complexities of urban transportation systems. SUMO has gained popularity among these simulators due to its open-source nature, flexibility, and extensibility [200]. SUMO provides a platform for researchers to simulate different traffic scenarios and evaluate the performance of various traffic management strategies [201]. TraCI is an integral component of SUMO, allowing for real-time interaction with the traffic simulation. With TraCI, researchers can dynamically modify the behavior of vehicles, traffic lights, and



other elements of the traffic environment, providing a powerful tool to investigate and optimize traffic flow [202].

This chapter proposes a new contribution to solving EVRP using SUMO simulation and the TraCI interface. The primary objective is to optimize traffic flow and determine the best route for each EV in the given scenario. Additionally, the TraCI model used in our approach can dynamically choose the best alternative route for EVs in case of heavy traffic in the main route. This could avoid unnecessary waiting time and improve the system's overall efficiency. Our approach considers the current state of the traffic, the battery capacity of EVs, and the availability of charging infrastructure, providing an adaptive and efficient solution to the EVRP.

This study tested the applied approach in a large-scale area in Apulia, Italy, effectively demonstrating its practicality and success in addressing real-world challenges. Our approach enabled vehicles to select the best route to reach their destinations, considering traveling time in varying traffic conditions. This demonstrates the practical benefits of our methodology in optimizing EV routing under diverse circumstances.

## **4.2 Background: Routing Optimization Review**

### **4.2.1 Electric Vehicle Routing Problem**

EVRP is an extension of the classic Vehicle Routing Problem (VRP) that incorporates the unique characteristics of EVs, such as limited driving range, battery charging requirements, and charging infrastructure availability. Indeed, EVRP has gained significant attention in recent years due to the growing adoption of EVs and the need to develop efficient routing strategies to support their widespread use. Hiermann et al. [203] introduce the Electric Fleet Size and Mix Vehicle Routing Problem with Time Windows and Recharging Stations (E-FSMFTW) for decision-making on fleet composition and vehicle routes. They develop a hybrid heuristic combining Adaptive Large Neighborhood Search and labeling procedures, demonstrating its effectiveness using a benchmark set for E-FSMFTW. Schiffer

et al. [204] present a location routing method that simultaneously addresses EVR and charging station siting to support the decisions of logistics fleet operators. It considers various real-world recharging constraints and alternative objective functions, such as minimizing distance, vehicle count, charging station count, and total costs. Goeke and Schneider [205] investigated routing a mixed fleet of EVs and other vehicles, further emphasizing the importance of considering charging infrastructure in routing decisions. María A. et al. [206] developed an optimization-based approach for the EVRP, incorporating smart charging methods to minimize charging/discharging costs. The model considers power grid limits and balancing needs to avoid exceeding maximum energy peaks.

### 4.2.2 Simulation of Urban Mobility

SUMO is a popular open-source software framework that enables modeling, simulation, and analysis of various transportation systems, including cars, buses, and trains [181]. SUMO has been used in numerous studies to investigate and evaluate traffic management strategies, road infrastructure design, and transportation planning [207][208]. The software's flexibility and versatility have made it popular for simulating a wide range of urban mobility scenarios, such as investigating the impacts of ITS on traffic flow and optimizing public transportation networks [209]. In recent years, SUMO has been enhanced with various advanced features, such as support for parallel processing and real-time traffic data integration, further expanding its capabilities [210][211]. Furthermore, SUMO has been integrated with other simulation tools like MATSim and Omnet++ to support modeling more complex and interconnected transportation systems [212] [213]. Yin, R et al. [214] proposed a simulation-based bi-level model for the continuous network design problem (CNDP). The approach integrates micro and macro traffic dynamics in SUMO and proposes lane width expansion as a decision variable. The flexibility and robustness of SUMO make it a valuable tool for studying and developing ITS that can effectively accommodate EVs in large-scale traffic scenarios.

### 4.2.3 Traffic Control Interface

TraCI is an interface that allows real-time interaction with SUMO simulations, enabling the dynamic manipulation of traffic elements such as vehicles, traffic lights, and routes. TraCI has been employed in various studies to evaluate and optimize traffic management strategies, such as developing adaptive signal control algorithms to improve intersection performance [215]. Another study by Yang et al. [216] integrated SUMO and middleware tool TraCI with NS3 to evaluate the performance of four different topology-based routing protocols for VANETs. Similarly, Kusic et al. [217] proposed an approach that uses TraCI as an interface to talk to outside controllers and ensure that the simulated traffic situation is constantly calibrated as actual traffic dynamics change.

### 4.2.4 Optimization Techniques for EVRP

Researchers have used many optimization techniques to solve the EVRP efficiently. Metaheuristics such as genetic algorithms, particle swarm optimization, and ant colony optimization are examples of these techniques. [218]. Laporte et al. [219] proposed a metaheuristic algorithm based on tabu search and guided local search to solve the EVRP with stochastic travel times. Sadati et al. [220] proposed a mixed integer linear programming model and a hybrid General Variable Neighborhood Search and Tabu Search approach to solve the Multi-Depot Green Vehicle Routing Problem. Bruglieri et al. [221] presented a variable neighborhood search method for the EVRP with charging stations, which achieves better performance than current state-of-the-art techniques. The present study builds upon the existing research on EVRP, SUMO, TraCI, and optimization techniques to develop a novel approach for optimizing EV routing in large-scale traffic scenarios. Our methodology addresses the challenges posed by EV routing and traffic optimization, contributing to the development of intelligent transportation solutions tailored to the unique characteristics and requirements of EVs. In addition to considering time windows as constraints, our approach also focuses on minimizing traveling time, which is the actual time it takes for a vehicle to travel from one location to another, considering various factors such as distance, speed limits, and traffic conditions. Minimizing

traveling time is often one of the primary objectives in routing problems, as it directly impacts the overall efficiency of the transportation system.

## 4.3 Modeling and Methodology

The EVRP is a well-known optimization problem in transportation planning, involving the determination of the optimal routes for a fleet of EVs to serve a set of customers while considering factors such as travel time, battery capacity, and charging infrastructure availability. The problem can be formulated as follows: Given a set of  $n$  Points of Interest (POI) and a set of  $m$  EVs with a limited driving range, find the optimal routes to serve all the customers while minimizing the total travel time and ensuring that the EVs can complete their tours without running out of battery power or delay. The problem must also consider the availability and location of charging infrastructure along the routes. This study proposes a new approach for solving the EVRP using the SUMO traffic simulation software and the TraCI interface. The proposed approach involves three main steps: (1) model initialization and input data preparation, (2) simulation-based optimization using TraCI, and (3) output analysis.

### 4.3.1 Model Initialization and Input Data Preparation

The first step involves initializing the SUMO simulation model and preparing input data for the EVRP. The SUMO model consists of a road network, a set of charging stations, a set of POI, a fleet of EVs that will serve the customers, and a fleet of other vehicles that represent the real traffic conditions on the network. The input data for the EVRP includes the customer locations, the EV driving range, battery capacity, the charging station locations, and the charging infrastructure availability and capacity.

### 4.3.2 Simulation-Based Rerouting Using TraCI

The second step involves the simulation-based optimization of the EVRP using the TraCI interface to communicate between SUMO simulation model and the optimization algorithm. The TraCI interface allows real-time interaction with the SUMO simulation, enabling the dynamic manipulation of traffic elements such as vehicles, roadside units, traffic lights, and routes.

The proposed model employed in this study consists of a set of steps written in Python on TraCi interface as follows:

- Import required libraries (os, sys, traci, sumolib).
- Set up the SUMO\_HOME environment variable if it is not present.
- Start SUMO and TraCI with necessary command-line arguments.
- Define time horizon and time slot duration.
- Define threshold battery capacity.
- Define the function to get available charging stations.
- Define a function to check traffic congestion.
- Define a function to get the least congested route.
- Get the list of charging stations.
- Loop through time slots and advance the simulation.
- Get the list of vehicles and their current battery capacity.
- Get the charging station status at the current time slot.
- Check the battery capacities of all vehicles and choose the shortest route based on battery capacity and charging station availability.
- Retrieve the route with the minimal congested route and update the vehicle's route to the least congested path.

- Update the charging station availability status.
- Save the simulation results and close the TraCI connection.

The code starts by setting up the SUMO environment and defining the time horizon, time slot duration, and battery capacity threshold. It then defines three main functions: `get_available_charging_stations`, `is_congested`, and `get_least_congested_route`. The first function returns a list of available charging stations at the current time slot, the second checks if a road segment is congested based on average speed and vehicle count, and the third finds the least congested route between two points. The simulation iterates through the defined time slots, updating the charging stations' status and obtaining each vehicle's current battery capacity. It then checks if the battery capacity of a vehicle is below the defined threshold. If so, it selects the shortest route to the nearest available charging station or, if none are available, the shortest route to the vehicle's destination. If the battery capacity exceeds the threshold, it selects the shortest route to the destination. Then, the code checks for traffic congestion on the selected route and finds the least congested alternative. It updates the vehicle's route to the least congested path and sets the route color based on the vehicle's battery capacity and the availability of charging stations. Finally, the code saves the simulation results in a CSV file and closes the TraCI connection.

### 4.3.3 Output Analysis

After running the simulation, the output data can be analyzed and visualized to gain insights into the system's behavior. The simulation output files include "tripinfo.xml," which provides information on each vehicle's trip; "summary.xml," which provides a summary of the simulation results; and "battery.xml," which provides information on the battery capacity of each EV at each time step. The tripinfo.xml file can calculate metrics such as travel time, waiting time, and distance traveled for each vehicle. The summary.xml file provides system-level metrics such as total travel time, total waiting time, and total distance traveled.

## 4.4 Case Study: The Southern Italian Region of Puglia

### 4.4.1 Description of the simulation model

A real case study was carried out in the southern Italian region of Puglia to analyze a network of customer nodes and their transportation requirements. The microscopic traffic simulation tool SUMO was utilized to simulate this scenario accurately. By replicating actual traffic conditions, a realistic environment was established to study the behavior of EVs in this network.

Each vehicle in the simulation had explicit definitions for its unique starting point and endpoint. These vehicles traveled independently across the network, taking into account road conditions and traffic flow. This approach facilitated an assessment of the efficiency and effectiveness of the transportation system in meeting customer demands.

Figure 4.1 showcases the SUMO simulation graphic interface, where the key components of the network are visualized. The Depot, represented by the light green color, serves as the starting point for the EVs. The customer nodes (CN), depicted in blue, indicate the locations where deliveries need to be made. The charging stations (CP), represented in yellow, are strategically placed throughout the network to facilitate the recharging of EV batteries.

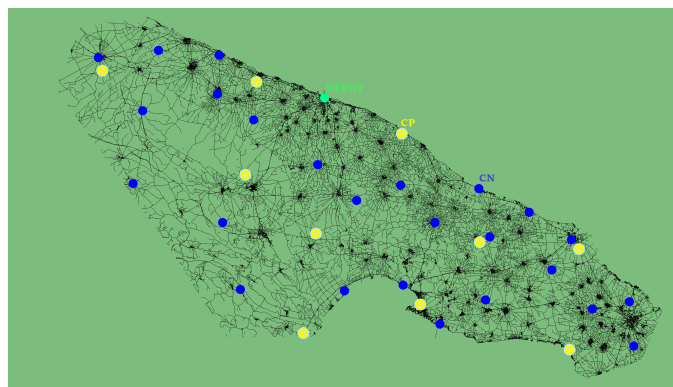


Figure 4.1: SUMO Simulation view of Region of Puglia

The node network itself consists of several interconnected nodes. At the core is the Depot Node, from which all EVs originate. There are 25 customer nodes (CN1..CN25) spread across the network, representing the destinations where freight needs to be delivered. Additionally, 10 charging station nodes (CP1..CP10) are strategically placed to provide charging facilities for the EVs. The distances between the connected pairs of nodes, measured in kilometers, are detailed in Table 4.1, and the location of connected nodes and charging points are shown in Figure 4.2.

Table 4.1: Distance Between the Nodes

	tdij (km)		
CN1	CN7	CN13	
	22	59	
CN2	DEPOT	CN5	CP2
	38	49	21
CN3	DEPOT	CP3	CP4
	72	63	48
CN4	CN14	CP3	
	29	31	
CN5	CN2	CN14	CP2
	49	27	53
CN6	CN12	CP4	CP5
	51	19	63
CN7	DEPOT	CP4	CP2
	64	30	38
CN8	DEPOT	CN9	
	59	53	
CN9	CN8	CN15	
	53	32	
CN10	DEPOT	CP4	CP5
	41	67	21
CN11	DEPOT	CN13	CP1
	61	97	51
CN12	CN6	CP3	CP4
	51	38	59
CN13	DEPOT	CN1	CP1
	103	59	73
CN14	CN4	CN5	
	29	27	
CN15	CN1	CP1	CP5
	32	66	73
CN16	CP5	CN8	CP8
	55	52	37
CN17	CP7		
	63		
	CN11	CP7	



CN18	40	6			
CN19	CN25				
	22				
CN20	DEPOT		CP5		CN25
	30		34		36
CN21	CN13		CN5	CN2	CP9
	56		46	23	34
CN22	CP3				
	56				
CN23	CN6				
	43				
CN24	CP3		CN12		
	30		23		
CN25	CP1	DEPOT	CN19	CP6	CN20
	29	31	21	16	36
CP1	CN11		CN13		CN15
	51		73		66
CP2	CN2		CN5		CN7
	21		53		38
CP3	CN3		CN4	CN12	
	63		31	38	
CP4	DEPOT	CN3	CN6	CN10	CN12
	73	48	19	67	59
CP5	DEPOT		CN6	CN10	
	51		63	21	
CP6	CN25			CN15	
	16			21	
CP7	CN18			CN17	
	6			63	
CP8	CN16			CN13	
	89			25	
CP9	CN21			CN4	
	34			66	
CP10	CN1			CN12	
	100			142	

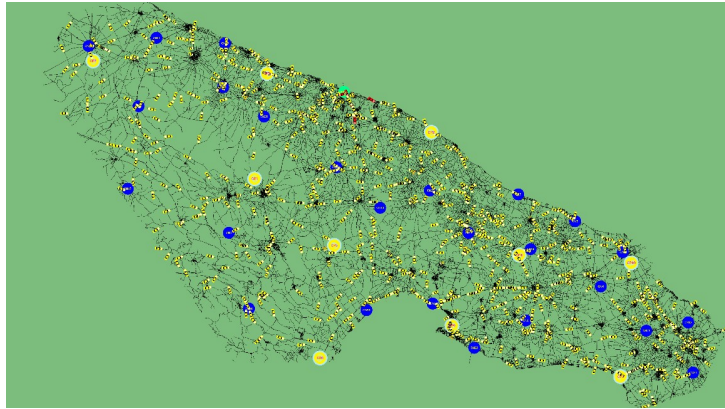


Figure 4.2: Traffic on connected nodes and charging points

The objective of this study was to efficiently meet customer demands using a fleet of 14 EVs (EV1...EV14). Each vehicle in the fleet was characterized by specific parameters, such as battery capacity, charging rate, and weight capacity, as outlined in Table 4.2. The EVs operated at an average speed of 100 km/h, aiming to deliver goods and fulfill customer requests on time.

Table 4.2: Distance Between the Nodes

	$B_k$	$C_k$	$Q_k$	$g_k$	$end_k$
EV1	3.8	300	58	43	CN1
EV2	5.5	350	100	22	CN15
EV3	4.4	400	80	22	CN6
EV4	2	250	52	22	CN12
EV5	3.4	450	52	43	CN4
EV6	3.3	600	60	22	CN14
EV7	8.3	300	100	43	CN7
EV8	2.2	450	58	43	CN23
EV9	3.5	350	80	22	CN19
EV10	4	400	52	43	CN17
EV11	3.8	250	90	22	CN24
EV12	3.1	450	52	43	CN5
EV13	4.8	600	80	22	CN22
EV14	2.6	300	100	43	CN4

Where  $B_k$  is the amount of time to charge EV fully  $k \in \mathcal{K}$  [h],  $C_k$  is the weight capacity of EV  $k \in \mathcal{K}$  [kg],  $Q_k$  is the battery capacity of EV  $k \in \mathcal{K}$  [kWh],  $g_k$  is the recharging rate of EV  $k \in \mathcal{K}$  [kW], and  $\text{end}_k$  is the end node of EV  $k \in \mathcal{K}$ .

To model the time dynamics of the system, we divided the 12-hour time horizon into 20-minute time slots, resulting in 36 time slots in total (3 per hour). Each time slot represented a specific period during which customer demands and charging station services could be initiated. By incorporating time frames for customers and charging stations, we ensured that services were provided within the designated time frames. Customers were also required to provide information about their freight demand and the estimated duration of service required at their respective nodes. To evaluate the efficacy of our approach, we conducted four distinct scenarios within the simulation framework. The first scenario utilized the default SUMO routing algorithm without considering traffic conditions. The second scenario also has the same default algorithm as the traffic of other vehicles during the simulation. The third scenario involved the implementation of the TraCi script, which incorporated traffic conditions. The fourth scenario included the TraCi script and a Rerouting algorithm to adapt dynamically to changing traffic conditions. In each scenario, we simulated a traffic scenario consisting of 2000 vehicles. These scenarios allowed us to assess the impact of our approach under various traffic conditions and determine its effectiveness in optimizing routing decisions.

## 4.5 Results and Discussion

The experiments were conducted using high-performance hardware, including an Intel processor I9 with a clock speed of up to 5.20 GHz, DDR4 64GB RAM, and an RTX 3090 24G GPU. This powerful computing setup ensured efficient processing and execution of the experiments, enabling us to gather accurate and reliable results. In our scenario, the travel of EVs was simulated using the SUMO simulation model, which integrated various parameters such as energy consumption, battery SoC, and delivery schedules. The EVs followed customized routes to efficiently reach the customer nodes (CNs) and complete their cargo deliveries while considering the availability of energy usage and charging points (CPs) along their routes. By

incorporating the Traci script with rerouting capabilities, the EVs in the simulation could dynamically adjust their routes in response to real-time traffic conditions. This algorithmic approach aimed to optimize energy usage and played a crucial role in minimizing waiting times caused by traffic congestion or blockages on the road, as shown in Figure 4.3. During the simulation, it was observed that using the default SUMO Algorithm with traffic, certain EVs, namely EV2, EV3, EV4, EV6, EV8, EV9, and EV11, encountered blockages and experienced delays in their travel due to high traffic or adverse road conditions. Even in the 3rd scenario, the basic Traci model experienced blockage for some EVs, namely EV3, EV6, EV8, and EV11, which also caused delays in travel time. However, thanks to the Traci script's rerouting capabilities in the fourth scenario, these EVs were redirected to alternative paths to bypass the congestion or blockages. As a result, they could continue their journeys without prolonged waiting times, ensuring smoother and more efficient travel.

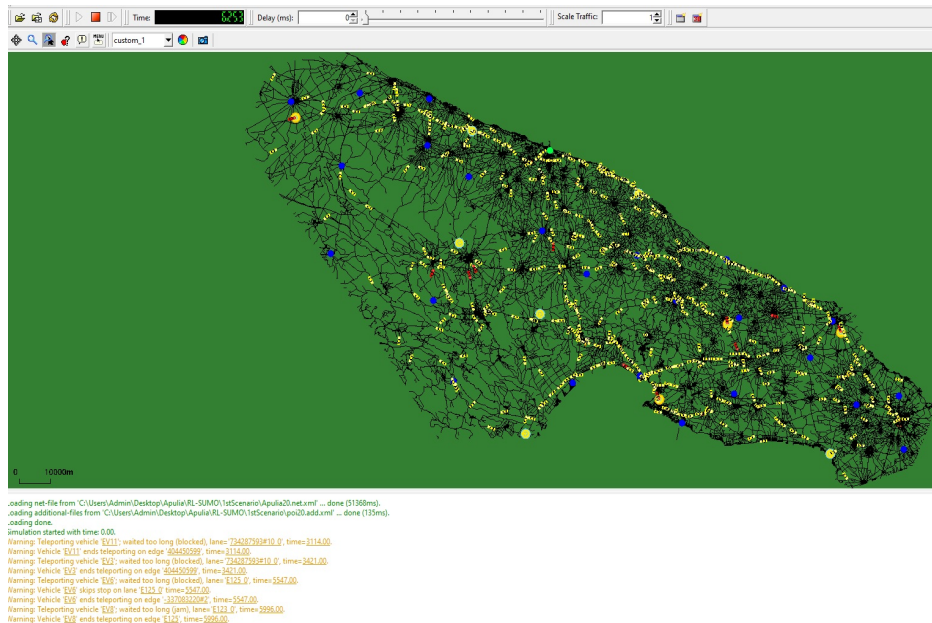


Figure 4.3: Blocked EVs in the waiting process

The Traci rerouting algorithm significantly reduced waiting times on the road. By avoiding areas with heavy traffic or incidents, the EVs could reach their destinations more quickly, improving overall travel time. Additionally, the algorithm

allowed the EVs to adapt their routes based on energy considerations and the availability of charging infrastructure. By considering the energy levels of the EVs and the proximity of charging stations, the algorithm guided the EVs to make optimal decisions regarding recharging or discharging their batteries, ensuring they remained within acceptable energy limits without compromising their delivery schedules. Integrating the Traci script with rerouting capabilities in the SUMO simulation offered multiple advantages. It enabled the EVs to dynamically respond to real-time traffic conditions, avoiding waiting times and optimizing travel routes. Simultaneously, the algorithm accounted for energy considerations, ensuring efficient usage and management of the EVs' battery resources. The results of our analysis were generated efficiently, with the outcomes compiled and presented in a concise format in Table 4.3. This table encompasses the total traveled time and distance for each investigated scenario. These metrics provide valuable insights into the performance and efficiency of our proposed approach.

Table 4.3: SCENARIOS RESULTS

	SUMO without traffic condition		SUMO with traffic condition		TraCi with traffic condition		TraCi with Rerouting	
	route Length(m)	duration (s)	route Length(m)	duration (s)	route Length(m)	duration (s)	route Length(m)	duration (s)
EV1	112848	4108	112848	4108	112848	4108	112848	4108
EV2	177361	7944	177361	8479	177361	7944	177361	7944
EV3	114169	6705	114169	7311	114169	7047	115743	6881
EV4	201936	9928	201936	10486	201936	9928	201936	9928
EV5	155567	8221	155567	8221	155567	8221	155567	8221
EV6	285311	19865	285311	20194	285311	19613	288155	19283
EV7	586229	29546	586229	29546	586229	29546	586229	29546
EV8	145830	8615	145830	8845	145830	8825	149512	8735
EV9	141339	7808	141339	8127	141339	7808	141339	7808
EV10	159699	7683	159699	7683	159699	7683	159699	7683
EV11	286680	14835	286680	17562	286680	17809	291719	17562
EV12	238759	14151	238759	14151	238759	14151	238759	14151
EV13	219753	11087	219753	11087	219572	11087	219572	11087
EV14	157981	8268	157981	8268	157981	8268	157981	8268

As traffic volume was limited to 2000 vehicles in the 2nd and 3rd scenarios, the overall average percentage of time reduction achieved by using the rerouting algorithm in the third scenario compared to the first scenario is 7.59%, compared to the second scenario is 2.63%. This indicates that the rerouting algorithm has

significantly reduced travel time across all EVs in the simulation compared to the scenario without the rerouting method.

In conclusion, the presented work introduces a novel solution for addressing EVRP using SUMO, Traci, and a rerouting algorithm. The simulation model effectively captured real-time traffic dynamics, allowing EVs to adapt their routes. Results indicated the efficacy of the Traci script with rerouting capabilities in reducing waiting times and optimizing energy usage for EVs within the network. Furthermore, comparing different traffic scenarios highlighted the impact of traffic conditions on travel time. The findings of this work contribute to the understanding of EVRP optimization, the role of real-time traffic conditions, and the potential of rerouting algorithms in improving the efficiency and performance of EVs.

# Chapter 5

## Intrusion Detection System for Vehicular Ad-Hoc Network

### 5.1 Introduction

The evolution in ICT and networks has accelerated ITS applications. With the application of information and communication technologies such as IoT, intelligent sensors, AI, big data, cloud computing, 5G networks, and DT, the new generation of automotive systems and assisted driving technologies are rapidly improving. However, while modern communication technologies bring users easy connection and convenience, they also bring increasingly severe cybersecurity threats to transportation networks and smart cities [222]. In recent years, and with the vast data availability, the increase in information security needs has attracted significant attention from researchers.

EVs and AVs have the characteristics of close integration of network information and physical components, especially with VSs technology. VS added an innovation service for the vehicle industry to enhance the privacy and security aspects [25]. Attacks on the ITS network will not only cause privacy leakage and economic losses but also endanger human life and even become a national public safety issue.

The new generation of automotive systems relies on two network types: in-vehicle networks and external networks, which allow the EVs to share communication bandwidth, storage, and computing resources to realize the collaborative processing of a large amount of real-time data. The development of ITS is a double-edged sword for both the external vehicle network and the in-vehicle network, which faces new problems and challenges. However, the current automotive network environment is full of issues and challenges regarding information security. These challenges mainly include an open external environment (multiple ways of wireless access), limited resources (bandwidth resources, computing power and energy, storage resources, cache, etc.), strict time requirements (real-time and schedulable analysis), and cost sensitivity brought by large-scale applications.

Vehicle-to-everything (V2X) technologies connect vehicles to the exterior world via external networks. Modern vehicles can communicate with other vehicles in different ways, such as V2V and V2I, thanks to V2X technology [223]. The ad-hoc mobile communication technology applied to vehicular network applications was introduced as VANET. Generally, researchers and industry agree that VANET holds great potential for implementing ITS, which would simultaneously improve safety and dependability on our increasingly crowded highways [224].

The communication medium in EV and AV is VANET, which comprises ad hoc infrastructure and mobile vehicles communicating via open wireless channels. Shared wireless medium, high mobility, and lack of centralized security services make VANET more susceptible to attacks. Using ML and data analytics in IDS model development to spot network threats is widely considered a new promising area in cybersecurity [225]. However, the current detection techniques have limited prevention as they are designed for a set of known attacks, which makes it difficult to avoid the recent attacks designed to penetrate existing security systems [100]. Therefore, the need for IDS has increased to protect V2X communication EV and AV from potential and updated attacks [226].

Authors in [227] proposed a Secure and Efficient Conditional Privacy-Preserving Authentication scheme in VANETs for resisting impersonation attacks and achieving better performance efficiency. The proposed model reduces the computation costs of signing and verifying the message simultaneously and the communication costs of the message size. Some authentication methods, such as the Intrusion Pre-



sensation System (IPS), detect unauthorized users and prevent them from accessing the network [228]. However, an IPS can be bypassed by a skillful intrusion. Therefore, an IDS can be the second layer of security to detect these types of intrusions. The role of an IDS is to continuously monitor network activity and collect relevant data for further analysis to make appropriate decisions in different situations.

Due to the growth of knowledge and technology, the problem of decision-making is becoming more complicated to handle; thus, developing new techniques to solve it is critical. Classifier combination in IDS is a promising direction in ML techniques. Today's main challenge in massively connected networks, including VANET, is keeping valuable information away from intruders and hackers. Moreover, ML algorithms can be used effectively if combined with predictive testing tools and Digital Twin technology in order to address the safety and security aspects of VANET [229]. Despite these threats, the developers of IDS make every effort to combat cyber-attacks. There are two types of IDS: misuse and anomaly detection. Misuse detection aims to detect instances of network intrusions by comparing current activities to an intruder's expected actions [230]. In anomaly detection, the typical traffic flow over a network is first established by the system administrator; any deviation from this baseline or pattern that does not match an expected normal state will be considered an anomaly [231].

IDS is installed inside network gateways to monitor the traffic on the external network [232]. Suspicious network activity is detected when abnormal traffic passes through gateways in external vehicular networks. Then, any packet transmitted in the VANET is collected by network packet taps and examined by the developed IDS before passing to the connected vehicle [233]. Traditional networks have utilized IDS-based techniques for years because they efficiently deal with network attacks. However, fundamental features of these networks, such as restriction in power supply, node storage, poor transmission range, and processing capacity [234], provide significant challenges to deploying IDS-based systems in highly mobile and delay-sensitive networks such as VANET. Due to these restrictions, classic IDS and other security solutions for wired/wireless networks cannot apply to VANETs [235]. VANET has distinct characteristics that must be dealt with and considered during the development of IDSs. The most important of these characteristics is the high mobility in the network and the new protocols used in

connecting VANET networks [236] since the constant movement in the mobility makes monitoring intrusions difficult, affecting the detection of malicious attacks.

On the other hand, VANET does not use traditional network protocols but deals with new and special protocols, such as 802.11p and Dedicated Short-Range Communication [237], as well as 5G communications. Moreover, 5G technology-enabled vehicular network and facilitates communication among VANET components. Authors in [238] proposed a fog computing-based authentication scheme to decrease performance overhead in 5G-enabled vehicular networks by applying one scalar multiplication operation of elliptic curve cryptography to prove information. The proposed model satisfies privacy-preserving and pseudonym authentication and identifies the common security attacks. With the development of communications protocols and the increasing sources of attacks, tremendous types of intrusions and attacks target VANET network. Therefore, developing an IDS for the VANET network requires special considerations to enhance the network's security and detect the new generation of attacks in time.

This chapter focuses on the information security threats VANET encountered in developing intelligent connected vehicles. In the related literature, many ML techniques are used in the classification stage, either individually or collectively. Feature Selection (FS) techniques are used to select the features that help get better results. Despite this, some ML techniques need considerable resources to give the desired results. Most of the current research used one type of FS method, either filter, wrapper, or embedded methods, which allows the model to select specific features that may not affect the results or are redundant.

To preserve the effort that the model requires in literature and achieve better results, we adopt ML techniques from the same category, i.e., Decision Tree-based models. We also use the hybrid FS method (Filter and Embedded methods) to assist the model in selecting relevant and significant features from the dataset without duplicating data. More precisely, in the IDS model, we propose a feature engineering model based on RF [239] and Fast Correlation-Based Filter [240] to reduce the dataset dimensionality and select the best features that enhance the detection rate and model accuracy. After feature engineering, we implement four ML models, each consisting of tree-based algorithms to detect unpredicted attacks and achieve precise results with higher accuracy and considering the alarming

rate. The proposed tree-based algorithms are XGBoost [241], RF, DTree, and EXT Classifiers [242]. The model is evaluated on the CICIDS2017 benchmark [243], considering parameters such as packet delivery, dropped and delays, and network latency. We used the stacking method to improve the proposed model's performance. Stacking is a common ensemble learning strategy involving using each algorithm's output labels as input to train a robust stack model, which is responsible for making the final prediction [244].

The main contributions of this chapter are the following: • a tree-based IDS methodology is introduced to accurately detect intrusions and identify 14 classes of attacks on VANET, such as DoS, DDoS, PortScan, and Botnet [245], by using multiple ML algorithms; • a feature engineering model is proposed based on RF and FCBF techniques to reduce the dataset dimensionality and select the best features that enhance the detection rate and model accuracy; • Stacking has been used to aggregate the ML algorithms by applying a parallel ensemble model, which allows for improved detection rate and efficiency and reduces training time.

The remaining sections of the chapter are structured as follows. The related works in VANET security, IDS, and type of attacks on VANET are recalled in Section 2, and the proposed model is presented in Section 3. Then, Sections 5 and 6 discuss the experiment, outcomes, and new directions in VANET security research.

## **5.2 Intrusion Detection System for VANET**

### **5.2.1 VANET Security & IDS**

Privacy and security are considered challenging issues in private and public transportation, especially with emerging technologies such as 5G-enabled vehicular networks. Authors in [246] introduced a modular square root-based to resist DoS attacks in 5G-enabled vehicular networks. The proposed scheme satisfies the following: authenticity of the source, integrity of the message, pseudonym privacy preservation, traceability, and revocability. Recent years have seen an increase in studies focusing on developing IDS for VANET and connected vehicles. Detecting

cyber-attacks on vehicular networks (V2V/V2I) is a hot topic in the academic community. Multicluster anomaly-based has been proposed by [235] to detect attacks in VANET. The Dolphin Swarm Algorithm optimizes the model, which can accurately identify various forms of cyber-attack. The results were compared with many existing models in different parameters, such as detection rate, detection time, and alarm rate. They concluded that the proposed multi-cluster model performs better on VANET than existing IDS.

Authors in [247] developed a classification framework to detect malicious attacks on internal vehicular networks with the help of hybrid algorithms, which are k-nearest neighbor (KNN) and SVM. The model is built based on KNN as a hybrid model with the support of SVM and utilizes the "DoS dataset" and the "fuzzy dataset" to simulate vehicle hacking. In order to effectively detect attacks on AVs' communication networks, authors in [248] presented an approach using a bidirectional Long Short-Term Memory (LSTM) architecture based on DL in addition to the conventional state-based approach. The developed framework is tested using two benchmark datasets: UNSWNB-15 for communications with external networks and the car hacking dataset for in-vehicle communications.

In the recently published work [249], a novel DL-based IDS was introduced to identify suspicious network behavior in In-Vehicle Networks, V2V, and V2I networks. Two benchmark datasets were used to assess the proposed IDS: the car hacking dataset for internal communication and the UNSWNB15 dataset for external communication networks. Moreover, in [155], the authors introduced a DL framework to detect in-vehicle attacks and external intrusions on VANET. The suggested model relies on LSTM and the gated recurrent unit. The model's performance was examined on a combined DDoS dataset for external communication and a car-hacking dataset for internal. The authors in [230] presented a hybrid IDS approach for VANET to optimize the performance of IDSs by combining RF classifier and a posterior detection method based on coresets to enhance the detection accuracy.

The most recent dataset, CICIDS2017, includes the cyber-attack scenarios, such as the most common types of assaults in VANET, such as DoS Slowloris and DDoS attacks. An anomaly-based IDS was proposed in [250] to detect CAN attacks using the adaptive cumulative sum method. Based on the changes in

statistical patterns, this method can efficiently identify intrusions with a minimal figure of delay. They tested how well the detection system worked by employing the CAN logs in the Car-Hacking Dataset. A unique hybrid IDS was proposed in [251], which accurately identifies the various cyber-attack forms, including those launched on internal and external vehicular networks. The model was evaluated and tested on the CICIDS2017 and CAN-intrusion datasets.

An effective two-layer IDS mechanism for vehicular communication networks was introduced in [252], which uses collaborative detection between two IDSs located on the vehicle and at the network's edge. The proposed model was tested on a widely used CICIDS2017 dataset for IDS systems. To detect Bot attacks on IoV, in [253], an ML model was developed representing these attacks in a VANET system. Moreover, they did not include any other attacks from the CICIDS2017 dataset. A network IDS based on a multi-layer presented in [254] perceptron to identify cyber-attacks on connected vehicles. The model has been evaluated on two versions of CICIDS2017 dataset and installed on a vehicle microprocessor.

Starting from the current research results, this chapter proposes a new methodology for intrusion detection based on the synergistic use of tree-based ML techniques. Moreover, the novel hybrid FS method to assist the intrusion detection model is presented for selecting relevant and significant features from the dataset without duplicating data. By such choices, we accurately identify VANET intrusions and current assaults like DoS, DDoS, Ports can, and Bot by obtaining improved results concerning the related literature considering basic indices such as accuracy and precision.

## 5.2.2 Attacks on Vehicular Networks

There are two types of attacks on the current vehicular networks: internal and external.

**Internal attacks** Currently, the primary defensive techniques against intrusions are firewalls, access control, and cryptography. These processes serve as the initial line of protection for connected vehicular networks. Cryptography provides secure communication, while access control was implemented for user authentication.

However, the level of internal security that they provide to in-vehicle systems is insufficient. A CAN is a bus standard for EVs, and it has been exposed to several attacks due to its broadcasting approach, data encryption needs, and insecure priority mechanism [255]. By inserting malicious messages into CAN packets, attackers may intrude on a system and have complete vehicle control in various ways, including gear and RPM spoofing. This type of assault is referred to a message injection attack and can result in legitimate nodes or vehicles. The most common kind of intra-vehicle assault is called a message injection attack [256].

**External attacks** VANET is a wireless mobile network that facilitates communication between Roadside Units, V2I, and V2V. The absence of firewalls and gateways in such wireless networks makes them vulnerable to attacks from anywhere within the radio range. Furthermore, unlike wired networks, an attacker does not need physical access to the car to launch an attack. Each vehicle can be compromised since they are exposed, and the cars can travel autonomously without any protection from attack [257]. As external communication employs a decentralized design, self-driving cars depend on the cooperation of other vehicles inside their radio coverage region. In VANETs, systems security measures such as encryption/decryption processes and digital signatures may reduce the number of possible vulnerabilities, serving as the first line of defense. Nevertheless, self-driving cars need a second layer of protection to detect and even identify unknown attacks, which the current system's security techniques cannot avoid.

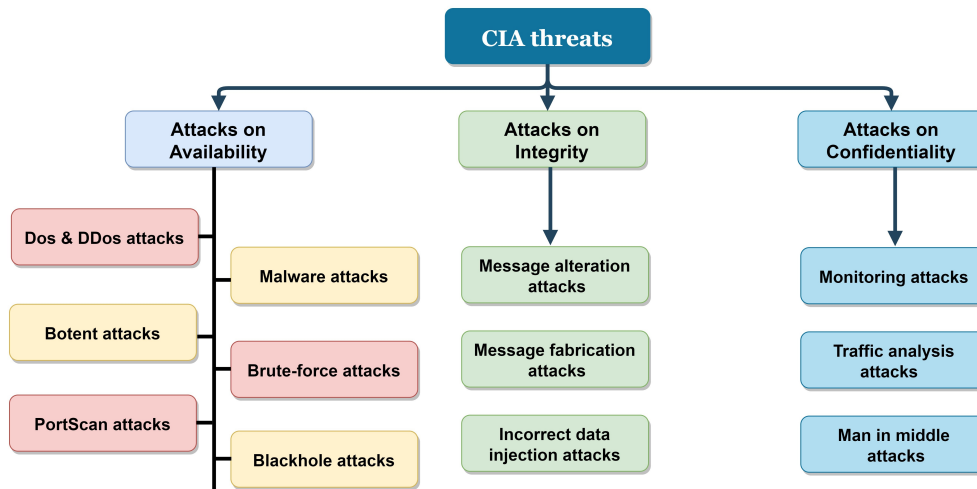


Figure 5.1: Cyber-attacks goals on VANET based on CIA triad.

Confidentiality, Integrity, and Availability, also known as the CIA triad, are highlighted as the three primary data and information security goals in any network, including VANET [258]. Figure 5.1 reports the cyber-attacks that can occur in VANET, divided into the different CIA attacks.

- **Attacks on Confidentiality:** This attack aims to breach the confidentiality of VANET and makes it possible for the attacker to listen to the conversation between nodes and try to steal the user credentials and other important information [259].
- **Attacks on Integrity:** The attacker hacks the network without changing the message's content and adding delay, which makes the user get the message later than expected. Therefore, information and messages must be sent to VANET users at the right time because breaching the network integrity results in disasters [260].
- **Attacks on Availability:** To accomplish the primary objective of VANET, it must always be accessible to allow users access to all applications and services. However, its main purpose will be useless if users cannot communicate across the network [261]. These attacks are associated with network

resources' accessibility, reliability, performance, and data processing. Network attacks have recently been a concern for many institutions that provide wired and wireless services, including government and private organizations. With the emergence of VANET and smart cars, there has been increased concern about these attacks, which may lead to tragic accidents and privacy breaches. Table 5.1 reports the common types of attacks on network availability.

Table 5.1: Common attacks on availability.

DoS	DoS attack is an attempt to disable a computer system or network so its intended users cannot access it.
DDoS	A DDoS attack attempts to block a server, service, or network's traffic by overloading the target and its surrounding infrastructure with an enormous amount of Internet traffic.
PortScan	Attackers send packets to specified ports on a host and find vulnerabilities and determine which service versions are running.
Brute-Force	Attackers use trial and error to get login credentials by cracking passwords and encryption keys.
Botnet	Attackers hack many connected vehicles and network components with one or more bot viruses and use this swarm of infected systems to launch different attacks on the connected systems.
Blackhole	Attack on communication protocol where the malicious node attempt to decrease the availability of VANET and block vehicles' communication
Syble	The attacker takes over the network and the service's system by operating many active fake identities.
Infiltration	Attackers disconnect hacked network systems and malicious implant nodes for future attacks
GPS Spoofing	It is an attack on GPS where the attacker sends a wrong geographic location to the victim node.



### 5.3 The Proposed IDS Model

The proposed system is an IDS framework focused on protecting the external vehicular network and built to obtain optimal results such as high accuracy and detection rate. We present a multi-level IDS model to classify normal and abnormal activities across VANET accurately. Figure 5.2 shows the proposed model flowchart and the methods used to build the IDS. The proposed IDS model begins with data pre-processing, including feature selection using RF and FCBF methods.

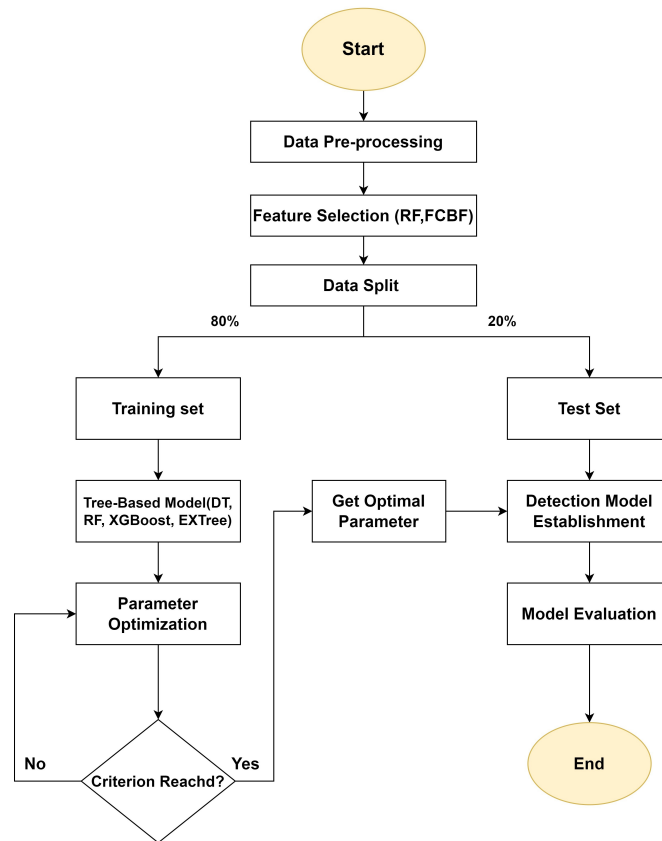


Figure 5.2: Flowchart of ID.

The data are then split into 80% training set and 20% test set. Tree-based algorithms (DTree, RF, XGBoost, EXT) are applied to the training set, and their parameters are optimized using HPO method. The detection model is established

and tested on the test set if the optimal parameters are found. Next, a stacking ensemble is implemented to combine the four tree-based algorithms to improve the detection efficiency. The model evaluation concludes the process. The model is evaluated on well-known data benchmarks (CICIDS2017) containing recent attacks with 80 network flow features. In the following section, we describe the system architecture in detail, including CICIDS2017 dataset, and explain the phases in the architecture of the proposed model.

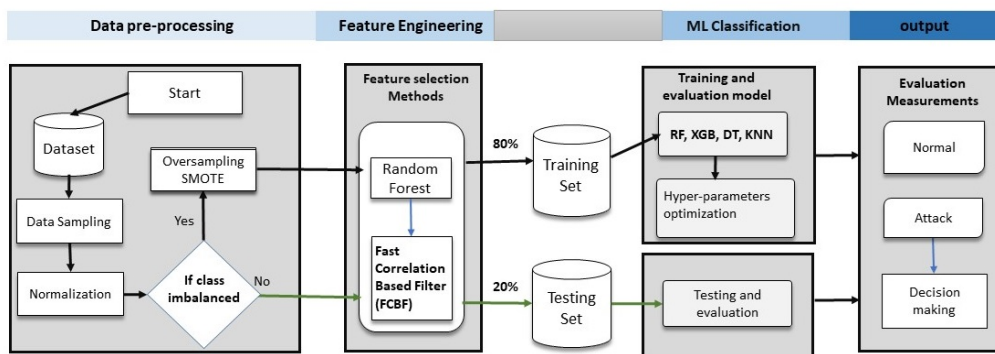


Figure 5.3: Multi-level IDS model architecture.

### 5.3.1 System architecture

In this section, the architecture of the proposed IDS model is described in detail and depicted in Figure 5.3. The framework is divided into four phases: data preparation, feature engineering, ML model, and evaluation. The first phase implements data pre-processing techniques such as data sampling, normalization, and data imbalance solutions like Synthetic Minority Over Sampling Techniques (SMOTE) [262]. The data should be prepared in order to create a representative subset for training and testing while avoiding outliers and class imbalance issues. The second phase implements feature selection techniques, i.e., RF and FCBF, that are used to remove unnecessary and redundant features, as well as low dimensionality and noisy features that might affect the accuracy of the prediction result.

Then, in the ML classification phase, the four tree-based ML methods are applied synchronously to detect unpredicted attacks and combined by a stacking

method to improve the attack detection accuracy further. In the final phase, the attack prediction accuracy is evaluated by standard performance metrics like accuracy, recall, precision, and score. The aim is to detect whether an attack occurs or does not occur and, consequently, whether a corrective action is needed or not.

### 5.3.2 Benchmark description

The CICIDS2017 dataset is used as a benchmark of network data from the real world devoted to IDS analysis. The dataset is created on an internet-connected testbed with a network infrastructure hosting various client devices and operating systems that simulate real-world scenarios. CICFlowmeter-V3.0 (Canadian Institute for Cybersecurity 2017) is applied to this dataset to extract features and labels, from which 80 network flow features have been extracted. The attacks in this dataset can break down into botnet, infiltration, brute force FTP/ SSH, DoS and DDoS attacks, heartbeat, and web, which are not found in any other datasets [243]. Using the B-Profile system, CICIDS2017 abstractly profiles how people interact with each other and simulates different multi-stage attack scenarios. These datasets differ because they are based on real-world data generated from wireless local area networks (WLANs) and mobile networks and can be trusted. On the other hand, others in [263] provided a comparative study of current datasets used to evaluate the efficiency of ML IDS techniques.

The study concluded that the CICIDS2017 dataset is built and designed to overcome all the limitations of other available datasets, such as the variety of communication protocols, network traffic scenarios, and new patterns of attack discovered in the vehicular network. To ensure the evaluation is accurate, the benchmark uses 11 criteria, such as total traffic and available protocols[264]. However, VANET datasets are unavailable for evaluating external vehicular network IDS for many reasons, such as data privacy, distribution, and availability. Despite that, WLANs and mobile networks are the communication techniques used in VANET. Therefore, all the attacks on such networks are similar to malicious intrusions executed on VANET. Since CICIDS2017 is the most comprehensive and illustrative available cyber-attack dataset, the proposed IDS uses such benchmark

data to evaluate the model's efficiency in detecting attacks on VANET. Furthermore, the types of attacks included in the CICIDS2017 dataset are reported in Table 2.

### 5.3.3 Data Pre-processing

The first phase in the data analysis process is data pre-processing. Various techniques, data sampling, normalization, and data imbalance solutions are used in this phase as described in the following sections.

**Normalization** Some features in an IDS available public dataset are nominal, while others are not normalized. Normalization of data is an essential pre-processing step for learners who learn from the statistical properties of features. Data normalization is the process of normalizing each attribute's value so that it falls within a specified range, preventing the influence of one attribute from dominating the effects of the others. Linear and non-linear techniques can be used to normalize feature-based data. In particular, Z-score normalization is used to normalize the numerical attributes between 0 and 1, minimize their values, and decrease the training processes [265]:

$$z = \frac{x - \mu}{\sigma}$$

where  $Z$  stands for the Z-score normalization,  $\sigma$  refers to the standard deviation of features,  $x$  stands for the values, and  $\mu$  the mean of the sample.

**Data encoding** Since most ML algorithms cannot work with string data type, the dataset is encoded using a label encoder. This encoder method helps to convert categorical features in the dataset into numerical features, which are accepted as input to the proposed ML model.

**Data sampling** It is challenging to train ML models on enormous amounts of network data, which requires repeatedly training an ML model. Data sampling is a standard ML approach that can create a portion of the original dataset to simplify the training phase and improve the performance of the model training process. The proposed model for data sampling uses the K-means-based algorithm to produce a highly representative portion of the original data (10% of the dataset is used in the

considered experiments). K-means seeks to find the optimal values for each of the distances between each data point and cluster's centroid by minimizing the sum of the squares of those distances [266]:

$$J = \sum_{j=1}^k \sum_{i=1}^n \left\| x_i^j - \mu_j \right\|^2$$

where  $J$  is the objective function,  $k$  is the number of clusters,  $n$  is the number of cases,  $x_i^j$  is the data point of case  $i$  in cluster  $j$  and  $\mu_j$  is centroid for cluster  $j$ . K-means is a widely used technique for sampling because of its simplicity and low computational requirements.

**SMOTE** It is used to solve the class imbalance problem, particularly oversampling. The oversampling issue occurs when the dataset classes are imbalanced regarding samples. In this case, SMOTE is applied to solve the oversampling problem to balance the class dimensions.

### 5.3.4 Feature Engineering

Feature engineering aims to select the best features from the pre-processed dataset, enhancing the IDS's accuracy and helping to detect any vulnerable attacks on the vehicle network at low cost. Using feature selection techniques helps to decrease the cost of the ML workflow in terms of time and resource usage. Additionally, it enables feature transformation and removing unnecessary features to improve the quality and accuracy of results. Generally, the quality of datasets would be enhanced for more precise and effective model learning by choosing the ideal set of features. This work implements a double feature selection method based on RF and FCBF before the training stage. In particular, RF is well known as the best method to select features with high accuracy and faster results. Anyway, RF method gives the same importance level to the relevant features, leading to redundancy in the selected features. Hence, we use the FCBF method to filter the relevant features and remove the redundancy in the final set of selected features. In the following, we describe in detail the RF and FCBF methods.

**RF method:** RF is an ensemble-learning method utilized for data classification and regression. The regression tree is a hierarchical arrangement of criteria or constraints gradually executed from the root to the tree's leaf. Trees are generated using chosen bootstrap samples and randomly picked n-estimator parameters in each node separation in the RF technique [239]. RF method delivers a unique validity and model interpretation estimate within ML approaches. In the first stage of feature selection, RF technique calculates the importance of a feature based on its ability to increase the pureness of the leaves. It selects the 38 features as shown in Table 5.2, and these features are fed to the second filtering method. Figure 4 explains how RF method works for class prediction and feature selection. In the scheme, RF randomly selects a subset of important features and builds multiple decision trees based on these subsets. Each tree provides its prediction, and the final class prediction is determined by majority voting, where the class with the most votes becomes the final prediction. This approach reduces overfitting and enhances the model's generalization in classification tasks.

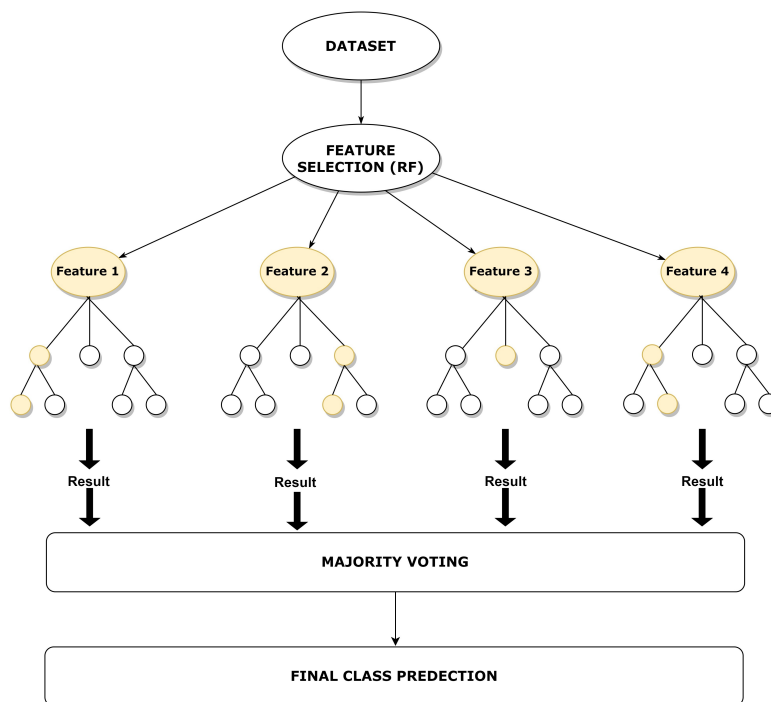


Figure 5.4: Random Forest structure.

**FCBF method:** FCBF is used in addition to the RF method since RF method removes the insignificant features to manage and decreases the time complexity but does not avoid feature redundancy. Indeed, the model accuracy is not optimized as many redundant and unneeded features remain in the dataset after RF application. The model's time and space complexity may also rise due to feature redundancy. Therefore, reducing duplicated features and computing the input correlation characteristics might positively impact the model's efficiency. Moreover, the FCBF technique is applied as a second FS technique because it shows excellent performance, especially on high-dimensional datasets, by reducing time complexity and removing redundant features while preserving important features.

Table 5.2: Selected features by RF.

No	Features	NO	Features
1	SubflowFwd Bytes	20	Fwd IAT Min
2	Total Length of Fwd Packets	21	Flow IAT Max
3	Total Length of Bwd Packets	22	InitWinbytesforward
4	SubflowBwd Bytes	23	Avg Bwd Segment Size
5	Flow IAT Std	24	Bwd Packet Length Max
6	Bwd Packets/s	25	Destination Port
7	Fwd Packet Length Std	26	Bwd Packet Length Std
8	InitWinbytesbackward	27	SYN Flag Count
9	Fwd IAT Total	28	Max Packet Length
10	Bwd Packet Length Min	29	Flow IAT Min
11	Flow Duration	30	Fwd Packet Length Mean
12	Fwd Packet Length Max	31	Average Packet Size
13	Bwd IAT Total	32	Packet Length Variance
14	Fwd IAT Mean	33	min_seg_size_forward
15	Fwd Header Length.1	34	Packet Length Std
16	Fwd Header Length	35	Packet Length Mean
17	Flow IAT Mean	36	Active Min
18	Fwd IAT Std	37	Total Fwd Packets
19	Bwd Packet Length Mean	38	Bwd Header Length

The first step in the FCBF process is to select a collection of features with a high degree of correlation with the FCBF class. After that, the resulting values are sorted into a list based on the feature importance. It uses heuristics to eliminate features that are not necessary and retain the features that are more important to the class. The proposed method selects the best 25 features in the dataset, as shown in Table 5.3, directly impacting the model accuracy.

Table 5.3: Selected features by FCBF.

No	Features	NO	Features
1	Total Length of Fwd Packets	14	Flow IAT Mean
2	SubflowFwd Bytes	15	Flow IAT Std
3	Bwd Packet Length Std	16	Flow IAT Max
4	Fwd IAT Std	17	Max Packet Length
5	Fwd IAT Min	18	Destination Port
6	Bwd Packets/s	19	Fwd Packet Length Std
7	Avg Bwd Segment Size	20	Fwd Header Length.1
8	Bwd Packet Length Mean	21	Total Length of Bwd Packets
9	Init_Win_bytes_backward	22	Fwd Header Length
10	Packet Length Std	23	Bwd Packet Length Max
11	Init_Win_bytes_forward	24	Flow Duration
12	Fwd IAT Total	25	min_seg_size_forward
13	Flow IAT Min	-	-

### 5.3.5 The ML model

According to Intel Corporation [267], four terabytes of data are generated daily due to the necessary technological improvement of connected vehicles. Strong IDS based on ML is often suggested to achieve vehicle security related to road and human life safety, protecting vehicular network communications. In this phase, we aim to examine the considered pre-processed dataset to detect the attacks by classifying the traffic as normal and abnormal. For that goal, we implement a multi-



ML algorithms model consisting of four tree-based algorithms to detect unpredicted attacks and achieve accurate results. The proposed tree-based algorithms are XGBoost, RF, DTree, and EXT classifiers.

The DTree algorithm is a typical ML technique that fits data into a tree structure in order to generate predictions. The depth of the tree, minimum weight, full sample nodes, minimum sample split, fraction leaf, and minimum sample leaf are some of the DTree hyper-parameters that need to be tuned. The DTree is the foundation of all the tree-based algorithms designed to increase model processing speed through parallelization. In addition, XGBoost algorithm structure is shown in Figure 5.5. It works by creating an ensemble of decision trees. The trees are constructed to correct the errors in the other trees, considering the previous prediction errors. The final prediction is made by combining the predictions of all the trees, with each tree's contribution weighted based on its performance. XGBoost is known for its speed and performance and is widely used for both classification and regression tasks. RF is also used for classification, and it is a tree-based algorithm that utilizes the majority voting rule to merge many different DTree classifiers.

On the other hand, EXT method combines several randomized decision trees that have been constructed using various subsets of a dataset. The reason behind selecting tree-based algorithms is that they make it possible to execute many tasks simultaneously, which notably minimizes the time required for model training and improves performance. Moreover, a tree-based algorithm determines the significance of features during the training phase. Since traffic data on the vehicular network is non-linear, tree-based algorithms are the best choice to deal with it.

Finally, in the ensemble stage, tree-based algorithms use randomization to encourage the development of a flexible ensemble model with higher generalizability on various domains compared to existing ML approaches. It is also essential to optimize each tree-based algorithm's parameters to avoid using default parameter values. We apply the HPO method to select the best hyperparameter configuration in each model. In addition, the stacking technique is used to ensemble the four tree-based algorithms. This technique is an ensemble ML method aiming at learning how to combine the predictions at best from several well-performing ML models [244].

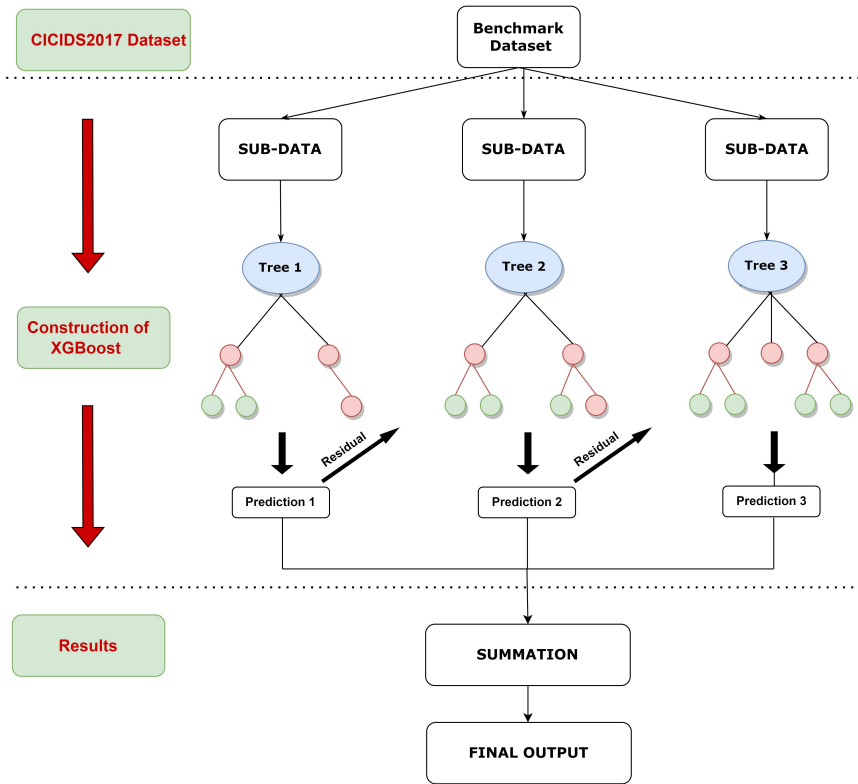


Figure 5.5: Structure of XGBoost Algorithm.

The Bayesian Optimization (BO) [268] technique is applied with the Tree Parzen Estimator (TPE) [269] in this phase. The BO is an iterative method frequently adopted for supporting HPO process. In particular, BO is used in ML to tune the hyper-parameters of a given model with good performance on a validation dataset. TPE is a sequential model-based optimization approach. These methods sequentially construct models to approximate hyperparameters performance based on historical measures and then consequently find new hyperparameters to evaluate.

In order to use TPE, the observation results are first separated into excellent and bad results using a pre-defined percentile  $y$ . Then, the two different sets of results are modeled using basic Parzen windows:

$$p(x | y, D) = \begin{cases} l(x), & \text{if } y < y^* \\ g(x), & \text{if } y > y^* \end{cases} .$$

Where  $y^*$  is the threshold value of the objective function,  $x$  is the proposed set of hyperparameters,  $y$  is the actual value of the objective function using hyperparameters  $x$ , and  $p(y | x)$  is the conditioned probability expressing the probability of  $y$  given  $x$ . In addition,  $g(x)$  and  $l(x)$  represent the probability of identifying hyper-parameters in areas that have performed well and areas that have performed poorly, respectively. BO-TPE can determine the ideal hyperparameter values by increasing the ratio  $l(x)/g(x)$ . The complexity of BO-TPE in terms of time is  $O(n \log n)$ , which is less than what the other methods require.

In conclusion, the proposed IDS model training process could be performed on an external server machine on VANET with high performance and computational speed. The model can also minimize the latency and fulfill the vehicle system's real-time requirements. The implemented multi-level IDS tree-based model that integrates feature selection technologies, ML algorithms, HPO-BO optimization, and stacking processes allows for optimizing the detection rate of attacks on VANET by outperforming existing approaches.

### 5.3.6 Evaluation methods

During the evaluation of an IDS, the detection rate, as well as the false-positive rate, are important criteria that should be taken into consideration. Many different measures may be utilized for IDS performance evaluation. Confusion matrix in Table 5.4 used as standard evaluation metrics, Table 5.5 presents accuracy, recall, precision, and f-score, which are metrics commonly used to evaluate the performance of the proposed model [270].

Table 5.4: Confusion Matrix (CM).

		Prediction	
		Intrusion	Legitimate
Actual	Intrusion	TP	FN
	Legitimate	FP	TN

Table 5.5: Selected features by FCBF.

Accuracy	The percentage of correctly predicted instances in the testing dataset	$Acc = \frac{TP+TN}{(TP+TN+FP+FN)}$
Precision	The number of true positives divided by the total of true positives and false positives	$Precision = \frac{TP}{TP+FP}$
Recall	The number of true positives divided by the total of true positives and false negatives	$Recall = \frac{TP}{TP+FN}$
F-score	The harmonic average of recall and precision knew	$FM = 2 \times \frac{precision \times recall}{precision+recall}$

Assessing the efficacy of our proposed intrusion detection model involves employing the confusion matrix, which is particularly useful in classification scenarios. This matrix precisely illustrates the correspondence between predicted and actual class outcomes, offering insights into the model's accuracy. Leveraging pairwise similarity metrics, the confusion matrix comprehensively analyzes class relationships, providing a nuanced evaluation of intrusion detection performance.

The confusion matrix has four alarm rates that have to be calculated during the evaluation process, providing the value of the evaluation metrics:

$$\text{True positive rate } TPR = \frac{TP}{TP+FN}$$

$$\text{True negative rate } TNR = 1 - FPR$$

$$\text{False positive rate } FPR = \frac{FP}{TN+FP}$$

$$\text{False negative rate } FNR = 1 - TPR$$

Where TP is the number of intrusions correctly detected, TN is the number of non-intrusions correctly detected, FP is the number of non-intrusions incorrectly detected, and FN is the number of intrusions incorrectly detected.

## 5.4 Experimental Setup

The model is developed in the Jupyter environment (Pimentel et al., 2021) and uses Python and its libraries, which support feature engineering, ML, and HPO, such as Pandas, Xgboost, Hyperopt, and Scikit-learn. The experiments are carried out on a Toshiba Satellite Pro with Core i5-1135G7 (CPU) (4.20 GHz) and 16GB of memory 3200MHz, Intel Iris Xe Graphics 8GB, and Windows 10 Home 64Bit. The experiment takes approximately 200 hours (around 9 days), and the model is run in different cases. First, we download CICIDS2017 dataset from the Canadian Institute for Cybersecurity website [271], the version for ML, which is in the form of a CS file in a compressed size of about 1GB. The data are divided into 8 CSV files; each file is generated on a different day and with different types of attacks.

The files were combined into one file to enable the model to read the data from a unified source, which is the best way to save time and reduce effort on the processor reading from one source. Then, a normalization process is performed for the data using Z-score. Data sampling is done using the K-mean technique, where the algorithm chooses the best 10% of the data with a size of 100 megabytes, representing the entire dataset. Successively, we apply feature selection methods: the first RF method selects 38 features, and then the second FCBF method selects 25 features. We ran the model seven times with different packs of features: 10, 15, 20, 25, 30, 35, and 38 features. We found that when selecting 25 features, the accuracy and evaluation metrics results are the best, as shown in Table 5.6. Figure 5.6 shows the results of testing the IDS model with different features.

Table 5.6: Model evaluation with a different set of features in CICIDS2017 dataset

	Accuracy	Precision	Recall	F1 Score
10 features	0.9959	0.9973	0.9967	0.9964
15 features	0.9962	0.9974	0.9969	0.9965
20 features	0.9967	0.9978	0.9971	0.9968
25 features	<b>0.99856</b>	<b>0.99849</b>	<b>0.9985</b>	<b>0.99851</b>
30 features	0.9981	0.9982	0.9977	0.9978
35 features	0.9980	0.9981	0.9979	0.9980
38 features	0.9975	0.9979	0.9976	0.99768

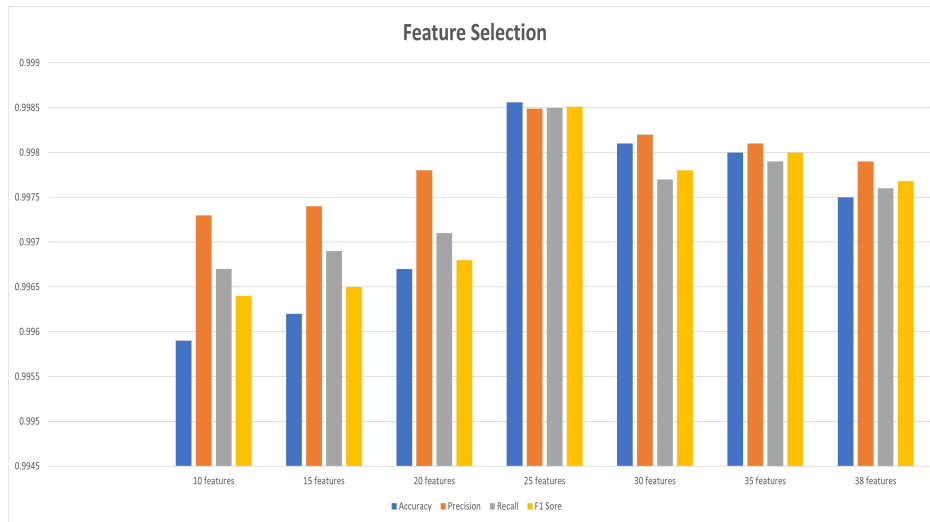


Figure 5.6: Feature selected performance with evaluation matrices.

An imbalance in the classes was observed, and the use of SMOTE is necessary to address the class imbalance issue. It works by generating synthetic samples for minority categories and avoiding bias in the ML model [262]. In the proposed ML model, each tree-based algorithm is executed under three different cases:

- 1) without the feature selection,
- 2) with the feature selection but without the optimization algorithm HPO, and
- 3) using feature selection techniques and HPO algorithms.

It is noticed that when FS and HPO are used, the model can detect intrusions with a higher accuracy of up to 98.86%, as shown in Table 5.7. Table 5.8 presents a comparison with outcomes achieved by other researchers employing diverse techniques on the same dataset. Moreover, Figure 5.7 shows the results obtained by the proposed model with various evaluation matrices.

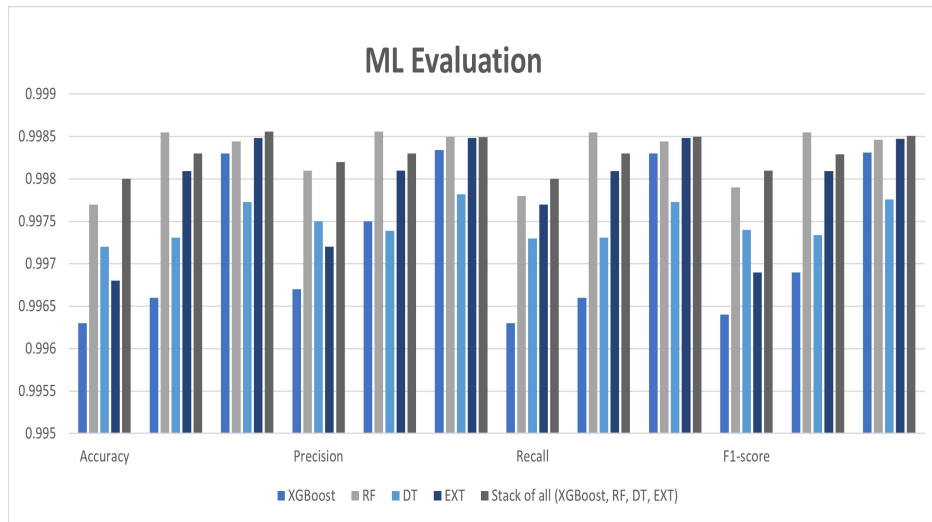


Figure 5.7: Performance of model on CICIDS2017 with different evaluation matrices.

Table 5.7: Model evaluation results in three different cases.

Evaluation	Accuracy			Precision			Recall			F1-score		
	Without	Without	With	Without	Without	With	Without	Without	With	Without	Without	With
ML technique	FS	HPO	HPO	FS	HPO	HPO	FS	HPO	HPO	FS	HPO	HPO
XGBoost	0.9963	0.9966	0.99830	0.9967	0.9975	0.99834	0.9963	0.9966	0.9983	0.9964	0.9969	0.9983
RF	0.9977	0.99855	0.99844	0.9981	0.99856	0.9985	0.9978	0.99855	0.9984	0.9979	0.9985	0.9984
DTree	0.9972	0.99731	0.99773	0.9975	0.99739	0.99782	0.9973	0.99731	0.9977	0.9974	0.9973	0.9977
EXT	0.9968	0.99809	0.99848	0.9972	0.99810	0.99848	0.9977	0.99809	0.9984	0.9969	0.9980	0.9984
Stack(RF, XGBoost, ,DTree, EXT)	0.9980	0.9983	0.99856	0.9982	0.9983	0.99849	0.9980	0.9983	0.9985	0.9981	0.9982	0.9985



Table 5.8: Comparison between our work and literature.

Paper	Dataset	Security threats	ML algorithm	Detection result
Alshammari et al.	car-hacking datasets CHD	DoS and Fuzzy attacks	KNN and SVM	Ac: 96 F1: 93
Izhar khan et al	UNSWNB-15 + CHD	DoS, zero-day, Reconnaissance and injections attacks, Gear spoofing, RPM spoofing, and Fuzzy attacks.	PCA -bidirectional LSTM	Ac: 99.11 Pr: 99.13 Re: 98.42 Fs: 99.09 Kp: 98.22
Javed Ashraf et al.	UNSWNB-15 + CHD	DoS, zero-day, Reconnaissance, and injections attacks, Gear spoofing, RPM spoofing, and Fuzzy attacks.	LSTM Autoencoder	Ac: 99 F1: 99
Safi Ullah et al.	CIC DoS CICIDS2017 CSE-CIC-IDS 2018 CHD	DDoS, fuzzing, and spoofing	(LSTM) and gated recurrent unit extra tree classifier (ETC) for FS SMOTE stratified random sampling (SRS)	Ac:99.5%
Bangui et al.	CICIDS2017	DDoS, DoS, Heartbleed, PortScan,Bot,FTP-Patator, SSH-Patator, Web Attack, Infiltration	- Random forest - Bisecting k-means - Weighted K-means - Coresets	Ac: 96.93 F1: 94.41 Pr: 98.5 Re: 90.4
Olufowobi et al.	Car-Hacking Dataset	DoS, Fuzzy Attack, Spoofing Attack	cumulative sum change-point detection algorithm	Ac: 96
Rosay et al.	CICIDS2017	DDoS, DoS, Heartbleed, PortScan,Bot,FTP-Patator, SSH-Patator, Web Attack, Infiltration	Multi-Layer Perceptron	Ac: 99
Proposed model	CICIDS2017	DoS Hulk, PortScan, DDoS, DoS GoldenEye ,FTP-Patator ,SSH-Patator , DoS slow loris, DoS Slowhttpstest, Bot, Web Attack, Brute Force, Web Attack XSS 'Infiltration 'Web Attack SQL Injection 'Heartbleed.	Multi-level Tree-based (DTree, RF, XGBoost, EXT), BO-TPE.	Ac: 99.86 Pr: 0.99849 Re: 0.9985. F1: 0.99851

## 5.5 Results and Discussion

The proposed IDS model shows its effectiveness in predicting different kinds of attacks on VANET. It is based on different strategic levels:

- Data processing is applied to purify the initial dataset from errors and missing data; Z-score is used for normalization, K-means-based algorithm is used for data sampling, and SMOTE technique for solving the imbalanced problem;
- Feature selection is applied to choose the features that can improve the accuracy of detecting intrusions. In particular, we propose a feature engineering model based on RF and FCBF to reduce the dataset dimensionality and select the best features that enhance the detection rate and model accuracy;
- ML algorithms are applied to discover gaps by classifying the normal and the abnormal traffic based on the classes available in the dataset. A multi-ML algorithms model is implemented consisting of four tree-based algorithms to detect unpredicted attacks and achieve accurate results;
- the prediction accuracy is evaluated by using standard metrics, i.e., accuracy, recall, precision, and f-score;
- To optimize the parameters of each tree-based algorithm, we apply HPO method, selecting the best hyper-parameter configuration in each model.

Although this field has received the attention of many researchers and several research works have been submitted in this context, a hybrid IDS framework can rarely obtain such a high prediction accuracy. Using two methods to filter the features led to selecting the essential features that help the model deal with a real dataset containing data similar to the real-time data.

Tree-based ML algorithms are used since they work perfectly on non-linear data and can perform many tasks simultaneously, which is needed during the training process. Moreover, we used the stacking technique to ensemble the four tree-based algorithms, and this technique is an ensemble ML method aiming to learn how to combine the predictions from tree-based models at best. In addition, tree-based model uses randomness during the building process, which helps to

construct a flexible model that can be generalized to any domain case in the future. During the feature selection phase, we carefully picked 25 features, resulting in impressive performance metrics. The proposed model achieved a 99.86% accuracy, with 99.85% precision, recall, and F1 Score, as detailed in Table 6.

We conducted three different experiments: one without feature selection, one without HPO, and one with HPO, for each algorithm separately. Then, we combined all the algorithms using the ensemble technique, referred to as the stacking method. The experiments with HPO showed better performance in various evaluation metrics. However, the stacking method stood out with the highest accuracy, precision, and recall, reaching 99.86%, 99.85%, and 99.85%, respectively, as indicated in Table 5.7.

Looking at these results and comparing them with previous studies in Table 5.8, our model excelled in detecting attacks, especially regarding the accuracy and other necessary metrics. This achievement is realized through combining feature selection techniques, effectively refining the selected features to optimize the model's outcomes. Additionally, the performed experiments are based on a subset comprising 10% of the CICIDS2017 dataset, containing 14 common attack types, enabling us to design a more comprehensive IDS for VANET that can detect attacks effectively.

To the best of our knowledge, integrating the selected FS techniques and detecting this number of attack classes for VANET has not been previously reported in the literature. There is a need to secure EV networks for internal or external communication, and developing such systems is essential in the future of smart cars. Correct information may save lives and prevent disasters, while lacking data or giving the wrong position may cause unfortunate accidents.

# Chapter 6

## Security and Intelligent Transportation Systems

### 6.1 Introduction

The development of ITS and the broader concept of Smart Cities have catalyzed significant advances in urban mobility, introducing disruptive changes. Central to this evolution are VANETs, which have become a pivotal element in facilitating V2V and V2I communications. These networks are instrumental in the design of intelligent and adaptable transportation systems [272]. As a specialized subset of mobile ad hoc networks, VANETs focus on improving inter-vehicle communication, optimizing routing protocols, and fostering a dynamic networking infrastructure [273].

VANETs have attracted considerable interest due to their potential to enhance road safety and traffic efficiency. Research in this domain spans various areas, including broadcasting, Quality of Service (QoS), routing, and security. The integration of network communication in EV and AV significantly enhances their situational awareness and decision-making capabilities [274]. In VANETs, each vehicle acts as a wireless receiver and transmitter, relaying data to nearby vehicles or infrastructure [275]. Standard protocols like IEEE 802.11p [276] and IEEE 1609.4 [277] are employed for inter-vehicle communication (IVC), with IEEE

802.11p being an adaptation of IEEE 802.11 [278], tailored for the dynamic and complex environments of VANETs. This protocol is also known as Wireless Access in Vehicular Environment (WAVE) [279]

Dedicated Short-Range Communications (DSRC) is a technology designed for short-range wireless communication, operating within the 5.9 GHz ITS band (5.85–5.925 GHz) [280]. It plays a crucial role in improving public and private safety by enabling effective V2I and V2V communications [237]. WAVE, which operates under the IEEE 802.11 standard, utilizes the DSRC band and is based on the IEEE P1609 family standards. This framework defines vehicular communications' structure, communication model, management, and security aspects, with key components including RSUs, Onboard Units (OBUs), and the WAVE interface. The WAVE protocol stack comprises the IEEE 1609 family, IEEE 802.11p, and the Society of Automotive Engineers (SAE). Illustrating the key components of the WAVE protocol architecture in Figure 6.1, these elements are briefly outlined as follows [281]:

- IEEE P1609.0 Draft Standard for WAVE - Architecture.
- IEEE 1609.1 Trial Use Standard for WAVE - Resource Manager.
- IEEE 1609.2 Trial Use Standard for WAVE - Security Services for Applications and Management Messages.
- IEEE 1609.3 Trial Use Standard for WAVE - Networking Services.
- IEEE 1609.4 Trial Use Standard for WAVW - Multi-Channel Operations.
- IEEE P1609.11 Over-the-Air Data Exchange Protocol for ITS.
- IEEE 802.11p Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications – Amendment: WAVE.

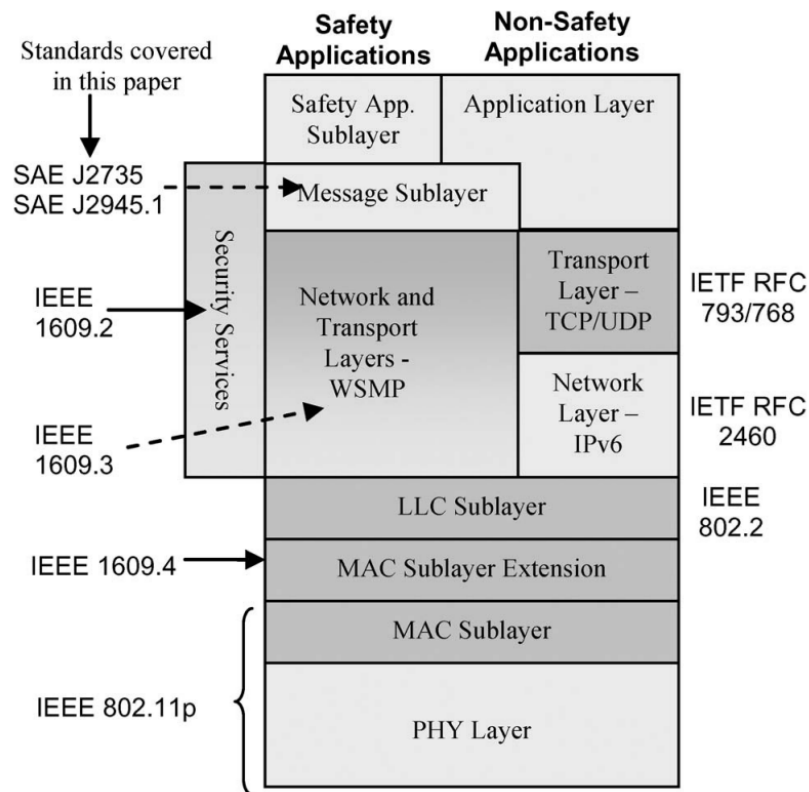


Figure 6.1: Architecture of DSRC communication.

The coming Section will explain the significance of WAVE protocol and the message types exchanged among the network in the context of the IEEE 802.11p protocol in VANETs.

In VANETs, continuous exchange of messages is vital to sharing crucial information. However, unauthorized access to these data introduces security risks and potential network vulnerabilities [282]. Current research is focused on improving security through AI-driven anomaly and IDS. These systems, deployed in vehicles and roadside units, scrutinize network behavior to identify patterns indicative of security threats, ensuring real-time monitoring and a robust network security strategy.

This research aims to simulate VANET communication protocols, study WAVE standards, and analyze message exchanges to bolster vehicle security and detect attacks using AI and ML tools. The approach involves integrating network and

traffic simulations into a unified framework, enabling the simulation of real networks, vehicle communications, WAVE message analysis, and developing security solutions based on simulation data.

A case study was conducted using the SUMO to model vehicle traffic, explicitly focusing on the real traffic dynamics of the Pasubio region of Bologna in Italy, as per a previous study by [283]. Additionally, the OMNeT++ software was used to simulate communication and network protocols [284], with the VEINS framework integrated into OMNeT++ to simulate the vehicular environment, linking traffic and network simulations into a cohesive model that accurately represents VANETs.

The sections of the chapter are organized as follows. Section 2 provides a general description of VANET security. Section 3 presents the IEEE 802.11p (WAVE) protocol. Section 4 describes simulation tools. Section 5 proposes a case study and discusses the study results, and Section 6 draws conclusions and considers future work.

## 6.2 Simulation Approach for VANET Security

VANETs usher in a new era of intelligent transportation by enabling vehicles to communicate wirelessly, enhancing road safety and traffic efficiency seamlessly. At the heart of this technological transformation are advancements such as IEEE 802.11p, which empowers VANETs to facilitate swift and real-time data exchange among vehicles [285]. This exchange of information fosters improved road safety and enable sophisticated features such as collision warnings and efficient traffic management systems [286].

While VANETs exhibit tremendous potential in revolutionizing how vehicles interact with each other and with infrastructure, they also present many challenges, particularly concerning preserving data integrity and security [287]. The complex interplay between vehicles and network components in VANETs necessitates robust mechanisms to safeguard against potential threats and unauthorized access. These challenges are exemplified in Fig. 6.1, underscoring the critical importance of addressing security concerns in the context of VANET applications.

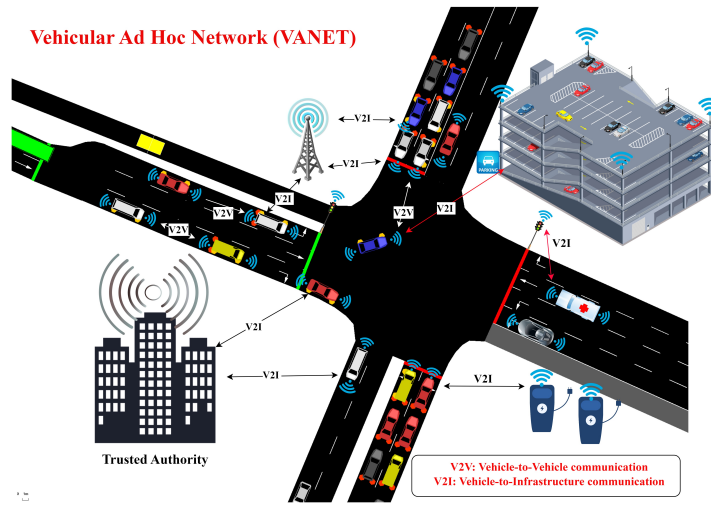


Figure 6.2: Vehicular Ad Hoc Networks (VANETs)

In the dynamic landscape of VANETs, the continuous and bidirectional flow of information between vehicles and infrastructure introduces vulnerabilities that must be addressed to ensure the data's reliability and trustworthiness [288]. Cybersecurity threats such as data tampering, eavesdropping, and unauthorized access pose significant risks that could compromise the integrity of communication within the VANET ecosystem [289].

Moreover, the need for real-time responsiveness and the decentralized nature of VANETs amplify the complexity of implementing robust security measures [272]. Striking a delicate balance between facilitating seamless communication and fortifying the network against potential breaches becomes paramount [290]. This intricate dance requires innovative solutions that leverage state-of-the-art encryption techniques, authentication protocols, and IDS [291].

As the deployment of VANETs becomes more widespread, researchers and practitioners are actively engaged in developing comprehensive security frameworks [292]. These frameworks aim to mitigate risks, ensuring that VANETs enhance road safety and traffic efficiency and operate within a secure and trustworthy communication environment. The ongoing pursuit of advancements in cryptographic algorithms, secure key management, and anomaly detection mecha-



nisms signifies a commitment to overcoming the challenges posed by the dynamic and interconnected nature of VANETs [293].

Securing VANETs is paramount for thwarting potential threats and upholding the trustworthiness of vehicular communication systems. Adopting robust security measures is essential to ensure the integrity and confidentiality of data transmitted within these dynamic networks. Encryption techniques play a pivotal role in safeguarding the content of exchanged information, shielding it from unauthorized access and tampering. Simultaneously, authentication mechanisms validate the identities of participating entities, establishing a foundation of trust within the VANET ecosystem.

To enhance VANET security further, anomaly detection systems have emerged as a vital component, enhancing the overall resilience of the network [294]. These systems scrutinize behavior patterns, promptly identifying deviations that may indicate potential security threats [295]. Extensive studies have delved into the multifaceted challenges and solutions for VANET security, contributing to the evolving landscape of secure vehicular communication [296].

One notable contribution is a comprehensive survey on AI techniques for VANET security, meticulously presented in [272]. This survey explores the diverse applications of AI in addressing the unique security challenges posed by VANETs, providing valuable insights into cutting-edge methodologies.

The work by [297] delves into the advances and challenges within VANETs, specifically focusing on security aspects. This research sheds light on the evolving nature of vehicular communication and the concurrent efforts to enhance security measures in this dynamic domain.

To pursue a robust IDS, [298] proposes the Secure and Private-Collaborative IDS (SP-CIDS). This innovative system aims to detect network attacks efficiently, thereby mitigating security concerns and ensuring the integrity of vehicular communication within VANETs.

Addressing the intricate challenge of message congestion and accurate attack detection, [299] introduces a novel IDS tailored for VANETs. This system not only contributes to the mitigation of potential security threats but also optimizes the network's performance by alleviating message congestion.

Notably, previous studies in the field have commonly utilized VANETs as dynamic wireless networks for vehicle communication. However, a notable gap has been identified concerning the lack of integrated simulation models for traffic and networking. This discrepancy motivated the present study to construct a holistic simulation model using SUMO and Omnet++, incorporating the WAVE protocol. The study analyzes the diverse messages exchanged between vehicles within this simulated environment, aiming to uncover anomalies that may signify potential security threats. The ensuing sections will delve into a detailed exploration of these messages and their implications in the context of VANET security.

## **6.3 The IEEE 802.11p - WAVE Protocol**

### **6.3.1 Dedicated Short-Range Communications**

Dedicated Short-Range Communications (DSRC) is a specialized communication service operating within the 5.9 GHz frequency band and purposefully designed for short-range wireless communication between devices. DSRC is an innovative solution that facilitates direct communication over short-to-medium distances, typically within a range of approximately 300 meters [237]. Functioning as an extension of Wi-Fi, DSRC marks a significant advancement, offering a means for devices to communicate directly without relying on intermediaries.

The distinctive feature of DSRC lies in its capability to enable seamless device-to-device data transmission. This direct communication is particularly advantageous when telecommunication infrastructure is limited or unavailable. By operating independently of external networks, DSRC ensures reliable and immediate connectivity between devices, thereby overcoming potential limitations posed by the absence of traditional communication infrastructure.

DSRC technology finds widespread application in vehicular networks, playing a pivotal role in V2V and V2I communication. In vehicular environments, where real-time communication is crucial for ensuring road safety and optimizing traffic flow, DSRC establishes itself as a critical enabler. Its short-range communication

capabilities make it well-suited for supporting communication between vehicles nearby, fostering a dynamic and responsive vehicular communication network.

### 6.3.2 Wireless Access in Vehicular Environments

IEEE 802.11p protocol, also known as the Wireless Access in Vehicular Environments (WAVE) protocol, stands as a pivotal standard within the realm of vehicular communication, operating seamlessly in the 5.9 GHz frequency band [300]. Conceived and developed by the Institute of Electrical and Electronics Engineers (IEEE), this protocol was tailored to meet the distinctive requirements of VANETs. WAVE provides a comprehensive framework to facilitate DSRC among vehicles [301]. This, in turn, enables direct and reliable information exchange, spanning distances from short to medium range. The widespread adoption of the IEEE 802.11p protocol marks a fundamental stride toward enhancing crucial aspects of modern transportation, including road safety, traffic efficiency, and the overall efficacy of ITS.

Within the context of VANETs leveraging the IEEE 802.11p protocol, three primary types of messages play instrumental roles in orchestrating effective communication. The Basic Safety Message (BSM), colloquially called "Beacon messages," constitutes the first message type [302]. These messages are dispatched regularly and convey essential vehicle information, including identification details, route specifics, and current location. The consistent dissemination of BSMs forms a foundational element in the continuous exchange of critical data among vehicles within the network.

The second message type, the WAVE Short Message (WSM), extends the functionality of BSMs by incorporating additional details about road conditions and specific requests. These requests may include route requests (RRQ), authentication certificates, and network verification. Unlike BSMs, WSM messages are not transmitted periodically but are dispatched selectively when necessary, contributing to the efficiency of the WAVE Short Message Protocols (WSMP) [303].

The third message type is the WAVE Service Advertisement (WSA), which incorporates Vendor Specific Action (VSA) within organizational networks to

broadcast service advertisements. WSA frames play a crucial role in announcing the availability of various services, issuing alerts, and conveying information related to parking, commercial purposes, and more. These messages can be transmitted across all communication channels or selectively directed to a specific service channel, providing flexibility within the VANET ecosystem.

In summary, the IEEE 802.11p /WAVE protocol lays the groundwork for efficient vehicular communication. The orchestrated transmission of BSMs, WSMs, and WSAs represents a sophisticated mechanism for enhancing communication reliability and supporting diverse functionalities critical for advancing ITS.

## 6.4 Simulation Tools for VANET

Simulation tools play an essential role in understanding and optimizing network dynamics. VANETs operate in diverse real-world environments. The simulation replicates these environments, incorporating road structures, traffic patterns, and urban layouts. This ensures a realistic representation to evaluate the network performance. Simulation tools facilitate scalability by comprising many vehicles and infrastructure elements in VANET. Communication protocols such as IEEE 802.11p can be enabled and evaluated in such simulation environments. Finally, security analysis is a critical issue for VANET.

Before implementing security measures in the real world, researchers can evaluate vulnerabilities, test countermeasures, and model and analyze security systems using simulation tools. Researchers use different tools for research purposes to evaluate vehicular networks, including SUMO, UNITY, OMESON, NS3, NS2, OMNeT++, OPNET, VISSM, and VISUM. This study uses the software and tools described in the following subsections. Many reasons drive the choice to use SUMO with OMNeT++ in this study. SUMO is designed to simulate complex traffic scenarios, which is crucial for studying realistic traffic.

Moreover, OMNeT++ is widely used for network simulations and allows researchers to model traffic and network communication in a unified simulation environment. VEINS framework within OMNeT++ is the key factor to this integration. It also provides the ability to simulate VANET protocols, which may not

be available in other simulation tools. Finally, all these tools are open-source tools that can be modified and extended based on specific simulation requirements.

### 6.4.1 SUMO, OMNeT++ and VEINS

SUMO is a powerful microscopic traffic simulator that simulates realistic automotive dynamics in urban environments. Due to its microscopic methodology, the system constructs individual models for each vehicle, capturing small and complex details of their movements. Integration of the software with other simulators, such as OMNeT++, provides researchers with a comprehensive platform for VANET simulations. Researchers frequently leverage SUMO's capabilities to analyze and optimize VANET behavior in urban scenarios, making it an important software in vehicular network research.

OMNeT++, a discrete-event simulation toolkit, is essential for network simulations. It is easier to model complex VANET communication protocols and behaviors with OMNeT++ due to its adaptability and ability to simulate different scenarios. OMNeT++ has provided several frameworks that can be integrated into the software so that the researcher can simulate all types of networks, including LTE and 5G, and VEINS framework dedicated to the vehicular environment. The research community prefers it because of its robust and extendable C++ architecture, which helps them investigate communication dynamics in complicated networks such as VANET.

VEINS is an open-source framework built to simulate vehicular networks in OMNeT++. It integrates seamlessly with SUMO through TraCi [182] to realistically represent urban and suburban mobility. VEINS supports multiple VANET communication protocols such as IEEE 802.11p, AODV, UDP, TCP, etc. Its accessibility helps researchers study vehicle communication, and the integration between VEINS and SUMO makes VANET simulations more realistic and accurate.

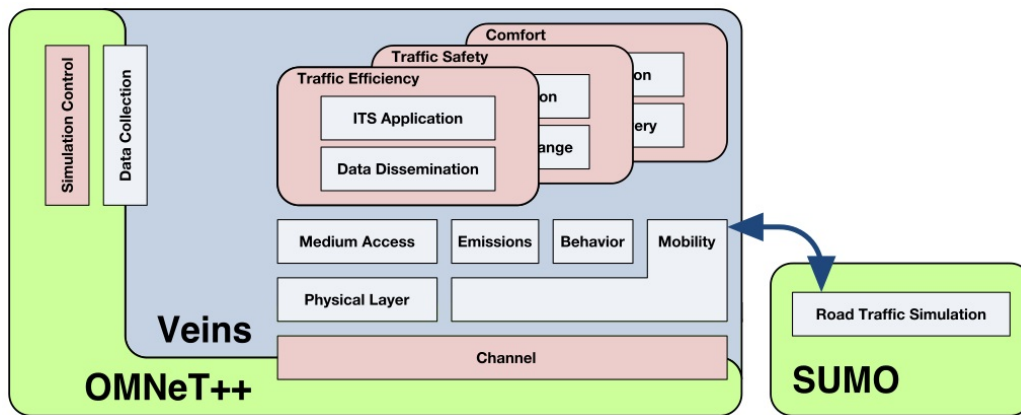


Figure 6.3: VEINS Architecture with SUMO and OMNeT++

Those simulators and tools were used in this study as they can provide realistic emulation and let us focus on the study objectives at hand rather than side concerns such as networks, signals, and traffic. Figure 6.3 shows the integration between the tools and how VEINS can be connected to SUMO and OMNeT++.

SUMO generally provides authentic mobility patterns and integrates real-world road networks and traffic conditions into simulations. On the other hand, OMNeT++ facilitates the modeling of communication protocols, adding a layer of realism to the exchange of messages in V2V and V2I communications. The integration process encounters challenges arising from the distinct paradigms employed by SUMO and OMNeT++. Real-time communication between these tools is enabled through the TraCI interface, embedded in the VEINS framework and represented in the `TraCiDemo11p.cc` and `TraCiDemo11p.h` files. An integration script was built, and the scenario functions within C++-based `TraCiDemo11p` files to execute code seamlessly in OMNeT++, as shown in Figure 6.4.

Furthermore, making these different simulations work together requires adjusting the parameters, data formats, nodes, RSUs, and communication setup, and all these parameters are adjusted in the `omnet.ini` file as shown in Figure 6.5. Despite integration challenges, the simulation successfully handled these intricacies, proving its strength in accurately representing different attack scenarios.

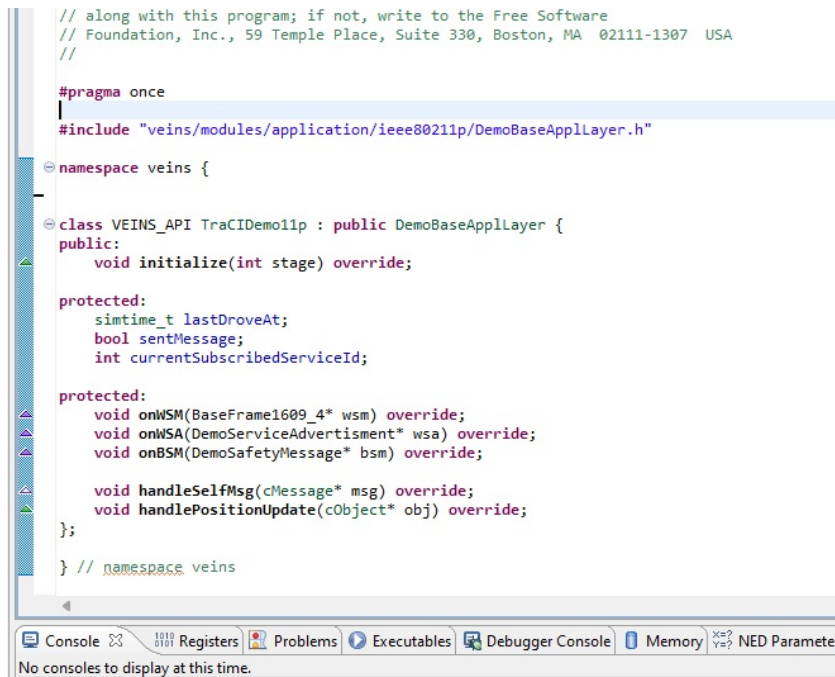
The image shows a code editor window with a light blue header bar. The code is written in C++ and is part of a namespace named 'veins'. It defines a class 'VEINS\_API TraCIDemo11p' which inherits from 'DemoBaseApplLayer'. The class has a public method 'initialize(int stage) override', a protected attribute 'simtime\_t lastDroveAt', and protected attributes 'bool sentMessage' and 'int currentSubscribedServiceId'. There are also protected methods for handling various messages: 'onWSM(BaseFrame1609\_4\* wsm) override', 'onWSA(DemoServiceAdvertisement\* wsa) override', 'onBSM(DemoSafetyMessage\* bsm) override', 'handleSelfMsg(cMessage\* msg) override', and 'handlePositionUpdate(cObject\* obj) override'. The code is enclosed in a namespace 'veins' block. The editor's status bar at the bottom shows 'No consoles to display at this time.' and various tool icons like Console, Registers, Problems, Executables, Debugger Console, and Memory.

Figure 6.4: TraCIDemo11p.h implementation in OMNeT++

Implementing the IEEE 802.11p protocol and the WAVE standard enables modeling specific attack scenarios, such as evading vehicles from charging stations and DDOS attacks. The flexibility of these tools has been demonstrated in the ability to simulate different types of attacks and analyze the messages exchanged between vehicles through the WAVE protocol to identify malicious traffic. The amalgamation of SUMO, OMNeT++, and VEINS provides a robust platform that harmoniously blends scalability and realism to help implement security solutions for vehicular networks.

## 6.5 Case Study: Bologna City - Pasubio Region

### 6.5.1 Simulation Description

This study used a realistic traffic simulation scenario from Bologna, specifically focusing on the Pasubio area during the city's peak traffic hour (8:00 - 9:00 a.m.) as shown in Figure 6.6. The simulation incorporates additional datasets provided

```

*.CS[*].appl.dataOnSch = true
*.CS[*].appl.beaconInterval = 60s
*.CS[*].appl.beaconUserPriority = 7
*.CS[*].appl.dataUserPriority = 5
*.CS[*].nic.phy80211p.antennaOffsetZ = 0 m

#####
#           11p specific parameters           #
#                                           #
#           NIC-Settings                     #
#####
*.connectionManager.sendDirect = true
*.connectionManager.maxInterfDist = 3600m
*.connectionManager.drawMaxIntfDist = true

*.*.nic.mac1609_4.useServiceChannel = true

*.*.nic.mac1609_4.txPower = 20mW
*.*.nic.mac1609_4.bitrate = 6Mbps
*.*.nic.phy80211p.minPowerLevel = -89dBm

*.*.nic.phy80211p.useNoiseFloor = true
*.*.nic.phy80211p.noiseFloor = -98dBm

*.*.nic.phy80211p.decider = xmldoc("config.xml")
*.*.nic.phy80211p.analogueModels = xmldoc("config.xml")
*.*.nic.phy80211p.usePropagationDelay = true

*.*.nic.phy80211p.antenna = xmldoc("antenna.xml", "/root/Antenna[@id='monopole']")
*.node[*].nic.phy80211p.antennaOffsetY = 0 m
*.node[*].nic.phy80211p.antennaOffsetZ = 1.895 m
*.hacker[*].nic.phy80211p.antennaOffsetY = 0 m
*.hacker[*].nic.phy80211p.antennaOffsetZ = 1.895 m

#####
#           App Layer                       #
#####

*.hacker[*].applType = "TraCIDemo11p"
*.hacker[*].appl.headerLength = 80 bit
*.hacker[*].appl.sendBeacons = true
*.hacker[*].appl.dataOnSch = true
*.hacker[*].appl.beaconInterval = 5s

*.node[*].applType = "TraCIDemo11p"
*.node[*].appl.headerLength = 80 bit
*.node[*].appl.sendBeacons = true
*.node[*].appl.dataOnSch = true
*.node[*].appl.beaconInterval = 1s

```

Figure 6.5: Sample of parameters setup in omnet.ini file

by the municipality of Bologna, including the positions and plans of traffic lights, the positions of the inductive loops, and measurements, among others. Initially, the scenario featured 3700 conventional vehicles. However, in order to align with the study objective of simulating the behavior of EV in the network, I adapted the vehicle specifications in SUMO to represent EVs and scaled the number down to 500. This adjustment aims to create a more realistic representation of EV communication within the network. The authenticity of the simulated traffic scenario is further enhanced by integrating real traffic data from Bologna into the OMNeT++ model.



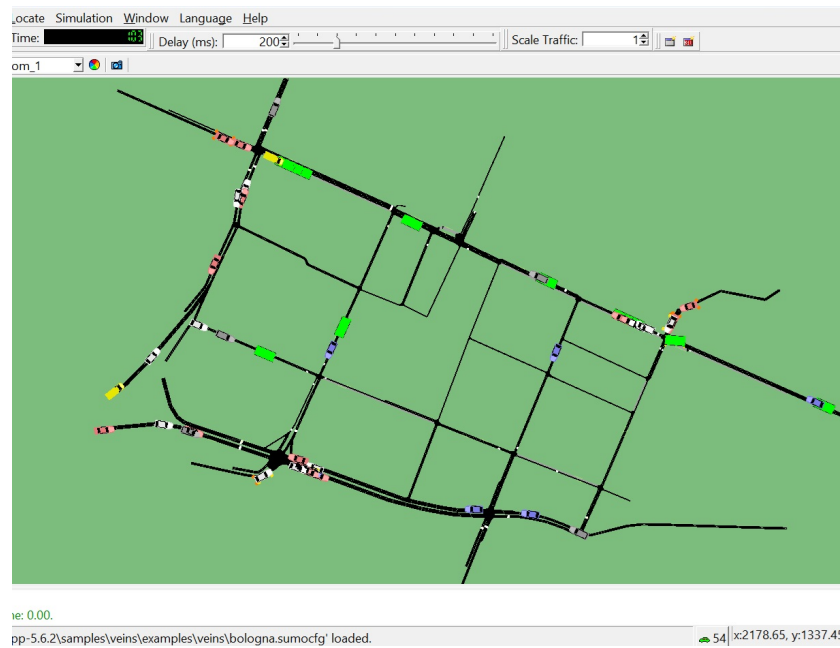


Figure 6.6: Realistic traffic environment in Bologna city.

The design of the network model in OMNeT++ is based on the SUMO simulation, and the network components are meticulously defined. As depicted in Figure 6.7, two RSUs were strategically positioned to cover a large map area. Additionally, three charging stations were incorporated on the map. These stations can send and receive beacons from RSUs, EVs, and other objects within the network's range, enhancing the simulation's realism and applicability.

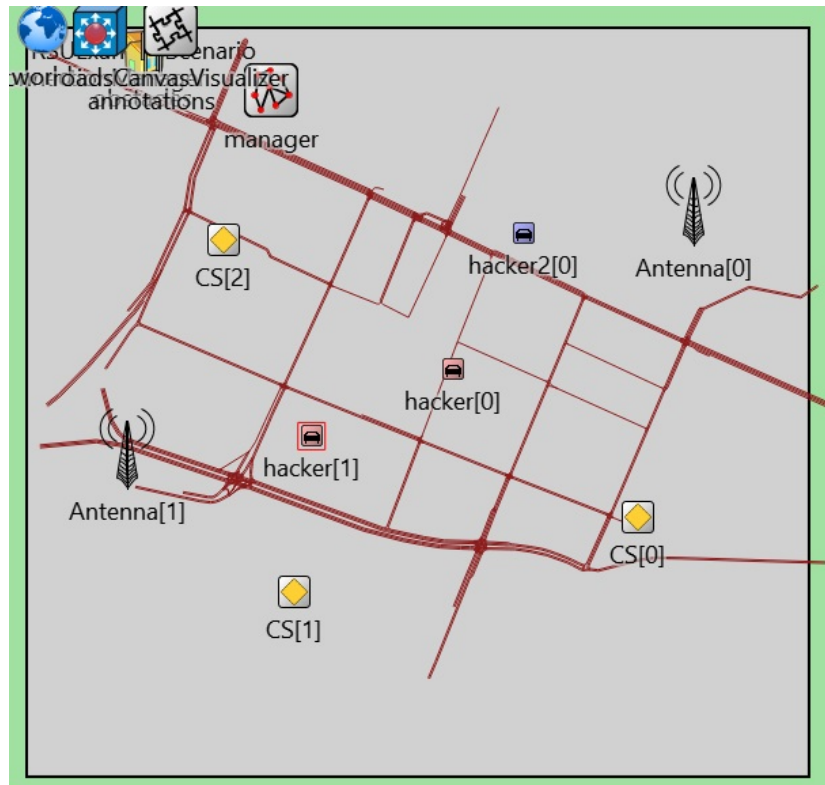


Figure 6.7: RSU and attackers position in OMNeT++

## 6.5.2 Attacks Simulation

The simulation, as illustrated in Figure. 6.7, introduced two vehicles designated as attackers. Having gained unauthorized access to the network through fake identities or other hacking methods, these attackers are positioned at fixed locations to initiate their attacks at varying intensities.

The first attacker, represented in red, has a relatively mild impact on the network. Their strategy involves exploiting the WAVE protocol's WSA messages, typically used for network advertisements. The attacker broadcasts false alerts about an accident in a specific location to neighboring vehicles every 60 seconds, as depicted in Figure 6.8. This misleading information misleads other vehicles, forcing them to reroute away from charging stations.

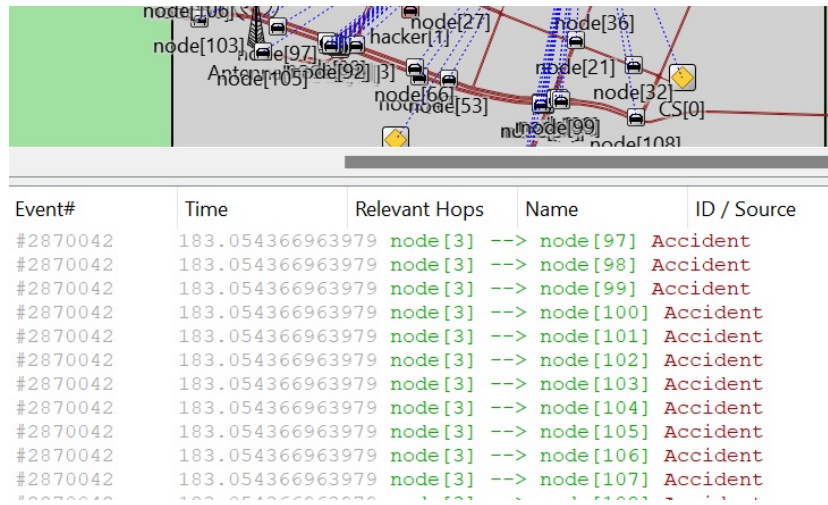


Figure 6.8: First attack: broadcast fake accident advertisement

While this attack does not entirely hinder the network, it disrupts the charging reservation system and the coordination between vehicles and charging stations. Consequently, many vehicles miss their opportunity to secure a charging slot at the desired time, potentially allowing the attacker to monopolize charging availability. Moreover, such attacks can undermine the credibility and reliability of service providers, leading to significant losses due to the compromised reservation system.

The second attacker, shown in blue in Figure 6.9, poses a more severe threat to the network. They aim to execute a DoS attack to disrupt the network and deny service to all antennas and vehicles. DoS attacks in VANETs involve malicious efforts to overload communication channels, leading to service unavailability or degradation. These attacks typically target the communication infrastructure, exploiting vulnerabilities to hinder the exchange of crucial safety and traffic information. In this case study, the attacker floods the RSUs with 1000 WSM messages per second, far exceeding the network's capacity of 300 messages per second. This deluge of messages overloads the communication channel, leading to the shutdown of the RSUs and rendering the entire network inoperative.

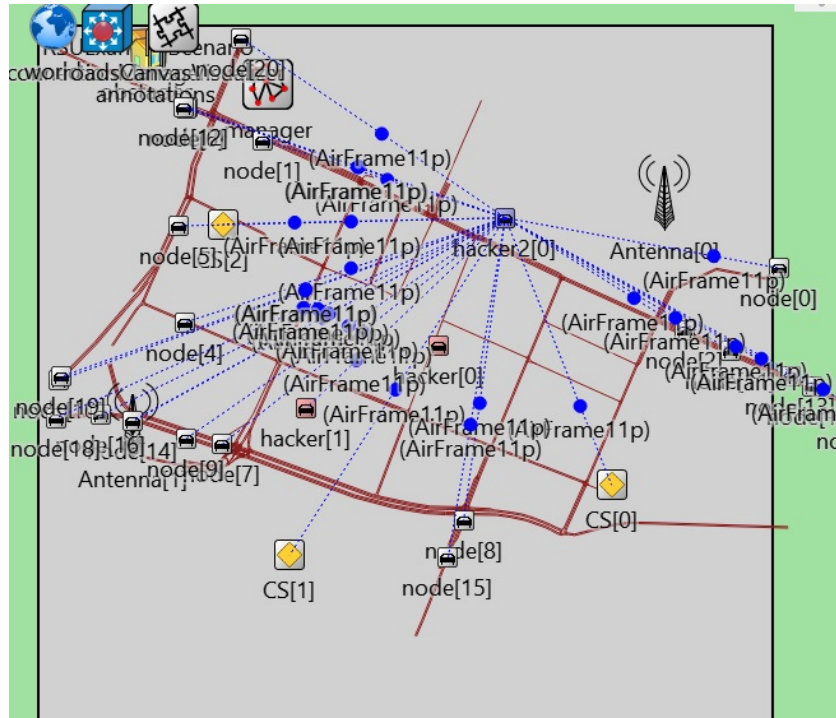


Figure 6.9: Second attack: DDoS attacks on the entire network

## 6.6 Result and Discussion

This study observed that the hackers, who had gained legitimate access to the network, could exploit WAVE message broadcasting. They misused this access by flooding the WSM with numerous requests, such as RRQ, verification certificates, and permissions, or using WSA to disseminate false advertisements. As a standard mode of communication within the VANET network, this form of broadcasting poses a challenge in distinguishing legitimate operations from malicious activities.

Due to the limited duration of the simulation and the number of vehicles involved, the simulation yielded a relatively small unlabeled dataset. Consequently, traditional machine learning techniques were not feasible for evaluating the simulation. Instead, I developed a simple Python script based on threshold analysis to identify anomalies in the dataset.

The simulation results, as shown in Table 6.1, are credible and applicable, making a significant contribution to the field of VANET simulation, particularly in network security. After simulating for 3600 seconds, critical data was captured, including the number of generated WSMs and WSAs and packets sent by the injected attackers. The data highlight the extensive number of DoS attacks initiated by Hacker Type2 and the volume of misleading messages propagated by Hacker Type1. Additional metrics such as TimesIntoBackoff, Channel Busy, and Total-BusyTime were also recorded, providing information on network behavior under attack conditions.

Table 6.1: Simulation Results

	Hacker Type1	Hacker Type2	Antena
generatedWSMs	0	3599900	0
generatedWSAs	2000	0	0
SendPackets	840	359985	3600
SlotsBackoff	1207	5400510	5312
Channel Busy	1.1433333E-5	0.01049956	1.05E-4
totalBusyTime	0.04116	37.798	0.378
totalTime	3600s	3600s	3598s

In future work, there is a plan to evolve this model to mirror real-world conditions more closely. This will involve incorporating a mix of regular and EVs, extending the simulation duration, and introducing a wider array of attack types. The focus will be replicating real-life events and potential faults within the simulation environment. This approach is expected to generate a larger dataset suitable for applying machine learning algorithms to detect and analyze network intrusions and attacks more effectively.

# Chapter 7

## Conclusion

This thesis marks a significant milestone in exploring ITS. It focuses on the integration of security measures and sustainability strategies to enhance the effectiveness of vehicular networks. This research underscores the imperative need to address the complex challenges inherent in modern transportation systems.

### **Comprehensive Review of Digital Twin Technologies:**

The foundational element of this thesis rests upon an exhaustive review of DT technologies in the context of ITS. This thorough exploration spans the inception of DT to its diverse applications, particularly within ITS and smart cities. The insights gained from this review serve as the underpinning framework for subsequent research and highlight the pivotal role of DT in providing real-time monitoring and response capabilities, thereby fortifying the overall cybersecurity framework.

### **Intrusion Detection Systems :**

IDS tailored for VANETs represents a substantial leap forward in ensuring the security of vehicular networks. Leveraging decision tree-based ML techniques, this IDS combines sophisticated feature selection methods and tree-based algorithms to achieve remarkable attack detection accuracy. Beyond its efficacy in fortifying vehicular network security, the proposed IDS aligns with sustainability goals by minimizing potential security breaches and the associated costs.

### **Routing Simulation and Optimization under Smart Charging Strategies:**

A holistic approach to sustainability within ITS is introduced by optimizing EV routing for logistics operations. Formulated as an integer linear programming problem, the optimization method minimizes charging/discharging costs while considering the shortest path for each EV. Integrating smart charging strategies contributes to sustainability by minimizing charging costs and reflects an innovative paradigm in optimizing logistics operations. Real-world case studies, including scenarios in the Puglia region (Italy), validate the effectiveness of the proposed optimization method.

#### **Simulation of Traffic Networks and Communication Protocols:**

A simulation-based study on traffic networks and communication protocols further enriches the contributions of this thesis. Employing a hybrid methodology that integrates SUMO, OMNeT++, and VEINS frameworks, the study comprehensively models and simulates interactions within the dynamic urban setting of Bologna, Italy. Emphasis is placed on examining different types of attacks on VANET networks through IEEE 802.11p protocol / WAVE standard messages. This simulation-based approach enhances vehicular network security and contributes to sustainability by ensuring the reliability and efficiency of communication protocols.

#### **Synthesis of Contributions and Future Perspectives:**

DT Technologies is the global framework that integrates and unifies many processes, including this thesis's diverse proposals and techniques. Rather than each section being independent work, they collectively constitute integral components within the broader concept of DT. Every section contributes to the comprehensive understanding and application of DT in the context of ITS.

The Comprehensive Review of DT Technologies sets the stage by providing a foundational understanding of their evolution and applications. Subsequently, the proposed IDS, routing optimization under smart charging strategies, and simulation of traffic networks and communication protocols are all intricately working within the DT paradigm. Each section represents a crucial facet of DT to simulate physical ITS networks, analyze outcomes, and store data for processing. The conclusions of these sections underscore its role as a holistic approach to enhancing VANET security by simulating, analyzing, and making decisions within the DT concept.

The future holds promising avenues for further exploration and refinement of the proposed methodologies. The ongoing advancement of technology, particularly in AI and simulation tools, provides opportunities for continual improvement and application in real-world scenarios. As we strive toward a more secure, efficient, and sustainable transportation ecosystem, the contributions of this thesis lay a solid foundation for continued progress in the dynamic field of ITS.



# References

- [1] David Jones, Chris Snider, Aydin Nassehi, Jason Yon, and Ben Hicks. Characterising the digital twin: A systematic literature review. *CIRP journal of manufacturing science and technology*, 29:36–52, 2020.
- [2] Bin He and Kai-Jian Bai. Digital twin-based sustainable intelligent manufacturing: A review. *Advances in Manufacturing*, 9:1–21, 2021.
- [3] Sagheer Khan, Tughrul Arslan, and Tharmalingam Ratnarajah. Digital twin perspective of fourth industrial and healthcare revolution. *Ieee Access*, 10:25732–25754, 2022.
- [4] Ramy Elsehrawy, Bimal Kumar, and Richard Watson. A digital twin uses classification system for urban planning & city infrastructure management. *Journal of Information Technology in Construction*, 26:832–862, 2021.
- [5] Xin Tong, Qiang Liu, Shiwei Pi, and Yao Xiao. Real-time machining data application and service based on imt digital twin. *Journal of Intelligent Manufacturing*, 31:1113–1132, 2020.
- [6] Andrey Rudskoy, Igor Ilin, and Andrey Prokhorov. Digital twins in the intelligent transport systems. *Transportation Research Procedia*, 54:927–935, 2021.
- [7] George Dimitrakopoulos and Panagiotis Demestichas. Intelligent transportation systems. *IEEE Vehicular Technology Magazine*, 5(1):77–84, 2010.
- [8] Jingwen Wu, Hua Liao, and Jin-Wei Wang. Analysis of consumer attitudes towards autonomous, connected, and electric vehicles: A survey in china. *Research in transportation economics*, 80:100828, 2020.
- [9] Daniel Sperling. *Future drive: Electric vehicles and sustainable transportation*. Island Press, 2013.
- [10] Daniel J Fagnant and Kara Kockelman. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, 77:167–181, 2015.

- [11] Pantelis Kopelias, Elissavet Demiridi, Konstantinos Vogiatzis, Alexandros Skabardonis, and Vassiliki Zafiropoulou. Connected & autonomous vehicles—environmental impacts—a review. *Science of the total environment*, 712:135237, 2020.
- [12] Santhanakrishnan Narayanan, Emmanouil Chaniotakis, and Constantinos Antoniou. Shared autonomous vehicle services: A comprehensive review. *Transportation Research Part C: Emerging Technologies*, 111:255–293, 2020.
- [13] S Muthuramalingam, A Bharathi, S Rakesh Kumar, N Gayathri, R Sathiyaraj, and B Balamurugan. Iot based intelligent transportation system (iot-its) for global perspective: A case study. *Internet of Things and Big Data Analytics for Smart Generation*, pages 279–300, 2019.
- [14] Zhi Cheng, Min-Seok Pang, and Paul A Pavlou. Mitigating traffic congestion: The role of intelligent transportation systems. *Information Systems Research*, 31(3):653–674, 2020.
- [15] Dalton Hahn, Arslan Munir, and Vahid Behzadan. Security and privacy issues in intelligent transportation systems: Classification and challenges. *IEEE Intelligent Transportation Systems Magazine*, 13(1):181–196, 2019.
- [16] Alexandros Nikitas, Kalliopi Michalakopoulou, Eric Tchouamou Njoya, and Dimitris Karampatzakis. Artificial intelligence, transport and the smart city: Definitions and dimensions of a new mobility era. *Sustainability*, 12(7):2789, 2020.
- [17] Jonathan Petit and Steven E Shladover. Potential cyberattacks on automated vehicles. *IEEE Transactions on Intelligent transportation systems*, 16(2):546–556, 2014.
- [18] A Haghi, D Ketabi, M Ghanbari, and H Rajabi. Assessment of human errors in driving accidents; analysis of the causes based on aberrant behaviors. *Life Science Journal*, 11(9):414–420, 2014.
- [19] Stefan K Gehrig and Fridtjof J Stein. Dead reckoning and cartography using stereo vision for an autonomous car. In *Proceedings 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human and Environment Friendly Robots with High Intelligence and Emotional Quotients (Cat. No. 99CH36289)*, volume 3, pages 1507–1512. IEEE, 1999.
- [20] Halabi Hasbullah, Irshad Ahmed Soomro, et al. Denial of service (dos) attack and its possible solutions in vanet. *International Journal of Electronics and Communication Engineering*, 4(5):813–817, 2010.
- [21] Jasmin Bharadiya. Artificial intelligence in transportation systems a critical review. *American Journal of Computing and Engineering*, 6(1):34–45, 2023.

- [22] Gayathri Chandrasekaran. Vanets: The networking platform for future vehicular applications. *Department of Computer Science, Rutgers University*, pages 45–51, 2008.
- [23] Hassnaa Moustafa and Yan Zhang. *Vehicular networks: techniques, standards, and applications*. Auerbach publications, 2009.
- [24] Xueqin Lü, Yinbo Wu, Jie Lian, Yangyang Zhang, Chao Chen, Peisong Wang, and Lingzheng Meng. Energy management of hybrid electric vehicles: A review of energy optimization of fuel cell hybrid power system based on genetic algorithm. *Energy Conversion and Management*, 205:112474, 2020.
- [25] Michele Roccotelli, Massimiliano Nolich, Maria Pia Fanti, and Walter Ukovich. Internet of things and virtual sensors for electromobility. *Internet Technology Letters*, 1(3):e39, 2018.
- [26] Anton Rassõlkin, Toomas Vaimann, Ants Kallaste, and Vladimir Kuts. Digital twin for propulsion drive of autonomous electric vehicle. In *2019 IEEE 60th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON)*, pages 1–4. IEEE, 2019.
- [27] Gartner top 10 strategic technology trends for 2019. <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019>. (Accessed on 12/16/2023).
- [28] Qinglin Qi and Fei Tao. Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. *Ieee Access*, 6:3585–3593, 2018.
- [29] Michael W Grieves. *Virtually intelligent product systems: Digital and physical twins*. 2019.
- [30] Eric J Tuegel, Anthony R Ingrassia, Thomas G Eason, S Michael Spottswood, et al. Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, 2011, 2011.
- [31] Stephan Weyer, Torben Meyer, Moritz Ohmer, Dominic Gorecky, and Detlef Zühlke. Future modeling and simulation of cps-based factories: an example from the automotive industry. *Ifac-Papersonline*, 49(31):97–102, 2016.
- [32] M Mazhar Rathore, Syed Attique Shah, Dharendra Shukla, Elmahdi Bentafat, and Spiridon Bakiras. The role of ai, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities. *IEEE Access*, 9:32030–32052, 2021.

- [33] Vyacheslav Kharchenko, Oleg Illiashenko, Olga Morozova, and Sergii Sokolov. Combination of digital twin and artificial intelligence in manufacturing using industrial iot. In *2020 IEEE 11th international conference on dependable systems, services and technologies (DESSERT)*, pages 196–201. IEEE, 2020.
- [34] Qiuchen Lu, Xiang Xie, Ajith Kumar Parlikad, and Jennifer Mary Schooling. Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance. *Automation in Construction*, 118:103277, 2020.
- [35] Wasim A Ali, Michèle Roccotelli, and Maria Pia Fanti. Digital twin in intelligent transportation systems: A review. In *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)*, volume 1, pages 576–581. IEEE, 2022.
- [36] Ziran Wang, Xishun Liao, Xuanpeng Zhao, Kyungtae Han, Prashant Tiwari, Matthew J Barth, and Guoyuan Wu. A digital twin paradigm: Vehicle-to-cloud based advanced driver assistance systems. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pages 1–6. IEEE, 2020.
- [37] Kazi Masudul Alam and Abdulmoteleb El Saddik. C2ps: A digital twin architecture reference model for the cloud-based cyber-physical systems. *IEEE access*, 5:2050–2062, 2017.
- [38] Ghanishtha Bhatti, Harshit Mohan, and R Raja Singh. Towards the future of smart electric vehicles: Digital twin technology. *Renewable and Sustainable Energy Reviews*, 141:110801, 2021.
- [39] Zhihan Lv, Yuxi Li, Hailin Feng, and Haibin Lv. Deep learning for security in digital twins of cooperative intelligent transportation systems. *IEEE transactions on intelligent transportation systems*, 23(9):16666–16675, 2021.
- [40] Sathish AP Kumar, R Madhumathi, Pethuru Raj Chelliah, Lei Tao, and Shangguang Wang. A novel digital twin-centric approach for driver intention prediction and traffic congestion avoidance. *Journal of Reliable Intelligent Environments*, 4:199–209, 2018.
- [41] Lidbe A. D. Lu W. Dasgupta S., Rahman M. and Jones S. A transportation digital-twin approach for adaptive traffic control systems. *Transportation Research Board 100th Annual Meeting and publication in Transportation Research Record*, arXiv:2109.10863, 2021.
- [42] Mohammad Aslani, Mohammad Saadi Mesgari, and Marco Wiering. Adaptive traffic signal control with actor-critic methods in a real-world traffic network with different traffic disruption events. *Transportation Research Part C: Emerging Technologies*, 85:732–752, 2017.

- [43] Hayato Shikata, Tetsuo Yamashita, Kouji Arai, Takayuki Nakano, Kenichi Hatanaka, and Hiroyuki Fujikawa. Digital twin environment to integrate vehicle simulation and physical verification. *SEI Technical Review*, 88:18–21, 2019.
- [44] Wenwen Wang, Jun Wang, Jinpeng Tian, Jiahuan Lu, and Rui Xiong. Application of digital twin in smart battery management systems. *Chinese Journal of Mechanical Engineering*, 34(1):1–19, 2021.
- [45] Yongkang Liu, Ziran Wang, Kyungtae Han, Zhenyu Shou, Prashant Tiwari, and John HL Hansen. Sensor fusion of camera and cloud digital twin information for intelligent vehicles. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 182–187. IEEE, 2020.
- [46] Yiwen Wu, Ke Zhang, and Yan Zhang. Digital twin networks: A survey. *IEEE Internet of Things Journal*, 8(18):13789–13804, 2021.
- [47] Yueyue Dai, Ke Zhang, Sabita Maharjan, and Yan Zhang. Deep reinforcement learning for stochastic computation offloading in digital twin networks. *IEEE Transactions on Industrial Informatics*, 17(7):4968–4977, 2020.
- [48] Aidan Fuller, Zhong Fan, Charles Day, and Chris Barlow. Digital twin: Enabling technologies, challenges and open research. *IEEE access*, 8:108952–108971, 2020.
- [49] Zhiheng Zhao, Leidi Shen, Chen Yang, Wei Wu, Mengdi Zhang, and George Q Huang. Iot and digital twin enabled smart tracking for safety management. *Computers & Operations Research*, 128:105183, 2021.
- [50] Mehdi Kherbache, Moufida Maimour, and Eric Rondeau. Network digital twin for the industrial internet of things. In *2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoW-MoM)*, pages 573–578. IEEE, 2022.
- [51] Maria Pia Fanti, Massimiliano Nolich, Michele Roccotelli, and Walter Ukovich. Virtual sensors for electromobility. In *2018 5th International Conference on Control, Decision and Information Technologies (CoDIT)*, pages 635–640. IEEE, 2018.
- [52] Giambattista Grusso, Giancarlo Storti Gajani, Fredy Ruiz, Juan Diego Valladolid, and Diego Patino. A virtual sensor for electric vehicles’ state of charge estimation. *Electronics*, 9(2):278, 2020.
- [53] Maria Pia Fanti, Agostino Marcello Mangini, and Michele Roccotelli. An innovative service for electric vehicle energy demand prediction. In *2020 7th International Conference on Control, Decision and Information Technologies (CoDIT)*, volume 1, pages 880–885. IEEE, 2020.

- [54] Hailin Feng, Dongliang Chen, and Zhihan Lv. Blockchain in digital twins-based vehicle management in vanets. *IEEE Transactions on Intelligent Transportation Systems*, 23(10):19613–19623, 2022.
- [55] Jiajie Guo, Muhammad Bilal, Yuying Qiu, Cheng Qian, Xiaolong Xu, and Kim-Kwang Raymond Choo. Survey on digital twins for internet of vehicles: Fundamentals, challenges, and opportunities. *Digital Communications and Networks*, 2022.
- [56] Tianle Zhang, Xiangtao Liu, Zongwei Luo, Fuqiang Dong, and Yu Jiang. Time series behavior modeling with digital twin for internet of vehicles. *EURASIP Journal on Wireless Communications and Networking*, 2019:1–11, 2019.
- [57] João A Peças Lopes, Filipe Joel Soares, and Pedro M Rocha Almeida. Integration of electric vehicles in the electric power system. *Proceedings of the IEEE*, 99(1):168–183, 2010.
- [58] Pal Varga, Jozsef Peto, Attila Franko, David Balla, David Haja, Ferenc Janky, Gabor Soos, Daniel Ficzer, Markosz Maliosz, and Laszlo Toka. 5g support for industrial iot applications—challenges, solutions, and research gaps. *Sensors*, 20(3):828, 2020.
- [59] Ali Gohar and Gianfranco Nencioni. The role of 5g technologies in a smart city: The case for intelligent transportation system. *Sustainability*, 13(9):5188, 2021.
- [60] Chunhua Hu, Weicun Fan, Elan Zeng, Zhi Hang, Fan Wang, Lianyong Qi, and Md Zakirul Alam Bhuiyan. Digital twin-assisted real-time traffic data prediction method for 5g-enabled internet of vehicles. *IEEE Transactions on Industrial Informatics*, 18(4):2811–2819, 2021.
- [61] Juan Deng, Qingbi Zheng, Guangyi Liu, Jielin Bai, Kaicong Tian, Changhao Sun, Yujie Yan, and Yitong Liu. A digital twin approach for self-optimization of mobile networks. In *2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 1–6. IEEE, 2021.
- [62] Rui Dong, Changyang She, Wibowo Hardjawana, Yonghui Li, and Branka Vucetic. Deep learning for hybrid 5g services in mobile edge computing systems: Learn from a digital twin. *IEEE Transactions on Wireless Communications*, 18(10):4692–4707, 2019.
- [63] Jithin Jagannath, Keyvan Ramezanzpour, and Anu Jagannath. Digital twin virtualization with machine learning for iot and beyond 5g networks: Research directions for security and optimal control. In *Proceedings of the 2022 ACM Workshop on Wireless Security and Machine Learning*, pages 81–86, 2022.

- [64] Self-driving car - wikipedia. [https://en.wikipedia.org/wiki/Self-driving\\_car](https://en.wikipedia.org/wiki/Self-driving_car). (Accessed on 12/16/2023).
- [65] Joaquín López, Pablo Sánchez-Vilariño, Rafael Sanz, and Enrique Paz. Efficient local navigation approach for autonomous driving vehicles. *IEEE Access*, 9:79776–79792, 2021.
- [66] Kichun Jo, Minchul Lee, Wontek Lim, and Myoung-ho Sunwoo. Hybrid local route generation combining perception and a precise map for autonomous cars. *IEEE Access*, 7:120128–120140, 2019.
- [67] Nijat Rajabli, Francesco Flammini, Roberto Nardone, and Valeria Vittorini. Software verification and validation of safe autonomous cars: A systematic literature review. *IEEE Access*, 9:4797–4819, 2020.
- [68] Haifei Yang, Changjiang Zheng, Yi Zhao, and Zhong Wu. Integrating the intelligent driver model with the action point paradigm to enhance the performance of autonomous driving. *IEEE Access*, 8:106284–106295, 2020.
- [69] Mehdi Masmoudi, Hamdi Friji, Hakim Ghazzai, and Yehia Massoud. A reinforcement learning framework for video frame-based autonomous car-following. *IEEE Open Journal of Intelligent Transportation Systems*, 2:111–127, 2021.
- [70] Flexsim 2022: Reinforcement learning + experimenter improvements | flexsim. <https://www.flexsim.com/fr/news/flexsim-2022-reinforcement-learning-experimenter-improvements/>. (Accessed on 12/16/2023).
- [71] Suchitra Venkatesan, Krishnan Manickavasagam, Nikita Tengenkai, and Nagendran Vijayalakshmi. Health monitoring and prognosis of electric vehicle motor using intelligent-digital twin. *IET Electric Power Applications*, 13(9):1328–1335, 2019.
- [72] Kaveh Sarrafan, Kashem M Muttaqi, and Danny Sutanto. Real-time state-of-charge tracking embedded in the advanced driver assistance system of electric vehicles. *IEEE Transactions on Intelligent Vehicles*, 5(3):497–507, 2020.
- [73] Akshay Malhotra, Giulio Binetti, Ali Davoudi, and Ioannis D Schizas. Distributed power profile tracking for heterogeneous charging of electric vehicles. *IEEE Transactions on Smart Grid*, 8(5):2090–2099, 2016.
- [74] Rui Wang, Dengqiang Wang, Lanhong Wang, and Yuming Dou. Design of remote monitoring system for new energy vehicles from the perspective of key technologies. In *2021 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, pages 463–467. IEEE, 2021.

- [75] Fazel Mohammadi and Mehrdad Saif. A comprehensive overview of electric vehicle batteries market. *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, page 100127, 2023.
- [76] Taesic Kim, Darshan Makwana, Amit Adhikaree, Jitendra Shamjibhai Vagdoda, and Young Lee. Cloud-based battery condition monitoring and fault diagnosis platform for large-scale lithium-ion battery energy storage systems. *Energies*, 11(1):125, 2018.
- [77] Tetsu Tanizawa, Takayuki Suzumiya, and Kazuto Ikeda. Cloud-connected battery management system supporting e-mobility. *Fujitsu Sci. Tech. J.*, 51(4):27–35, 2015.
- [78] Koko Friansa, Irsyad Nashirul Haq, Bening Maria Santi, Deddy Kurniadi, Edi Leksono, and Brian Yulianto. Development of battery monitoring system in smart microgrid based on internet of things (iot). *Procedia engineering*, 170:482–487, 2017.
- [79] Weihan Li, Monika Rentemeister, Julia Badeda, Dominik Jöst, Dominik Schulte, and Dirk Uwe Sauer. Digital twin for battery systems: Cloud battery management system with online state-of-charge and state-of-health estimation. *Journal of energy storage*, 30:101557, 2020.
- [80] Morsy Nour, Sayed M Said, Abdelfatah Ali, and Csaba Farkas. Smart charging of electric vehicles according to electricity price. In *2019 international conference on innovative trends in computer engineering (ITCE)*, pages 432–437. IEEE, 2019.
- [81] Rachana Vidhi and Prasanna Shrivastava. A review of electric vehicle lifecycle emissions and policy recommendations to increase ev penetration in india. *Energies*, 11(3):483, 2018.
- [82] Onur Alparslan, Shin’ichi Arakawa, and Masayuki Murata. Next generation intra-vehicle backbone network architectures. In *2021 IEEE 22nd International Conference on High Performance Switching and Routing (HPSR)*, pages 1–7. IEEE, 2021.
- [83] Martin Smuts, Brenda Scholtz, and Janet Wesson. Issues in implementing a data integration platform for electric vehicles using the internet of things. In *Internet of Things. Information Processing in an Increasingly Connected World: First IFIP International Cross-Domain Conference, IFIPIoT 2018, Held at the 24th IFIP World Computer Congress, WCC 2018, Poznan, Poland, September 18-19, 2018, Revised Selected Papers 1*, pages 160–177. Springer, 2019.
- [84] Maanak Gupta and Ravi Sandhu. Authorization framework for secure cloud assisted connected cars and vehicular internet of things. In *Proceedings of*



- the 23rd ACM on symposium on access control models and technologies*, pages 193–204, 2018.
- [85] Shah Khalid Khan, Nirajan Shiwakoti, Peter Stasinopoulos, and Yilun Chen. Cyber-attacks in the next-generation cars, mitigation techniques, anticipated readiness and future directions. *Accident Analysis & Prevention*, 148:105837, 2020.
- [86] Sadeq Almeaibed, Saba Al-Rubaye, Antonios Tsourdos, and Nicolas P Avdelidis. Digital twin analysis to promote safety and security in autonomous vehicles. *IEEE Communications Standards Magazine*, 5(1):40–46, 2021.
- [87] Lulu Guo, Bowen Yang, and Jin Ye. Enhanced cyber-physical security of steering stability control system for four-wheel independent drive electric vehicles. In *2020 IEEE Transportation Electrification Conference & Expo (ITEC)*, pages 1240–1245. IEEE, 2020.
- [88] Ponnuru Raveendra Babu, Ruhul Amin, Alavalapati Goutham Reddy, Ashok Kumar Das, Willy Susilo, and YoungHo Park. Robust authentication protocol for dynamic charging system of electric vehicles. *IEEE Transactions on Vehicular Technology*, 70(11):11338–11351, 2021.
- [89] Abdollah Kavousi-Fard, Tao Jin, Wencong Su, and Navid Parsa. An effective anomaly detection model for securing communications in electric vehicles. *IEEE Transactions on Industry Applications*, 2020.
- [90] Violeta Damjanovic-Behrendt. A digital twin-based privacy enhancement mechanism for the automotive industry. In *2018 International Conference on Intelligent Systems (IS)*, pages 272–279. IEEE, 2018.
- [91] Yilong Hui, Xiaoqing Ma, Zhou Su, Nan Cheng, Zhisheng Yin, Tom H Luan, and Ye Chen. Collaboration as a service: Digital-twin-enabled collaborative and distributed autonomous driving. *IEEE Internet of Things Journal*, 9(19):18607–18619, 2022.
- [92] How big data analytics is optimizing electric vehicles chargers? <https://soulpageit.com/big-data-analytics-for-optimizing-electric-vehicles-chargers/>. (Accessed on 12/16/2023).
- [93] Michela Longo, Federica Foiadelli, and Wahiba Yaïci. Electric vehicles integrated with renewable energy sources for sustainable mobility. *New trends in electrical vehicle powertrains*, 10:203–223, 2018.
- [94] Habiballah Rahimi-Eichi and Mo-Yuen Chow. Big-data framework for electric vehicle range estimation. In *IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society*, pages 5628–5634. IEEE, 2014.

- [95] Hua Cai, Xiaoping Jia, Anthony SF Chiu, Xiaojun Hu, and Ming Xu. Siting public electric vehicle charging stations in beijing using big-data informed travel patterns of the taxi fleet. *Transportation Research Part D: Transport and Environment*, 33:39–46, 2014.
- [96] R Sreedhar and K Karunanithi. Withdrawn: Design, simulation analysis of universal battery management system for ev applications, 2021.
- [97] Xiaolin Tang, Tong Jia, Xiaosong Hu, Yanjun Huang, Zhongwei Deng, and Huayan Pu. Naturalistic data-driven predictive energy management for plug-in hybrid electric vehicles. *IEEE Transactions on Transportation Electrification*, 7(2):497–508, 2020.
- [98] Wenchang Li, Tao Chen, Jinghua Guo, and Jin Wang. Adaptive car-following control of intelligent electric vehicles. In *2018 IEEE 4th International Conference on Control Science and Systems Engineering (ICCSSE)*, pages 86–89. IEEE, 2018.
- [99] Jinghua Guo, Yugong Luo, Chuan Hu, Chen Tao, and Keqiang Li. Robust combined lane keeping and direct yaw moment control for intelligent electric vehicles with time delay. *International Journal of Automotive Technology*, 20:289–296, 2019.
- [100] Wasim A Ali, KN Manasa, Malika Bendeche, Mohammed Fadhel Aljunaid, and P Sandhya. A review of current machine learning approaches for anomaly detection in network traffic. *Journal of Telecommunications and the Digital Economy*, 8(4):64–95, 2020.
- [101] George Loukas, Eirini Karapistoli, Emmanouil Panaousis, Panagiotis Sariannidis, Anatolij Bezemskij, and Tuan Vuong. A taxonomy and survey of cyber-physical intrusion detection approaches for vehicles. *Ad Hoc Networks*, 84:124–147, 2019.
- [102] Yi Zeng, Meikang Qiu, Dan Zhu, Zhihao Xue, Jian Xiong, and Meiqin Liu. Deepvcm: A deep learning based intrusion detection method in vanet. In *2019 IEEE 5th intl conference on big data security on cloud (BigDataSecurity), IEEE intl conference on high performance and smart computing.(HPSC) and IEEE intl conference on intelligent data and security (IDS)*, pages 288–293. IEEE, 2019.
- [103] Erfan A Shams, Ahmet Rizer, and Ali Hakan Ulusoy. Trust aware support vector machine intrusion detection and prevention system in vehicular ad hoc networks. *Computers & Security*, 78:245–254, 2018.
- [104] Junwei Liang, Maode Ma, Muhammad Sadiq, and Kai-Hau Yeung. A filter model for intrusion detection system in vehicle ad hoc networks: A hidden markov methodology. *Knowledge-Based Systems*, 163:611–623, 2019.

- [105] Hichem Sedjelmaci, Sidi Mohammed Senouci, and Mosa Ali Abu-Rgheff. An efficient and lightweight intrusion detection mechanism for service-oriented vehicular networks. *IEEE Internet of things journal*, 1(6):570–577, 2014.
- [106] Benjamin Fraser, Saba Al-Rubaye, Sohaib Aslam, and Antonios Tsourdos. Enhancing the security of unmanned aerial systems using digital-twin technology and intrusion detection. In *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*, pages 1–10. IEEE, 2021.
- [107] Chuanchao Gao, Heejong Park, and Arvind Easwaran. An anomaly detection framework for digital twin driven cyber-physical systems. In *Proceedings of the ACM/IEEE 12th International Conference on Cyber-Physical Systems*, pages 44–54, 2021.
- [108] Christian Gehrmann and Martin Gunnarsson. A digital twin based industrial automation and control system security architecture. *IEEE Transactions on Industrial Informatics*, 16(1):669–680, 2019.
- [109] Yan Xu, Yanming Sun, Xiaolong Liu, and Yonghua Zheng. A digital-twin-assisted fault diagnosis using deep transfer learning. *Ieee Access*, 7:19990–19999, 2019.
- [110] Ron Snijders, Paolo Pileggi, Jeroen Broekhuijsen, Jacques Verriet, Marco Wiering, and Koen Kok. Machine learning for digital twins to predict responsiveness of cyber-physical energy systems. In *2020 8th workshop on modeling and simulation of cyber-physical energy systems*, pages 1–6. IEEE, 2020.
- [111] Andrea Castellani, Sebastian Schmitt, and Stefano Squartini. Real-world anomaly detection by using digital twin systems and weakly supervised learning. *IEEE Transactions on Industrial Informatics*, 17(7):4733–4742, 2020.
- [112] Zhen Tu, Liang Qiao, Robert Nowak, Haibin Lv, and Zhihan Lv. Digital twins-based automated pilot for energy-efficiency assessment of intelligent transportation infrastructure. *IEEE Transactions on Intelligent Transportation Systems*, 23(11):22320–22330, 2022.
- [113] Rodolfo I Meneguette, R De Grande, and AA Loureiro. Intelligent transport system in smart cities. *Cham: Springer International Publishing*, 2018.
- [114] Ayyoub Lamssaggad, Nabil Benamar, Abdelhakim Senhaji Hafid, and Mounira Msahli. A survey on the current security landscape of intelligent transportation systems. *IEEE Access*, 9:9180–9208, 2021.

- [115] Julie Harvey and Sathish Kumar. A survey of intelligent transportation systems security: challenges and solutions. In *2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing,(HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, pages 263–268. IEEE, 2020.
- [116] Karen Scarfone and Peter Mell. Intrusion detection and prevention systems. In *Handbook of Information and Communication Security*, pages 177–192. Springer, 2010.
- [117] Asmaa Shaker Ashoor and Sharad Gore. Importance of intrusion detection system (ids). *International Journal of Scientific and Engineering Research*, 2(1):1–4, 2011.
- [118] Marek Pawlicki, Michał Choraś, and Rafał Kozik. Defending network intrusion detection systems against adversarial evasion attacks. *Future Generation Computer Systems*, 110:148–154, 2020.
- [119] I-Hsien Liu, Cheng-Hsiang Lo, Ta-Che Liu, Jung-Shian Li, Chuan-Gang Liu, and Chu-Fen Li. Ids malicious flow classification. *J. Robotics Netw. Artif. Life*, 7(2):103–106, 2020.
- [120] James P Anderson. Computer security threat monitoring and surveillance. *Technical Report, James P. Anderson Company*, 1980.
- [121] Dhruva Kumar Bhattacharyya and Jugal Kumar Kalita. *Network anomaly detection: A machine learning perspective*. Crc Press, 2013.
- [122] EL MOSTAPHA CHAKIR, Mohamed Moughit, and Youness Idrissi Khamlichi. An effective intrusion detection model based on svm with feature selection and parameters optimization. *Journal of Theoretical & Applied Information Technology*, 96(12), 2018.
- [123] Yi Yi Aung and Myat Myat Min. An analysis of k-means algorithm based network intrusion detection system. *Advances in Science, Technology and Engineering Systems Journal*, 3(1):496–501, 2018.
- [124] Shruti Kapil and Meenu Chawla. Performance evaluation of k-means clustering algorithm with various distance metrics. In *2016 IEEE 1st international conference on power electronics, intelligent control and energy systems (ICPEICES)*, pages 1–4. IEEE, 2016.
- [125] Meriem Kherbache, David Espes, and Kamal Amroun. An enhanced approach of the k-means clustering for anomaly-based intrusion detection systems. In *2021 International Conference on Computing, Computational Modelling and Applications (ICCMA)*, pages 78–83. IEEE, 2021.

- [126] H Kamal Idrissi, Z Kartit, A Kartit, and M El Marraki. Ckmsa: an anomaly detection process based on k-means and simulated annealing algorithms. *International Review on Computers and Software (IRE-COS)*, 11(1):42–48, 2016.
- [127] Elham Besharati, Marjan Naderan, and Ehsan Namjoo. Lr-hids: logistic regression host-based intrusion detection system for cloud environments. *Journal of Ambient Intelligence and Humanized Computing*, 10:3669–3692, 2019.
- [128] Reehan Ali Shah, Yuntao Qian, Dileep Kumar, Munwar Ali, and Muhammad Bux Alvi. Network intrusion detection through discriminative feature selection by using sparse logistic regression. *Future Internet*, 9(4):81, 2017.
- [129] Souhail Meftah, Tajjeeddine Rachidi, and Nasser Assem. Network based intrusion detection using the unsw-nb15 dataset. *International Journal of Computing and Digital Systems*, 8(5):478–487, 2019.
- [130] BS Sharmila and Rohini Nagapadma. Intrusion detection system using naive bayes algorithm. In *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, pages 1–4. IEEE, 2019.
- [131] Amjad Mehmood, Mithun Mukherjee, Syed Hassan Ahmed, Houbing Song, and Khalid Mahmood Malik. Nbc-maids: Naïve bayesian classification technique in multi-agent system-enriched ids for securing iot against ddos attacks. *The Journal of Supercomputing*, 74:5156–5170, 2018.
- [132] Kehe Wu, Zuge Chen, and Wei Li. A novel intrusion detection model for a massive network using convolutional neural networks. *Ieee Access*, 6:50850–50859, 2018.
- [133] Ahmad Shokoohsaljooghi and Hamid Mirvaziri. Performance improvement of intrusion detection system using neural networks and particle swarm optimization algorithms. *International Journal of Information Technology*, 12:849–860, 2020.
- [134] Minh Tuan Nguyen and Kiseon Kim. Genetic convolutional neural network for intrusion detection systems. *Future Generation Computer Systems*, 113:418–427, 2020.
- [135] Randy Paffenroth, Kathleen Kay, and Les Servi. Robust pca for anomaly detection in cyber networks. *arXiv preprint arXiv:1801.01571*, 2018.
- [136] Zhida Li, Ana Laura Gonzalez Rios, Guangyu Xu, and Ljiljana Trajković. Machine learning techniques for classifying network anomalies and intrusions. In *2019 IEEE international symposium on circuits and systems (ISCAS)*, pages 1–5. IEEE, 2019.

- [137] Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali A Ghorbani. A detailed analysis of the kdd cup 99 data set. In *2009 IEEE symposium on computational intelligence for security and defense applications*, pages 1–6. Ieee, 2009.
- [138] John McHugh. Testing intrusion detection systems: a critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by lincoln laboratory. *ACM Transactions on Information and System Security (TISSEC)*, 3(4):262–294, 2000.
- [139] Ankit Thakkar and Ritika Lohiya. A review of the advancement in intrusion detection datasets. *Procedia Computer Science*, 167:636–645, 2020.
- [140] Karin Mascher, Stefan Laller, and Philipp Berglez. Hybrid autoencoder for interference detection in raw gnss observations. In *Proceedings of the 36th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2023)*, pages 3744–3758, 2023.
- [141] Hanan Hindy, David Brosset, Ethan Bayne, Amar Kumar Seeam, Christos Tachtatzis, Robert Atkinson, and Xavier Bellekens. A taxonomy of network threats and the effect of current datasets on intrusion detection systems. *IEEE Access*, 8:104650–104675, 2020.
- [142] Ansam Khraisat, Iqbal Gondal, Peter Vamplew, and Joarder Kamruzzaman. Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, 2(1):1–22, 2019.
- [143] Terry Brugger. Kdd cup’99 dataset (network intrusion) considered harmful. *KDnuggets newsletter*, 7(18):15, 2007.
- [144] Nour Moustafa and Jill Slay. Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set). In *2015 military communications and information systems conference (MilCIS)*, pages 1–6. IEEE, 2015.
- [145] Ali Shiravi, Hadi Shiravi, Mahbod Tavallaee, and Ali A Ghorbani. Toward developing a systematic approach to generate benchmark datasets for intrusion detection. *computers & security*, 31(3):357–374, 2012.
- [146] Iman Sharafaldin, Arash Habibi Lashkari, Saqib Hakak, and Ali A Ghorbani. Developing realistic distributed denial of service (ddos) attack dataset and taxonomy. In *2019 International Carnahan Conference on Security Technology (ICCST)*, pages 1–8. IEEE, 2019.
- [147] Faisal Hussain, Syed Ghazanfar Abbas, Muhammad Husnain, Ubaid U Fayyaz, Farrukh Shahzad, and Ghalib A Shah. Iot dos and ddos attack detection using resnet. In *2020 IEEE 23rd International Multitopic Conference (INMIC)*, pages 1–6. IEEE, 2020.

- [148] Iman Sharafaldin, Arash Habibi Lashkari, and Ali A Ghorbani. Intrusion detection evaluation dataset (cic-ids2017). *Proceedings of the of Canadian Institute for Cybersecurity*, 2018.
- [149] Ids 2018 | datasets | research | canadian institute for cybersecurity | unb. <https://www.unb.ca/cic/datasets/ids-2018.html>.
- [150] Jiyeon Kim, Yulim Shin, Eunjung Choi, et al. An intrusion detection model based on a convolutional neural network. *Journal of Multimedia Information System*, 6(4):165–172, 2019.
- [151] Iman Sharafaldin, Amirhossein Gharib, Arash Habibi Lashkari, and Ali A Ghorbani. Towards a reliable intrusion detection benchmark dataset. *Software Networking*, 2017(1):177–200, 2017.
- [152] Robert Koch, Mario Golling, and Gabi Dreo Rodosek. Towards comparability of intrusion detection systems: New data sets. In *TERENA Networking Conference*, volume 7, 2014.
- [153] Iman Sharafaldin, Arash Habibi Lashkari, and Ali A Ghorbani. Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSp*, 1:108–116, 2018.
- [154] R Vijayanand, D Devaraj, and B Kannapiran. Intrusion detection system for wireless mesh network using multiple support vector machine classifiers with genetic-algorithm-based feature selection. *Computers & Security*, 77:304–314, 2018.
- [155] Safi Ullah, Muazzam A Khan, Jawad Ahmad, Sajjad Shaukat Jamal, Zil e Huma, Muhammad Tahir Hassan, Nikolaos Pitropakis, Arshad, and William J Buchanan. Hdl-ids: a hybrid deep learning architecture for intrusion detection in the internet of vehicles. *Sensors*, 22(4):1340, 2022.
- [156] Sugandh Seth, Kuljit Kaur Chahal, and Gurvinder Singh. A novel ensemble framework for an intelligent intrusion detection system. *IEEE Access*, 9:138451–138467, 2021.
- [157] M Shahid Anwer and Chris Guy. A survey of vanet technologies. *Journal of Emerging Trends in Computing and Information Sciences*, 5(9):661–671, 2014.
- [158] Ns3 vanet projects - ns3 simulator. <https://ns3-code.com/ns3-vanet-projects/>. (Accessed on 01/22/2024).
- [159] Kakan Chandra Dey, Anjan Rayamajhi, Mashrur Chowdhury, Parth Bhavsar, and James Martin. Vehicle-to-vehicle (v2v) and vehicle-to-infrastructure (v2i) communication in a heterogeneous wireless network—performance evaluation. *Transportation Research Part C: Emerging Technologies*, 68:168–184, 2016.

- [160] Mustafa Maad Hamdi, Lukman Audah, Sami Abduljabbar Rashid, Alaa Hamid Mohammed, Sameer Alani, and Ahmed Shamil Mustafa. A review of applications, characteristics and challenges in vehicular ad hoc networks (vanets). In *2020 international congress on human-computer interaction, optimization and robotic applications (HORA)*, pages 1–7. IEEE, 2020.
- [161] Muhammad Sameer Sheikh and Jun Liang. A comprehensive survey on vanet security services in traffic management system. *Wireless Communications and Mobile Computing*, 2019:1–23, 2019.
- [162] Z Li, Z Wang, and C Chigan. Security of vehicular ad hoc networks in intelligent transportation systems. *Wireless Technologies for Intelligent Transportation Systems*, pages 133–74, 2009.
- [163] David Chaum and Eugène Van Heyst. Group signatures. In *Advances in Cryptology—EUROCRYPT’91: Workshop on the Theory and Application of Cryptographic Techniques Brighton, UK, April 8–11, 1991 Proceedings 10*, pages 257–265. Springer, 1991.
- [164] Ronald L Rivest, Adi Shamir, and Yael Tauman. How to leak a secret. In *Advances in Cryptology—ASIACRYPT 2001: 7th International Conference on the Theory and Application of Cryptology and Information Security Gold Coast, Australia, December 9–13, 2001 Proceedings 7*, pages 552–565. Springer, 2001.
- [165] Kewei Sha, Yong Xi, Weisong Shi, Loren Schwiebert, and Tao Zhang. Adaptive privacy-preserving authentication in vehicular networks. In *2006 First International Conference on Communications and Networking in China*, pages 1–8. IEEE, 2006.
- [166] John R Douceur. The sybil attack. In *International workshop on peer-to-peer systems*, pages 251–260. Springer, 2002.
- [167] Gongjun Yan Yan, Gyanesh Choudhary, Michele C Weigle, and Stephan Olariu. Providing vanet security through active position detection. In *Proceedings of the fourth ACM international workshop on Vehicular ad hoc networks*, pages 73–74, 2007.
- [168] Frank Kargl, Elmar Schoch, Björn Wiedersheim, and Tim Leinmüller. Secure and efficient beaconing for vehicular networks. In *Proceedings of the fifth ACM international workshop on Vehicular Inter-NETworking*, pages 82–83, 2008.
- [169] Nikodin Ristanovic, Panos Papadimitratos, George Theodorakopoulos, Jean-Pierre Hubaux, and Jean-Yves Leboudec. Adaptive message authentication for vehicular networks. In *Proceedings of the sixth ACM international workshop on Vehicular InterNETworking*, pages 121–122, 2009.



- [170] S Kanthimathi and P Jhansi Rani. An efficient packet dropping attack detection mechanism in wireless ad-hoc networks using ecc based aodv-aco protocol. *Wireless Networks*, pages 1–13, 2022.
- [171] Jemal Abawajy, Guojun Wang, Laurence T Yang, and Bahman Javadi. Trust, security and privacy in emerging distributed systems, 2016.
- [172] Marcela Mejia, Nestor Pena, Jose L Munoz, and Oscar Esparza. A review of trust modeling in ad hoc networks. *Internet Research*, 19(1):88–104, 2009.
- [173] Imran Memon, Riaz Ahmed Shaikh, and Hidayatullah Shaikh. Dynamic pseudonyms trust-based model to protect attack scenario for internet of vehicle ad-hoc networks. *Multimedia Tools and Applications*, pages 1–32, 2023.
- [174] Sohan Gyawali, Yi Qian, and Rose Qingyang Hu. A privacy-preserving misbehavior detection system in vehicular communication networks. *IEEE transactions on Vehicular Technology*, 70(6):6147–6158, 2021.
- [175] Shrikant Tangade, Sunilkumar S Manvi, and Pascal Lorenz. Decentralized and scalable privacy-preserving authentication scheme in vanets. *IEEE Transactions on Vehicular Technology*, 67(9):8647–8655, 2018.
- [176] Dakshnamoorthy Manivannan, Shafika Showkat Moni, and Sherali Zeadally. Secure authentication and privacy-preserving techniques in vehicular ad-hoc networks (vanets). *Vehicular Communications*, 25:100247, 2020.
- [177] Saad Ali Alfadhli, Songfeng Lu, Kai Chen, and Meriem Sebai. Mfspv: A multi-factor secured and lightweight privacy-preserving authentication scheme for vanets. *IEEE Access*, 8:142858–142874, 2020.
- [178] Chaker Abdelaziz Kerrache, Carlos T Calafate, Juan-Carlos Cano, Nasreddine Lagraa, and Pietro Manzoni. Trust management for vehicular networks: An adversary-oriented overview. *IEEE Access*, 4:9293–9307, 2016.
- [179] Robert E Shannon. Introduction to the art and science of simulation. In *1998 winter simulation conference. proceedings (cat. no. 98ch36274)*, volume 1, pages 7–14. IEEE, 1998.
- [180] Sara Imene Boucetta, Youcef Guichi, and Zsolt Csaba Johanyák. Review of mobility scenarios generators for vehicular ad-hoc networks simulators. In *Journal of Physics: Conference Series*, volume 1935, page 012006. IOP Publishing, 2021.
- [181] Michael Behrisch, Laura Bieker, Jakob Erdmann, and Daniel Krajzewicz. Sumo—simulation of urban mobility: an overview. In *Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation*. ThinkMind, 2011.

- [182] Axel Wegener, Michał Piórkowski, Maxim Raya, Horst Hellbrück, Stefan Fischer, and Jean-Pierre Hubaux. Traci: an interface for coupling road traffic and network simulators. In *Proceedings of the 11th communications and networking simulation symposium*, pages 155–163, 2008.
- [183] What is omnet++? <https://omnetpp.org/intro/>. (Accessed on 01/02/2024).
- [184] Veins. <https://veins.car2x.org/>. (Accessed on 01/02/2024).
- [185] Christoph Sommer, David Eckhoff, Alexander Brummer, Dominik S Buse, Florian Hagenauer, Stefan Joerer, and Michele Segata. Veins: The open source vehicular network simulation framework. *Recent Advances in Network Simulation: The OMNeT++ Environment and its Ecosystem*, pages 215–252, 2019.
- [186] Maria Pia Fanti, Agostino Marcello Mangini, Michele Roccotelli, Massimiliano Nolich, and Walter Ukovich. Modeling virtual sensors for electric vehicles charge services. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3853–3858. IEEE, 2018.
- [187] Ángel Felipe, M Teresa Ortuño, Giovanni Righini, and Gregorio Tirado. A heuristic approach for the green vehicle routing problem with multiple technologies and partial recharges. *Transportation Research Part E: Logistics and Transportation Review*, 71:111–128, 2014.
- [188] Jane Lin, Wei Zhou, and Ouri Wolfson. Electric vehicle routing problem. *Transportation research procedia*, 12:508–521, 2016.
- [189] J Paz, Mauricio Granada-Echeverri, and J Escobar. The multi-depot electric vehicle location routing problem with time windows. *International journal of industrial engineering computations*, 9(1):123–136, 2018.
- [190] A Hecker and R Wies. Charging infrastructure for evs in beijing: A spatial analysis from real customer data at two districts. In *2015 International Conference on Connected Vehicles and Expo (ICCVE)*, pages 336–341. IEEE, 2015.
- [191] Ilker Kucukoglu, Reginald Dewil, and Dirk Cattrysse. The electric vehicle routing problem and its variations: A literature review. *Computers & Industrial Engineering*, 161:107650, 2021.
- [192] Omid Sadeghian, Arman Oshnoei, Behnam Mohammadi-Ivatloo, Vahid Vahidinasab, and Amjad Anvari-Moghaddam. A comprehensive review on electric vehicles smart charging: Solutions, strategies, technologies, and challenges. *Journal of Energy Storage*, 54:105241, 2022.

- [193] Merve Keskin, Bülent Çatay, and Gilbert Laporte. A simulation-based heuristic for the electric vehicle routing problem with time windows and stochastic waiting times at recharging stations. *Computers & Operations Research*, 125:105060, 2021.
- [194] Oussama Sbayti and Khalid Housni. A new routing method based on ant colony optimization in vehicular ad-hoc network. *Statistics, Optimization & Information Computing*, 12(1):167–181, 2024.
- [195] Niklas Åkerblom, Yuxin Chen, and Morteza Haghiri Chehreghani. Online learning of energy consumption for navigation of electric vehicles. *Artificial Intelligence*, 317:103879, 2023.
- [196] Alejandro Montoya, Christelle Guéret, Jorge E Mendoza, and Juan G Villegas. The electric vehicle routing problem with nonlinear charging function. *Transportation Research Part B: Methodological*, 103:87–110, 2017.
- [197] Merve Keskin and Bülent Çatay. Partial recharge strategies for the electric vehicle routing problem with time windows. *Transportation research part C: emerging technologies*, 65:111–127, 2016.
- [198] Samuel Pelletier, Ola Jabali, and Gilbert Laporte. 50th anniversary invited article—goods distribution with electric vehicles: review and research perspectives. *Transportation science*, 50(1):3–22, 2016.
- [199] Chungmok Lee. An exact algorithm for the electric-vehicle routing problem with nonlinear charging time. *Journal of the Operational Research Society*, 72(7):1461–1485, 2021.
- [200] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. Recent development and applications of sumo-simulation of urban mobility. *International journal on advances in systems and measurements*, 5(3&4), 2012.
- [201] Mingfeng Shang and Raphael E Stern. Impacts of commercially available adaptive cruise control vehicles on highway stability and throughput. *Transportation research part C: emerging technologies*, 122:102897, 2021.
- [202] David ZW Wang, Haoxiang Liu, and Wai Y Szeto. A novel discrete network design problem formulation and its global optimization solution algorithm. *Transportation Research Part E: Logistics and Transportation Review*, 79:213–230, 2015.
- [203] Gerhard Hiermann, Jakob Puchinger, Stefan Ropke, and Richard F Hartl. The electric fleet size and mix vehicle routing problem with time windows and recharging stations. *European Journal of Operational Research*, 252(3):995–1018, 2016.

- [204] Maximilian Schiffer and Grit Walther. The electric location routing problem with time windows and partial recharging. *European journal of operational research*, 260(3):995–1013, 2017.
- [205] Dominik Goeke and Michael Schneider. Routing a mixed fleet of electric and conventional vehicles. *European Journal of Operational Research*, 245(1):81–99, 2015.
- [206] María A del Cacho Estil-les, Maria Pia Fanti, Agostino M Mangini, and Michele Roccotelli. Electric vehicles routing including smart-charging method and energy constraints. In *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)*, pages 1735–1740. IEEE, 2022.
- [207] Hashmatullah Sadid, Moeid Qurashi, and Constantinos Antoniou. Simulation-based optimization of autonomous driving behaviors. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, pages 4101–4108. IEEE, 2022.
- [208] Yaofeng Song, Han Zhao, Ruikang Luo, Liping Huang, Yicheng Zhang, and Rong Su. A sumo framework for deep reinforcement learning experiments solving electric vehicle charging dispatching problem. *arXiv preprint arXiv:2209.02921*, 2022.
- [209] Lara Codecá, Raphaël Frank, Sébastien Faye, and Thomas Engel. Luxembourg sumo traffic (lust) scenario: Traffic demand evaluation. *IEEE Intelligent Transportation Systems Magazine*, 9(2):52–63, 2017.
- [210] Matt Franchi, Rebecca Kahn, Mashrur Chowdhury, Sakib Khan, Ken Kennedy, Linh Ngo, and Amy Apon. Webots. hpc: A parallel simulation pipeline for autonomous vehicles. In *Practice and Experience in Advanced Research Computing*, pages 1–4. 2022.
- [211] Tania Iram, Jawwad Shamsi, Usama Alvi, Saif ur Rahman, and Muhammad Maaz. Controlling smart-city traffic using machine learning. In *2019 International Conference on Frontiers of Information Technology (FIT)*, pages 203–2035. IEEE, 2019.
- [212] Anum Mushtaq, Irfan Ul Haq, Muhammad Azeem Sarwar, Asifullah Khan, Wajeeha Khalil, and Muhammad Abid Mughal. Multi-agent reinforcement learning for traffic flow management of autonomous vehicles. *Sensors*, 23(5):2373, 2023.
- [213] Maytheewat Aramrattana, Tony Larsson, Jonas Jansson, and Arne Nåbo. A simulation framework for cooperative intelligent transport systems testing and evaluation. *Transportation research part F: traffic psychology and behaviour*, 61:268–280, 2019.

- [214] Ruyang Yin, Zhiyuan Liu, and Nan Zheng. A simulation-based model for continuous network design problem using bayesian optimization. *IEEE Transactions on Intelligent Transportation Systems*, 23(11):20352–20367, 2022.
- [215] Muzamil Eltejani Mohammed Ali, Akif Durdu, Seyit Alperen Çeltek, and Alper Yilmaz. An adaptive method for traffic signal control based on fuzzy logic with webster and modified webster formula using sumo traffic simulator. *IEEE Access*, 9:102985–102997, 2021.
- [216] Yue Yang, Si-Yuan Hao, and Ha-Bin Cai. Comparison and evaluation of routing protocols based on a collaborative simulation using sumo and ns3 with traci. In *2016 International Conference on Information System and Artificial Intelligence (ISAI)*, pages 253–257. IEEE, 2016.
- [217] Krešimir Kušić, Rene Schumann, and Edouard Ivanjko. Building a motorway digital twin in sumo: real-time simulation of continuous data stream from traffic counters. In *2022 International Symposium ELMAR*, pages 71–76. IEEE, 2022.
- [218] Seval Ene, İlker Küçükoğlu, Aslı Aksoy, and Nursel Öztürk. A hybrid metaheuristic algorithm for the green vehicle routing problem with a heterogeneous fleet. *International Journal of Vehicle Design*, 71(1-4):75–102, 2016.
- [219] Gilbert Laporte, Francois Louveaux, and Hélène Mercure. The vehicle routing problem with stochastic travel times. *Transportation science*, 26(3):161–170, 1992.
- [220] Mir Ehsan Hesam Sadati and Bülent Çatay. A hybrid variable neighborhood search approach for the multi-depot green vehicle routing problem. *Transportation Research Part E: Logistics and Transportation Review*, 149:102293, 2021.
- [221] Maurizio Bruglieri, Ferdinando Pezzella, Ornella Pisacane, and Stefano Suraci. A variable neighborhood search branching for the electric vehicle routing problem with time windows. *Electronic Notes in Discrete Mathematics*, 47:221–228, 2015.
- [222] Chen Ma. Smart city and cyber-security; technologies used, leading challenges and future recommendations. *Energy Reports*, 7:7999–8012, 2021.
- [223] Omar Y Al-Jarrah, Carsten Maple, Mehrdad Dianati, David Oxtoby, and Alex Mouzakitis. Intrusion detection systems for intra-vehicle networks: A review. *IEEE Access*, 7:21266–21289, 2019.

- [224] Elias C Eze, Sijing Zhang, and Enjie Liu. Vehicular ad hoc networks (vanets): Current state, challenges, potentials and way forward. In *2014 20th international conference on automation and computing*, pages 176–181. IEEE, 2014.
- [225] MohammadNoor Injadat, Abdallah Moubayed, Ali Bou Nassif, and Abdallah Shami. Machine learning towards intelligent systems: applications, challenges, and opportunities. *Artificial Intelligence Review*, 54:3299–3348, 2021.
- [226] Zeinab El-Rewini, Karthikeyan Sadatsharan, Daisy Flora Selvaraj, Siby Jose Plathottam, and Prakash Ranganathan. Cybersecurity challenges in vehicular communications. *Vehicular Communications*, 23:100214, 2020.
- [227] Mahmood A Al-Shareeda, Mohammed Anbar, Selvakumar Manickam, and Iznan H Hasbullah. Se-cppa: A secure and efficient conditional privacy-preserving authentication scheme in vehicular ad-hoc networks. *Sensors*, 21(24):8206, 2021.
- [228] Hristos Giannopoulos, Alexander M Wyglinski, and Joseph Chapman. Securing vehicular controller area networks: An approach to active bus-level countermeasures. *IEEE Vehicular Technology Magazine*, 12(4):60–68, 2017.
- [229] Wasim A Ali, Maria Pia Fanti, Michele Roccotelli, and Luigi Ranieri. A review of digital twin technology for electric and autonomous vehicles. *Applied Sciences*, 13(10):5871, 2023.
- [230] Hind Bangui, Mouzhi Ge, and Barbora Buhnova. A hybrid machine learning model for intrusion detection in vanet. *Computing*, 104(3):503–531, 2022.
- [231] Hamideh Baharlouei, Adetokunbo Makanju, and Nur Zincir-Heywood. Exploring realistic vanet simulations for anomaly detection of ddos attacks. In *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)*, pages 1–7. IEEE, 2022.
- [232] Ulf E Larson, Dennis K Nilsson, and Erland Jonsson. An approach to specification-based attack detection for in-vehicle networks. In *2008 IEEE Intelligent Vehicles Symposium*, pages 220–225. IEEE, 2008.
- [233] Donghao Zhou, Zheng Yan, Yulong Fu, and Zhen Yao. A survey on network data collection. *Journal of Network and Computer Applications*, 116:9–23, 2018.
- [234] Wenliang Fu, Xin Xin, Ping Guo, and Zhou Zhou. A practical intrusion detection system for internet of vehicles. *China Communications*, 13(10):263–275, 2016.

- [235] Sparsh Sharma and Ajay Kaul. A survey on intrusion detection systems and honeypot based proactive security mechanisms in vanets and vanet cloud. *Vehicular communications*, 12:138–164, 2018.
- [236] K Indira and E Christal Joy. Energy efficient ids for cluster-based vanets. *Asian Journal of Information Technology*, 14(1):37–41, 2015.
- [237] John B Kenney. Dedicated short-range communications (dsrc) standards in the united states. *Proceedings of the IEEE*, 99(7):1162–1182, 2011.
- [238] Badiea Abdulkarem Mohammed, Mahmood A Al-Shareeda, Selvakumar Manickam, Zeyad Ghaleb Al-Mekhlafi, Abdulrahman Alreshidi, Meshari Alazmi, Jalawi Sulaiman Alshudukhi, and Mohammad Alsaffar. Fc-pa: fog computing-based pseudonym authentication scheme in 5g-enabled vehicular networks. *IEEE Access*, 11:18571–18581, 2023.
- [239] Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.
- [240] Lei Yu and Huan Liu. Feature selection for high-dimensional data: A fast correlation-based filter solution. In *Proceedings of the 20th international conference on machine learning (ICML-03)*, pages 856–863, 2003.
- [241] Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, et al. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4, 2015.
- [242] Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely randomized trees. *Machine learning*, 63:3–42, 2006.
- [243] Iman Sharafaldin, Arash Habibi Lashkari, and Ali A Ghorbani. A detailed analysis of the cicids2017 data set. In *Information Systems Security and Privacy: 4th International Conference, ICISSP 2018, Funchal-Madeira, Portugal, January 22-24, 2018, Revised Selected Papers 4*, pages 172–188. Springer, 2019.
- [244] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25, 2012.
- [245] M Uma and Ganapathi Padmavathi. A survey on various cyber attacks and their classification. *Int. J. Netw. Secur.*, 15(5):390–396, 2013.
- [246] Mahmood A Al-Shareeda and Selvakumar Manickam. Msr-dos: Modular square root-based scheme to resist denial of service (dos) attacks in 5g-enabled vehicular networks. *IEEE Access*, 10:120606–120615, 2022.

- [247] Abdulaziz Alshammari, Mohamed A Zohdy, Debatosh Debnath, and George Corser. Classification approach for intrusion detection in vehicle systems. *Wireless Engineering and Technology*, 9(4):79–94, 2018.
- [248] Izhar Ahmed Khan, Nour Moustafa, Dechang Pi, Waqas Haider, Bentian Li, and Alireza Jolfaei. An enhanced multi-stage deep learning framework for detecting malicious activities from autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 23(12):25469–25478, 2021.
- [249] Javed Ashraf, Asim D Bakhshi, Nour Moustafa, Hasnat Khurshid, Abdullah Javed, and Amin Beheshti. Novel deep learning-enabled lstm autoencoder architecture for discovering anomalous events from intelligent transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, 22(7):4507–4518, 2020.
- [250] Habeeb Olufowobi, Uchenna Ezeobi, Eric Muhati, Gaylon Robinson, Clinton Young, Joseph Zambreno, and Gedare Bloom. Anomaly detection approach using adaptive cumulative sum algorithm for controller area network. In *Proceedings of the ACM Workshop on Automotive Cybersecurity*, pages 25–30, 2019.
- [251] Li Yang, Abdallah Moubayed, and Abdallah Shami. Mth-ids: A multitiered hybrid intrusion detection system for internet of vehicles. *IEEE Internet of Things Journal*, 9(1):616–632, 2021.
- [252] Parya Haji Mirzaee, Mohammad Shojafar, Hamidreza Bagheri, Tsz Hin Chan, Haitham Cruickshank, and Rahim Tafazolli. A two-layer collaborative vehicle-edge intrusion detection system for vehicular communications. In *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, pages 1–6. IEEE, 2021.
- [253] Kiran Aswal, Dinesh C Dobhal, and Heman Pathak. Comparative analysis of machine learning algorithms for identification of bot attack on the internet of vehicles (ioV). In *2020 International Conference on Inventive Computation Technologies (ICICT)*, pages 312–317. IEEE, 2020.
- [254] Arnaud Rosay, Florent Carlier, and Pascal Leroux. Feed-forward neural network for network intrusion detection. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pages 1–6. IEEE, 2020.
- [255] Mabrouka Gmiden, Mohamed Hedi Gmiden, and Hafedh Trabelsi. An intrusion detection method for securing in-vehicle can bus. In *2016 17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, pages 176–180. IEEE, 2016.
- [256] Jiajia Liu, Shubin Zhang, Wen Sun, and Yongpeng Shi. In-vehicle network attacks and countermeasures: Challenges and future directions. *IEEE Network*, 31(5):50–58, 2017.



- [257] Mohammed Erritali and Bouabid El Ouahidi. A review and classification of various vanet intrusion detection systems. *2013 National Security Days (JNS3)*, pages 1–6, 2013.
- [258] Quentin Covert, Dustin Steinhagen, Mary Francis, and Kevin Streff. Towards a triad for data privacy. 2020.
- [259] Kuan Zhang, Xiaohui Liang, Rongxing Lu, and Xuemin Shen. Sybil attacks and their defenses in the internet of things. *IEEE Internet of Things Journal*, 1(5):372–383, 2014.
- [260] Ram Shringar Raw, Manish Kumar, and Nanhay Singh. Security challenges, issues and their solutions for vanet. *International journal of network security & its applications*, 5(5):95, 2013.
- [261] Irshad Ahmed Sumra, Halabi Bin Hasbullah, and Jamalul-lail Bin AbManan. Attacks on security goals (confidentiality, integrity, availability) in vanet: a survey. In *Vehicular Ad-hoc Networks for Smart Cities: First International Workshop, 2014*, pages 51–61. Springer, 2014.
- [262] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- [263] Wasim A Ali, P Sandhya, Michele Roccotelli, and Maria Pia Fanti. A comparative study of current dataset used to evaluate intrusion detection system. *International Journal on Engineering Applications*, 10(5), 2022.
- [264] Zachariah Pelletier and Munther Abualkibash. Evaluating the cic ids-2017 dataset using machine learning methods and creating multiple predictive models in the statistical computing language r. *Science*, 5(2):187–191, 2020.
- [265] Nanyi Fei, Yizhao Gao, Zhiwu Lu, and Tao Xiang. Z-score normalization, hubness, and few-shot learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 142–151, 2021.
- [266] John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1):100–108, 1979.
- [267] Newsroom home. <https://www.intel.com/content/www/us/en/newsroom/home.html#gs.iwfv6a>. (Accessed on 01/17/2024).
- [268] James Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for hyper-parameter optimization. *Advances in neural information processing systems*, 24, 2011.

- [269] Ian Dewancker, Michael McCourt, and Scott Clark. Bayesian optimization for machine learning: A practical guidebook. *arXiv preprint arXiv:1612.04858*, 2016.
- [270] Cyril Goutte and Eric Gaussier. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *European conference on information retrieval*, pages 345–359. Springer, 2005.
- [271] Ids 2017 | datasets | research | canadian institute for cybersecurity | unb. <https://www.unb.ca/cic/datasets/ids-2017.html>. (Accessed on 01/17/2024).
- [272] Abir Mchergui, Tarek Moulahi, and Sherali Zeadally. Survey on artificial intelligence (ai) techniques for vehicular ad-hoc networks (vanets). *Vehicular Communications*, 34:100403, 2022.
- [273] Issam W Damaj, Dina K Serhal, Lama A Hamandi, Rached N Zantout, and Hussein T Mouftah. Connected and autonomous electric vehicles: Quality of experience survey and taxonomy. *Vehicular Communications*, 28:100312, 2021.
- [274] Alberto Marroquin, Marco Antonio To, Cesar A Azurdia-Meza, and Sandy Bolufé. A general overview of vehicle-to-x (v2x) beacon-based cooperative vehicular networks. In *2019 IEEE 39th Central America and Panama Convention (CONCAPAN XXXIX)*, pages 1–6. IEEE, 2019.
- [275] Razvan-Gabriel Lazar and Constantin-Florin Caruntu. Evaluation of channel congestion for v2v safety communication in multiplatooning applications. In *2023 International Symposium ELMAR*, pages 133–136. IEEE, 2023.
- [276] *IEEE Std 802.11p-2010 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, IEEE Std 802.11n-2009, and IEEE Std 802.11w-2009)*, pages 1–51, 2010.
- [277] Ieee standard for wireless access in vehicular environments (wave) – multi-channel operation. *IEEE Std 1609.4-2016 (Revision of IEEE Std 1609.4-2010)*, pages 1–94, 2016.
- [278] Ieee 802.11-2016 - 802.11-2016 - ieee standard for information technology—telecommunications and information exchange between systems local and metropolitan area networks—specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications | joinup.
- [279] David Eckhoff and Christoph Sommer. A multi-channel ieee 1609.4 and 802.11 p edca model for the veins framework. In *Proceedings of 5th ACM/ICST international conference on simulation tools and techniques for communications, networks and systems: 5th ACM/ICST international workshop on OMNet++ (Desenzano, Italy, 19-23 March, 2012)*. OMNeT, 2012.

- [280] Christian Cseh. Architecture of the dedicated short-range communications (dsrc) protocol. In *VTC'98. 48th IEEE Vehicular Technology Conference. Pathway to Global Wireless Revolution (Cat. No. 98CH36151)*, volume 3, pages 2095–2099. IEEE, 1998.
- [281] Le Ou-Yang, Hong Yan, and Xiao-Fei Zhang. A multi-network clustering method for detecting protein complexes from multiple heterogeneous networks. *BMC bioinformatics*, 18(13):23–34, 2017.
- [282] Felipe Cunha, Leandro Villas, Azzedine Boukerche, Guilherme Maia, Aline Viana, Raquel AF Mini, and Antonio AF Loureiro. Data communication in vanets: Protocols, applications and challenges. *Ad hoc networks*, 44:90–103, 2016.
- [283] Laura Bieker, Daniel Krajzewicz, Antonio Pio Morra, Carlo Michelacci, and Fabio Cartolano. Traffic simulation for all: a real world traffic scenario from the city of bologna. In *Modeling Mobility with Open Data: 2nd SUMO Conference 2014 Berlin, Germany, May 15-16, 2014*, pages 47–60. Springer, 2015.
- [284] András Varga and Rudolf Hornig. An overview of the omnet++ simulation environment. In *1st International ICST Conference on Simulation Tools and Techniques for Communications, Networks and Systems*, 2010.
- [285] Florian Hagenauer, Takamasa Higuchi, Onur Altintas, and Falko Dressler. Efficient data handling in vehicular micro clouds. *Ad Hoc Networks*, 91:101871, 2019.
- [286] Eduard Zadobrischi and Mihai Dimian. Vehicular communications utility in road safety applications: A step toward self-aware intelligent traffic systems. *Symmetry*, 13(3):438, 2021.
- [287] Megha V Kadam, Hemant B Mahajan, Nilesh J Uke, and Pravin R Futane. Cybersecurity threats mitigation in internet of vehicles communication system using reliable clustering and routing. *Microprocessors and Microsystems*, 102:104926, 2023.
- [288] Ekaterina S Stolyarova, Denis M Shiryaev, Andrei G Vladyko, and Mikhail V Buinevich. Vanet/its cybersecurity threats: Analysis, categorization and forecasting. In *2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIconRus)*, pages 136–141. IEEE, 2018.
- [289] Farhan Ahmad, Asma Adnane, and Virginia NL Franqueira. A systematic approach for cyber security in vehicular networks. *Journal of Computer and Communications*, 4(16):38–62, 2016.

- [290] Shi Dong, Huadong Su, Yuanjun Xia, Fei Zhu, Xinrong Hu, and Bangchao Wang. A comprehensive survey on authentication and attack detection schemes that threaten it in vehicular ad-hoc networks. *IEEE Transactions on Intelligent Transportation Systems*, 2023.
- [291] Muhammad Usama, Ubaid Ullah, and Ahthasham Sajid. Cyber attacks against intelligent transportation systems. In *Cyber Security for Next-Generation Computing Technologies*, pages 190–230. CRC Press, 2024.
- [292] Safras Iqbal, Peter Ball, Muhammad H Kamarudin, and Andrew Bradley. Simulating malicious attacks on vanets for connected and autonomous vehicle cybersecurity: a machine learning dataset. In *2022 13th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, pages 332–337. IEEE, 2022.
- [293] Erasmo Notaro. *Simulating Malicious Attacks on VANETs for Connected and Autonomous Vehicles*. PhD thesis, Politecnico di Torino, 2023.
- [294] Ghayth ALMahadin, Yassine Aoudni, Mohammad Shabaz, Anurag Vijay Agrawal, Ghazaala Yasmin, Esraa Saleh Alomari, Hamza Mohammed Ridha Al-Khafaji, Debabrata Dansana, and Renato R Maaliw. Vanet network traffic anomaly detection using gru-based deep learning model. *IEEE Transactions on Consumer Electronics*, 2023.
- [295] Julia Silva Weber, Tiago Ferreto, and Nur Zincir-Heywood. Exploring anomaly detection techniques for enhancing vanet availability. In *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, pages 1–7. IEEE, 2023.
- [296] Pavan Kumar Pandey, Vineet Kansal, and Abhishek Swaroop. Security challenges and solutions for next-generation vanets: An exploratory study. In *Role of Data-Intensive Distributed Computing Systems in Designing Data Solutions*, pages 183–201. Springer, 2023.
- [297] Elias C Eze, Si-Jing Zhang, En-Jie Liu, and Joy C Eze. Advances in vehicular ad-hoc networks (vanets): Challenges and road-map for future development. *International Journal of Automation and Computing*, 13:1–18, 2016.
- [298] Gunasekaran Raja, Sudha Anbalagan, Geetha Vijayaraghavan, Sudhakar Theerthagiri, Saran Vaitangarukav Suryanarayan, and Xin-Wen Wu. Sp-cids: Secure and private collaborative ids for vanets. *IEEE Transactions on Intelligent Transportation Systems*, 22(7):4385–4393, 2020.
- [299] Junwei Liang, Jianyong Chen, Yingying Zhu, and Richard Yu. A novel intrusion detection system for vehicular ad hoc networks (vanets) based on differences of traffic flow and position. *Applied Soft Computing*, 75:712–727, 2019.

- 
- [300] Fabio Arena, Giovanni Pau, and Alessandro Severino. A review on iee 802.11 p for intelligent transportation systems. *Journal of Sensor and Actuator Networks*, 9(2):22, 2020.
- [301] Ruobing Jiang and Yanmin Zhu. *Wireless Access in Vehicular Environment*, pages 1–5. Springer International Publishing, Cham, 2019.
- [302] Federico Poli. *Vehicular communications: from DSRC to Cellular V2X*. PhD thesis, Politecnico di Torino, 2018.
- [303] Damian Vertal, Ivan Baronak, and Jiri Hosek. Options to broadcast information in vanet. *Advances in Electrical and Electronic Engineering*, 20(2):185–192, 2022.