

**A SURVEY: CYBER SECURITY IN CONNECTED AND AUTOMATED
VEHICLES**

(BAĞLANTILI VE OTONOM ARAÇLARDA SİBER GÜVENLİK ÇALIŞMASI)

by

Ayşegül KANDEFER, B.S.

Thesis

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

in

COMPUTER ENGINEERING

in the

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

of

GALATASARAY UNIVERSITY

Supervisor: Prof. Dr. Tankut ACARMAN

November 2024

ACKNOWLEDGEMENTS

I would like to sincerely thank Prof. Dr. Tankut Acarman for his guidance and support throughout my thesis journey. His expertise and encouragement have been invaluable, and I'm truly grateful for his mentorship. I also want to express my heartfelt thanks to my family for standing by me every step of the way. Their support, patience, and belief in me have been my biggest source of motivation and strength.

November 2024

Ayşegül KANDEFER

TABLE OF CONTENTS

LIST OF SYMBOLS	vii
LIST OF FIGURES	viii
LIST OF TABLES	ix
ABSTRACT.....	xi
ÖZET	xii
1. INTRODUCTION	1
1.1 Automated Vehicles	4
1.1.1 Benefits and Usage Areas of Automated Vehicles	4
1.1.2 Automated Vehicles Levels.....	5
1.2 Automated Vehicle Architecture	7
1.3 Main Components of Automated Vehicles	7
1.3.1 Automotive Control Systems	7
1.3.2 Automated Driving System Components	8
1.3.3 V2X Communication Technologies	9
1.4 Automated and Connected Vehicle Regulations and Standards	9
1.4.1. R155 Cybersecurity	10
1.4.2. R156 Software Update	11
1.4.3. R157 Automated Lane Keeping System	12
1.4.4. Cybersecurity Standards.....	12
1.5 In-Vehicle Network Protocols: CAN, Ethernet, and FlexRay	13
1.5.1. Controller Area Network (CAN).....	13
1.5.2. Ethernet	14
1.5.3. FlexRay	14
2. CONNECTED & AUTOMATED VEHICLE’S CYBERSECURITY REVIEW	16
2.1 Research Framework and Methodology	16
2.1.1. Research Questions	16

2.1.2. Search Process	16
2.1.3. Cybersecurity Keywords and Trends in Connected and Autonomous Vehicles	17
2.1.4. Datasets and Their Applications in Vehicle Network Security Research ...	22
2.2 State-of-the-Art Cybersecurity in Connected and Automated Technology	25
2.2.1. Vector, Kaspersky & Symantec: Services and Contributions to Cybersecurity	28
2.2.2. Emerging Technologies on Automotive Cyber-Security	35
2.2.3. Cyber Security Risks in Autonomous Vehicles	39
2.2.4. Cyber Security Severity Classes in Autonomous Vehicles	40
2.2.5. Cyber Attack Detection Tools	41
3. LITERATURE REVIEW	46
3.1 Attacks on Driving System Components	46
3.1.1 Manipulation-Based Attacks	47
3.1.2 Identity-Based Attacks	55
3.1.3 Service-Based Attacks	58
3.1.4 Software-Based Attacks	63
3.2 Defense of Connected and Automated Vehicles	71
3.2.1 Defense Frameworks for Automated Vehicles	72
3.2.2 Intrusion Detection System for CAN	74
3.3 HCRL Car Hacking & Survival Analysis Dataset Referenced Papers' Review .	74
4. CASE STUDY OF HCLR CAR HACKING DATASET	149
4.1 Dataset Instruction	149
4.2 Dataset Definition	149
4.3 Dataset Analysis by Vehicles	151
4.4 Machine Information.....	160
4.5 Application of machine Learning Algorithms	161
4.5.1 Hyundai Dataset Application Results.....	162
4.5.2 Kia Dataset Application Results	167
4.5.3 Chevrolet Dataset Application Results	171

5. CONCLUSION AND FUTURE WORKS.....	178
REFERENCES.....	181
BIOGRAPHICAL SKETCH.....	208



LIST OF SYMBOLS

3GPP	: The Third Generation Partnership Project
5G	: The Fifth Generation
5G NR	: 5G New Radio
AI	: Artificial Intelligence
AV	: Automated Vehicle
CAN	: Controller Area Network
CAV	: Connected and Automated Vehicles
C-ITS	: Cooperative Intelligent Transport Systems
C-V2X	: Cellular Vehicle-To-Everything
DoS	: Denial of Service
ECU	: Electronic Control Unit
GPS	: Global Positioning System
IDS	: Intrusion Detection System
IVN	: In-vehicle Network
LTE	: Long-Term Evolution
ML	: Machine Learning
NR	: New Radio
V2X	: Vehicle-To-Everything
V2N	: Vehicle-To-Network
V2I	: Vehicle-To-Infrastructure
V2D	: Vehicle-To-Device
V2P	: Vehicle-To-Pedestrian
V2V	: Vehicle-To-Vehicle

LIST OF FIGURES

Figure 1.1: Road Vehicle Taxonomy.....	4
Figure 1.2: Autonomous Vehicle Architecture.....	7
Figure 2.1: Publication Counts by Year – Cybersecurity OR Connected and Automated Vehicles	18
Figure 2.2: Publication Counts by Year	19
Figure 2.3: Years When Keywords Appeared	21
Figure 3.1: Attacks on Driving System Components	47
Figure 3.2: Manipulation-Based Attacks Papers	49
Figure 3.3: Identity-Based Attacks Papers	57
Figure 3.4: Service-Based Attacks Papers	60
Figure 3.5: Software-Based Attacks Papers	65

LIST OF TABLES

Table 1.1: Autonomous Vehicles Levels	5
Table 1.2: Comparison of CAN, Ethernet, and FlexRay Protocols	15
Table 2.1: Overview of Cybersecurity Threats, Anomaly Detection Techniques, and Datasets in Vehicular Networks	25
Table 2.2: Features of the Companies	31
Table 2.3: Comparison of Industry and Academia	35
Table 2.4: Definition of Emerging Systems	37
Table 2.5: Examples of Threats and Risk Values	39
Table 2.6: Cyber Security Severity Classes	40
Table 2.7: Tools to Use for Detecting Cyber Attacks	44
Table 3.1: A Comparative Overview of Attacks on Driving System Components	68
Table 3.2: A Comprehensive Summary of Research Conducted on In-Vehicle Networks	99
Table 4.1: Dataset Distribution and Balance Status by Vehicles	153
Table 4.2: Hyundai Dataset's First 10 row for Flooding Detection	155
Table 4.3: Hyundai Dataset's First 10 Row for Fuzzing Detection.....	157
Table 4.4: Hyundai Dataset's First 10 Row for Malfunction Detection	159
Table 4.5: Hyundai Dataset Algorithm Performance Results (60-40).....	162
Table 4.6: Hyundai Dataset Algorithm Performance Results (70-30).....	162
Table 4.7: Hyundai Dataset Algorithm Performance Results (80-20).....	163
Table 4.8: Runtime Costs for Hyundai Dataset Algorithm Results (60-40).....	164

Table 4.9: Runtime Costs for Hyundai Dataset Algorithm Results (70-30).....	164
Table 4.10: Runtime Costs for Hyundai Dataset Algorithm Results (80-20).....	165
Table 4.11: Kia Dataset Algorithm Performance Results (60-40).....	166
Table 4.12: Kia Dataset Algorithm Performance Results (70-30).....	166
Table 4.13: Kia Dataset Algorithm Performance Results (80-20).....	169
Table 4.14: Runtime Costs for Kia Dataset Algorithm Results (60-40).....	170
Table 4.15: Runtime Costs for Kia Dataset Algorithm Results (70-30).....	170
Table 4.16: Runtime Costs for Kia Dataset Algorithm Results (80-20).....	171
Table 4.17: Chevrolet Dataset Algorithm Performance Results (60-40).....	173
Table 4.18: Chevrolet Dataset Algorithm Performance Results (70-30).....	173
Table 4.19: Chevrolet Dataset Algorithm Performance Results (80-20).....	174
Table 4.20: Runtime Costs for Chevrolet Dataset Algorithm Results (60-40).....	175
Table 4.21: Runtime Costs for Chevrolet Dataset Algorithm Results (70-30).....	176
Table 4.22: Runtime Costs for Chevrolet Dataset Algorithm Results (80-20).....	176

ABSTRACT

The rapid development of autonomous vehicle technologies has made the security requirements of these systems more complex and has made it necessary to take security measures against cybersecurity threats in in-vehicle communication networks. In this thesis, a literature review is presented that comprehensively examines academic and industrial studies in the field of cybersecurity for autonomous and connected vehicles. The main purpose of the study is to address anomaly detection methods developed to detect security vulnerabilities in autonomous vehicles and to comprehensively present academic studies evaluating the role of these methods in determining security vulnerabilities. The differences between academic and industrial approaches in the field of cybersecurity are discussed in detail and an overview is provided regarding developments in the fields of generative artificial intelligence, machine learning, artificial intelligence and anomaly detection. In addition, various types of attacks such as fuzzy, malfunction and flooding, which are widely covered in the literature on cybersecurity vulnerabilities in autonomous vehicles and carry potential risks, are discussed and the prevention methods used against such attacks are evaluated. In the continuation of the study, various machine learning algorithms are divided into different test percentages and run on the relevant data set and presented in tables. These analyses were conducted using the In-Vehicle Intrusion Detection System Dataset, and a detailed assessment is provided on effective measures that can be taken against cyber-attacks. This study aims to make a comprehensive contribution to cyber security research in the field of autonomous and connected vehicles, and to constitute an important reference in the literature on the detection of security vulnerabilities in autonomous vehicles.

Keywords: Cyber-security, Automated Vehicles, Anomaly Detection, Machine Learning, In-Vehicle Intrusion Detection System

ÖZET

Otonom araç teknolojilerinin hızla gelişmesi, bu sistemlerin güvenlik gereksinimlerini daha karmaşık hale getirmiş ve araç içi haberleşme ağlarında siber güvenlik tehditlerine karşı güvenlik önlemleri alınmasını zorunlu kılmıştır. Bu tezde, otonom ve bağlantılı araçlar için siber güvenlik alanında yapılan akademik ve endüstriyel çalışmaları kapsamlı bir şekilde inceleyen bir literatür taraması sunulmaktadır. Çalışmanın temel amacı, otonom araçlarda güvenlik açıklarını tespit edebilmek için geliştirilen anomali tespit yöntemlerini ele almak ve bu yöntemlerin güvenlik açıklarının belirlenmesindeki rolünü değerlendiren akademik çalışmaları kapsamlı bir şekilde sunmaktır. Siber güvenlik alanında akademik ve endüstriyel yaklaşımlar arasındaki farklar ayrıntılı bir şekilde ele alınmakta ve generatif yapay zekâ, makine öğrenimi, yapay zekâ ve anomali tespiti alanlarındaki gelişmelere yönelik bir bakış sunulmaktadır. Ayrıca, otonom araçlarda siber güvenlik kırılganlıklarına dair literatürde geniş yer bulan ve potansiyel risk taşıyan fuzzy, malfunction ve flooding gibi çeşitli saldırı türleri ele alınmakta ve bu tür saldırılara karşı kullanılan önleme yöntemleri değerlendirilmektedir. Çalışmanın devamında çeşitli makine öğrenmesi algoritmaları farklı test yüzdelere ayrılıp ilgili veri seti üzerinde çalıştırılmış ve tablolar halinde sunulmuştur. Bu analizler, In-Vehicle Intrusion Detection System Veri seti kullanılarak yapılmış olup, siber saldırılara karşı alınabilecek etkin önlemler üzerinde detaylı bir değerlendirme sunulmaktadır. Bu çalışma, otonom ve bağlantılı araçlar alanında siber güvenlik araştırmalarına kapsamlı bir katkı sunmayı hedeflemekte olup, otonom araçlarda güvenlik açıklarının tespitine yönelik literatürde önemli bir referans oluşturmayı amaçlamaktadır.

Anahtar Kelimeler: Siber Güvenlik, Otonom Araçlar, Anomali Tespiti, Makine Öğrenimi, Araç İçi İzinsiz Giriş Tespit Sistemi

1. INTRODUCTION

Connected and automated cars have been recently penetrated in our daily transportation routines as a result of electrification of mass production vehicles and emerging car software and on-board computing technology. Traveling with cars operated as robotics in a wild traffic environment leaves us vulnerable to cyberattacks or at risk of losing control of the vehicle due to outside interference, raising concerns about safety. These cars have a long history that begins in the early years of automobiles. In 1925, the first autonomous car was observed, and the vehicle was operated by a remote control (Wikipedia, 2024). The same expectations have been placed on automotive technologies as a result of technology's rapid progress and return to comfort in human existence. Travel and wheeled conveyance in history began with the introduction of wheels. It persisted in the 19th and 20th centuries with the advent of oil and steam as energy sources. The steam vehicle is on display at the Arts et Metiers Museum in Paris. It was constructed in 1769 by Nicolas Joseph Cugnot and is regarded as one of the first instances of a vehicle. The car was taken to the army storehouse because it did not spark any curiosity during that time (Bellu, 1998; Vers et al., 2019). Ivan Kulibin invented the vehicle in 1791 and included characteristics including a gearbox, brakes, flywheel, and bearings. But at that point, the government decided this vehicle had no future, and the project was abandoned. Examples of steam-powered vehicles were constructed by Walter Hancock, Richard Trevithick, Josef Bozek, and Oliver Evans (Devichnick, 2017; Gridin, 2017). In the 1860s, internal combustion engines were created. The engines with this configuration operate on kerosene. Karl Benz created the first 4-stroke carbureted engine in 1884. The 4-stroke Benz engine was coupled to a 4-person vehicle in 1889 by Rene Pabhard and Emile Levassor (Eckermann, 2001; Milev et al., 2019). This vehicle was upgraded with several inventions, and quicker, more luxurious, and more elegantly designed cars were produced. Examples include valve systems, brake systems, steering systems, water

cooling systems for engines, spark plug ignition systems, and designs that shield car occupants from rain and dust (Coulibaly, 2007; Taymans et al., 2020).

Racing has accelerated automotive development even further. The first known motor race was the 1894 Paris-Rouen. Developments in suspension, steering, aerodynamics, wheels, braking systems, and engine performance have accelerated since speed and vehicle control are the primary priorities in racing. Francis Houdina uses a remote control to operate the first recognized and respectable autonomous vehicle (Esteban, 2019; Puiboube, 2000). General Motors introduced radio signals in 1956 to allow the Firebird model car to go autonomously on the highway (Kendi, 2017). Autonomous cars did not advance to human-controlled driving until 2004 because to technological shortcomings and the growth of the automotive industry, Miller (2014).

The advancement of electronic sensors has influenced the development of autonomous cars. Developed from the 1970s to the 2000s for airborne submarine detection, vision systems, particularly Lidar Sensors, found application in land mapping, forest and water research, and autonomous vehicle applications starting in the 2000s. In 1978, the US Department of Defense began using a Global Positioning System (GPS), another sensor utilized in autonomous cars (Congress, 2018). 24 satellites are positioned in the space to use it. Mining and agriculture are two industries that employ GPS navigation (Haider & Khalid, 2017; GPS.gov, 2020).

The OSU Demo '97 Vehicle project showcased significant advancements in autonomous vehicle technologies. Ohio State University demonstrated two fully automated cars interacting with a manually driven vehicle, highlighting a scenario with mixed traffic. Unique technologies such as radar-reflective stripes and vision-based steering systems were employed, providing a cost-effective and user-transparent alternative to traditional methods (Ozguner et al., 1997).

The DARPA Grand Challenge was declared in 2004. Car manufacturers (General Motors, Ford, Volkswagen, Toyota, and Volvo) started using Google to test their automobiles in 2010 when the company started working in this area (Patterson et al., 2017). In 2013, both level 1 and level 2 cars were used in traffic testing. 2014 saw tests of level 3 and level 4

vehicles running from San Francisco to Seattle (Tunali, 2019). This paper looks at the vulnerabilities that can be exploited in level 5 fully autonomous cars, as well as possible preventative measures and future research directions.

While security flaws might make it impossible or dangerous to deploy autonomous cars, they can also result in a loss of control over them. The first cyberthreat that springs to mind and is caused by cyber security flaws is the possibility that an uncontrollably moving automobile may gravitate toward a certain coordinate. Cyberattacks might result in the automobile being stolen or its occupants being arrested. The automobile and/or its occupant suffering harm is the second concern that comes to mind in the event that the system malfunctions. For instance, attacker-initiated buffer overflow attacks may render system sensors unusable. Issues include improperly recognizing traffic signals, deactivating the braking system, and/or being unable to brake appropriately might arise using this cyberattack technique. Additionally, any system designed and utilized for comfort, including air conditioning, heating, windows, wheels, suspensions, radios, and seat warmers, might be used accidentally.

The structural classification of road vehicles is illustrated in Figure 1.1 (Khrapova, Mariia, 2017). The figure categorizes road vehicles into two main groups: motor vehicles and attached vehicles. Motor vehicles are further subdivided into categories such as passenger cars, buses, trucks, special vehicles, and tractors, while attached vehicles include trailers, semi-trailers, and sidecars. This classification provides a systematic structure that highlights the diversity of road vehicles, allowing for the differentiation of their transport purposes, technical characteristics, and security requirements. The "passenger cars" category, in particular, holds significant relevance in terms of security risks for autonomous and connected vehicles. Within the scope of this study, datasets and literature specific to passenger cars were analyzed to examine potential cybersecurity threats that this type of vehicle may face and the methods used to detect these threats. This diagram supports the rationale for focusing solely on passenger cars in this research, as it emphasizes the unique security needs of each vehicle category, underscoring the importance of targeted cybersecurity measures for passenger vehicles.

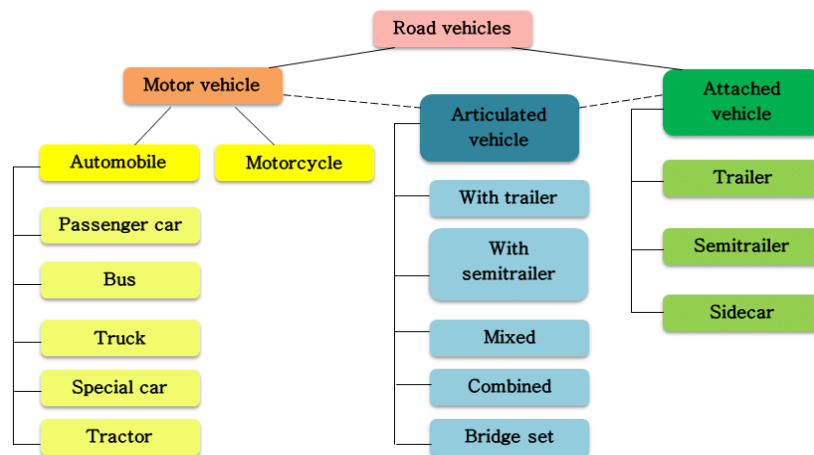


Figure 1.1 - Road Vehicle Taxonomy (Khrapova, Mariia, 2017)

1.1 Automated Vehicles

Since they first appeared in our lives in the 19th century, automobiles have evolved in every way. The increasing need for more comfortable and equipped automobiles has led to a widespread manufacture and usage of autonomous vehicles. The fully autonomous automotive category, or level 5, is the sixth category in which these vehicles are now analyzed in accordance with SAE J3016 criteria (European Union Agency for Network and Information Security (ENISA), 2017). There is a reduction in traffic time. Personal security is the topic that has to be considered right now.

Because of this, vulnerabilities are exposed by the system's interface with the outside world and the internet requirement for completely autonomous cars. In essence, autonomous cars are equipped with wireless communication sensors, vision sensors, position sensors, and vehicle part inspection sensors. It has many other characteristics with non-autonomous cars as well, like as mechanical components like wheels, suspension, steering, brakes, and interior car equipment. An autonomous car to be used now has almost this feature, even if there may be modifications, additions, and lightening to these elements in the future to enable simplicity of use (Molla & Electronik, 2018).

1.1.1 Benefits and Usage Areas of Autonomous Vehicles

Ninety percent of road accidents are the consequence of human mistake, according to statistical research (TUIK, 2018). It was intended to limit driver usage and traffic law infractions in order to eliminate these mistakes.

According to Pancorbo Crespo et al. (2019), autonomous cars are those that are utilized in situations where it is not desirable for accidents caused by human mistake to occur, such as in big regions, farming, and by people who have health issues and find it difficult to operate normal vehicles.

1.1.2 Autonomous Vehicles Levels

Design modifications in autonomous vehicles will also result from the goal to gradually convert automobiles to electric vehicles. These modifications entail the integration of robust and compact mechanical components with tiny models of electrical and digital components. There are presently six distinct levels as a result of the effort done in this direction. (ENISA, 2017; Feng, 2019; Ustam, 2020) refers to these as Level 0, Level 1, Level 2, Level 3, Level 4, and Level 5.

Table 1.1 - Autonomous Vehicles Levels (Lee, Hyo-Keun, 2022)

NHTSA Level	SAE Level	Definition of Autonomous Level		Characteristics
0	0	No automation		No support
1	1	Driver assistance		Provides driving information, generates alerts, and supports some controls
2	2	Automated	Partial automation	Automates some of the vehicle controls according to driver selection in special circumstances
3	3	Automated	Conditional automation	Automates all vehicle controls, driver manually/automatically selected
4	4	Autonomous	High automation	The vehicle can drive on its own in all traffic conditions. Commercialization distance, legal/institutional problem resolution required
4	5	Autonomous	Full automation	The vehicle can drive on its own in all traffic situations without any legal or institutional problems

Level 0 (No Driving Automation) refers to vehicles where the driver has full control at all times, with no automation features beyond basic alerts or warnings. There is no impact on use control from the sensors or cameras on, around, or within the car. The driver alone is in charge of the operation.

Level 1 (Driver Assistance) refers to vehicles with a single automated system for driver assistance, such as adaptive cruise control or lane-keeping assist, while the driver remains responsible for all other aspects of driving. Although some driving assistance is available,

the driver maintains control of the vehicle. Some of the impacts are headlamp angle sensor, adaptive cruise control, and braking-induced wheel stoppage.

Level 2 (Partial Driving Automation) refers to vehicles equipped with advanced systems that can control both steering and acceleration/deceleration simultaneously under certain conditions, while the human driver must remain engaged and monitor the environment at all times. Although the driver controls how the car runs and drives ahead, the driving system can take over control of the vehicle. They are made with the expectation that the driver will step in at any time. At this level, vehicles are designed to be driven on straight, mostly straight routes with little turns, where the vehicle brain can handle control. The driver is still required to handle the steering wheel in these types of cars.

Level 3 (Conditional Driving Automation) refers to vehicles that can perform all driving tasks under specific conditions, but the driver must be ready to take control when the system requests intervention. They have extra sensors in addition to level 2 vehicle sensors. The vehicle is in complete control of these cars. The driver is alerted in advance when there is anything on the road that requires their attention, including neighboring cars, traffic warning signs, lights, and pedestrians.

Level 4 (High Driving Automation) refers to vehicles capable of performing all driving tasks autonomously in specific conditions or predefined areas, without requiring driver intervention, but human control may still be optional. The driver is not expected to exert control during this trip; it is conceivable for the vehicle to arrive at the target destination without the driver's intervention. The car is self-parking. These cars are made to go as securely as possible in accordance with the information received from the brain sensors since they are driven by system sensors to ensure that occupants reach their destination. In the event of heavy traffic, these cars may also securely transport people to their destinations by charting other routes.

Level 5 (Full Driving Automation) refers to vehicles capable of performing all driving tasks under all conditions without any human intervention, eliminating the need for a driver altogether. Vehicles that meet Level 5 requirements have every necessary sensor installed. Without assistance from a person, it can drop off passengers at the appropriate

location. Level 5 cars are anticipated to be on the road by 2030, based on forecasts (Molla & Electronics, 2018).

1.2. Automated Vehicle Architecture

The equipment in the car is composed of electronic components at least as much as mechanical ones. These include computer vision systems, brakes, steering, speed control units, lidar sensors, infrared cameras, GPS, inertial measurement devices, wheel encoders, odometry, and radar (José E. Naranjo, 2009). The sensors that are employed also change based on the degree of vehicle autonomy. In general, the cars' architectural designs are quite similar to one another. (Narajo et al., 2009; İspir, 2019).

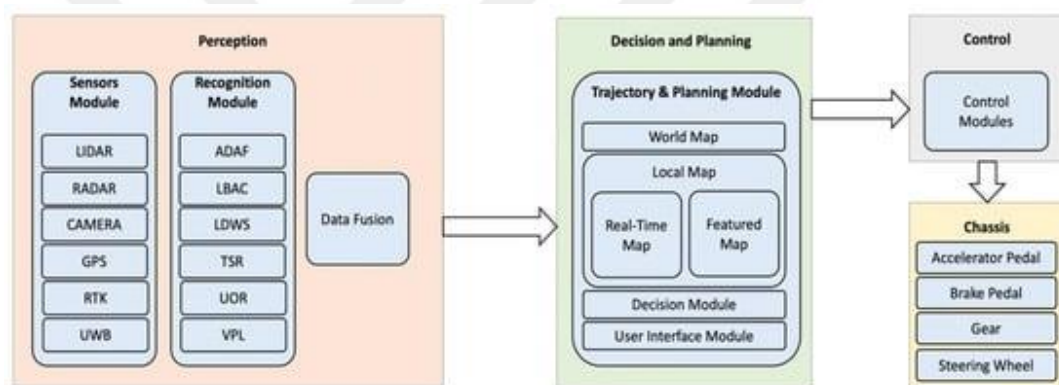


Figure 1.2 -Automated Vehicle Architecture (Algarni et al., 2022)

1.3. Main Components of Automated Vehicles

Numerous vulnerabilities related to sensors, electrical control units, and various connection procedures have been found by researchers. It was determined that the following three crucial elements were required for our examination of the security characteristics of driverless vehicles: components of automated driving systems, automobile control systems, and V2X.

1.3.1. Automotive Control Systems

Modern vehicle functionality and security depend heavily on automotive control systems, which include software algorithms integrated with automotive-grade hardware called electronic control units. They regulate crucial elements like engine optimization, transmission control, and safety features like Electronic Stability Control (ESC) and Anti-lock Braking Systems (ABS). Sensor integration, sensor fusion and decision-making algorithms are mandatory as cars get closer to being able to drive themselves, emphasizing how important it is to comprehend and enhance these systems for driving safety.

ECUs optimize subsystems to assure overall performance, efficiency, and safety, making today's cars as mechatronic devices having interfaces of smart things unlike the legacy immune automotive technology. For the advancement of automotive electronic control technology, it is essential to comprehend the function of ECUs throughout vehicle subsystems. Advanced Driver Assistance Systems (ADAS) are facilitated by specialist modules such as Adaptive Cruise Control (ACC) and Lane Departure Warning (LDW), which are in addition to Powertrain Control Modules (PCMs) that regulate the engine and gearbox.

1.3.2. Automated Driving System Components

Emerging sensor technology such as Laser Imaging Detection and Ranging (LIDAR), cameras, Radio Detection and Ranging (RADAR), central computers, and ultrasonic sensors are propelling the car industry's shift to autonomous driving. LIDAR sensors employ laser technology to generate intricate three-dimensional maps for accurate object recognition, whilst cameras offer visual information processed by sophisticated computer vision algorithms. GPS broadcasts global localization data for navigation, whereas RADAR sensors employ radio waves to identify objects, especially in inclement weather. For efficient navigation, the main computer combines sensor data, runs intricate algorithms, and regulates vehicle actuators. Ultrasonic sensors, on the other hand, are essential for jobs like parking assistance because of their short-range detecting capabilities, which improve accuracy in small areas. Their smooth integration should transform how a vehicle and its occupants interact, enhancing efficiency and safety.

Nonetheless, a variety of security threats confront automated cars, such as physical access attacks that jeopardize hardware functionality, data privacy issues brought on by mass data collection, software flaws that allow for unauthorized control, service interruptions that impact navigation, identity risks that target authentication systems, and manipulation-based attacks that distort perception and decision-making. In order to mitigate these complex threats and guarantee the security and safety of automated cars in practical situations, strong defenses are necessary.

1.3.3. V2X Communication Technologies

In the realm of intelligent transportation systems, V2X communication technologies offer seamless integration of fast and lossless connectivity, enabling communication among cars and their surroundings. The network (V2N), pedestrians (V2P), other vehicles (V2V), and infrastructure (V2I) are all encompassed by the "X" in V2X. Wireless communication protocols, of which Dedicated Short-Range Communication (DSRC) and C-V2X are the two main standards, provide the foundation for V2X communication.

DSRC, which is based on IEEE 802.11p, allows direct vehicle-to-vehicle and vehicle-to-infrastructure connections by providing low-latency, high-data-rate communication in the 5.9 GHz frequency range (IEEE, 2010). It also conforms to the European Telecommunications Standards Institute (ETSI) ITS G5 standard, which defines access layer protocols for safety-critical vehicular communication systems (ETSI, 2020). Nevertheless, C-V2X offers improved scalability, dependability, and the possibility of over-the-air updates by combining direct short-range and wide-area communication through the use of cellular networks.

1.4. Automated and Connected Vehicle Regulations and Standards

At the national and international levels, the regulatory environment pertaining to automated, autonomous, and connected cars is presently experiencing substantial growth. Global laws, such as UN Regulation No. 157 on Automated Lane-Keeping Systems (ALKS), are currently being drafted by organizations like the United Nations Economic Commission for Europe (UNECE). Governments throughout the world are developing

frameworks at the national and regional levels that include liability issues, cybersecurity precautions, safety regulations, and data privacy policies. These rules focus on consumer education, insurance framework adaptation to the particular hazards connected with autonomous driving, testing and certification procedures, and the various issues provided by the deployment of advanced automotive technology. In order to ensure the safe, secure, and responsible integration of automated, autonomous, and connected cars into the global transportation ecosystem, the regulatory environment is evolving in parallel with technology.

1.4.1. R155 Cybersecurity

The World Forum for Harmonization of Vehicle Regulations (WP.29), a working group under the Sustainable Transport Division of the United Nations Economic Commission for Europe (UNECE), developed UN Regulation No. 155 (UN R155). WP.29 incorporates new technology into existing regulatory frameworks to ensure that future car manufacturing will produce safer and more environmentally friendly automobiles. In January 2021, it carried out UN R155 as part of its duties.

Vehicles must have a cybersecurity management system (CSMS) installed, under UN R155. To put it briefly, a CSMS makes sure that cybersecurity procedures and safeguards are appropriately implemented throughout the vehicle's life cycle and development process.

UN R155 provides broad, objective-based guidelines for determining if a cybersecurity system is sufficiently implemented and if a CSMS is in place. The necessity to carry out a comprehensive risk assessment and the requirement to recognize and react to cyberattacks at every stage of a vehicle's life cycle are the main obstacles this rule presents to businesses.

In pursuit of robust cybersecurity measures within the automotive industry, UN R155 incorporates Annex 5, which provides a comprehensive overview of 69 potential attack vectors or risks. These risks span various focal areas that manufacturers are mandated to address in order to fortify the security of their vehicles. These areas encompass back-end servers, where threats involve potential abuse of privileges by staff and unauthorized

internet access to servers. The communication channels within a vehicle are identified as susceptible to risks such as spoofing messages, code injection, and interception of information. Risks associated with update procedures include the manipulation of software prior to an update process and the potential for denial-of-service (DoS) attacks that could hinder the rollout of updates. The human factor is recognized as a source of risk, with a focus on errors stemming from non-compliance with defined security procedures and susceptibility to tactics facilitating cyberattacks. External connectivity risks are outlined concerning a vehicle's interaction and communication with its external environment, encompassing potential attacks on sensors, external interfaces, and remote functions. Additionally, threats related to data or code encompass unauthorized access to personal information, falsification of vehicle data, and the introduction of malicious software. Lastly, the sufficiency of vulnerability hardening is highlighted, emphasizing potential vulnerabilities that could be exploited if protective measures prove insufficient, involving compromises to cryptographic technologies, hardware, and software. The delineation of these focus areas underscores the intricate and multifaceted nature of cybersecurity considerations in the automotive industry, necessitating a comprehensive and systematic approach to risk mitigation and the enhancement of vehicle security

1.4.2. R156 Software Update

UN Regulation 156 encompasses uniform provisions for vehicle software updates and software update management systems. It mandates original equipment manufacturers (OEMs) to establish, at an organizational level, a suite of core processes. These processes involve configuration control to record relevant hardware and software versions for a specific vehicle type, encompassing integrity validation data for the software. Additionally, OEMs must identify and track changes in software and hardware on a vehicle as per specific UN regulations. The regulation emphasizes the verification and validation of software on vehicle components, identification of interdependencies among systems, especially concerning software updates, and confirmation of compatibility with target vehicles. Furthermore, OEMs are tasked with assessing the potential impact of software updates on Type Approvals, legally defined parameters, vehicle safety, and safe driving. The regulation also emphasizes the importance of informing consumers about

updates, documenting all processes for inspection during audits, and ensuring the cybersecurity of software updates before their implementation into a vehicle.

1.4.3. R157 Automated Lane Keeping System

Vehicle "Automated Lane Keeping Systems" (ALKS) are covered under UN Regulation 157. With the help of an automated lateral and longitudinal control system (ALKS), a driver can stop actively controlling the vehicle under certain circumstances. The technical criteria and performance standards for ALKS are outlined in UN R157, which guarantees the systems' dependable and safe functioning.

UN R157 may address important topics such as backup circumstances, minimum and maximum operating speeds, interventions, system activation requirements, and other safety-related characteristics. This pertains to the autonomous driving technology and is a component of the larger regulatory framework that the United Nations Economic Commission for Europe (UNECE) created.

To get up-to-date and comprehensive details on UN Regulation 157 and its particular requirements, it is advisable to visit the official UNECE website or confer with pertinent automotive regulatory authorities. The most recent information on regulations in this area may be obtained in official regulatory announcements and papers, which are updated and revised periodically.

1.4.4. Cybersecurity Standards

Establishing rules is the initial step in assisting organizations in protecting themselves. An international standard (ISO/IEC 15408) for computer security certification is known as the "common criteria" or ISO (Common Criteria for Information Technology Security Evaluation). However, much like quality, cybersecurity must be taken into account in every organizational process. The goal of an Information Security Management System (ISMS), a management system like to those suggested by other ISO standards like ISO 9000 and ISO 14000, is to manage information security risks and controls inside an organization. This is explained by the family of ISO 27000 standards. Deliberately

placing information security under direct management control is a fundamental tenet of ISO/IEC 27000 standards.

A cybersecurity process structure and guidelines are provided by SAE (Society of Automotive Engineering) Recommended Practice J3061 (S.V.E.S.S. Committee, 2016). It outlines a comprehensive lifecycle process framework that can be customized and used to the development procedures of any firm. Following the security by design methodology, its primary objective is to enable businesses to integrate cybersecurity considerations into physical vehicle systems from the idea stage through manufacture, operation, servicing, and decommissioning.

It is goal-based rather than prescriptive, offers high-level guiding concepts, and offers details on current procedures and techniques. It is also intended to serve as a basis for the creation of further standards.

In order to create a globally recognized standard for cybersecurity engineering that is appropriate for the automotive sector and takes a risk-based approach, SAE and ISO have partnered to create ISO/SAE 21434 (International Organization for Standardization, 2018).

1.5. In-Vehicle Network Protocols: CAN, Ethernet, and FlexRay

1.5.1. Controller Area Network (CAN)

Controller Area Network is a protocol developed by Bosch in the late 1980s for in-vehicle data communication (Bosch, 1991). CAN is widely accepted as a standard communication protocol in the automotive industry due to its low cost, robustness, and reliability. The CAN protocol enables data transfer between critical vehicle components such as engine control, transmission, and braking systems. Although the CAN bus operates at relatively low data rates (up to 1 Mbps), it employs an asynchronous communication structure that prevents collisions during data transmission. CAN messages are transmitted based on a specific priority order, meaning high-priority messages are sent before lower-priority ones. This prioritization feature provides a

reliable and continuous communication solution within in-vehicle systems, despite the limited data speed.

1.5.2. Ethernet

Ethernet, a protocol widely used in computer networks, has recently been integrated into in-vehicle data communication. Ethernet's high data capacity is suitable for data-intensive applications such as infotainment systems, Advanced Driver Assistance Systems (ADAS), and autonomous driving. Traditionally, Ethernet supports data rates ranging from 100 Mbps to 1 Gbps, offering a significantly higher data transfer rate than CAN. In Ethernet-based systems, the high volume of data can be processed and transferred quickly to meet the demands of real-time applications. This protocol allows in-vehicle components such as high-resolution cameras, radar systems, and lidar sensors to perform synchronized and instant data transfers. However, the higher cost of Ethernet-based communication infrastructure limits its use to certain applications, primarily those requiring substantial data throughput (IEEE, 2018).

1.5.3. FlexRay

FlexRay is another in-vehicle network protocol designed to enable high-speed, deterministic, and reliable data transmission. The primary advantage of FlexRay lies in its flexibility in synchronization and timing, which supports a data transmission rate of up to 10 Mbps, addressing situations where CAN is insufficient in speed. FlexRay is particularly suitable for applications requiring high reliability, such as autonomous driving technologies, chassis control systems, and advanced safety functions. FlexRay also supports multiple data channels, enabling simultaneous data transmission and fault tolerance to enhance the overall system performance and reliability. Its deterministic structure ensures that in-vehicle systems operate in a synchronized and consistent manner, which is critical for safety systems where high speed and accuracy are required (FlexRay Consortium, 2015).

The following Table 1.2. summarizes the key features of CAN, Ethernet, and FlexRay protocols in a comparative format:

Table 1.2. – Comparison of CAN, Ethernet, and FlexRay Protocols (CAN: Bosch, 1991; Ethernet: IEEE, 2018; FlexRay: FlexRay Consortium, 2015)

Feature	CAN	Ethernet	FlexRay
Data Rate	Maximum 1 Mbps	100 Mbps - 1 Gbps	Maximum 10 Mbps
Applications	Engine control, braking systems, ABS	Infotainment, ADAS, data-intensive systems	Autonomous driving, sensor data, chassis control, active safety
Communication Structure	Message-based, prioritized ordering	Packet-based	Synchronized, multi-channel
Cost	Low	High	Medium
Reliability	High	Moderate	Very High
Synchronization	Not synchronized	Synchronized	Provides high-level synchronization
Deterministic Communication	Limited	None	Supported
Advantages	Low cost, reliability	High speed	High speed and reliability, deterministic
Disadvantages	Low speed, limited bandwidth	High cost	Complex infrastructure and costly

2. CONNECTED & AUTOMATED VEHICLE'S CYBERSECURITY REVIEW

2.1. Research Framework and Methodology

This section outlines the research framework and methodology used to explore cybersecurity challenges in connected and autonomous vehicles. It begins with the formulation of research questions, followed by a detailed description of the search process. Additionally, it examines key cybersecurity trends and relevant datasets, highlighting their applications in vehicle network security research.

2.1.1. Research Questions

This study is designed to investigate the cybersecurity threats and solutions within CAV networks, guided by the following research questions:

Q1: What cybersecurity threats and attack types are identified for connected and autonomous vehicles?

Q2: Which artificial intelligence and machine learning algorithms are used to detect these attacks?

Q3: What solutions are proposed in the literature to address cybersecurity threats in connected and autonomous vehicles?

These questions help structure the literature review and direct the analysis toward the key components of vehicle network security.

2.1.2. Search Process

A systematic literature review was conducted to identify studies relevant to cybersecurity in connected and autonomous vehicles, with a particular focus on articles citing the paper Anomaly Intrusion Detection Method for Vehicular Networks Based on Survival Analysis (Han

et al, 2018). This approach enabled the selection of sources that either directly address or build upon the methodology and findings presented in this paper.

Keywords and Databases: The search included terms such as "connected vehicles security," "anomaly detection in autonomous vehicles," "survival analysis in intrusion detection," and "cyber threats in CAVs." Key academic databases like IEEE Xplore, ScienceDirect, and Google Scholar were utilized to retrieve relevant publications.

Inclusion and Exclusion Criteria: Studies that explicitly referenced survival analysis in the context of intrusion detection and CAV security were prioritized. Research lacking empirical analysis or methodological rigor was excluded, focusing instead on peer-reviewed journal articles, conference proceedings, and reputable technical reports.

Selection Process: The initial selection involved screening abstracts and keywords for relevance. Detailed review and selection yielded articles that provided specific insights into attack types, detection algorithms, and mitigation approaches, which were then analyzed in relation to Q1, Q2, and Q3.

2.1.3. Cybersecurity Keywords and Trends in Connected and Automated Vehicles

In this study, a systematic literature review was conducted to explore cybersecurity in connected and autonomous vehicles by using specific keywords and databases. Keywords such as "connected vehicles security," "anomaly detection in autonomous vehicles," "survival analysis in intrusion detection," and "cyber threats in CAVs" were selected to guide the literature search. These keywords were chosen to align with the core research questions of the study and to narrow the scope, enabling more focused and relevant results.

During the data collection process, leading academic databases such as IEEE Xplore, ScienceDirect, and Google Scholar were utilized. These databases were selected for their comprehensive collections of peer-reviewed publications and their strong reputation in relevant fields. The search aimed to investigate general trends in the field as well as to focus on more specific subtopics.

The findings from the keyword research highlight both the overall trends in publication volume and the growing interest in specific areas of the field. The Figure 2.1., demonstrates the

publication counts by year using the broad keywords "Cybersecurity OR Connected and Automated Vehicles." A significant increase can be observed starting from 2018, reaching its peak in 2023 with 232 publications. This reflects the growing interest in cybersecurity alongside the advancement of connected and autonomous vehicle technologies.

In conclusion, the results of this process reveal not only the growing volume of publications in the field but also the deepening focus on more specialized topics. This trend underscores the increasing urgency and significance of addressing cybersecurity challenges in connected and autonomous vehicles. The insights gained through this review provide a strong foundation for further analysis and innovation in the field, paving the way for more advanced solutions and applications in the years to come.

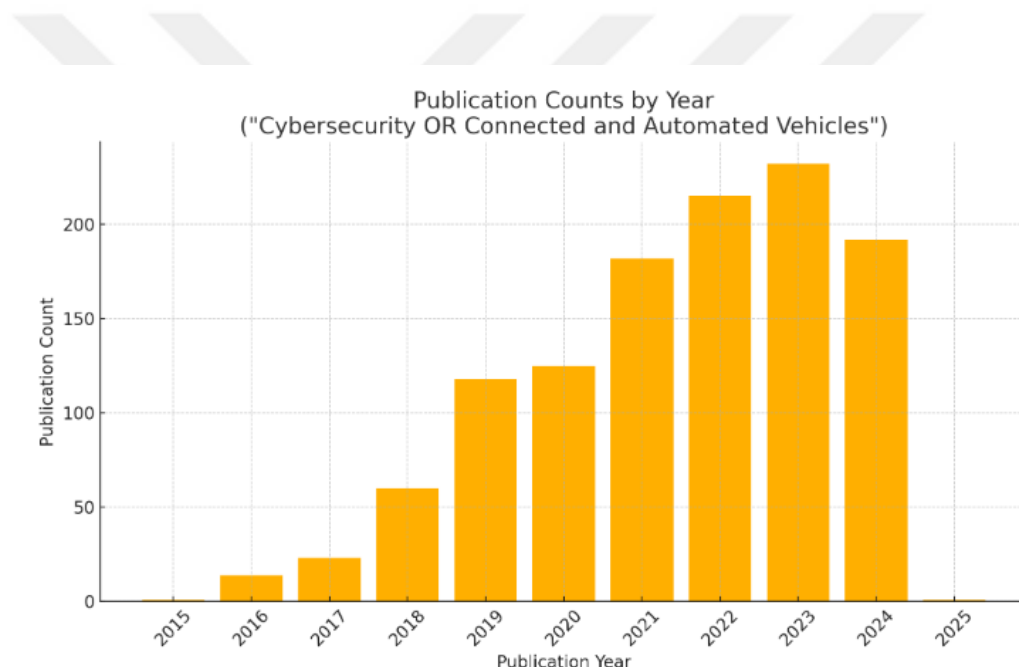


Figure 2.1. – Publication Counts by Year – Cybersecurity OR Connected and Automated Vehicles

The Figure 2.2., on the other hand, reflects the results of searches conducted with more focused keywords. It reveals a steady increase in publication volume starting in 2020, with a notable peak of 20 publications in 2024. This indicates that research in this field is becoming increasingly specialized, focusing on advanced methodologies and specific topics.

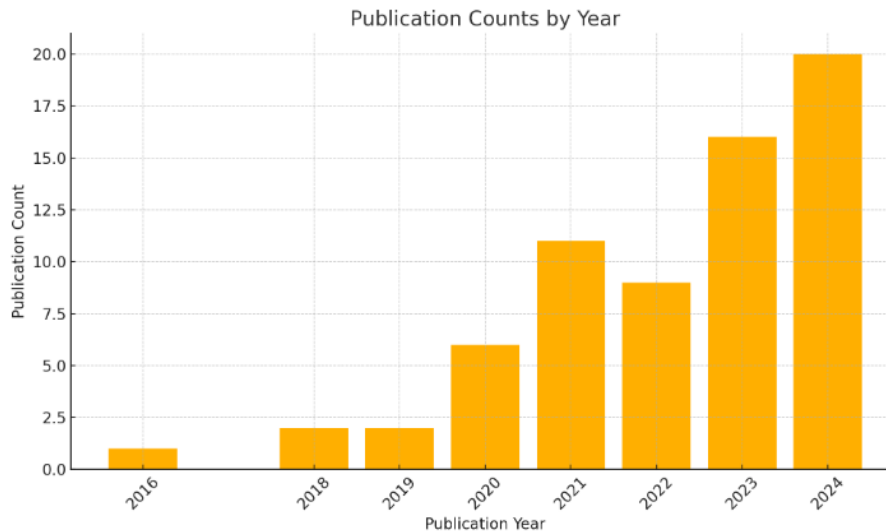


Figure 2.2. – Publication Counts by Year

The cybersecurity landscape in connected and autonomous vehicles has evolved significantly from 2019 to 2024. The focus has shifted toward protecting vehicle networks, utilizing advanced algorithms, and addressing emerging threats. Below, Figure 2.3 present a comprehensive overview of the key topics and trends shaping this field.

Figure 2.3 illustrates the distribution of research focus areas in cybersecurity for connected and autonomous vehicles over the years 2019 to 2024. The figure highlights when specific keywords gained prominence, providing a chronological view of trends in the field.

Intrusion Detection System appears consistently across all years from 2019 to 2024, underscoring its foundational role in vehicular cybersecurity. As a critical mechanism for identifying and mitigating threats, IDS remains a cornerstone of research and development in connected vehicle security. This persistent focus indicates that real-time threat detection and anomaly identification are vital for maintaining secure vehicular networks.

The Controller Area Network emerges as a key topic in 2021 and 2022. This emphasis aligns with a period of intensified efforts to address vulnerabilities within CAN, which serves as a primary communication protocol for in-vehicle networks. Research during these years likely aimed at fortifying CAN against attacks such as spoofing and injection, which exploit its legacy architecture.

Machine Learning begins to gain prominence in 2022 and continues to feature in 2023 and 2024. This reflects a shift toward leveraging ML techniques to enhance intrusion detection systems and predict potential threats. The application of ML represents a transition from traditional security mechanisms to more adaptive, intelligent solutions capable of handling the complexity of modern vehicular systems.

Cybersecurity becomes a focal point in 2023 and 2024, reflecting a broader approach to vehicular security. This shift indicates increased awareness of the interconnected nature of modern vehicles and the need for comprehensive strategies that address data integrity, privacy, and resilience against sophisticated attacks. The emphasis on cybersecurity highlights the importance of addressing challenges beyond specific network protocols or attack types.

In-Vehicle Network security receives attention in multiple years, particularly in 2020, 2021, and 2023. This recurring focus reflects ongoing efforts to secure internal vehicle communications, which are essential for the safe and reliable operation of connected and autonomous vehicles. IVN-related research likely addresses challenges related to latency, scalability, and real-time threat detection.

Overall, Figure 2.3 reveals the key shifts in research priorities from 2019 to 2024. The field has transitioned from traditional methods like IDS and CAN to advanced techniques like machine learning, cybersecurity, and holistic system solutions. This evolution demonstrates the need for scalable, adaptive, and proactive approaches to address the growing complexity of threats in the connected and autonomous vehicle ecosystem. These trends provide crucial guidance for future research directions. It also captures a clear evolution in research priorities, transitioning from foundational technologies like IDS and CAN to advanced techniques such as machine learning and holistic cybersecurity frameworks. This progression demonstrates how the field has adapted to emerging challenges, focusing on scalability, transparency, and robustness in securing connected and autonomous vehicles.

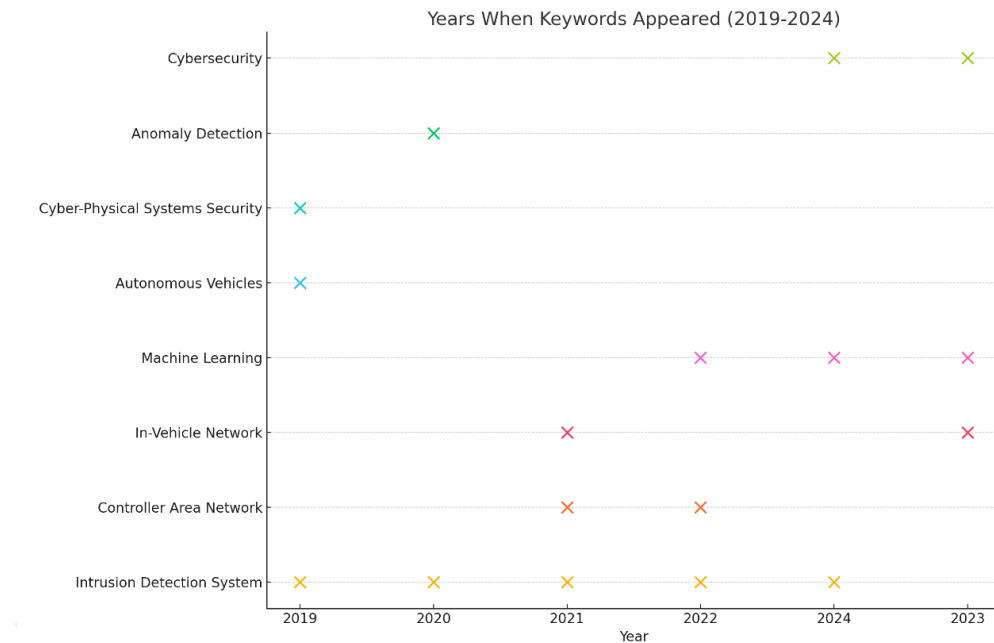


Figure 2.3. – Years When Keywords Appeared (2019-2024)

The visualization also highlights a transition from foundational topics like IDS and CAN to more sophisticated and interdisciplinary approaches, such as integrating cybersecurity with AI-driven solutions. This progression reflects the industry's response to the increasing complexity of threats in connected and autonomous vehicles.

Intrusion Detection System is consistently highlighted across all years, proving itself to be a fundamental technology in vehicular cybersecurity. Early research on IDS primarily focused on static, rule-based approaches. Over time, IDS systems have evolved into more dynamic structures, supported by machine learning techniques, enabling real-time detection of complex threats. The continuous emphasis on IDS reflects the growing complexity of cyber threats and the necessity to secure vehicle communication systems.

Controller Area Network emerged as another key focus area in 2021 and 2022. Widely used as the primary communication protocol in vehicles, CAN is inherently vulnerable to cyberattacks due to its lack of built-in security mechanisms. Research during these years concentrated on strengthening CAN against spoofing, injection, and replay attacks. Techniques such as encryption, hybrid security models, and anomaly detection have been widely explored to enhance the robustness of CAN-based networks.

Machine Learning gained significant attention starting in 2022 and continued to be a major focus through 2024. The integration of ML into cybersecurity marks a shift from static security mechanisms to adaptive and intelligent models. These techniques, including algorithms like Random Forest and LSTM, have enhanced IDS performance, enabling anomaly detection, threat prediction, and automated response capabilities.

In-Vehicle Network was notably emphasized in 2020, 2021, and 2023. This focus reflects the increasing complexity of securing internal communication systems within vehicles. IVN research has targeted real-time threat detection and network efficiency improvements, ensuring the reliability and continuity of in-vehicle systems.

Cybersecurity gained prominence in 2023 and 2024, emphasizing a broader, more comprehensive approach. Research during these years addressed not only network security but also data privacy, secure cloud integration, and systemic threat modeling. This shift highlights the need for holistic security strategies that consider the interconnected nature of modern vehicle systems.

2.1.4. Datasets and Their Applications in Vehicle Network Security Research

In CAV network security research, both real-world and synthetic datasets are essential for validating intrusion detection models. The following datasets, each with specific applications, serve as benchmarks in evaluating security measures.

Datasets collected from real vehicles provide invaluable insights into the behavior of vehicular systems under normal and abnormal conditions, offering a foundation for developing and testing cybersecurity and anomaly detection techniques.

HCRL Car Hacking and Survival Analysis Datasets: The HCRL dataset plays a critical role in vehicular cybersecurity research by simulating various flood attack types that target the CAN bus network. One of the key attacks in this dataset is the Flooding Attack, where multiple sender ECUs transmit messages simultaneously to a single receiver ECU. This attack exploits the CAN protocol's prioritization mechanism, where lower CAN IDs are assigned higher priority. As a result, ECU communication becomes limited, disrupting normal driving operations and

compromising critical safety systems like airbags that may fail to activate in time (Han et al., 2018).

The Fuzzy Attack involves the random generation of CAN IDs (ranging from 0x000 to 0x7FF) and data fields. Both valid and non-existent CAN IDs are used in this attack, creating chaos in the network. The payloads often include values like FFFF FFFF FFFF FFFF, which result in severe communication breakdowns between ECUs. Such attacks may cause anomalies such as flashing warning lights or uncontrollable acceleration.

The Malfunction Attack targets specific CAN IDs by manipulating their data fields (0x00 or random values) while simultaneously injecting random CAN IDs. This causes the vehicle to exhibit abnormal behavior, such as irregular engine RPM or false overheating alerts. For example, a manipulated message for the engine coolant system could falsely indicate overheating, creating unnecessary concern for the driver.

Real ORNL Automotive Dynamometer (ROAD) CAN Intrusion Dataset: The ROAD dataset offers real-world CAN bus data tailored for time-based anomaly detection research and highlights flood attack scenarios in detail. The Fuzzing Flood attack in this dataset involves the high-frequency injection of random CAN messages with maximum payloads such as FFFFFFFFFFFFFFFFFF. This disrupts normal communication flows, causing physical issues like unresponsive gas pedals or flashing warning lights. Messages such as FFF#0000000000000000 are used as filler messages to simulate a stream of random, high-priority data (Hollifield et al., 2021).

The Targeted Flood attack focuses on specific CAN IDs by injecting fabricated messages repeatedly, effectively overriding legitimate signals. For instance, targeting the speedometer CAN ID might cause it to display a constant maximum speed, misleading the driver and creating safety risks.

The Fabrication Flood attack injects counterfeit messages that mimic legitimate ones, saturating the CAN network and delaying critical safety responses such as airbag deployment. These scenarios demonstrate the vulnerabilities in vehicular communication systems and provide a robust platform for testing the effectiveness of intrusion detection systems (IDS).

(Blevins et al., 2021) compared time-based attack detection methods on this dataset, demonstrating that distribution-agnostic techniques, especially binning, excelled in precision-recall analysis.

VeReMi Dataset: The VeReMi dataset specializes in misbehavior detection in V2X communication, focusing on position spoofing attacks in vehicular ad hoc networks (VANETs). The Position Spoofing Flood involves continuously broadcasting fabricated location data to mislead navigation systems and cooperative driving applications. This can result in vehicles deviating from their intended routes or engaging in unnecessary braking (Sharma and Jaekel, 2021).

The Constant Position Flood attack repeatedly transmits static but false location data. This disrupts systems like platooning by simulating a stationary vehicle, leading to breakdowns in traffic flow and increased collision risks. Similarly, the Eventual Stop Flood manipulates speed data to simulate a gradual stop, causing unnecessary braking by trailing vehicles and increasing the likelihood of rear-end collisions. These scenarios underline the need for robust position falsification detection mechanisms in modern V2X systems.

Synthetic datasets are artificially generated datasets designed to simulate real-world vehicular network scenarios. They are commonly used for testing and evaluating cybersecurity and anomaly detection techniques when access to real-world data is limited or impractical.

CARACAS Dataset: The CARACAS dataset is a synthetic resource designed to simulate flood attack scenarios in Battery Electric Vehicles (BEVs). The Torque Control Flood repeatedly injects fabricated torque signals, disrupting normal acceleration and braking functions. For example, during cruise control, this attack may cause abrupt deceleration, destabilizing the vehicle and endangering passengers (Kirdi et al., 2024).

The dataset also includes the Signal Saturation Flood, which overwhelms the CAN network with excessive signals, saturating its bandwidth. This saturation delays legitimate messages, impacting critical systems such as steering and braking. The Frequency-Based Flood injects CAN messages at irregular intervals, causing desynchronization among ECUs. This leads to issues in time-sensitive systems such as adaptive cruise control. These scenarios make the

CARACAS dataset an essential tool for studying the impact of cyberattacks on BEV dynamics and testing the performance of IDS solutions.

CANGen Framework: CANGen leverages deep generative models to produce synthetic CAN traces tailored to flood attack scenarios. The Payload Flood involves injecting identical or randomized payloads at high frequencies, saturating the network and delaying critical responses such as emergency braking or airbag deployment (Yin et al., 2024).

The Message Injection Flood targets specific CAN IDs, repeatedly injecting counterfeit messages that override legitimate signals and desynchronize ECU communication. Additionally, the Frequency Flood uses irregular injection patterns to destabilize communication timing, disrupting distributed systems such as anti-lock braking or traction control. CANGen's ability to generate high-fidelity synthetic data for these scenarios makes it a pivotal tool in advancing IDS research.

2.2. State-of-the-Art Cybersecurity in Connected and Automated Technology

This section delves into the current advancements in cybersecurity for connected and automated technology, highlighting the contributions of leading organizations like Vector, Kaspersky, and Symantec. It explores emerging technologies in automotive cybersecurity, identifies potential risks and severity classes associated with autonomous vehicles, and examines the tools and methodologies used for detecting cyberattacks. Together, these insights provide a comprehensive overview of the state-of-the-art measures implemented to protect connected and autonomous systems

Table 2.1. provides a simplified overview of cybersecurity threats, anomaly detection techniques, and datasets in vehicular networks. The studies are grouped into categories and subcategories, with each paper listed under its relevant section.

Table 2.1 - Overview of Cybersecurity Threats, Anomaly Detection Techniques, and Datasets in Vehicular Networks

Category	Subcategory	Articles	Methodology	Evaluation Results
Cybersecurity Threats	V2X Communication Security	(Timčenko and Boštjančič Rakas, 2024)	Proposed a secure V2X communication protocol using blockchain to enhance trust and integrity.	Achieved 98% data integrity verification rate in simulated V2X environments.
Cybersecurity Threats	In-Vehicle Security Vulnerabilities	(Kim et al., 2020)	Designed a CAN-bus anomaly detection mechanism leveraging entropy-based metrics.	Detected 92% of injection attacks with low false positive rates.
Cybersecurity Threats		(Khatri et al., 2021)	Investigated ECU vulnerabilities with fuzz testing and proposed a mitigation framework.	Identified vulnerabilities in 85% of tested ECUs and validated the proposed fixes.
Anomaly Detection	Machine Learning Approaches	(Xiao et al., 2019)	Implemented an SVM-based anomaly detection model for vehicle networks.	Achieved 95% accuracy with low latency in CAN-bus datasets.
Anomaly Detection		(Garg et al., 2019)	Utilized random forest classifiers to detect anomalies in vehicular communication.	Obtained 93% precision and 89% recall on synthetic datasets.
Anomaly Detection		(Hossain et al., 2020)	Proposed a deep autoencoder model for real-time anomaly detection.	Detected anomalies with 90% F1 score on public automotive datasets.
Anomaly Detection	Time-Based Methods	(Khullar et al., 2020)	Developed a survival analysis-based anomaly detection system for vehicular networks.	Achieved significant improvement in detecting time-based attacks, reducing detection delays by 30%.

Category	Subcategory	Articles	Methodology	Evaluation Results
Anomaly Detection	Hybrid Approaches	(Anjum et al., 2021)	Combined neural networks with rule-based systems for comprehensive anomaly detection.	Achieved 96% detection rate with a 10% false alarm reduction.
Datasets	Datasets for Autonomous Vehicles	(Verma et al., 2024)	Introduced a dataset tailored for autonomous vehicle attack scenarios.	Dataset validated for reproducibility and used in multiple anomaly detection frameworks.
Datasets	Attack Scenario Datasets	(Swessi and Idoud, 2021)	Provided a dataset including detailed scenarios for flooding and spoofing attacks.	Widely cited for testing anomaly detection algorithms, with high utility in machine learning training.
Algorithm Development	Algorithm Development	(Rajapaksha et al., 2024)	Proposed lightweight algorithms for in-vehicle intrusion detection systems.	Improved computational efficiency by 20% compared to traditional methods.
Algorithm Development	Deep Learning Methods	(Kim et al., 2020)	Implemented a CNN-LSTM hybrid for advanced attack detection in vehicular networks.	Achieved a detection accuracy of 98% on the benchmark dataset.
Security Metrics	Development of New Metrics	(Bouchouia et al., 2024)	Proposed a comprehensive framework for evaluating attack resilience in CAV networks.	Demonstrated the metric's ability to identify weaknesses in 10 different algorithms.
Security Metrics	Evaluation of Classical Metrics	(Dahiya, 2022)	Evaluated the effectiveness of legacy metrics in	Found that classical metrics remain effective for 70% of

Category	Subcategory	Articles	Methodology	Evaluation Results
			contemporary vehicular networks.	traditional attack vectors.
Attack Detection	Classification of Attack Types	(Anjum et al., 2021)	Categorized attack types in vehicular environments using clustering algorithms.	Provided a taxonomy used in 15 subsequent studies for framework development.
Intrusion Detection Systems	Machine Learning-Based IDS	(Hossain et al., 2020)	Applied a gradient boosting approach for vehicular IDS.	Shown a 95% detection rate with a low false positive rate of 2%.
Intrusion Detection Systems	Hybrid IDS Models	(Kang et al., 2021)	Proposed an IDS combining ML and rule-based approaches for detecting complex attack types.	Demonstrated effectiveness in detecting multi-stage attacks with an accuracy of 97%.
Advanced Techniques	Quantum and Deep Learning Integration	(Salek et al., 2023)	Explored quantum-enhanced deep learning for vehicular anomaly detection.	Reduced training time by 40% while maintaining high detection accuracy.
Advanced Techniques	Adversarial Training Techniques	(Marchiori and Conti, 2024)	Implemented adversarial training to harden models against evasion attacks.	Improved model robustness with a 20% reduction in successful adversarial attacks.

2.2.1 Vector, Kaspersky & Symantec: Services and Contributions to Cybersecurity

In today's rapidly digitalizing world, cybersecurity has become indispensable for both businesses and individuals. The increase in digital threats, diversification of attack methods, and the stringent security standards required by modern technologies have further heightened

the demand for comprehensive cybersecurity solutions. Industry leaders like Vector Consulting Services, Kaspersky, and Symantec play critical roles by addressing complex cybersecurity issues with a wide range of security products and services. This section examines these three companies' contributions to cybersecurity, focusing on their services and impact on the industry.

Vector Consulting Services is renowned for its engineering solutions and security validation methods, particularly in cyber-physical systems and CAV. The company's formal traceability model enhances the reliability of validation processes by systematically linking each security requirement to specific test cases. This model enables the application of security validation procedures in complex, high-risk systems. For example, a gray-box penetration test conducted on an insulin pump demonstrates Vector's approach, which involves linking each requirement to test cases, providing both forward and backward traceability. This systematic framework ensures that security requirements are directly connected to validation processes, enabling comprehensive identification of potential vulnerabilities (Vector Informatik, 2024).

In the gray-box penetration test on the insulin pump, Vector used system architecture information to evaluate both internal and external threats, providing a balanced security analysis. This approach identified critical security vulnerabilities that could threaten the functionality of the device, highlighting the importance of a layered security strategy for cyber-physical devices. Vector also applied a multi-layered security analysis method to validate security at each component, subsystem, and system level, using Key Performance Indicators to complete a rigorous validation process. Through its traceability model and meticulous testing methods, Vector detected all high-risk vulnerabilities with minimal test cases, thereby optimizing the security testing process and enhancing system resilience.

Founded in 1997 in Russia, Kaspersky has built a strong reputation for its antivirus solutions and endpoint security services, serving a global user base of both individual and corporate clients. Although it has faced restrictions in certain regions due to security concerns, Kaspersky continues to maintain its global presence. Kaspersky's offerings include antivirus and internet security tools, which are highly valued for their user-friendly interfaces and reliable protection against threats like malware and ransomware. For corporate clients, Kaspersky provides endpoint protection solutions to secure devices, networks, and data, offering robust threat detection and incident response capabilities (Kaspersky Lab., 2024).

A central element of Kaspersky's cybersecurity strategy is its multi-layered security approach, which combines multiple levels of protection to defend against complex threats. The first layer provides fast, reliable malware detection through masks and hashes.

The second layer employs emulation to run suspicious code in an isolated environment, effectively protecting against web-based threats.

The third layer, a classic detection routine, enables Kaspersky Lab experts to deliver decryption tools and unpackers directly to users for added protection against ransomware.

In the fourth layer, machine-learning models on the client side provide generalization capabilities to detect unknown threats, even when database updates are delayed.

The fifth layer leverages big data and cloud-based detection through the Kaspersky Security Network, ensuring rapid response to emerging threats while minimizing false positives.

The sixth layer uses heuristics based on execution logs, allowing instant data backup and automated rollback for immediate malware neutralization.

Finally, the seventh layer gathers real-time behavioral insights on files, using deep learning models to identify malicious files with minimal data analysis. This multi-layered approach not only enhances detection accuracy but also delivers comprehensive protection for both individual and corporate users by minimizing threat persistence and ensuring a high detection rate, even without frequent updates.

Founded in 1982 in the United States, Symantec initially began as a general software company before focusing on antivirus solutions and security software. Over time, Symantec has established expertise in endpoint security, data loss prevention (DLP), information protection, network security, and threat intelligence. In 2019, Broadcom acquired Symantec, which now operates under the Broadcom brand to provide enterprise-level security solutions. Symantec's endpoint security solutions protect corporate devices from threats like malware and ransomware by utilizing artificial intelligence and machine learning, making it a trusted choice in enterprise security (Symantec Corporation., 2024).

Symantec also provides data loss prevention (DLP) tools that focus on safeguarding sensitive information from unauthorized access. These tools monitor data transfers, enforce security policies, and protect data integrity. Its Advanced Threat Protection (ATP) solutions leverage advanced analytics and threat intelligence to respond to targeted attacks swiftly, giving security analysts the tools to mitigate complex threats effectively. Additionally, Symantec's integrated cybersecurity platforms enable centralized security management within organizations, allowing for efficient management of security protocols and rapid response to security incidents. Through these services, Symantec continues to address the complex security needs of large organizations, reinforcing its leadership position in enterprise security.

Vector, Kaspersky, and Symantec each offer distinct strengths in the cybersecurity sector. Vector specializes in traceability and multi-level validation for high-risk, complex systems, particularly within the realm of cyber-physical and connected environments. Kaspersky's multi-layered security approach provides robust antivirus solutions and threat intelligence, delivering effective protection for both individuals and organizations. Symantec focuses on enterprise security, excelling in data protection, endpoint security, and integrated platforms for large organizations. Collectively, these companies meet the evolving demands of cybersecurity, contributing to a safer digital landscape through innovative and resilient solutions.

The Table-2.2 below shows a comparison of various features of the companies:

Table 2.2 – Features of the Companies (Vector Informatik, 2024, Kaspersky Lab., 2024, Symantec Corporation., 2024)

Features	Vector Consulting Services	Kaspersky	Symantec
Target Customer Base	Primarily corporate, with a focus on CAV and cyber-physical systems	Individual and corporate	Corporate
Core Service Areas	Security validation, traceability, CAV security	Antivirus, endpoint security, multi-	Endpoint security, data loss prevention,

		layered security model	ATP, integrated platforms
Key Security Approach	Formal traceability model, multi-layered validation	Multi-layered security, user and cloud-based analytics	Integrated security platform, advanced threat analytics
Security Layers	Traceability for requirements, layered security testing	Seven-layer security model (hashing, emulation, cloud, etc.)	Multi-level security with centralized management
Highlighted Security Testing	Gray-box penetration tests, internal and external threat analysis	Emulation, machine learning, cloud-based analytics	ATP, centralized management tools, DLP for data protection
Advanced Threat Protection	Threat analytics, CAV-specific security measures	Machine learning and cloud detection, real-time monitoring	AI and ML-backed ATP, integrated threat detection
Key Strength	Traceability model, expertise in CAV security	Strong antivirus and endpoint security, multi-layered approach	Comprehensive enterprise security, extensive integrated solutions
Weaknesses	Limited focus to specific industries	Restrictions in certain countries due to security concerns	Primarily enterprise-focused, lacks individual user orientation

Academic and industrial sectors adopt distinct approaches to addressing cybersecurity challenges in autonomous vehicles, each contributing unique methodologies and perspectives to the field. The security of autonomous vehicles has become an increasingly important area of focus in both academic and industrial spheres. While academic research tends to strengthen theoretical foundations by exploring new security protocols and algorithms, industrial companies transform these theoretical foundations into practical, market-driven solutions. In this context, cybersecurity solutions offered by companies like Vector Consulting Services, Kaspersky, and Symantec combine the knowledge base of the academic world with the real-world needs of industry.

Industrial firms, particularly in the field of autonomous vehicle security, focus on developing solutions that are directly applicable in real-world scenarios. For example, Vector Consulting Services provides specialized solutions for cyber-physical systems and autonomous vehicle security through its traceability model and security validation processes. This practical approach is designed to establish a security validation structure resilient to real-world threats. Kaspersky, with its user-friendly, multi-layered security model, offers solutions that optimize endpoint security and rapid threat response. Symantec has established a solid presence in data security and advanced threat protection by developing integrated platforms tailored to corporate security needs.

In contrast, academic research generally aims at long-term theoretical advancements. Topics such as threat detection using machine learning, cryptographic security protocols, and mathematical modeling of vulnerabilities in cyber-physical systems are prominent in academic research on autonomous vehicle security. These theoretical studies serve as a guide for the industry to develop more advanced security solutions in the future.

Industrial companies prioritize real-world testing to ensure the reliability of their solutions. For instance, Kaspersky uses big data from user devices to analyze and detect threats in real-time, effectively validating its multi-layered security model in practice. Vector tests its security solutions for autonomous vehicles through various penetration tests and multi-layered validation analyses. Similarly, Symantec tests its integrated platforms for large corporate clients using centralized management tools.

In academia, testing and validation processes are primarily conducted through simulations, laboratory experiments, and theoretical analysis. For instance, researchers test machine learning-based threat detection models on datasets through simulations or analyze the vulnerabilities of autonomous systems theoretically. Findings from academic research provide valuable insights for industrial companies to strengthen their testing processes.

Collaboration and joint research between academic and industrial sectors play a crucial role in advancing cybersecurity solutions for autonomous vehicles, combining theoretical insights with

practical applications. Industrial companies and the academic world collaborate extensively, especially in complex areas such as autonomous vehicle security. Kaspersky partners with universities and research laboratories to develop machine learning-based threat detection algorithms, while Symantec undertakes projects with academic institutions to enhance its advanced threat analysis. These collaborations facilitate the transfer of academic findings into practical applications, particularly in artificial intelligence-based threat detection and cybersecurity for cyber-physical systems.

In academia, researchers find opportunities to transform theoretical findings into practical applications through collaboration with industry. For instance, the theoretical discovery of security vulnerabilities can guide the industry in developing solutions to address these weaknesses. New algorithms, security protocols, and AI-based security systems developed in academia can be tested and integrated into real-world applications by industrial companies.

In terms of the impact of academic research on industry, fundamental research conducted in academia forms the basis for future industrial developments. Studies on quantum encryption, biometric security, AI-assisted threat detection, and mathematical modeling of cybersecurity strategies specific to autonomous vehicle security contribute to the industry's ability to offer more resilient and innovative solutions. Thus, companies like Kaspersky may further develop their machine learning models; Vector may enhance its validation processes in autonomous systems, and Symantec could offer more integrated solutions within its corporate security platforms.

In conclusion, the differences and collaborations between academic research and industrial applications foster continuous advancements in cybersecurity solutions for autonomous vehicles. Industrial companies transform the theoretical foundations provided by academic research into practical applications, while the academic world shapes its studies based on industry needs, creating a sustainable information cycle between the two fields. This mutual interaction provides both parties with the opportunity to develop innovative and resilient solutions, ultimately enhancing security in autonomous vehicles.

Table-2.3 in the below highlights the core differences and complementary aspects of the approaches taken by industry and academia in cybersecurity for autonomous vehicles.

Table 2.3 – Comparison of Industry and Academia

Aspect	Industry	Academia
Goal Orientation	Focuses on practical, market-driven solutions tailored for real-world applications in autonomous vehicle security.	Aims at long-term theoretical advancements, developing new algorithms and protocols to strengthen core concepts.
Approach to Security Solutions	Uses frameworks like multi-layered models and traceability for robust, deployable cybersecurity solutions.	Explores innovative methods, such as machine learning and cryptography, to address emerging security challenges.
Testing and Validation	Prioritizes real-world testing through penetration tests, multi-layered validation, and big data analysis.	Relies on simulations, laboratory experiments, and theoretical modeling for validating security frameworks.
Collaboration with Other Sectors	Partners with academia for access to cutting-edge research and to refine machine learning algorithms.	Collaborates with industry to bring theoretical findings to practical use and access real-world data resources.
Impact on Future Developments	Transforms academic findings into practical, scalable applications for current market needs	Provides foundational research that guides the industry in developing advanced security solutions over time
Application of AI and Machine Learning	Implements AI for real-time threat detection and endpoint security.	Develops new AI models and explores their theoretical applications for identifying vulnerabilities.
Focus of Research	Concentrates on user-oriented solutions like endpoint security and integrated platforms for corporate needs.	Investigates broader security theories, such as decentralized systems, blockchain, and federated learning.

2.2.2. Emerging Technologies on Automotive Cyber-Security

In automotive cybersecurity, emerging technologies—particularly generative AI, machine learning, and artificial intelligence—are playing a crucial role in enhancing threat detection and prevention. These advanced approaches are especially valuable in autonomous vehicles, where

they help establish more effective mechanisms to detect and respond to potential security breaches within complex network architectures. Generative AI, for instance, is employed to generate realistic attack scenarios, thereby testing in-vehicle security systems. Machine learning accelerates anomaly detection in critical network components like the CAN bus, identifying unusual patterns that may indicate threats. AI provides a supporting framework, enabling real-time decision-making capabilities for vehicles to respond more swiftly to potential attacks. Within this context, anomaly detection is a key component of vehicle security, continuously refined to better address various cyber threats, serving as a proactive measure to strengthen the resilience of autonomous vehicle systems.

Generative AI: In recent years, Generative AI has garnered significant interest in security applications due to its data generation capabilities. Techniques such as Generative Adversarial Network can be utilized to produce realistic attack scenarios to test the robustness of intrusion detection systems in vehicular networks. For example, GAN-based systems can simulate different types of attacks on the CAN bus, generating adversarial samples that provide diverse training data to strengthen IDS models against a broader range of threats. Other research has also explored data generation techniques to improve IDS performance, with findings showing that realistic simulations of varied attack types enable more flexible and comprehensive training datasets.

Machine Learning: Machine learning plays a pivotal role in anomaly detection within vehicular networks, particularly for detecting anomalies in CAN bus traffic. For instance, recent studies demonstrate the effectiveness of machine learning algorithms like Random Forest, Support Vector Machines, and neural networks in detecting anomalies such as impersonation, DoS, and fuzzy attacks. Models combining recurrent neural networks with attention mechanisms provide high detection accuracy with low computational costs, making them suitable for real-time applications. Supervised learning models, which rely on labeled data, are particularly effective in environments where sufficient attack data is available. Unsupervised methods, such as clustering, are suitable for real-time anomaly detection in dynamic vehicular environments, as they can identify outliers without explicit labels.

Artificial Intelligence: Artificial Intelligence extends beyond machine learning, encompassing rule-based and knowledge-based systems that are essential for real-time decision-making and threat response in vehicular networks. AI algorithms are used to process data from sensors and

network sources, laying the foundation for IDS to detect and respond to security threats effectively. For example, hybrid AI approaches, which combine rule-based detection with machine learning, have demonstrated notable improvements in accuracy and speed for detecting anomalies within CAN bus traffic. These applications highlight AI's broader role in supporting secure and intelligent operations within interconnected vehicle systems.

Anomaly Detection: Anomaly detection is a key component in securing vehicular networks, focusing on identifying unusual patterns that could indicate cyber threats. Statistical methods, such as z-score analysis, provide a foundational approach to detecting outliers in vehicular data, while advanced machine learning techniques offer flexibility for more complex detection tasks. For instance, studies have implemented machine learning models to effectively detect injection attacks, such as flooding and fuzzing, in CAN networks, utilizing techniques like cosine similarity and frequency-based data mining for real-time anomaly identification. Additionally, deep learning models like LSTMs have shown high accuracy in detecting DoS and spoofing attacks by leveraging temporal patterns in CAN bus messages, allowing the system to capture dependencies in sequential data that are common in in-vehicle communications.

This expanded literature review provides a comprehensive theoretical foundation for your thesis, defining the roles of each area in cybersecurity applications and summarizing related work that supports the research framework of your study.

Table 2.4 shows definition of these systems:

Table 2.4. – Definition of Emerging Systems

Category	Definition	Key Techniques	Applications in Vehicular Security	Example Studies
Generative AI	AI focused on generating realistic content based on input data, useful for creating training	GANs, VAEs, Diffusion Models, Transformers	Generating adversarial examples to test IDS; simulating various attack types to enhance training	(Seo et al., 2023) (Salek et al., 2023)

	data and testing IDS robustness.		data for CAN bus security.	
Machine Learning	AI that enables models to learn patterns in data and make predictions without explicit programming.	Supervised (e.g., Random Forest, SVM), Unsupervised (e.g., Clustering), Deep Learning (RNNs)	Detecting impersonation, DoS, and fuzzy attacks; enabling real-time anomaly detection in dynamic vehicular networks; optimizing computational costs for large data processing.	(Hossain et al., 2020) (Xiao et al., 2024)
Artificial Intelligence	Broad AI encompassing machine learning, rule-based, and knowledge-based systems, used for decision-making and real-time threat response.	Rule-Based Systems, Knowledge-Based Systems, Hybrid AI	Real-time processing of CAN bus data; enhancing IDS performance through hybrid models combining rule-based and machine learning techniques.	(Kim et al., 2020) (Kang et al., 2021)
Anomaly Detection	Techniques focused on identifying outliers or unusual patterns that may indicate cyber threats in vehicular networks.	Statistical Methods (e.g., z-score, box plot), Machine Learning (e.g., RNNs, Autoencoders)	Detecting injection attacks (e.g., flooding, fuzzing); enabling real-time CAN bus security; using temporal patterns in data for sequence-based attacks like DoS and spoofing.	(Kwak et al., 2020) (Hossain et al., 2020)

2.2.3 Cyber Security Risks in Autonomous Vehicles

Cyberattacks can prevent autonomous cars from using their sensors. The system might be permanently damaged by this approach. The goal of technology is to reduce the number of mishaps that result from human use. But cyberthreats have resulted from this. Cyberattacks appear to have a high probability of causing mishaps (El-Rewini et al., 2020; Schmittner et al., 2016). Owner-owned GPS sensors have a unique security risk called GPS Spoofing, which can make user usage ineffective. The system may interpret its present or destination position differently when erroneous signals are received (Çuhadar, 2017). Accidents that could happen as a result of steering are another consequence. Redirects may be controlled by using shell codes. The goal of the assault is to take control of the user's rights. Some system codes are maintained in assembly language by Shell Code. Modifications to the codes will result in issues with how the system functions. If these codes are not changed, the system will always be vulnerable to outside control. The user will lose control of the brakes and steering. There will be an uncontrollable increase in airbag, lighting, and other protective device systems (Taeiagh & Lim, 2019). Until the system software is reinstalled, other people will be in charge of managing the system. The potential attacks and the effects they provide are indicated in Table 2.5.

Table 2.5 - Examples of Threats and Risk Values

Possible Attacks	Threats	Effect
Known exploits or remote control in OS	Rootkit or Trojan	Loss of ECU control
Known exploits or BufferOverflow in OS	Deleting Software Tools	Reducing ECU control
Man-in-the-middle Attack	Secretly Listening to the Password to be Used	Obtaining the Connection That Must Be
Compromise and DDoS Attack between ECU and Web Server	Stopping the system to interrupt incoming data	Stopping the ECU
Man-in-the-middle attack in communication	-Sending manipulative data to Web Server	-Preset maintenance data causes unnecessary maintenance
	- Unauthorized Access to Transferred Data	- Specific Access Usage Data

Malware and Sending data to too many ECUs at the same time	Temporarily disables OS or applications	It causes decrease in system functions and decrease in storage space
GPS Spoofing	Changing incoming GPS data	Causes changes in the route to be determined
ARP Spoofing	Keeps the attacker's identity secret	The system cannot correctly identify the attacker

2.2.4 Cyber Security Severity Classes in Autonomous Vehicles

Table 2.6 categorizes the assaults according to the level of victimization that the users would experience. The severity class is supposed to indicate the extent of losses to be incurred. The degree of weight directly relates to the loss that will occur. Losses to users' life safety are categorized by security, losses to personal information by privacy, losses to the budget by cost, and system effect by performance.

Table 2.6 - Cyber Security Severity Classes

Severity Classes	Safety	Confidentiality	Financial	Operational
0	No injuries	No unauthorized access to data	NO financial loss	No impact on operational performance
1	Light or moderate injuries	Configuration data only	Low-level loss	Impact not discernible to operator
2	Severe injuries or moderate injuries for multiple vehicles	Partial data (access to a single update or one application)	Moderate loss	Low losses for multiple vehicles. Operator aware of performance degradation. Indiscernible impacts for multiple vehicles.
3	Life threatening or fatal injuries	Access to complete data.	Heavy loss Moderate losses for multiple vehicles	Significant impact on performance. Noticeable impact for

				multiple vehicles.
4	Life threatening or fatal injuries for multiple vehicles	Access to data from multiple ECUs in the vehicle	Heavy losses for multiple vehicles	Significant impact on multiple vehicles.

2.2.5 Cyber Attack Detection Tools

Several safeguards are in place to prevent cyberattacks. The general measures to be concentrated on will be the detection tools to be employed in the system as a consequence of the research in this study. The system can be made safer by exterior protection mechanisms in addition to the detecting tools that are to be utilized within it. The report goes into great depth on detection techniques. Inspections will be made easier by autonomous systems that have incorporated inspection tools. Regular system audits are necessary to stop cyberattacks. Port scans make these checks feasible (Marquez, 2010; Technical & Str, 2019). In this research, we will talk about port scans. As the attacks are thought to be about gaining control, our investigation will show that control is lost by keeping an eye on how the motions of the vehicles alter. NMAP, Maltego, and Metasploit are the scanning methods that will be applied in this case (Kennedy et al., n.d.; Singh, 2013).

It is necessary to remove malicious malware from the system. Using Metasploit, malware may be detected. Inspections and preventative measures about the vulnerabilities of the cars may be carried out using the metasploit database, which includes all of the system vulnerabilities. IP can be used to establish wireless connections in the event that the assaults do not require a physical connection. Because of the movements of IPs that are known not to exist in the system, control over IP movements can reveal the identity of the attacker. Nmap is the tool to use. All of the user IPs on the network that the machine is linked to are provided by Nmap. The packets that the system will expose foreign IPs to may cause them to be disabled among the IPs that need to be obtained.

We have instruments that will alert us to potential attack points in our system. Our tools may be used to identify and fix system vulnerabilities. By utilizing the existing detection and control tools, it can facilitate system security. Table 3.4 lists the names of our attack detection instruments, their detection techniques, and the information they give us. Throughout the investigation, Nmap, Maltego, and Metasploit were looked at.

Nmap (Network Mapper) is a powerful and widely used open-source tool for network discovery and security auditing. Developed by Gordon Lyon, it is designed to scan and map networks, identifying devices, open ports, operating systems, and services running on a network. Nmap is essential for network administrators, cybersecurity professionals, and ethical hackers to assess network security and detect vulnerabilities. The systems include 65,532 widely recognized ports. Users can access helpful external connections and equipment that will make their lives easier using these ports. NMAP is used to scan for cyberattacks utilizing these connections. In the event that no specific port is given, NMAP searches the 1000 most popular ports. Here, the status of connecting to the system with a foreign, unknown IP is verified as a consequence of scanning external connections and IP addresses. Since every IP address in the system is examined by the scanning method to be employed, identifying the system is also crucial.

In computer networks, ports are closed systems that are used for data transmission and communication. Three fundamental criteria are often used to evaluate a port's condition. The protocol in use is indicated by the port number. The term "service" describes the port-running service names.

"Open", "Closed", and "Filtered" are the three primary categories used to describe port status. A port that is in the open condition might be used to initiate communication since it is considered accessible. The port could be open if it is closed, but it cannot be changed.

A port that has a "filtered status" indicates that a filtering mechanism has blocked it. It is unclear in this instance whether the port is open or closed. The port can be accessed even if it is unfiltered, but no information about its status is available.

The port may be open or filtered when it is in the Open Filtered condition, however the issue has not been fixed. Because open ports might result in unauthorized access or security breaches, this information is crucial to network security and monitoring procedures.

Maltego is a powerful open-source intelligence (OSINT) and graphical link analysis tool used for data gathering, relationship mapping, and investigation of entities such as people, organizations, domains, IP addresses, and social networks. Developed by Paterva, Maltego is widely utilized in cybersecurity, threat intelligence, and digital forensics. Understanding the systems that are going to be attacked is crucial. By gathering both active and passive data on the systems, one must have sufficient knowledge about them. Information that is fixed online for the system about which data collection requests are made is known as passive information. We can obtain this information via search engines, social networking sites, whois and DNS query websites, and websites that gather personal data. IP is the acquisition of DNS and subdomain names through servers and service providers; it is active information (Çelik, 2020). The tool that gives us access to the system's active and passive data is called Maltego. Paterva created it using the Java programming language. There are both paid and free versions available. Open source material is gathered and analyzed by security specialists as well as individuals who wish to advance in this sector (Andress & Winterfeld, 2011; Hai-Jew, 2014). gives the system's IP address and connection details. Domain names and URL addresses can be obtained through scanning all connected systems, intermediaries, connections that provide IP addresses, network connections, receivers to which active data is sent, open source information about senders from which active data is received, all information that the system can access and transmit to the system, websites, IPv4 addresses, e-mail addresses, and so on. With the activation code that may be found on their websites, it is functional (Agus & Pratama, 2019; Ağyol, 2020; Hai-Jew, 2014; Maltego.com, 2023).

Metasploit is a widely used open-source framework for penetration testing and vulnerability assessment. Developed by H.D. Moore and maintained by Rapid7, it provides a comprehensive platform for discovering, exploiting, and validating security vulnerabilities in networks, systems, and applications. Developed since 2003, this open-source platform facilitates the identification and management of security flaws within the system. Perl is used to develop Metasploit 2.0, while Ruby is used to code Metasploit 3.0. It displays the security flaws in programs or services that are operating on computers. On its framework, it has more than 1400 exploits, 800 auxiliary modules, 400 payloads, 30 encoders, and 8 nobs. Existing tools may be

used for vulnerability discovery, attack, and eavesdropping (Bagggett, 2008; Singh, 2013). The database used by Metasploit Framework is unique. There is information accessible, such as the vulnerabilities' descriptions and methods of usage. The explanation of the apertures' function also makes it easier to utilize them. Certain modules make advantage of it, including those that gather data, identify vulnerabilities, and allow users to pass through security barriers without requiring permission. A complete framework with a variety of attack methodologies is Metasploit. The features and tools needed for attacks are included in this framework. Attack techniques known as exploits can be used to turn off a service by taking advantage of holes in the target system. By taking advantage of all vulnerabilities that are deemed faults or do not function in the system, exploits are utilized to render the system useless.

To get over security points, the encoder carries out the required encryption procedures. This is done in an attempt to impede detection of the assault. Attack scripts designed to listen to system connections are called payloads. It can monitor and listen to system connections, including speakers, cameras, and Bluetooth connections. It can also make system modifications after an assault.

Prior to the assault, auxiliary attacks are employed to gather system information. This category includes tools like vulnerability scans, port scans, and service discovery. Bits known as Not Operations (Nops) occupy memory space in order to identify memory locations. It is employed to trick systems and make assault detection challenging. You can alter the program's flow and get to any random location using it.

The following are unique file system characteristics of Metasploit: Data, which Metasploit may handle and modify; The Framework system's primary structural library is called Lib; add-ons or plugins that utilise every feature available in the Framework; Tools, tools that execute commands via the command line; outside, outside sources, outside software, outside source codes; Documentation: Records that provide information for the Framework.

Table 2.7 - Tools to Use for Detecting Cyber Attacks

Attack Detection Tools	Detection Methods	Given Data
NMAP	By scanning the ports connected to IPs,	It gives that Open, Closed, Filtered and Unfiltered

	information about port openings is obtained. Ports where attacks can be made can be detected.	values depending on the openness of the ports.
Maltego	It maps the IP routes by mapping the ports to which the system is connected in the IP routes.	It presents the connection paths that the system has. All external connections are represented by IP paths.
Metasploit	It enables listening to the system listening paths in the database and the paths that can be provided to the system.	The vulnerabilities of the system are presented.



3. LITERATURE REVIEW

This section compiles literature studies from past years to the present. In the first part of the literature review, studies examining the impact of cyber-attacks on connected and autonomous vehicles in the field of machine learning are discussed. These studies are categorized under the main heading: Autonomous Driving Components. The first section includes a compilation of articles covering attacks occurring within these three main areas. In the second part of the literature review, the paper titled "Anomaly intrusion detection method for vehicular networks based on survival analysis" by Mee Lan Han et al. is examined. This paper, along with studies referencing it, is analyzed in detail within this thesis, and the conducted research is explained. Finally, the differences between studies in industry and academia are presented from various perspectives, highlighting the distinct approaches of each field.

3.1. Attacks on Driving System Components

Connected and autonomous vehicles have the potential to significantly enhance driving safety and user experience through their innovative features. However, these technological advancements also create opportunities for various cybersecurity threats that can compromise the digital security of these vehicles. The networked structure of vehicles makes them vulnerable to different types of attacks, which can hinder safe operation and pose serious safety risks for drivers and passengers. In cybersecurity literature, attacks on vehicles are categorized into different classifications based on their goals, targets, and methods. The primary categories commonly addressed in automotive cybersecurity include Manipulation-Based Attacks, Identity-Based Attacks, Service-Based Attacks, and Software-Based Attacks. This classification enables a better understanding of how different types of attacks impact vehicle systems and helps in developing more effective security measures against such threats. Below is a tree diagram, Figure 3.1, that illustrates the types of attacks in each category in detail:

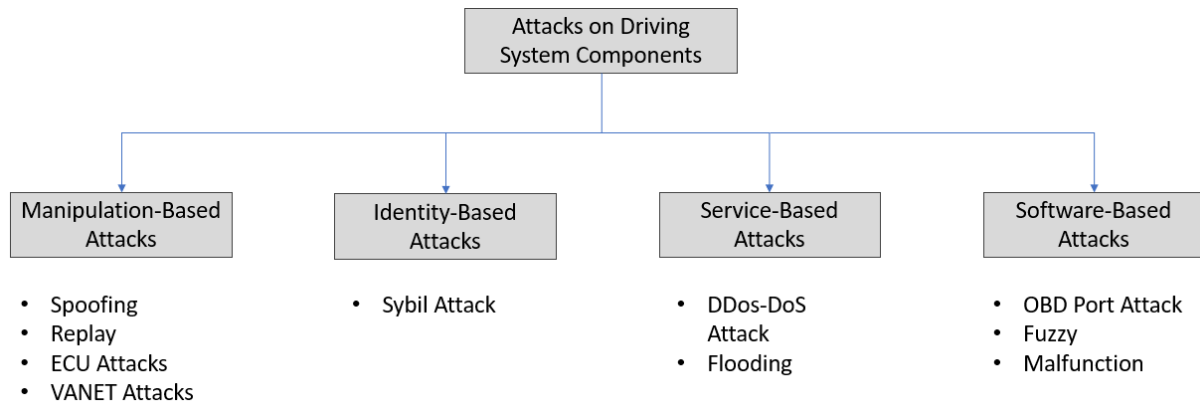


Figure 3.1 - Attacks on Driving System Components

3.1.1. Manipulation-Based Attacks

Manipulation-based attacks are categorized by their fundamental aim to alter a vehicle's data or communication pathways, ultimately disrupting system integrity and reliability. These attacks specifically target the inputs and outputs of a system, attempting to modify, replace, or corrupt data in order to generate false information or provoke unplanned behaviors that the system was not designed to handle. By altering sensitive data or corrupting sensor outputs, these attacks can create serious safety risks for connected and autonomous vehicles, which rely heavily on accurate data for safe operation. This category of attacks includes methods like Spoofing, Replay attacks, ECU attacks, and VANET attacks, each of which manipulates vehicle data in a distinctive manner.

GPS Spoofing is placed in this category because it directly feeds incorrect data to the GPS system, deceiving the vehicle into interpreting a false location. By corrupting the GPS input, this manipulation can misdirect navigation systems, causing the vehicle to navigate to unintended destinations.

LiDAR spoofing is a cyberattack that targets the LiDAR sensor systems used in autonomous vehicles to detect and interpret surrounding objects and environmental features. In this attack, an external source sends false laser pulses to the vehicle's LiDAR system, mimicking real-world obstacles or manipulating distance data. By altering the LiDAR readings, attackers can deceive the vehicle into perceiving nonexistent obstacles or misinterpreting the size and position of actual objects. This manipulation can cause the vehicle to take unnecessary evasive actions, such as sudden braking or steering adjustments, potentially leading to accidents or operational

disruptions. LiDAR spoofing is particularly dangerous because autonomous systems heavily rely on accurate sensor data to make real-time driving decisions, meaning that even slight inaccuracies can have critical safety implications. As autonomous vehicles become more widespread, mitigating LiDAR spoofing threats requires advanced sensor validation techniques and robust security protocols to maintain data integrity.

Replay attacks are included in manipulation-based attacks because they deceive the system into accepting outdated or previously valid data as if it were new, leading to inappropriate system responses. This type of manipulation can lead a vehicle to act based on inaccurate information, potentially jeopardizing safety if, for example, the replayed data involves an old location or control command.

ECU attacks disrupt the proper functioning of the vehicle's control units, which are critical for managing various vehicle subsystems, including braking, steering, and acceleration. By manipulating ECU data, attackers can cause malfunctions or incorrect vehicle behavior, such as activating or deactivating essential functions unexpectedly. This type of manipulation poses significant risks, as it directly interferes with the core functions necessary for vehicle control.

Lastly, VANET attacks manipulate the data distributed within vehicle networks. By injecting or altering information within the network, these attacks can mislead other vehicles, disrupting cooperative driving efforts and potentially leading to unsafe driving conditions. By spreading false information, attackers can cause multiple vehicles to react inappropriately, which is particularly dangerous in environments that rely on synchronized vehicle communication.

Each of these attacks falls under the manipulation-based category because they alter data or communication flows integral to the vehicle's correct functioning. These disruptions threaten not only data integrity but also the broader operational safety of connected and autonomous vehicles, as they rely on accurate and reliable data to make real-time decisions.

Figure 3.2 shows the related articles of manipulation-based attacks.

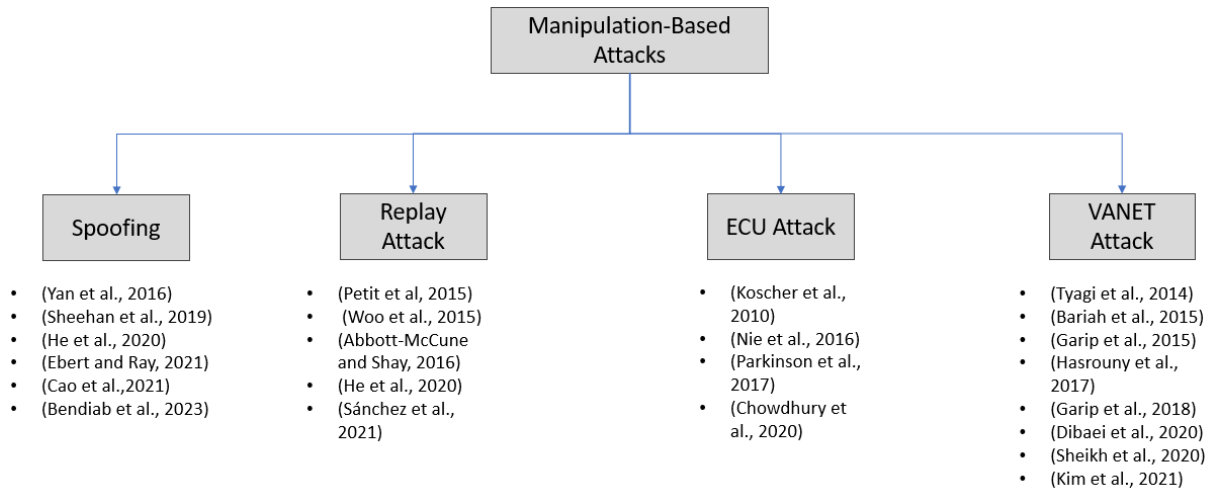


Figure 2.2 – Manipulation-Based Attacks Papers

Sheehan et al. present a Bayesian Network (BN) model to classify cyber-risks for connected and autonomous vehicles (CAVs), integrating data from the National Vulnerability Database (NVD) and the Common Vulnerability Scoring System (CVSS). The model addresses the lack of historical cyber-attack data for CAVs by applying probabilistic assessment techniques, enabling risk prediction for vehicle components such as GPS systems. Through a case study on GPS vulnerabilities, the authors demonstrate how cryptographic measures can reduce cyber-risk scores, making the model a valuable tool for manufacturers, insurers, and regulators to assess and manage cyber-risks in CAV systems proactively (Sheehan et al., 2019).

Cao et al. demonstrate the vulnerability of LiDAR sensors in autonomous vehicles to spoofing attacks by presenting a novel laser-based spoofing approach on moving targets. Using an automated tracking system integrated with a pan-tilt laser spoofer, they successfully injected up to 100 fake points into the LiDAR data, creating non-existent obstacles in the vehicle's perception system. This attack exploits moving LiDARs, which previous studies had only tested on static systems. The study highlights the potential dangers of such attacks, which could cause sudden emergency responses like braking or steering in autonomous vehicles, emphasizing the need for improved sensor security in dynamic environments. (Cao et al., 2021).

Bendiab et al. analyze the security and privacy challenges faced by autonomous vehicles, proposing a combined approach that integrates blockchain and artificial intelligence (AI) as a solution. The authors classify prevalent threats, such as spoofing, data manipulation, and denial-of-service attacks, and highlight how blockchain's secure, decentralized ledger can enhance

data integrity, while AI improves threat detection through real-time analytics and anomaly recognition. They identify research gaps and suggest that future work should focus on the seamless integration of these technologies, aiming to build resilient and transparent systems that safeguard intelligent transportation networks (Bendiab et al., 2023).

El-Rewini et al. provide an in-depth review of cybersecurity threats targeting the sensing layer in modern vehicles, emphasizing the vulnerabilities of vehicle dynamics and environmental sensors. As vehicles increasingly rely on sensors for critical functions such as adaptive cruise control and lane-keeping, they face threats like spoofing, jamming, and denial-of-service attacks. The authors explore attacks on various sensors, including LiDAR, radar, and GPS, and discuss potential countermeasures, such as signal modulation and sensor fusion. They argue for advanced security frameworks leveraging machine learning, IoT, and blockchain technologies to secure vehicular sensor data, given the rising adoption of autonomous and connected vehicles (El-Rewini et al., 2020).

Yan et al. investigate the vulnerabilities of sensors used in autonomous vehicles, focusing on ultrasonic sensors, millimeter-wave radars, and cameras. Using a Tesla Model S as a test platform, they conduct jamming and spoofing attacks to disrupt these sensors without physical contact. The researchers demonstrate that low-cost, off-the-shelf hardware can cause temporary "blindness" in sensors, potentially leading to dangerous situations such as unexpected braking or steering errors. To mitigate these risks, they propose both hardware and software countermeasures aimed at improving sensor resilience against contactless attacks, emphasizing the importance of secure sensor design for autonomous vehicle safety (Yan et al., 2016).

Woo, Jo, and Lee demonstrate a practical long-range wireless attack on a connected car using a malicious smartphone app. This app allows an attacker to access the vehicle's CAN network remotely via the driver's smartphone, manipulating vehicle functions by sending unauthorized CAN data frames. The authors also propose a security protocol to mitigate these vulnerabilities, focusing on real-time data encryption and authentication compatible with CAN's limited data payload. Their findings underscore the critical need for robust cybersecurity measures in the design of connected vehicle systems to prevent remote intrusions that compromise vehicle safety (Woo et al., 2015).

Abbott-McCune and Shay present a study on hacking techniques and the development of a benchtop simulator for the modern automotive Controller Area Network (CAN). Their research emphasizes the vulnerabilities of the CAN bus, a primary communication protocol in vehicles, which was originally designed without security due to assumptions of network isolation. By capturing and replaying network traffic, they demonstrate how malicious actors could manipulate various in-vehicle functions, such as brakes and steering. The benchtop simulator offers a safe environment for testing these vulnerabilities, enabling further research into securing CAN networks from cyber-physical attacks (Abbott-McCune and Shay, 2016)

He et al. examines cybersecurity challenges in the Over-The-Air (OTA) update process for intelligent connected vehicles, focusing on risks such as data tampering and transmission interception. The authors outline the OTA system architecture, including both service platform and vehicle components, and propose a multi-layered security strategy addressing authentication, data integrity, and encrypted communication. Using a specialized darkroom environment, they develop a comprehensive test method that includes threat modeling, risk assessment, and penetration testing to evaluate OTA update security. This framework aims to improve OTA cybersecurity and ensure safe remote software updates for connected vehicles (He et al., 2020).

Sánchez et al. address the detection of replay attacks in autonomous vehicles, emphasizing the limitations of traditional linear approaches due to the nonlinearities inherent in vehicle dynamics. They propose a method utilizing a bank of quadratic parameter varying (QPV) observers, each designed to be insensitive to replay attacks on specific sensor channels. This observer-based approach enables the system to detect and localize replay attacks effectively, using simulation results to validate the decision algorithm's accuracy. The study underscores the importance of QPV frameworks in enhancing the resilience of autonomous vehicles against sophisticated cyber threats (Sánchez et al., 2021).

Important ECUs in connection to vehicle security were thoroughly examined by Koscher et al., in their research in (Koscher & Karl, 2010). The authors conducted a thorough investigation of the security of the Controller Area Network, identifying issues such inadequate access control, a dearth of authentication fields, and susceptibility to denial-of-service attacks. Utilizing techniques including CAN packet sniffing, target probing, reverse engineering, and fuzzing, the researchers validated their attack plan. Their findings were noteworthy since they showed that fuzzing is a powerful method for identifying CAN vulnerabilities.

In (Salfer & Mardin, 2014), Salfer and Mardin discussed the automation of vulnerability assessment by putting forth a systematic and automated measurement technique based on software pictures and ECU development data. This methodology included characteristics such as penetration testing and code review into metrics that were designed for easy evaluation. A diagnostic protocol, driver-based user interfaces, external interfaces such as USB, and low-level hardware interfaces were all included in the attack interface.

Shukla investigated the vulnerabilities of popular ECU software in a follow-up study (Shukla & Siddharth, 2016), outlining how to target networks like CAN and reviewing security updates currently implemented for the tri-core architecture. Exploiting ECU software weaknesses was the focus of his study, which focused on the Automotive Open System Architecture (AUTOSAR) standard for intelligent mobility. Shukla's study uses a tri-core microcontroller to assess vulnerabilities in order to find and exploit more weaknesses.

Nie et al. present a remote attack on the Tesla Model S, successfully achieving control over the vehicle's in-vehicle systems in both Parking and Driving modes. This complex attack chain exploits vulnerabilities in Tesla's Wi-Fi and cellular interfaces, leveraging weaknesses in the vehicle's browser to gain initial access. Through privilege escalation and bypassing AppArmor restrictions, the attackers accessed and injected malicious CAN messages into the CAN bus, enabling unauthorized manipulation of critical vehicle functions like braking and acceleration. Tesla responded promptly with an over-the-air update, implementing code-signing protections and strengthening AppArmor policies to mitigate similar attacks. This research underscores the necessity for robust, multi-layered cybersecurity measures in connected vehicles (Nie et al., 2016).

Three steps were identified by Halahan and Chen in research on the potential for automotive hacking (Halahan & Chen, 2017): remote compromise, message injection through physical or cyber components, and allowing the ECU to do desired tasks. Indirect physical access, long- and short-range wireless communications, and dongle-based attacks were some of the hacking methods employed.

Chowdhury et al. conduct a detailed analysis of the cybersecurity threats facing self-driving cars, emphasizing real-life attacks and their associated vulnerabilities. The authors categorize potential cyber-attacks, including remote intrusions, malware, man-in-the-middle, denial-of-

service, ransomware, spoofing, and Sybil attacks, each posing unique threats to autonomous vehicle functionality and safety. Through case studies and examples, they highlight incidents where hackers exploited weaknesses in systems like ECUs, GPS, and sensors, resulting in unauthorized control over crucial vehicle functions. The authors also review countermeasures, such as cryptographic protocols, secure communication architectures, and government policies, recommending robust, proactive security solutions to mitigate these threats and ensure safe deployment of autonomous vehicles (Chowdhury et al., 2020).

Tyagi et al. investigates security threats within Vehicular Ad Hoc Networks, focusing on the unique vulnerabilities and risks associated with vehicle-to-vehicle and vehicle-to-infrastructure communications. VANETs, intended to enhance road safety and driving comfort, are susceptible to a variety of cyber attacks, such as jamming, impersonation, and bogus information attacks, which can significantly disrupt network functionality and safety. The authors examine these security issues in depth, highlighting how VANETs are affected by their dynamic topology, frequent node changes, and the continuous need for data integrity. To address these challenges, the paper suggests improving network architectures with robust security features and offers recommendations for enhancing data protection and network resilience (Tyagi et al., 2014)

Bariah et al. provide a comprehensive survey of recent advancements in VANET security, analyzing the unique vulnerabilities and threats posed by the network's dynamic and decentralized structure. They categorize the main security requirements for VANETs—such as authentication, integrity, confidentiality, and non-repudiation—that are critical for ensuring reliable and secure vehicle-to-vehicle and vehicle-to-infrastructure communications. The authors explore various types of attacks, including Sybil, jamming, and GPS spoofing, and examine prominent security mechanisms like Public Key Infrastructure (PKI) and ID-based cryptosystems. Emphasizing the need for robust security solutions, simulation tools, and comprehensive frameworks, they underscore the importance of addressing these vulnerabilities to ensure secure VANET deployment for future intelligent transportation systems. (Bariah et al., 2015)

Hasrouny et al. explore security challenges and solutions within Vehicular Ad Hoc Networks, a critical technology aiming to enhance traffic safety and efficiency through vehicle-to-vehicle and vehicle-to-infrastructure communication. The authors categorize VANET security

requirements, such as authentication, confidentiality, integrity, and non-repudiation, and highlight prominent attack types, including Sybil, denial of service, and GPS spoofing, which threaten network reliability and user safety. Their study examines current security architectures and protocols, evaluates various solutions against established criteria, and identifies significant open issues, including trust evaluation, revocation processes, and self-organization challenges in dynamic environments, thus underscoring areas for further research in VANET security frameworks. (Hasrouny et al., 2017)

Garip et al. (2018) introduce BOTVEILLANCE, a vehicular botnet surveillance attack targeting pseudonymous systems in Vehicular Ad Hoc Networks. By leveraging botnet-enabled vehicles, the attack enables long-range, global tracking of target vehicles without additional hardware, bypassing traditional pseudonym-changing defenses. Through cooperative tracking and secret inter-vehicle communication, the botnets track target vehicles across 85% of their route, accurately predicting destinations in 90% of cases. The authors suggest enhancements to existing VANET location privacy measures, highlighting the need for robust, adaptive security solutions to counter such sophisticated surveillance threats (Garip et al., 2018)

Sheikh et al. provide a comprehensive survey on security and privacy issues in vehicular ad hoc networks and vehicular cloud computing. They discuss various threats and attack vectors, including location privacy issues, trust management challenges, and communication vulnerabilities, which affect VANETs and VCC systems. The authors elaborate on cryptographic and trust-based solutions, emphasizing the need for robust security frameworks to protect sensitive vehicular data and ensure secure vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. They also present the evolution of VANETs into VCC, proposing future research directions to address security gaps and enhance privacy measures in vehicular networks (Sheikh et al., 2020)

Dibaei et al. provide a comprehensive overview of cybersecurity challenges in the realm of intelligent connected vehicles. As advancements in automation and connectivity expose these vehicles to various cyber threats, including in-vehicle and vehicle-to-everything communication attacks, this study systematically categorizes these threats and identifies major types, such as data theft, hijacking, and denial-of-service. The authors explore four primary defensive approaches—cryptography, network security, software vulnerability detection, and malware detection—examining each for effectiveness in safeguarding ICVs. They highlight

critical vulnerabilities, particularly those amplified by 4G and 5G technologies, and recommend future research directions, including lightweight authentication, software-defined security, and enhanced deep learning applications, to enhance vehicle resilience against evolving cyber threats (Dibaei et al., 2020).

Kim et al. provide a systematic review of cybersecurity vulnerabilities and defense mechanisms for autonomous vehicles, analyzing attack types across key components such as autonomous control systems, driving components, and vehicle-to-everything communication. They categorize attacks into those targeting the electronic control units, sensors, and communication protocols, highlighting threats like sensor spoofing, denial-of-service, and remote ECU manipulation. To counteract these threats, the authors examine security architectures, intrusion detection systems, and anomaly detection based on artificial intelligence and big data techniques. Emphasizing the rapid evolution of smart cities, they advocate for integrating advanced AI-based defenses within intelligent transportation frameworks to improve resilience and safety in autonomous vehicle networks (Kim et al., 2021).

3.1.2. Identity-Based Attacks

Identity-based attacks focus on exploiting or manipulating identity information within a system to gain unauthorized access or privileges. These attacks target the core of system authentication and authorization processes, attempting to compromise the security framework by either creating false identities or altering legitimate identity data. By simulating valid credentials or creating multiple fake identities, identity-based attacks expose critical vulnerabilities in the system's ability to distinguish between authentic and malicious actors. This category of attacks is particularly harmful in connected vehicle environments, where identity authentication ensures secure communication between vehicles and infrastructure, thus directly affecting both system integrity and safety.

A prominent example of this category is the Sybil Attack. In a Sybil Attack, an attacker creates multiple fake identities within a network, allowing them to gain undue control over network activities. This proliferation of false identities can lead to malicious actors obtaining increased privileges within the network, which undermines trust in network data. For example, in a vehicle network, a Sybil Attack could enable a single attacker to influence network communications as if they were multiple independent entities. This could compromise decision-

making processes within a vehicle network, affecting traffic patterns, vehicle spacing, or route selection. Because these fake identities can appear to be legitimate, the network struggles to detect and eliminate them, posing serious security risks.

Eavesdropping is a passive cyberattack in which an unauthorized entity intercepts and monitors communication between two or more parties without altering the transmitted data. This type of attack exploits the lack of encryption or insufficient security measures within the communication channels, allowing attackers to capture sensitive information such as authentication credentials, location data, and personal information. In the context of connected and autonomous vehicles, eavesdropping poses significant risks as it can expose critical data shared between vehicles, infrastructure, or cloud services. This breach of confidentiality may enable attackers to analyze vehicle movements, user behaviors, or system protocols, potentially facilitating more sophisticated attacks like man-in-the-middle or replay attacks. As vehicle networks increasingly rely on real-time data exchange for safe and efficient operations, protecting against eavesdropping is essential to ensure data privacy and prevent unauthorized access to sensitive information. Robust encryption protocols and secure communication standards are necessary to mitigate eavesdropping risks and maintain the integrity of vehicle communication systems.

These attacks underscore the need for robust authentication mechanisms in vehicle networks. Without effective identity verification, connected and autonomous vehicles are susceptible to deceptive tactics that can lead to serious safety risks. By compromising the system's ability to accurately identify authorized entities, identity-based attacks create openings for unauthorized access and control. As connected vehicles rely heavily on accurate identity and location data for network coordination, identity-based attacks directly threaten the secure functioning of these systems, putting both vehicle and driver safety at risk.

Figure 3.3 shows the related articles of identity-based attacks:



Figure 3.3 – Identity-Based Attacks Papers

Brooks et al. addressed evolving security challenges in modern automotive systems, particularly in the context of increasing connectivity and embedded computing. Traditional anti-theft measures, such as alarms and immobilizers, are now supplemented by advanced communication protocols for safe driving and vehicle-to-vehicle interactions, which expose vehicles to cyber threats. The study categorizes various attacks, including physical breaches, wireless exploits, and cryptographic attacks on ECU systems. Additionally, it discusses vulnerabilities in VANETs, highlighting the potential for Sybil attacks. The authors call for improved cryptographic protocols and comprehensive security strategies to address these threats and secure the automotive ecosystem (Brooks et al., 2009).

Baza et al. propose a Sybil attack detection scheme for VANETs, which combines proof-of-work (PoW) and location verification to prevent malicious vehicles from masquerading as multiple nodes. In their approach, roadside units (RSUs) issue time-stamped location proofs to vehicles, creating unique anonymous trajectories. Each vehicle must solve a computational puzzle (PoW) before receiving a new proof, limiting its ability to generate multiple fake identities in areas with low RSU density. The event manager uses these trajectories and a graph-based maximum clique algorithm to detect Sybil nodes. Extensive experiments indicate that this scheme effectively reduces false negatives while maintaining low computation and communication overhead, improving the reliability of traffic management systems (Baza et al., 2020).

Bernardini et al. analyze the security and privacy challenges in vehicular communications, focusing on the vulnerabilities introduced by the increasing connectivity of modern cars. The authors categorize security requirements such as authentication, integrity, and confidentiality while examining specific challenges across intra-vehicle, inter-vehicle, and gateway communications. They highlight the risks posed by vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, which can be exploited for attacks like eavesdropping, spoofing, and denial of service. Bernardini et al. advocate for robust security measures, including encryption, authentication protocols, and standardized frameworks, to address the growing threat landscape in connected automotive systems (Bernardini et al., 2017)

Sommer et al. introduce an automotive portscanner designed to support penetration testing by detecting ECUs and identifying network protocols in vehicle systems. Their tool automates the process of extracting transport and diagnostic protocols, enhancing black-box penetration testing by allowing testers to gather detailed information about ECUs, services, and subservices. This functionality is extended to include protocol-specific use cases, such as fuzz testing, vulnerability scanning, and exploitation of vehicle networks. The portscanner's automatic detection of CAN-based protocols like ISO-TP and UDS enables more efficient and comprehensive security assessments for modern automotive networks (Sommer et al., 2019)

Ebert and Ray discuss the application of Grey-Box Penetration Testing (GBPT) in automotive cybersecurity, emphasizing its effectiveness in identifying vulnerabilities by combining elements of white-box and black-box testing. They argue that GBPT, which includes knowledge of system architecture and threat models, is more efficient and offers broader coverage than traditional black-box testing, which lacks architectural insights. Their method incorporates Threat Analysis and Risk Assessment (TARA) to prioritize high-risk areas, resulting in fewer test cases, reduced false positives, and better traceability. They advocate for GBPT as a robust testing approach, essential for ensuring cybersecurity in increasingly connected and autonomous vehicle systems (Ebert and Ray, 2021).

3.1.3. Service-Based Attacks

Service-based attacks focus on disrupting or obstructing the functionality of critical services within a system by overwhelming network resources or exceeding system capacity. These

attacks interfere with service availability, targeting the network's ability to reliably deliver essential functions. By deliberately consuming bandwidth or processing power, service-based attacks create conditions in which systems struggle to respond to legitimate requests, rendering them ineffective or even entirely inoperable. This category of attacks is especially problematic for connected and autonomous vehicles, which rely on uninterrupted service availability for real-time decision-making, communication, and navigation. When essential services are disrupted, vehicles may experience significant delays in response times, potentially endangering driver and passenger safety.

A prominent example within this category is the Distributed Denial of Service and Denial of Service attacks. In these attacks, an attacker floods the network with an overwhelming amount of traffic, either from multiple distributed sources in a DDoS attack or a single source in a DoS attack. The resulting overload consumes the system's available bandwidth and processing capacity, leading to service disruptions or complete inaccessibility. In a connected vehicle context, a DDoS attack could prevent a vehicle from communicating with traffic management systems or neighboring vehicles, which could impact route optimization, safety alerts, or cooperative driving maneuvers.

Flooding attacks represent another form of service-based disruption by sending an excessive number of requests to a target system or network. Unlike DDoS and DoS attacks that aim to disable the network entirely, flooding attacks specifically target service slowdowns by overloading the system's request-handling capacity. In a vehicle network, flooding attacks can interfere with essential communications, causing delays in service responses. For instance, a vehicle that relies on real-time data from surrounding vehicles or infrastructure may face delays or inaccurate data due to flooding interference, which may affect its ability to respond appropriately to dynamic traffic conditions.

Flooding attacks in vehicular networks aim to overwhelm the communication system, particularly the CAN bus, by sending an excessive volume of messages in a short period. This high message volume consumes network bandwidth and processing resources, disrupting normal communication between vehicle components and potentially leading to delays or complete denial of critical operations, such as braking or steering. In a flooding attack, legitimate messages are often delayed or discarded due to the network congestion caused by the flood of false requests, which can compromise vehicle safety and functionality. Detecting

flooding attacks typically involves monitoring traffic patterns for sudden surges in message volume and employing machine learning models or threshold-based methods to identify anomalies. Effective countermeasures focus on rate-limiting, prioritizing legitimate traffic, and reinforcing IDS to handle abnormal traffic loads (Han et al., 2018).

Service-based attacks pose unique challenges in the context of connected and autonomous vehicles, where continuous service availability is essential. These attacks interrupt vital communication and decision-making processes, resulting in reduced operational reliability and increased risk during vehicle operation. By obstructing the proper functioning of critical services, service-based attacks endanger the safety of vehicle systems, as they compromise the system's ability to respond quickly and accurately to real-world conditions.

Figure 3.4 shows the related articles of service-based attacks:

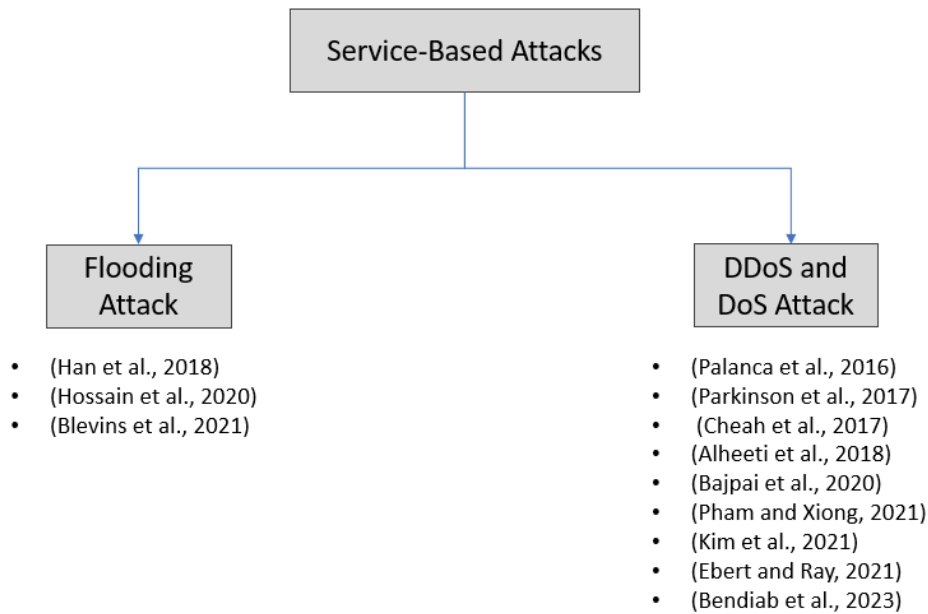


Figure 3.4 – Service-Based Attacks Papers

Han et al. propose an anomaly intrusion detection method for vehicular networks based on survival analysis, focusing on identifying malicious CAN messages without requiring semantic knowledge of CAN IDs. Using normal and abnormal driving data from three vehicle models, they assess detection performance under three attack scenarios: flooding, fuzzy, and

malfunction. Results indicate high detection accuracy and low computational costs, suggesting the method's suitability for real-time processing. The study emphasizes that long-cycle CAN IDs impact detection accuracy and concludes that effective intrusion detection in vehicle networks can significantly enhance driver safety and vehicle security (Han et al., 2018).

Hossain et al. propose a LSTM-based intrusion detection system for detecting attacks on the CAN bus in modern vehicles. Given the CAN protocol's lack of built-in security features, the authors create a dataset by capturing normal CAN traffic and injecting Denial of Service , Fuzzing, and Spoofing attacks. Their LSTM-based model achieves a high detection accuracy of 99.995%, outperforming the Survival Analysis method. By employing optimal hyperparameters, the model effectively identifies attack patterns, demonstrating the utility of deep learning for CAN bus cybersecurity and the importance of fine-tuning for real-time vehicular attack detection (Hossain et al., 2020).

Blevins et al. present a benchmark study on time-based IDSs for the CAN protocol used in vehicles. They evaluate four IDS methods using the ROAD dataset, which contains real, non-simulated attacks with physically verified impacts. Their findings indicate that "distribution-agnostic" methods, which detect anomalies in distribution-related statistics, outperform "distribution-based" methods by at least 55% in precision-recall area under the curve. Additionally, they introduce an aftermarket OBD-II plugin capable of deploying the best-performing IDS method in real-time, providing an efficient and accessible solution for enhancing vehicular cybersecurity (Blevins et al., 2021).

Palanca et al. introduce a stealthy, selective link-layer denial-of-service attack against automotive networks using the CAN bus protocol. The authors exploit the CAN protocol's vulnerability, specifically its error handling and fault confinement mechanisms, to prevent targeted ECUs from communicating without injecting any detectable frames, rendering conventional frame-level intrusion detection systems ineffective. The attack is demonstrated on a modern vehicle, showcasing its feasibility and minimal cost requirements. Palanca et al. emphasize the urgent need for improved CAN network security measures, as the proposed attack poses serious risks to the safety and reliability of automotive systems (Palanca et al., 2016).

Parkinson et al. examines the cybersecurity risks facing autonomous and connected vehicles, emphasizing the increasing vulnerability as vehicle connectivity and automation evolve. Through an extensive literature review, the authors highlight the main categories of threats, such as in-vehicle and V2X attacks, data privacy issues, and various denial-of-service attacks. These threats expose vehicles to risks like hijacking, tracking, and unauthorized data access. Moreover, the study identifies significant gaps in knowledge, stressing the need for advanced cryptographic techniques, rigorous monitoring systems, and enhanced resilience measures to safeguard both vehicle functionality and user privacy. The authors advocate for a proactive approach, suggesting that future research should address these vulnerabilities to ensure a secure foundation for the development of autonomous vehicle technologies (Parkinson et al., 2017).

Cheah et al. propose a systematic approach to evaluate the security of the automotive Bluetooth interface, focusing on vulnerabilities in modern vehicles that could be exploited for cyberattacks. They develop a framework using attack tree modeling and penetration testing to assess potential weaknesses in Bluetooth connections within infotainment systems. Their proof-of-concept tool reveals multiple vulnerabilities, including legacy pairing mechanisms and open communication channels, which could allow unauthorized data access or denial-of-service attacks. The authors emphasize the need for updated security practices, including improved Bluetooth protocols and user alerts, to enhance vehicular cybersecurity (Cheah et al., 2017).

Alheeti et al. introduce an intelligent security system for autonomous vehicles that leverages infrared sensors to enhance vehicular network security. This system, based on Integrated Circuit Metrics (ICMetrics) technology, uses unique identifiers created from the bias values of infrared sensors, making it effective in detecting abnormal or malicious behavior within Vehicular Ad Hoc Networks. The proposed ICMetrics-based Intrusion Detection System (ICMetric-IDS) is shown to be efficient in identifying both internal and external threats, including Sybil and denial-of-service attacks, through a process that combines feature extraction, preprocessing, and anomaly detection. Simulation results confirm the system's accuracy and low false positive rate, highlighting its potential to secure the external communications of self-driving and semi-autonomous vehicles (Alheeti et al., 2018).

Bajpai et al. examine the feasibility of ransomware attacks targeting the In-Vehicle Infotainment (IVI) systems in modern automobiles. They identify the expanding threat surface due to increased connectivity and demonstrate how automotive systems, particularly IVIs, are

vulnerable to various ransomware tactics. Through simulations, they show that attackers could exploit configuration errors, remote login services, and known vulnerabilities in IVIs to achieve denial-of-data, denial-of-service, and denial-of-privacy attacks. Their experiments on a QNX-based IVI system highlight potential consequences, such as data encryption and resource exhaustion attacks, which could immobilize IVI functionality until a ransom is paid. The authors call for stronger security measures in automotive design to counter these evolving threats (Bajpai et al., 2020).

Pham and Xiong provide a comprehensive survey of security attacks and defense techniques for connected and autonomous vehicles, drawing on 189 research papers from 2000 to 2020. They classify attacks based on access requirements and attack motives, with a focus on targeted CAV components such as electronic control units, GPS, LiDAR, radar, and camera systems. The authors highlight vulnerabilities in vehicle-to-vehicle and vehicle-to-infrastructure communication channels, which expose CAVs to threats like spoofing, jamming, and denial of service. In addition, they discuss various defensive approaches, including anomaly-based intrusion detection systems, encryption methods, and redundant sensor information, while emphasizing gaps between academic research and industrial applications. The study underscores the need for robust security solutions to protect CAVs against emerging threats and improve the resilience of intelligent transportation systems (Pham and Xiong, 2021).

Bendiab et al. examine the security and privacy challenges faced by autonomous vehicles, proposing an integrated approach using blockchain and artificial intelligence to address these issues. They categorize threats such as data manipulation, spoofing, and denial-of-service attacks, noting that blockchain's secure, decentralized ledger can ensure data integrity, while AI enhances real-time threat detection through pattern recognition. The authors highlight current gaps in research and suggest future work should focus on the seamless integration of these technologies to build resilient, transparent autonomous vehicle systems that ensure both security and privacy within intelligent transportation networks (Bendiab et al., 2023).

3.1.4. Software-Based Attacks

Software-based attacks exploit vulnerabilities in a system's software to execute malicious actions, compromise security, or disrupt normal operations. By targeting weak points in the code or protocols, these attacks pose a significant threat to connected and autonomous vehicles,

which increasingly rely on complex software systems for navigation, communication, and control. Software-based attacks are particularly dangerous because they can directly alter or damage critical functions within the vehicle, affecting both operational integrity and passenger safety. This category of attacks includes various methods that leverage software flaws to gain unauthorized access or disrupt system processes, highlighting the importance of robust security measures in vehicle software design.

The OBD Port Attack is another form of software-based attack specific to vehicles. By accessing the vehicle's diagnostic port, attackers can gather information or interfere with the vehicle's software. This port is generally intended for maintenance and diagnostic purposes, but when exploited, it can provide direct access to the vehicle's systems, allowing attackers to modify critical functions or collect sensitive data.

Fuzzy and Malfunction attacks target the protocols and data flows within the software, causing the system to behave abnormally or generate erroneous outputs. By manipulating or corrupting the data that the system relies on, these attacks disrupt regular operations, potentially leading to hazardous scenarios. In connected vehicles, where seamless communication between subsystems is vital, fuzzy and malfunction attacks can compromise the system's reliability, affecting safety-critical functions like braking, steering, or acceleration.

In order to interfere with regular communication and introduce weaknesses, fuzzy attacks target vehicular networks by introducing erroneous or random data into the system, particularly the CAN bus. In order to confuse the system and maybe lead to incorrect behavior in vehicle functions, these attacks entail sending unexpected or illegitimate signals that differ from the usual structure or expected values. By creating odd patterns, fuzzy attacks can overwhelm an IDS and make it difficult to discern between malicious and legitimate activity. Fuzzy attack detection frequently calls for sophisticated anomaly detection methods that keep an eye out for anomalous data patterns, like machine learning models that can recognize atypical message forms and frequencies that may indicate system manipulation or intrusion (Han et al., 2018).

Malfunction attacks in vehicular networks target the integrity and functionality of vehicle systems by injecting incorrect or deceptive messages into the CAN bus to simulate faulty behavior. These attacks aim to interfere with vehicle operations by making specific components appear malfunctioning or behaving erratically, which can confuse drivers and potentially lead

to hazardous situations. For example, a malfunction attack might trigger erroneous signals that cause warning lights to flash, disrupt braking controls, or alter speed readings, compromising driver trust and response time. Detecting malfunction attacks involves monitoring data consistency and cross-referencing sensor values to identify mismatches or unusual deviations. Techniques such as machine learning and rule-based systems help identify patterns that indicate tampering or false data, enabling the intrusion detection system to flag potentially harmful manipulations and maintain system integrity (Han et al., 2018).

Software-based attacks expose the vulnerabilities within vehicle software, underscoring the need for effective countermeasures such as regular updates, secure coding practices, and robust testing protocols. Without comprehensive software security, connected and autonomous vehicles remain vulnerable to malicious activities that can compromise vehicle functionality and passenger safety, emphasizing the crucial role of cybersecurity in modern automotive design.

Figure 3.5 shows the related articles of software-based attacks:

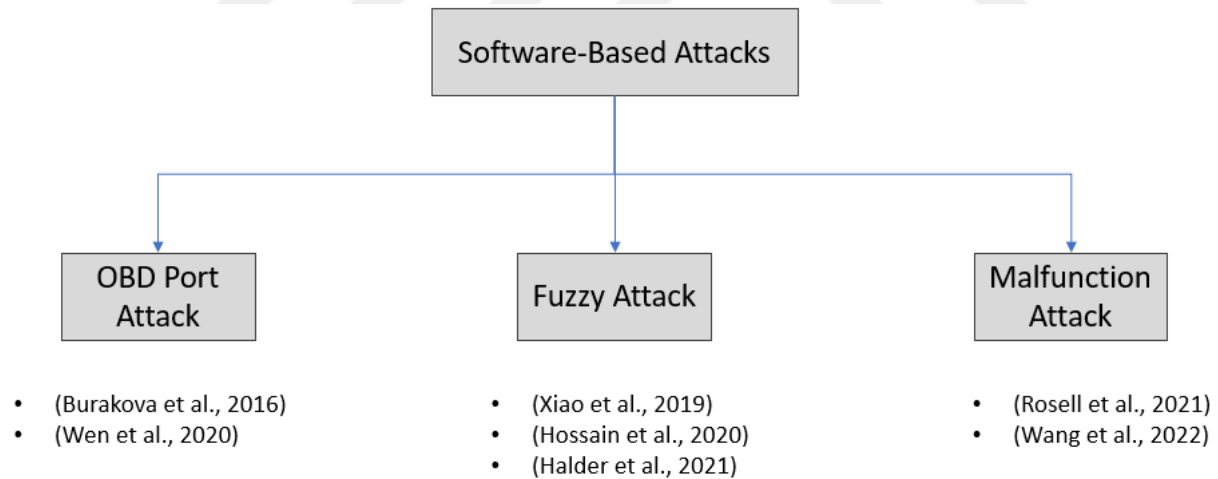


Figure 3.5 – Software-Based Attacks Papers

Burakova et al. explore cybersecurity vulnerabilities in heavy vehicles utilizing the SAE J1939 standard, a protocol commonly implemented across U.S. heavy vehicle industries. Through testing on a semi-truck and a school bus, the researchers demonstrate how this open standard enables critical attacks, including controlling the engine's RPM, disabling acceleration, and even deactivating the engine brake. These attacks reveal that access through the vehicle's OBD

port can compromise key systems, highlighting the lack of security measures in J1939 compared to proprietary automotive standards. The authors suggest that widespread use of this protocol across industries could expose many vehicles to similar risks, urging a reevaluation of security practices in the heavy vehicle sector (Burakova et al., 2016).

Wen et al. conduct a comprehensive vulnerability analysis of wireless OBD-II dongles, revealing significant security risks for vehicles. Utilizing DONGLESCOPE, an automated tool designed to dynamically test OBD-II dongles, they identified five major vulnerabilities, four of which were previously undiscovered. These vulnerabilities expose OBD-II dongles to a range of over-the-air attacks that could compromise user privacy, vehicle property, and even safety. The authors illustrate potential attacks, such as unauthorized engine control, data leakage, and denial-of-service, while also recommending security improvements like enhanced authentication protocols to mitigate these threats (Wen et al., 2020).

Wen et al. perform a comprehensive vulnerability analysis on wireless OBD-II dongles, identifying significant cybersecurity risks that these devices introduce as they interface with in-vehicle networks. Using an automated tool named DONGLESCOPE, they analyze 77 dongles, revealing five major vulnerabilities, four of which are newly discovered. The study categorizes vulnerabilities across the broadcast, connection, and communication stages, demonstrating how attackers can exploit these dongles for attacks such as unauthorized CAN bus access, remote vehicle control, and data theft. Wen et al. underscore the need for improved security measures, including enhanced authentication protocols, to safeguard vehicle systems from these emerging IoT-based threats (Wen et al., 2020).

Xiao et al. address the cybersecurity challenges of in-vehicle networks in the context of the Internet of Vehicles, proposing a lightweight, machine learning-based anomaly detection framework called SIMATT-SECCU. This framework combines a Security Control Unit (SECCU) with a Simplified Attention (SIMATT) model to achieve high detection accuracy with minimal computational costs, suitable for resource-limited vehicle electronics. By leveraging real data from normal vehicle operation for training, SIMATT-SECCU demonstrates robustness across various attack types, including denial of service, fuzzy, and impersonation attacks, while outperforming traditional models such as LSTM and GRU in both accuracy and efficiency. The study highlights the importance of lightweight and efficient anomaly detection for the evolving IoV ecosystem, where intelligent vehicles face growing security threats (Xiao et al., 2019).

Hossain et al. propose a LSTM-based IDS for detecting various attack types, such as DoS, Fuzzing, and Spoofing, on the CAN bus in modern vehicles. The CAN protocol lacks built-in security features, making it vulnerable to these attacks. In their approach, the authors collect CAN traffic data from a Toyota hybrid car, injecting attacks to create a dataset for model training and testing. Their LSTM model achieves a high detection accuracy of 99.9949% with low false positive and negative rates, outperforming traditional methods. The study highlights the importance of fine-tuning hyperparameters for optimal IDS performance, demonstrating the model's effectiveness in identifying malicious messages in real-time (Hossain et al., 2020).

Halder et al. propose a Clock Offset-based Intrusion Detection System (COIDS) for in-vehicle networks, utilizing clock offset as a unique identifier to monitor and detect intrusions on the CAN bus. COIDS calculates the clock offset of transmitter ECUs to establish a baseline of normal clock behavior. When deviations from this baseline are detected using the Cumulative Sum (CUSUM) method, COIDS identifies the intrusion and determines the intrusion time with a sequential change-point detection technique. To mitigate power overheads, the authors develop a cooperative game model that optimizes COIDS' active time on resource-constrained ECUs. Through extensive simulations, they demonstrate COIDS' effectiveness in detecting various CAN bus attacks, such as DoS and impersonation, with minimal power consumption (Halder et al., 2021).

Rosell et al. propose a frequency-based data mining approach to enhance in-vehicle network intrusion detection for the CAN. Given CAN's limitations in supporting authenticity and integrity, they develop a machine learning framework that relies on aggregated frequency features, bypassing proprietary CAN data encodings. Utilizing datasets such as Survival Analysis, they evaluate their approach using various attacks like flooding, fuzzing, and malfunction. Their study shows that random forests, support vector machines, and linear models effectively identify anomalies, with random forests achieving the highest accuracy. This frequency-based method offers an efficient, scalable solution for real-time CAN intrusion detection (Rosell et al., 2015).

Wang et al. propose an effective multi-view statistical graph learning approach called StatGraph for intrusion detection in ICVs CAN. The method constructs timing correlation and coupling relationship graphs to model short- and long-term message correlations, allowing fine-grained

detection of sophisticated attacks, such as masquerade attacks, often overlooked by conventional methods. StatGraph's graph-based approach improves detection granularity and performance, demonstrating robust generalization across multiple CAN datasets, including new, unexplored attacks. This novel approach offers enhanced detection accuracy and computational efficiency, making it suitable for real-time vehicular security applications (Wang et al., 2023).

Table 3.1. shows a comparative overview of Attacks on Driving System Components in survey literature:

Table 3.1 - A Comparative Overview of Attacks on Driving System Components

Literature	Attack Interface	Attack Type	Vulnerabilities	Impact
(Brooks et al., 2009)	Wireless Networks	Sybil Attack	Identity validation flaw	Fake identities disrupt network
(Koscher et al., 2010)	CAN, OBD	ECU Attack	Authentication and access control gaps	Malicious ECU manipulation
(Tyagi et al., 2014)	V2V Vectors	VANET Attack	Weak routing protocols	Traffic disruption or accidents
(Bariah et al., 2015)	V2V Vectors	Sybil Attack, VANET Attack	Weak routing protocols	Traffic disruption or accidents
(Garip et al., 2015)	CAN, Wireless Networks	VANET Attack	Weak routing protocols	Traffic disruption or accidents
(Petit et al., 2015)	Sensors	Replay Attack, Spoofing	Sensor data validation flaws	Sensor spoofing

(Woo et al., 2015)	OBD, CAN	DoS, Replay Attack	Resource exhaustion vulnerabilities	System crash or denial service
(Abbott-McCune and Shay, 2016)	OBD, CAN	DoS, Replay Attack	Resource exhaustion vulnerabilities	System crash or denial service
(Nie et al., 2016)	CAN	ECU Attack	Authentication flaws	Malicious ECU behavior
(Palanca et al., 2016)	CAN	DoS	Resource exhaustion vulnerabilities	Network unavailability
(Burakova et al., 2016)	CAN, OBD	OBD Port Attack	Port access misconfigurations	System compromise via OBD
(Yan et al., 2016)	Sensors	Spoofing	Signal spoofing vulnerability	GPS navigation manipulation
(Hasrouny et al., 2017)	V2V, V2I	VANET Attack	Weak routing protocols	Traffic disruption
(Cheah et al., 2017)	Bluetooth	DoS	Bluetooth communication insecurity	Loss of communication
(Parkinson et al., 2017)	OBD, V2V	DoS, ECU Attack	Authentication and access flaws	Malicious ECU commands
(Bernardini et al., 2017)	CAN, V2V, V2I	Eavesdropping	Encrypted communication absence	Data interception
(Garip et al., 2018)	V2V	VANET Attack	Weak routing protocols	Traffic disruption
(Alheeti et al., 2018)	VANET	DDoS, DoS	Resource exhaustion vulnerability	System slowdown or

				denial of service
(Han et al., 2018)	CAN	Flooding Attack	Flooding vulnerability in network protocols	Network flooding
(Sheehan et al., 2019)	CAN	Spoofing	Signal spoofing	GPS navigation manipulation
(Sommer et al., 2019)	In-vehicle Network	Eavesdropping	Unencrypted data flow vulnerability	Sensitive data interception
(Xiao et al., 2019)	CAN	Fuzzy Attack	Fuzzing susceptibility	Unintended vehicle commands
(Chowdhury et al., 2020)	Bluetooth, Wi-Fi, OBD	ECU Attack	Authentication and access control gaps	ECU manipulation
(He et al., 2020)	GNSS	Spoofing, Replay Attack	Signal replay vulnerability	Navigation system compromise
(Dibaei et al., 2020)	V2V, V2I	VANET Attack	Weak routing protocols	Traffic disruption
(Wen et al., 2020)	OBD	Spoofing, Eavesdropping, OBD Port Attack	OBD access misconfigurations	System compromise via OBD
(Baza et al., 2020)	VANET	Sybil Attack	Identity validation flaw	Network disruption
(Bajpai et al., 2020)	IVI	DoS	Resource exhaustion vulnerabilities	System unavailability
(Sheikh et al., 2020)	V2V	VANET Attack	Weak routing protocols	Traffic disruption

(Hossain et al., 2020)	CAN	Fuzzy Attack, Flooding Attack	Fuzzing and flooding susceptibility	System slowdown or crash
(Ebert and Ray, 2021)	Ethernet, IVI	Spoofing, DoS, Eavesdropping	Signal spoofing and resource exhaustion	Sensitive data manipulation
(Cao et al., 2021)	CAN	Spoofing	Signal spoofing vulnerability	Vehicle spoofing
(Sánchez et al., 2021)	CAN	Replay Attack	Signal replay vulnerability	Navigation manipulation
(Kim et al., 2021)	Wi-Fi, CAN, OBD, V2X	DoS, VANET Attack	Signal and routing protocol flaws	V2N disruptions
(Pham and Xiong, 2021)	OBD, CAN	DoS	Resource exhaustion vulnerabilities	Denial of service
(Blevins et al., 2021)	CAN	Flooding Attack	Flooding vulnerability in protocols	System crash
(Halder et al., 2021)	CAN	Fuzzy Attack	Fuzzing susceptibility	ECU malfunction
(Rosell et al., 2021)	CAN	Malfunction Attack	Hardware malfunction risks	ECU malfunction
(Wang et al., 2022)	CAN	Malfunction Attack	Hardware malfunction risks	Navigation and communication compromise
(Bendiab et al., 2023)	Sensors, CAN	DoS, Spoofing	Signal spoofing and resource exhaustion	Sensitive data manipulation

3.2 Defense of Connected and Automated Vehicles

Two primary types of defensive literature are elaborated in the field of CAV's security: security architecture and intrusion detection systems.

3.2.1. Defense Frameworks for Automated Vehicles

Initially, every autonomous vehicle component's security was examined. Over time, studies on malware and VANETs have been conducted in addition to the several aspects of CAN and ECU security that have been examined.

The security architecture in relation to the CAN and ECU has been the subject of several scholarly investigations. In 2008, Oguma et al. introduced a sophisticated security architecture designed to provide safe communication within a car (Bendiab et al, 2023). The inapplicability of remote authentication systems in real-world vehicular applications was addressed by the authors, who acknowledged that in order to prevent ECU manipulation, they are not feasible to implement in time-constrained automobile systems. Trusted platform modules are used in these systems. By offering a creative proof-based security architecture for in-vehicle communications, they adeptly addressed the flexibility requirements associated with software configuration authentication, encrypted communications, and switching.

An approach was introduced by Nilsson et al. in 2008 with the intention of counterbalancing the results of Oguma et al. (Nilsson et al., 2009). This method effectively delayed the delivery of data by utilizing a complex message authentication code. The authentication factor is delayed because the message with the code is computed as a composite element of subsequent messages that are broadcast in conjunction with the subsequent message. The proposed method considers the limitations of the ECU, the dynamics of real-time traffic inside the vehicle network, and the detailed CAN frame structure.

According to Groll & Ruland in (Groll et al, 2009), the CAN's lack of confidentiality and reliability puts in-vehicle communications at significant risk for security breaches. They pushed for the formation of a reliable organization for confidential communication between the many components of an automobile and underlined the need for a practical solution. Additionally, the authors stressed that manufacturers were required to present a signed certificate attesting to their involvement in a secret communication group.

Herrewewege et al. introduced CANAuth, a novel authentication method, in (Herrewewege et al, 2011). Its purpose was specifically broadcast communication. The possibility of achieving compatibility without requiring modifications to the existing nodes is a significant advantage of this concept. The inherent challenge of including defense mechanisms, particularly authentication, within the limited payload size of a CAN message prompted the authors to utilize the CAN + protocol, extending it to the available bits. After that, they modified the protocol to include a hash-based Message Authentication Code (HMAC), which offered a useful examination of the adversarial model, potential threats, and associated transmission overheads. With the help of this proposed method, ECUs may defend themselves against various injection attacks.

Hazem presented a straightforward message source authentication system created especially for CAN in (Hazem et al, 2012). The protocol outlines intricate safeguards that shield nations from potential intruders. Channel setup, data interchange, synchronization, startup, and chain updating are some of these phases. Although there has not been any real-world testing on the protocol's suitability for commercial vehicles, it was successfully implemented in a microcontroller. Additionally, the authors noted how well their protocol performed with single-pair Ethernet and FlexRay.

In order to identify inserted CAN signals, Wang & Sawhney developed the VeCure security framework (Wang et al., 2018), which monitors the delay from transmitter to receiver. The authors reported a minimal process delay of 50 μ s when doing the proof of concept on an embedded device. The framework's effectiveness against injection and replay threats was validated using real-world data experiments.

Groza & Murvay in (Groza et al, 2012) presented an improved broadcast protocol for the CAN that operates without the need for pre-shared, public, or private keys. In this protocol, the key is changed following each broadcast, or each time a single message is transmitted across the CAN bus. The researchers adjusted a number of parameters in their integrated test-bed to increase communication effectiveness.

In the context of an automobile network, Woo et al. (Woo et al., 2019) introduced a moving target protection method that makes use of a network address shuttle for the CAN. The CAN ID shuffling method is the name given to this technology. This approach serves as a practical

security precaution by making it difficult for attackers to get critical information, including the CAN ID, data frame saving cycle, and data field kinds. Furthermore, the proposed approach is considered viable as it may be implemented without needing modifications to specific fields in the May standard or the present protocol structure.

3.2.2 Intrusion Detection System for CAN

The primary goal of research on CAN intrusion detection systems for automated cars is to identify and prevent cyberattacks. Müter and Asaj presented an information-theoretic method for calculating message entropy in order to detect attacked ECUs in (Müter & Asaj, 2011). Ling and Feng presented a technique for identifying CAN message frequency abnormalities in (Ling & Feng, 2012), while Studnia et al. proposed using gateways as IDS (Studnia et al., 2018). Song et al. provided a simple technique for detecting injection attempts by examining the gap between CAN messages, whereas Boudguiga et al. in (Boudguiga et al., 2016) suggested using microcontrollers with hardware security modules. Cho and Shin used message arrival timings to demonstrate ECU fingerprinting in (Cho & Shin, 2016), whereas Gmiden et al. suggested a basic IDS based on timing differences between CAN messages in (Gmiden et al., 2016).

IDS detection techniques were examined by Lokman et al. in (Lokman et al., 2019) and categorized into four groups: frequency-based, hybrid, machine-learning-based, and statistical. In order to strengthen CAN network security in autonomous cars against cyberattacks, Olufowobi et al. developed a method to detect and fix CAN signal corruption and employed intrusion detection systems (IDS) to isolate unusual nodes. In contrast, Luo and Hou recommended a hardware-based firewall in their paper (Luo & Hou, 2021). These apps strengthen the defenses against CAN network breaches in driverless cars.

3.3 HCRL Car Hacking and Survival Analysis Dataset Referenced Papers' Review

In recent years, the automotive industry has undergone significant transformation due to the increasing levels of connectivity and automation in vehicles. This transformation necessitates integrating advanced communication systems, such as the CAN, which, while enhancing vehicle functionality, also exposes them to various cybersecurity threats. The literature reveals a growing body of research focused on detecting different types of cyberattacks that may occur

within in-vehicle networks. This literature review reflects ongoing efforts to develop effective and innovative solutions to enhance the security of automotive systems, ensuring the safety of drivers, passengers, and pedestrians alike. A chronological examination of these papers can be found below.

(Venkatesan and Raja, 2018) implemented a hybrid system for vehicular communication networks using DNN and the Zone Hierarchical Link State (ZHLS) protocol. Their approach focuses on optimizing routing between vehicles and infrastructure, reducing packet loss and improving the reliability of data transmission. The hybrid system was validated through MATLAB simulations, demonstrating significant improvements in communication efficiency and reduced delays within the vehicular network.

(Garg et al., 2019) propose a Software Defined Networking-based framework to enhance security and privacy in 5G-enabled vehicular networks. Their approach leverages elliptic curve cryptography for mutual authentication and a tensor-based dimensionality reduction method for intrusion detection. Through simulations using NS3, SUMO, and SPAN, the model demonstrates improved detection rates, lower false positives, and optimized communication overhead, making it a robust solution for modern vehicular networks.

(Ying et al., 2019) analyze clock skew-based IDS for CAN, focusing on a cloaking attack where an adversary manipulates message inter-transmission times to emulate clock skew and bypass detection. The authors develop formal models to predict the success rate of such attacks and validate them through experiments on both prototype and real vehicle platforms. Their findings demonstrate the vulnerability of existing IDS models to clock skew manipulation and emphasize the need for more robust detection mechanisms.

(Xiao et al., 2019) present a lightweight machine learning-based anomaly detection framework called SIMATT-SECCU, designed to enhance the security of in-vehicle networks. The system focuses on minimizing computational costs while maintaining high detection accuracy. By combining recurrent neural networks with attention models, this framework demonstrates superior performance compared to other classical models like LSTM and GRU in detecting anomalies in CAN bus systems, proving particularly effective against impersonation, DoS, and fuzzy attacks.

(Xiao et al., 2019) propose a robust and self-evolving IDS for in-vehicle networks by leveraging spatiotemporal information and time series prediction. The system, based on a Convolutional LSTM (ConvLSTM) model, processes CAN bus traffic in real-time and detects intrusions by analyzing time and space information from ECUs. The IDS continuously learns during the vehicle's operation, adapting to environmental changes like weather or road conditions, which enhances its ability to detect both known and unknown attack models effectively.

(Park and Choi, 2020) propose a machine learning-based malware detection system designed for self-driving vehicles. By analyzing network traffic from Android-based devices connected to vehicular networks, they develop a framework that detects adware and general malware with high accuracy. The study compares six machine learning algorithms, with Random Forest performing best in terms of detection accuracy and speed, making it suitable for real-time malware detection in automotive environments.

(Khullar et al., 2020) propose a decentralized communication architecture for Flying Ad Hoc Networks (FANETs) using blockchain technology. The system employs Practical Byzantine Fault Tolerance (PBFT) for consensus and a gossip protocol for message transmission between nodes. The architecture is designed to be computationally efficient and scalable, making it suitable for various applications, including smart city services, military operations, and healthcare delivery in remote areas. Experimental results show that the proposed system maintains consistent throughput and low latency as the network size increases.

(Park and Choi, 2020) propose a multi-labeled hierarchical classification (MLHC) model for anomaly detection in in-vehicle networks, specifically targeting message injection attacks in CAN bus traffic. By using machine learning algorithms, the model classifies the presence and type of attacks with high accuracy and speed. Compared to traditional methods, the MLHC model demonstrates superior performance in real-time detection while minimizing false positives and misclassification of attack types, making it highly suitable for real-time vehicular network environments.

(Nasri et al., 2020) provide a comprehensive overview of the reliability issues and challenges facing CAN in future intelligent transportation systems. The paper discusses various fault diagnosis methods, electromagnetic immunity concerns, and CAN payload expansion strategies, with a focus on how these issues impact the reliability of in-vehicle network control

systems (IVNCS). The authors also explore hardware-in-the-loop (HIL) and real-time analysis techniques as key validation tools for ensuring the safety and reliability of automated vehicular systems.

(Kwak et al., 2020) propose a cosine similarity-based anomaly detection method for CAN to identify injection attacks such as flooding, fuzzing, and malfunction attacks. By analyzing self-similarity in CAN bus traffic, the method achieves high accuracy in detecting attacks with minimal computational overhead, making it suitable for real-time detection in both driving and stationary vehicle conditions. The study shows that cosine similarity can effectively differentiate normal and attack states, particularly in the KIA Soul and Hyundai Sonata vehicles.

(Hossain et al., 2020) propose an LSTM-based IDS for detecting attacks on in-vehicle CAN bus communications. The authors generate their own dataset, focusing on Denial of Service, Fuzzing, and Spoofing attacks, and demonstrate that their LSTM model achieves an impressive detection accuracy of 99.995%. The model is shown to outperform traditional survival analysis techniques, making it a robust solution for real-time vehicular network security.

(Kim et al., 2020) review various cybersecurity threats and defense mechanisms for autonomous vehicles, focusing on vulnerabilities in three key areas: the autonomous control system, driving system components, and V2X communications. The paper emphasizes the increasing importance of machine learning and artificial intelligence in anomaly detection and IDS. By analyzing 151 papers, the authors provide a comprehensive assessment of the current and future challenges in securing autonomous vehicle networks, particularly in smart city environments.

(Blevins et al., 2021) benchmark four time-based IDS against the ROAD dataset, a publicly available dataset containing real (non-simulated) stealthy CAN bus attacks. The study compares methods like mean inter-message time and kernel density estimation for detecting timing anomalies in CAN messages. Results show that simple, distribution-agnostic methods outperform more complex distribution-based methods, particularly in precision-recall curves, providing a clear direction for improving vehicular network security through lightweight IDS deployment.

(Kang et al., 2021) describe their experience hosting the Car Hacking: Attack & Defense Challenge, where participants developed IDS and attack methods for the CAN bus, a standard in-vehicle communication protocol. The competition involved real-time injection and detection of attacks on a Hyundai Avante CN7, with participants using machine learning and rule-based models to identify threats. The results demonstrated that rule-based and tree-based models, such as Random Forest, were particularly effective in detecting anomalies within CAN traffic.

(Halder et al., 2021) propose a clock offset-based Intrusion Detection System (COIDS) for monitoring CAN traffic and detecting intrusions such as DoS, impersonation, and fuzzy attacks. The system uses the clock offset of ECUs to fingerprint the source of messages and employs a cooperative game model to reduce power consumption in resource-constrained environments. The authors demonstrate that COIDS detects intrusions more effectively than traditional methods while optimizing energy efficiency.

(Khatri et al., 2021) review the security issues in in-vehicle networks, focusing on vulnerabilities within CAN, FlexRay, and automotive Ethernet protocols. They propose enhanced countermeasures using a hybrid blockchain framework to address these challenges. The paper provides a comprehensive survey of existing security threats and IDS while highlighting the limitations of current approaches, emphasizing the need for blockchain integration to enhance vehicular network security.

(Anjum et al., 2021) propose an anomaly detection system for in-vehicle networks using the Extreme Gradient Boosting Machine (XGBoost) algorithm. The model is tested on real CAN bus data from Hyundai Sonata, Kia Soul, and Chevrolet Spark vehicles, under various attack scenarios such as Denial-of-Service (DoS), fuzzy, spoofing, and malfunction attacks. The results show that XGBoost outperforms traditional models, achieving over 99% accuracy, with reduced false-positive and false-negative rates, making it an effective solution for real-time vehicular security.

(Mehedi et al., 2021) propose a deep transfer learning-based IDS for electric vehicular networks. The system leverages a LeNet architecture to detect anomalies in the CAN bus protocol, using a dataset of real-time vehicular data. The proposed IDS significantly improves detection accuracy compared to traditional machine learning and deep learning models,

demonstrating better performance for real-time vehicular security against attacks such as flooding, fuzzing, and spoofing.

(Aliyu et al., 2021) propose a blockchain-based federated learning framework for an IDS in software-defined vehicular networks. By leveraging federated learning, their system allows for secure and decentralized model training without sharing sensitive in-vehicle data. The use of blockchain ensures data integrity, while a random forest algorithm enhances attack detection accuracy. The system is tested on a real-world dataset, demonstrating high detection rates for attacks such as DoS, Fuzzy, and Impersonation.

(Fu et al., 2021) introduce IoV-BERT-IDS, a hybrid IDS designed for Internet of Vehicles environments using large language models (LLMs). The system employs a semantic extractor to convert network traffic into contextual semantics and uses the BERT model to pre-train and fine-tune for in-vehicle and extra-vehicle network security. The IDS achieves high accuracy across multiple datasets, such as CICIDS and IVN-IDS, showing strong generalization across different vehicle models and network environments.

(Nakamura et al., 2021) propose a supervised-unsupervised hybrid approach for detecting in-vehicle network attacks across multiple car models. By using a pretrained model from historical vehicle data and combining it with unsupervised learning on the normal data from the target vehicle, the system can effectively detect and classify attacks even when limited data is available for the new vehicle model. The method demonstrates high adaptability and accuracy across various vehicle types, making it suitable for real-world vehicular security applications.

(Akşehir and Akleylek, 2021) review deep learning-based IDS aimed at securing the CAN in vehicles. The paper systematically analyzes various studies on IDS methods, datasets, selected attributes, and the types of attacks they address. By comparing these methods, the authors highlight the superiority of deep learning techniques such as LSTM in detecting network anomalies and propose further enhancements for improving CAN security against modern cyber threats.

(Sharma and Jaekel, 2021) introduce a machine learning-based misbehavior detection system for VANETs using consecutive Basic Safety Messages (BSMs). By analyzing data from multiple BSMs, the authors aim to detect position falsification attacks more effectively. Their

approach improves upon traditional methods by considering temporal data and using machine learning algorithms such as K-Nearest Neighbor and Random Forest, achieving better detection accuracy across different attack types, particularly in dynamic vehicular environments.

(Alkhatib et al., 2021) propose a deep learning-based IDS for SOME/IP protocol in automotive Ethernet networks. The system uses Recurrent Neural Networks to detect and classify intrusions, such as errors on events and missing responses. The authors generate a labeled dataset specific to the SOME/IP protocol and achieve high F1 scores and AUC values, demonstrating the effectiveness of RNNs in identifying in-vehicle network intrusions. This system shows promise for enhancing security in modern automotive Ethernet-based communication systems.

(Rosell et al., 2021) propose a frequency-based data mining approach to enhance in-vehicle network IDS. The authors focus on improving machine learning-based IDS performance by leveraging frequency-based feature representation rather than signal-based methods. This approach is validated using various real-world datasets, including CAN bus data, and tested with machine learning algorithms like Random Forest and Support Vector Machines. Their method shows significant improvements in detecting flooding, fuzzing, and malfunction attacks in a computationally efficient manner.

(Han et al., 2021) propose an event-triggered interval-based anomaly detection and attack identification method for in-vehicle networks, specifically focusing on the Controller Area Network. By examining the time intervals between CAN messages, the system detects anomalies and classifies attack types such as flooding, fuzzy, and replay attacks with high accuracy. The method utilizes machine learning models, particularly tree-based classifiers, to achieve up to 99% detection accuracy, making it a highly effective solution for real-time vehicular security.

(Swessi and Idoudi, 2021) provide a comparative review of security threat datasets used in vehicular networks, emphasizing the role of machine learning-based IDS. They categorize datasets into inter-vehicle and in-vehicle types, discussing popular datasets such as ROAD, VeReMi, and Car-Hacking. The paper highlights the challenges of dataset availability and the importance of using realistic, labeled datasets for training effective IDS models. The authors

conclude with recommendations on the best datasets for vehicular network security research, particularly in addressing common attacks like DoS, fuzzy, and replay.

(Kang et al., 2021) propose a Hybrid MR-IDS, a machine learning and ruleset-based IDS for in-vehicle networks using CAN bus data. The system combines the advantages of rule-based detection and random forest models to improve the accuracy of detecting both known and unknown attack patterns. The paper also introduces five attack scenarios, including spoofing and replay attacks, to test the system's performance. The Hybrid MR-IDS achieved a 99% detection accuracy, demonstrating its effectiveness in real-time vehicular security.

(Balaji and Ghaderi, 2021) introduce NeuroCAN, a deep learning-based anomaly detection system for in-vehicle Controller Area Networks. The system focuses on detecting contextual anomalies by utilizing Long Short-Term Memory units and learning the spatiotemporal correlations among sensor data on the CAN bus. NeuroCAN achieves over 95% detection accuracy and performs significantly better than existing baseline models by analyzing sensor data collectively rather than in isolation, which helps in identifying stealthy attacks that might evade traditional detection systems.

(Mowla et al., 2022) propose a dynamic voting-based explainable IDS for in-vehicle networks using Random Forest models. Their system integrates explainable AI (XAI) techniques, such as SHAP and LIME, to provide transparency in the decision-making process of IDS models. The IDS was evaluated on publicly available datasets like the Survival Analysis Dataset and ORNL, achieving high accuracy and F1-scores in detecting CAN bus attacks such as flooding and fuzzy attacks. The explainable aspect helps cybersecurity analysts understand and trust the model's decisions.

(Paddock, 2022) discusses best practices for law enforcement when investigating automotive accidents potentially caused by cyberattacks on in-vehicle networks. The study explores how law enforcement officers conduct preliminary cybersecurity forensic analysis, emphasizing the importance of training and access to relevant data such as Event Data Recorder (EDR) and Cybersecurity Threat Intelligence reports. The research highlights the gaps in technical understanding and resources that hinder effective analysis of cyber vulnerabilities in vehicle systems, particularly concerning the Controller Area Network.

(Karopoulos et al., 2022) provide a comprehensive meta-taxonomy of in-vehicle IDS, consolidating various classifications from previous surveys. The authors categorize IDS approaches based on factors such as detection method, deployment strategy, and evaluation metrics. The study also examines publicly available datasets and simulators used to evaluate IDS performance in vehicular networks, identifying future challenges and trends in this field. This work serves as a foundational reference for researchers aiming to classify and evaluate in-vehicle IDS systems effectively.

(Alamleh, 2022) introduces a benchmarking framework for evaluating IDS classifiers based on multi-criteria analysis. The study develops a comprehensive approach by integrating security and performance evaluation criteria, using methods like Fuzzy Delphi and multi-criteria decision-making (MCDM) techniques, including VIKOR and TOPSIS. The framework allows developers to select the most suitable IDS classifier based on both objective and subjective evaluation, providing a systematic method for improving IDS accuracy and efficiency in real-world applications.

(Sankaranarayanan et al., 2022) propose a cluster-based attack prevention algorithm for autonomous vehicles using machine learning techniques. The system integrates Convolutional Neural Networks with a clustering algorithm to detect and prevent cyberattacks on vehicular networks. The model shows significant improvements in detecting cluster-based attacks like spoofing and replay attacks, while optimizing energy consumption and stability in real-time operations. This approach demonstrates the potential of using machine learning for effective anomaly detection in the complex communication environments of autonomous vehicles.

(Ma et al., 2022) propose a GRU-based lightweight IDS for CAN bus networks, focusing on real-time detection with minimal computational overhead. The system employs a feature extraction algorithm designed for low-power embedded devices and demonstrates high classification accuracy in detecting DoS, fuzzy, and spoofing attacks. Tested on NVIDIA Jetson Xavier NX embedded devices, the model shows superior performance compared to LSTM-based models, achieving real-time detection with lower energy consumption.

(Lundberg, 2022) explores how to increase the trustworthiness of AI-based IDS for in-vehicle networks by using Explainable AI (XAI) methods. The study focuses on mitigating the interpretability issues of complex AI models like Random Forests and Deep Neural Networks

employed in detecting anomalies in CAN bus traffic. By integrating SHAP (SHapley Additive ExPlanations) values and custom visualization techniques, the research demonstrates that explainability enhances trustworthiness, helping users better understand the model's decisions and improving the overall reliability of the IDS.

(Lin et al., 2022) propose a deep learning-based IDS for in-vehicle networks, focusing on the Controller Area Network. The system employs the VGG16 classifier to detect cyberattacks, such as DoS, fuzzing, and spoofing, using data from the Hacking and Countermeasure Research Lab (HCRL). The authors compare the performance of VGG16 with other machine learning models like XGBoost, showing that their approach achieves high accuracy, precision, and recall, making it an effective solution for detecting in-vehicle network anomalies.

(Lundberg et al., 2022) conduct an experimental analysis of an IDS using eXplainable AI techniques to improve the trustworthiness of AI-based anomaly detection in vehicular networks. Their system employs the SHAP (SHapley Additive exPlanations) method to provide visual explanations for detected anomalies in CAN bus traffic, increasing transparency and trust in the system. Through expert surveys, they find that XAI techniques significantly improve trust and understanding in AI-based IDS systems, enhancing their applicability in real-world vehicular cybersecurity.

(Andreica et al., 2022) compare the performance of deploying IDS for in-vehicle networks on traditional ECUs versus modern Android head units. The study highlights that while Android devices benefit from higher computational power and cloud connectivity, they may be more vulnerable to attacks compared to in-vehicle ECUs. The authors implement a two-stage IDS with machine learning algorithms and test them on real-world CAN bus data, demonstrating the potential of Android head units to offer more advanced IDS capabilities with cloud support.

(Aliyu et al., 2022) investigate the use of a Blockchain-Based Federated Forest (BFF) system in detecting adversarial attacks in in-vehicle networks. Their work focuses on integrating a statistical adversarial detector to improve the security of machine learning-based IDS. By detecting adversarial examples and augmenting the BFF-IDS model with new adversarial classes, the system significantly improves its defense against previously unknown attacks, demonstrating its potential in securing vehicle networks against sophisticated evasion attacks.

(Dahiya, 2022) introduces a hybrid classifier strategy optimized with tuned training weights for detecting DDoS attacks in 5G networks. The system integrates Bidirectional Long Short-Term Memory (BiLSTM) with an optimized Deep Belief Network (DBN) to enhance detection precision. The classifier utilizes a novel optimization algorithm combining shark smell and electric fish models, which improves accuracy, sensitivity, and other performance metrics, achieving a detection accuracy of 98.59%. This approach demonstrates superior performance in identifying DDoS attacks compared to conventional deep learning models.

(Han et al., 2022) present TOW-IDS, an Intrusion Detection System for Automotive Ethernet based on three overlapped wavelets and deep convolutional neural networks. The system effectively detects anomalies in Automotive Ethernet networks by processing network traffic data through wavelet transformations. By leveraging protocols such as Audio-Video Transport Protocol (AVTP) and CAN, the method reduces the computational complexity while maintaining high detection accuracy, outperforming traditional models like ResNet and EfficientNet in both time-cost and efficiency.

(Patel et al., 2022) propose a deep learning and blockchain-based framework for detecting malware in autonomous vehicles. The system leverages the ResNet50V2 architecture to convert network traffic into grayscale images for malware detection, achieving a detection accuracy of 97.56%. The framework, called AVBlock, also integrates blockchain technology to securely store and manage vehicle data, ensuring that only verified, malware-free information is propagated within the vehicular network. This approach enhances the security and trustworthiness of autonomous vehicle operations.

(Rajapaksha et al., 2022) present a context-aware IDS called CAN-CID, designed to secure in-vehicle CAN bus communications. The system uses a combination of Gated Recurrent Units (GRUs) and a time-based model to predict the next CAN message and detect anomalies in real-time. By using benign data for training, the IDS avoids the need for attack data collection, making it more efficient in detecting various types of attacks, such as fabrication and masquerade attacks. CAN-CID achieves an F1-score of over 99% for multiple attack scenarios, demonstrating its effectiveness in securing CAN bus systems.

(Gundu and Maleki, 2022) present an IDS for securing the CAN in connected and autonomous vehicles using supervised machine learning approaches. Their work focuses on time intervals

between CAN messages as a key feature for identifying anomalies. The study compares Random Forest, K-Nearest Neighbor, and XGBoost algorithms, with RF achieving the highest accuracy at 97%. The inclusion of time interval features significantly improves the detection performance, demonstrating that feature engineering is critical in enhancing machine learning-based IDS systems.

(Madhav et al., 2023) introduce IMPROVE, an intelligent machine learning-based portable, reliable, and optimal verification system for future vehicles. The framework consists of two modules: the first focuses on securing the CAN through machine learning-based intrusion detection, while the second module monitors data transfers between external devices and the vehicle. The system employs an ensemble of machine learning algorithms, including Random Forest, AdaBoost, and Artificial Neural Networks, to detect and block malicious behavior in real-time. The IMPROVE system demonstrates high accuracy in preventing both internal and external cyberattacks on autonomous vehicles.

(Rajapaksha et al., 2023) provide a comprehensive survey of AI-based IDS for in-vehicle networks, with a specific focus on the CAN bus protocol. The study classifies IDS approaches based on machine learning and deep learning techniques, identifying key datasets and features used in the detection of cyberattacks, such as DoS, fuzzing, and message injection attacks. The paper also highlights the challenges of implementing AI models in vehicular networks and discusses future research directions for enhancing IDS performance.

(Khan et al., 2023) propose an IDS for in-vehicle networks using the statistical characteristics of attack behaviors to detect anomalies in CAN bus traffic. The system focuses on efficiently detecting DoS, fuzzy, and merge attacks by calculating optimized threshold values for various window sizes. The proposed IDS achieves high accuracy and minimizes false positives by exploiting malicious patterns in real-world CAN bus datasets. Experimental results validate its effectiveness in detecting complex attacks with a low misclassification rate.

(Park et al., 2023) propose a graph-based Intrusion Detection and Classification System (G-IDCS) for securing in-vehicle CAN. By integrating threshold-based detection with machine learning classifiers, the system effectively reduces the number of CAN messages required for intrusion detection while maintaining high accuracy. The G-IDCS method improves detection

rates for various attack types, including DoS, fuzzing, and replay attacks, making it a robust solution for real-time vehicular network security.

(Suhail et al., 2023) propose ENIGMA, an explainable digital twin security solution for cyber-physical systems (CPS). The framework leverages gamification techniques to train security analysts in detecting and mitigating cyberattacks on CPS through Digital Twins (DTs). ENIGMA incorporates explainable AI to enhance trust and transparency in AI-driven security decisions, using SHAP values to explain the contributions of various features to attack predictions. This approach allows analysts to take informed security measures based on understandable AI decisions, improving the overall security of CPS.

(Nguyen et al., 2023) propose a Transformer-based Attention Network (TAN) for intrusion detection in in-vehicle networks using the CAN bus. The system employs a multi-class classification approach to detect various attacks, including DoS, fuzzy, spoofing, and replay attacks. By leveraging transformers, the model overcomes the limitations of RNN-based systems in feature extraction and achieves superior performance, particularly in detecting replay attacks without the need for message-level labeling. The system also demonstrates the potential of transfer learning for improving detection across different car models.

(Kim et al., 2023) propose an anomaly detection method based on multiple LSTM-Autoencoder models for in-vehicle networks, specifically focusing on the CAN protocol. Their system uses different network features, such as transmission intervals and payload value changes, to detect anomalies by analyzing each feature separately with an LSTM-Autoencoder model. The system achieves a 99% detection rate in real-world vehicle network traffic experiments, proving effective at identifying various attacks like flooding, fuzzing, and denial of service with high precision.

(Salek, 2023) explores the use of generative neural networks to develop defense methods against cyberattacks in CAVs. The study presents a hybrid quantum-classical Restricted Boltzmann Machine (RBM)-based framework for detecting intrusions in in-vehicle networks and a Generative Adversarial Network (GAN)-based method for defending traffic sign classification systems from adversarial attacks. The hybrid RBM system achieved an average intrusion detection accuracy of 98%, while the GAN-based method outperformed traditional defense methods in mitigating adversarial attacks on traffic sign classifiers.

(Xiao et al., 2023) introduce a novel feature extraction framework for in-vehicle networks intrusion detection using a Graph Node Attention Network (GNAT). Their method converts CAN packets into graphs based on temporal adjacency, allowing the GNAT to identify malicious packets by analyzing graph topologies. This approach improves the detection accuracy of the Random Forest classifier used to identify attacks, such as denial of service and fuzzy attacks, and enhances the system's generalization capability across different vehicle brands without extensive retraining.

(Hoang and Kim, 2023) propose a supervised contrastive ResNet model combined with transfer learning for detecting intrusions in CAN bus traffic. The model effectively classifies multiple types of attacks, such as DoS, fuzzy, and spoofing attacks, using two real-world car datasets. By leveraging transfer learning, the system adapts to new car models without requiring large datasets, reducing false negatives by five times compared to cross-entropy loss. The model shows exceptional performance on the Car Hacking and Survival datasets, achieving the highest F1 scores for CAN bus intrusion detection.

(Salek et al., 2023) propose a hybrid approach for intrusion detection in in-vehicle CAN by combining a Classical Convolutional Neural Network with a Quantum Restricted Boltzmann Machine (RBM). The CCNN is responsible for feature extraction, while the quantum RBM handles CAN image reconstruction for classification. The hybrid approach demonstrates higher accuracy and recall compared to a classical-only approach, achieving detection accuracies of up to 98.3% in various attack scenarios such as fuzzy attacks. This study highlights the potential of quantum computing to enhance intrusion detection in vehicular networks.

(Wang et al., 2023) propose a multi-view statistical graph learning approach for fine-grained intrusion detection in in-vehicle networks, particularly focusing on detecting masquerade attacks on the CAN bus system. The method, named StatGraph, generates two statistical graphs: the Timing Correlation Graph (TCG) and the Coupling Relationship Graph (CRG), which analyze both temporal and contextual correlations of CAN messages. The use of lightweight graph convolutional networks (GCN) for detecting intrusions significantly improves detection accuracy, achieving over 99.4% on real-world datasets, including the Car Hacking and ROAD datasets. This method demonstrates enhanced detection performance for complex and subtle attacks, such as masquerade attacks, that were previously undetectable by traditional methods.

(Wu et al., 2023) propose ConvIDS, a ConvLSTM-based intrusion detection system for in-vehicle networks, specifically the CAN bus. Their method transforms CAN data into sequential images and applies ConvLSTM to detect anomalies such as fuzzing, malfunction, and replay attacks. The system outperforms traditional LSTM and Multilayer Perceptron (MLP) models, achieving high accuracy across all metrics, including F1-score and Area Under Curve (AUC). ConvIDS demonstrates significant improvements in real-time detection, making it suitable for modern vehicular networks.

(Batista de Melo and Dutt, 2023) introduce LOCoCAT, a low-overhead classification system for detecting and classifying CAN bus attack types in real-time. The system groups malicious messages into blocks and extracts relevant features to quickly identify the type of attack using lightweight machine learning classifiers. LOCoCAT achieves up to 99.16% accuracy within the first 50 milliseconds of the attack, allowing vehicles to react promptly and maintain safety. The system is tested on various CAN datasets, demonstrating low latency and minimal resource consumption, making it suitable for embedded systems in vehicles.

(Seo et al., 2023) introduce a GAN-based approach to generate adversarial attacks capable of bypassing machine learning-based IDS in in-vehicle networks. By exploiting the vulnerabilities of IDS models like Multi-Layer Perceptron (MLP), Random Forest, and Logistic Regression, the GAN successfully reduces the detection accuracy to below 50%. The study emphasizes the need for more robust ML-based IDS to handle adversarial attacks and highlights the effectiveness of adversarial examples in exposing security weaknesses in vehicular networks.

(Cobilean et al., 2023) propose a transformer-based neural network called CAN-Former IDS for anomaly detection in CAN communications. The system simultaneously analyzes the sequence of CAN IDs and message payload values, using self-supervised learning to identify anomalous behavior. The method does not rely on labeled data and is tested on the Survival Analysis dataset, which includes data from three vehicle models. CAN-Former IDS achieves high precision and recall in detecting anomalies, particularly in malfunction attacks, demonstrating its effectiveness in securing in-vehicle networks.

(Barletta et al., 2023) explore the application of artificial intelligence and quantum machine learning in automotive cybersecurity, particularly focusing on supporting developers during the

automotive software development lifecycle. The study emphasizes that traditional machine learning techniques, though effective in identifying in-vehicle network attacks, are often too resource-intensive for embedded systems. To address this, the authors propose a hybrid approach combining traditional and quantum machine learning algorithms to optimize attack detection, while providing a decision-support framework for developers to integrate security solutions more effectively during the design and development phases of vehicle systems.

(Lampe and Meng, 2023) introduce the can-train-and-test dataset to address the shortage of adequate CAN bus datasets for IDS development and evaluation. The dataset includes CAN data from four different vehicles (Chevrolet and Subaru) and provides both attack-free and attack-laden traffic captures, covering nine unique attack types, such as DoS, fuzzing, and RPM spoofing. This dataset is particularly useful for evaluating the generalizability of IDS models across different vehicles and attacks, making it a valuable resource for automotive security research.

(Mavikumbure et al., 2023) propose DAdAE, a Domain Adversarial Autoencoder-based framework for anomaly detection in in-vehicle CAN bus networks. The system leverages domain adversarial training to detect anomalies across different car models without requiring labeled data. DAdAE demonstrates superior performance compared to traditional models, achieving high F1 scores in detecting attacks like flooding, fuzzy, and malfunction attacks. The system effectively generalizes across vehicle types, such as Hyundai Sonata, Kia Soul, and Chevrolet Spark, making it a powerful tool for securing diverse vehicular networks.

(Haddaji et al., 2023) evaluate existing datasets used for enhancing the security of intra-vehicular networks, focusing on datasets that support Artificial Intelligence solutions for anomaly detection and intrusion prevention. The paper provides a comparative study of prominent datasets such as Car-Hacking, OTIDS, and Survival, detailing their strengths, limitations, and applicability to various attack types, including DoS, fuzzy, and masquerade attacks. The authors also discuss preprocessing techniques and challenges in using these datasets to develop robust AI-based IDS for in-vehicle networks.

(Kohli et al., 2023) provide a detailed analysis of intrusion detection techniques for securing 6G applications, with a focus on vehicular networks. The paper highlights the increased security challenges posed by 6G's high speed and large data volumes, emphasizing the need for

advanced machine learning algorithms to detect and mitigate intrusions. The authors compare several intrusion detection methods, including signature-based, anomaly-based, and machine learning-based detection systems, demonstrating their effectiveness in preventing cyberattacks such as Denial of Service and fuzzy attacks in vehicular environments.

(Kumar et al., 2023) propose a machine learning-based approach for detecting replay attacks in Vehicular Ad Hoc Networks, specifically focusing on Basic Safety Messages (BSMs). The study demonstrates that the proposed model significantly outperforms existing methods in identifying replay attacks, where a malicious node rebroadcasts BSMs with false sender information. Using the Vehicular Reference Misbehavior (VeReMi) dataset, the system achieves higher detection accuracy by incorporating timestamp and velocity parameters, thus improving the security of vehicular communication systems.

(Monge and Andel, 2023) analyze the vulnerabilities of the CAN in modern vehicles and propose a conglomerate IDS to mitigate remote attacks. The system identifies anomalies in CAN bus traffic, particularly those originating from wireless interfaces, and blocks malicious signals before they can cause substantial damage. The paper emphasizes the critical need for IDS solutions to enhance vehicle security in the face of growing cyber threats, particularly given the increasing reliance on wireless communication technologies in vehicles.

(Zenden et al., 2023) investigate the resilience of machine learning-based IDS for automotive networks, particularly in defending against adversarial attacks. The study evaluates four different IDS models, including DNN, CNN, and LSTM-based systems, using adversarial samples generated through techniques like Fast Gradient-Sign Method (FGSM). The findings demonstrate that most of the studied IDS models are vulnerable to adversarial attacks, though adversarial training can enhance their resilience. The research emphasizes the need for further advancements in IDS robustness to ensure secure automotive networks.

(Rai et al., 2023) provide a comprehensive taxonomy of attacks in in-vehicle networks, categorizing them based on both physical and remote access points. The paper identifies common attack surfaces such as the OBD-II port, USB ports, and remote interfaces like cellular data and GPS systems. The authors discuss the increasing vulnerabilities in modern connected vehicles and emphasize the need for more robust security measures to protect in-vehicle communication systems. By analyzing a wide range of attack vectors, the paper offers insights

into how attackers can exploit different components, from ECUs to infotainment systems, to compromise vehicle safety and security.

(Alaya et al., 2024) propose an ontological anomaly detection approach (OADA) for vehicular ad hoc networks, utilizing hierarchical temporal memory (HTM) to identify anomalies such as network scans, DNS tunnel attacks, and telemetry data discrepancies. By organizing network traffic properties into instant description windows, the OADA system effectively captures and analyzes traffic patterns. The study demonstrates the system's capability to detect various types of anomalies in real-time using emulated communication datasets, highlighting its effectiveness in enhancing vehicular network security.

(Nandy et al., 2024) conduct a comprehensive review of security attacks and IDS in vehicular networks, particularly focusing on VANETs. The review highlights various attack vectors such as DoS, message suppression, and spoofing, and evaluates multiple IDS techniques like anomaly-based, signature-based, and hybrid models. The paper emphasizes the importance of adaptive, resource-efficient IDS solutions that can respond to the dynamic nature of VANETs, suggesting future directions for improving security in intelligent transportation systems.

(Verma et al., 2024) introduce a comprehensive guide to publicly available CAN IDS datasets and present the ROAD dataset, which consists of over 3.5 hours of CAN data from a single vehicle. The guide categorizes attacks such as fabrication, suspension, and masquerade, and provides quality analyses of existing datasets. The ROAD dataset includes real fuzzing, fabrication, and advanced attacks, providing researchers with a valuable resource for testing CAN IDS methods in real-world scenarios.

(Lampe and Meng, 2024) introduce the can-train-and-test dataset, specifically curated for the development and evaluation of automotive IDS. The dataset includes CAN data from four different vehicle models, containing both attack-free data and nine unique types of attacks, such as DoS and gear spoofing. Collected during live, on-the-road experiments, the dataset allows researchers to test IDS performance across varying vehicles and attacks. By offering labeled data and attack diversity, can-train-and-test addresses limitations in existing datasets, making it a valuable resource for machine learning-based IDS research.

(Rajapaksha et al., 2024) introduce CAN-MIRGU, a comprehensive dataset designed to evaluate IDS within in-vehicle networks, specifically targeting the CAN bus protocol. The dataset includes both benign traffic and multiple attack types, such as Denial of Service, replay, fuzzing, and spoofing attacks. Collected under real-world driving conditions, CAN-MIRGU is intended to simulate realistic scenarios in connected and autonomous vehicles. The dataset serves as a benchmark for testing and improving IDS models, offering detailed labeling for enhanced analysis of different attack types.

(Quadar et al., 2024) discuss the challenges, opportunities, and future research trends in the use of IDS for automotive Ethernet networks. The study highlights the increasing security risks posed by integrating Ethernet into vehicle communication systems, particularly for advanced applications such as autonomous driving and infotainment systems. The paper emphasizes the need for scalable IDS designs capable of handling real-time data traffic while ensuring high detection accuracy. Additionally, it identifies gaps in the availability of robust datasets and testbeds for evaluating IDS solutions, urging collaboration between automakers and researchers to address these challenges.

(Salek et al., 2024) introduce a hybrid quantum-classical framework for intrusion detection in the CAN of vehicles. This framework combines a classical neural network for feature extraction with a quantum Restricted Boltzmann Machine (RBM) for intrusion detection using CAN images. By leveraging quantum computing, the system significantly improves detection accuracy and recall compared to classical-only models, achieving up to 98.3% accuracy in detecting various types of CAN bus attacks, such as fuzzy attacks. The study demonstrates the potential of quantum-based approaches in enhancing vehicular cybersecurity.

(Timčenko and Boštjančič Rakas, 2024) analyze the security and privacy challenges in V2X communication systems. Their research identifies various cyber threats, including denial of service and data spoofing attacks, that can jeopardize the integrity and confidentiality of vehicular communication. By utilizing the ns-3 simulation environment, they examine the effectiveness of different cybersecurity measures aimed at enhancing the resilience of intelligent transportation systems. The findings highlight the necessity of robust security frameworks to ensure reliable operations within V2X scenarios.

(Fang et al., 2024) propose a novel windowed Hamming distance-based IDS for the CAN bus, addressing security vulnerabilities inherent in automotive ECUs. Their method utilizes sliding windows to analyze time series data, comparing newly generated data windows with recorded benchmarks using Hamming distance to detect anomalies. Experimental results show that this approach achieves high detection accuracies of 99.67% for DoS attacks and 98.66% for fuzzing attacks, significantly outperforming traditional methods. The study emphasizes the efficiency of their system, achieving a detection time of 0.37 ms per frame, making it suitable for real-time applications.

(Haddaji et al., 2024) investigate security threat datasets in both intra- and inter-vehicular environments, highlighting the importance of robust security in vehicular networks. Their survey provides a comprehensive overview of various artificial intelligence approaches aimed at enhancing vehicular security by analyzing large datasets for early threat detection. The authors introduce an innovative taxonomy to classify relevant datasets, which aids researchers in selecting the most appropriate data for specific security scenarios. Additionally, the study emphasizes the necessity of diverse and publicly available datasets to effectively address the unique security challenges posed by modern vehicular networks.

(Gao et al., 2024) develop a multi-attack intrusion detection model specifically designed for Controller Area Network with Flexible Data Rate (CAN-FD) messages. The proposed model comprises two sub-models: the Anomaly Data Detection Model (ADDM) for identifying anomalous frames and the Anomaly Classification Detection Model (ACDM) for classifying the types of anomalies. ADDM utilizes LSTM layers to capture temporal patterns in CAN-FD frame data, while ACDM incorporates an attention mechanism to enhance the classification of various attack types, including DoS, Fuzzing, and Spoofing. The experimental results indicate that their model significantly improves detection accuracy compared to existing methods, achieving an accuracy enhancement of up to 1.44% over other advanced models. The study underscores the model's effectiveness and applicability in real-time vehicular cybersecurity scenarios.

(Sundfeldt and Widstam, 2024) develop a machine learning-based IDS for in-vehicle communication using deep neural networks, specifically targeting the CAN bus. Their research emphasizes the necessity for robust cybersecurity measures in connected vehicles, in light of regulations like UN Regulation No. 155. They introduce a tool to generate synthetic attack data

by analyzing normal traffic patterns from real driving sessions and augmenting this data with labeled attack datasets. The study evaluates the performance of two deep learning models—Convolutional Neural Networks and Long Short-Term Memory for detecting various types of cyber-attacks, including Masquerade, Suspension, and Replay attacks. Experimental results indicate that the LSTM model shows a slight edge in detecting more complex attacks, while both models achieve high accuracy, highlighting the importance of employing machine learning techniques for enhancing vehicular cybersecurity.

(Marchiori and Conti, 2024) propose CANEDERLI, a novel framework that evaluates the impact of adversarial training and transferability on IDSs for CAN. They address the vulnerabilities of existing data-driven IDSs, which are susceptible to adversarial evasion attacks. The authors implement a realistic threat model, analyzing diverse state-of-the-art attacks and demonstrating strong transferability properties among various methodologies. Their research introduces an adaptive online adversarial training technique, which outperforms traditional fine-tuning methods with F1 scores reaching 0.941. By making their framework publicly available, they aim to assist researchers and practitioners in assessing the resilience of IDSs against adversarial threats, contributing to the security of vehicular networks.

(Kushardianto et al., 2024) propose a 2-step deep learning framework for anomaly detection in vehicular networks. Their approach focuses on V2V communication, utilizing two distinct classifiers trained on aggregated datasets to enhance detection accuracy. The framework first identifies whether a vehicle is legitimate or an attacker using a binary classifier and then classifies the type of attack using a second classifier. By employing deep learning techniques, including Deep Belief Networks (DBN), Long Short-Term Memory, and Random Forest, the study demonstrates improved performance in recognizing various attack types compared to traditional single-step detection methods. The authors emphasize the framework's effectiveness in real-time applications, achieving notable detection rates while addressing the challenges posed by dynamic vehicular environments.

(Kidmose and Meng, 2024) investigate and evaluate several automotive intrusion detection datasets, focusing on the HCRL Car Hacking dataset, the HCRL Survival Analysis dataset, and the newly curated can-train-and-test dataset. Their research aims to provide a comprehensive comparison of these datasets to better support the development of effective IDS in automotive networks. They highlight the need for sufficient training and testing data, particularly for

machine learning-based approaches. Through experiments, the authors analyze the impacts of features such as timestamps and data field subdivision on the performance of various machine learning models. The results demonstrate that the can-train-and-test dataset offers significant advantages in terms of size and class imbalance, enabling more effective training and evaluation of IDSs. The study concludes that can-train-and-test not only contains a greater volume of data but also provides unique insights, making it a valuable resource for future automotive cybersecurity research.

(Ahmad et al., 2024) conduct a comprehensive investigation into AI-based anomaly identification techniques for vehicle communication protocol systems. The authors examine the evolution of anomaly detection strategies implemented from January 2018 to January 2024, highlighting the vulnerabilities within the CAN systems that can lead to data breaches. They classify various identification methodologies and address the privacy concerns associated with AI designs. The paper discusses the prerequisites for developing effective AI-based anomaly detection approaches and identifies research opportunities and challenges within the domain. By providing an extensive overview of existing strategies, the study aims to assist researchers and automotive manufacturers in enhancing the security of CAN systems against potential cyber threats.

(Kirdi et al., 2024) present CARACAS, a vehicular architecture designed for detailed simulation of CAN attacks. This paper emphasizes the growing vulnerability of modern vehicles, particularly Battery Electric Vehicles (BEVs), to network attacks. The authors propose a novel approach that utilizes simulation frameworks, specifically Simulink, to generate synthetic datasets of CAN messages, allowing researchers to analyze the effects of cyber threats on vehicular systems. CARACAS integrates a BEV dynamic model, a CAN bus model, and a CAN attack injection simulator, enabling the assessment of attack scenarios on vehicle torque control in different driving conditions. The results illustrate the framework's capability to produce realistic CAN traffic and malicious signals, contributing to improved cybersecurity measures in automotive networks.

(Nathi, 2024) presents a study on misbehavior detection in vehicular ad hoc networks (VANETs) using unsupervised algorithms. The research addresses the critical challenge of position falsification attacks, where vehicles transmit inaccurate location information through Basic Safety Messages (BSMs), potentially leading to dangerous traffic situations. Due to the

difficulty in obtaining labeled datasets in real-world scenarios, the author compares the performance of four unsupervised machine learning models for detecting these attacks using the publicly available Vehicular Reference Misbehavior (VeReMi) dataset. The findings indicate that the proposed models effectively identify position falsification, demonstrating the potential of unsupervised learning techniques in enhancing the security and reliability of VANETs.

(Maliha and Bhattacharjee, 2024) propose a unified time series analytics-based intrusion detection framework specifically designed for CAN attacks. Their framework aims to detect multiple attack types, including Denial of Service (DoS), Fuzzy, and Impersonation attacks. By analyzing the timing information of CAN packets over partitioned time windows, the authors construct a low-dimensional time series latent space representing benign CAN network behavior. They implement a two-tier anomaly detection model that tracks both short-term and long-term deviations within this latent space to establish a stateful representation. The study introduces a preference-based multi-objective optimization approach for tuning hyperparameters, balancing trade-offs between false alarm counts, detection times, and missed detection rates. Experimental results demonstrate the effectiveness of the framework using real benign and attack datasets from a Kia Soul vehicle, showcasing its superiority over existing detection methods.

(Almehdhar et al., 2024) provide a comprehensive survey on advanced IDS for intelligent vehicle networks, with a specific focus on deep learning and machine learning approaches. The paper highlights the increasing cybersecurity threats faced by modern vehicles due to their interconnected nature, particularly within the CAN protocol. The authors propose a novel taxonomy categorizing IDS techniques into conventional machine learning, deep learning, and hybrid models, emphasizing their effectiveness in detecting various cyber threats, including spoofing and denial-of-service attacks. The review covers a wide range of AI algorithms, including reinforcement learning and generative adversarial networks, and discusses emerging technologies such as federated learning to enhance IDS robustness. Furthermore, the authors identify key limitations in current methodologies and suggest potential future research directions to improve real-time data analysis and cross-layer security measures.

(Bouchouia et al., 2024) propose new cybersecurity metrics for evaluating AI-based IDS specifically designed for in-vehicle networks. The authors emphasize the limitations of existing

metrics, including the Common Vulnerability Scoring System (CVSS), in addressing the unique challenges faced by automotive environments. They introduce a refined metric that considers the severity of attacks to provide a more accurate representation of an IDS's effectiveness. Through their analysis, the authors demonstrate how these new metrics can enhance the assessment of machine learning models deployed in vehicle networks, thereby improving overall security posture. The study highlights the importance of tailored evaluation frameworks to address the growing cybersecurity threats in autonomous vehicles and proposes future work to expand the metric set further.

(Yin et al., 2024) present CANGen, a practical framework for generating synthetic CAN traces using deep generative models (DGMs). The study addresses the challenge of obtaining high-quality, diverse CAN traces necessary for various automotive security applications, such as anomaly detection and simulation. CANGen efficiently tackles the issues of generation and evaluation by implementing an end-to-end, open-source framework. The authors conduct an empirical evaluation demonstrating that CANGen can produce high-fidelity synthetic CAN traces that meet domain-specific requirements. Additionally, the paper highlights the importance of adaptive preprocessing strategies to ensure the compatibility of different CAN trace formats and introduces a suite of fidelity metrics for comprehensive evaluation. This work aims to enhance the resources available for the automotive security community by providing realistic CAN datasets for testing and development.

(Han et al., 2024) propose an anomaly intrusion detection method for vehicular networks based on survival analysis. The study addresses the growing vulnerability of vehicles due to the convergence of in-vehicle networks and wireless communication technologies, which increases the risk of malicious attacks on ECUs. The authors aim to identify malicious CAN messages without requiring prior knowledge of CAN ID functions. Their method analyzes normal and abnormal driving data collected from different vehicle types to evaluate the performance of the proposed model across three attack scenarios: flooding, fuzzy, and malfunction attacks. The results indicate that the survival analysis model effectively distinguishes between normal and anomalous behavior in the CAN network, demonstrating high detection accuracy and low computational cost, which are essential for real-time applications in vehicle security.

(Santa Barletta et al., 2024) introduce Black Mamba, a multi-class Random Forest-based IDS specifically designed for the CAN in vehicles. The study addresses the growing cybersecurity

threats faced by modern vehicles, particularly with the integration of advanced ECUs. Black Mamba aims to detect various cyber-attacks, including flooding, fuzzy, and malfunction attacks, in real time. The authors utilize a Random Forest classifier to process CAN messages, leveraging features from a state-of-the-art dataset called the Survival Analysis Dataset. The experimental setup involves using an Arduino-based system to simulate attack messages while a Raspberry Pi analyzes the incoming data to determine if an attack is occurring. The results indicate high accuracy and precision in identifying different attack types, showcasing the effectiveness of implementing a multi-class IDS within vehicles to enhance cybersecurity.

Table 3.2 provides a comprehensive summary of research conducted on in-vehicle networks, detailing the methods employed, publication venues, and authors of each study in chronological order. Furthermore, it includes citations to the seminal work by Han et al., titled "Anomaly Intrusion Detection Method for Vehicular Networks Based on Survival Analysis," highlighting its influence in the context of IDS development and security threat mitigation in vehicular networks.

Table 3.2 - A Comprehensive Summary of Research Conducted on In-Vehicle Networks

Article Title	Date	Publication Place	Authors	Techniques Used	Summary	Article Citation	Performance
An Implementation of Invanet Hybrid Architecture and Hybrid Routing Protocol Using Deep Neural Network	Dec 2018	International Journal of Mechanical Engineering and Technology (IJMET)	M. Venkatesan et al.	Deep Neural Network (DNN), Zone Hierarchical Link State Protocol (ZHLS)	This paper proposes a method to implement a hybrid architecture and routing protocol in in-vehicle networks. Using Zone Hierarchical Link State Protocol (ZHLS) and DNN, it is aimed to find the shortest path between nodes in different regions and ensure reliable communication. MATLAB simulations show that the proposed methods provide effective communication and packet delivery rate.	It is referenced in the context of based intrusion detection system (IDS) methods on CAN bus and focuses on the detection of security threats through survival analysis. The reference is made as an evaluation to increase the effectiveness of IDS methods.	N/A
SDN-Based Secure and Privacy-Preserving Scheme for Vehicular Networks: A 5G Perspective	Sept 2019	IEEE Transactions on Vehicular Technology	Sahil Garg et al.	Elliptic Curve Cryptography (ECC), Fuzzy C-Means (FCM) Clustering, Tensor-based Dimensiona	It proposes a software-defined networking (SDN) based framework to enhance security and privacy in 5G-enabled vehicular networks. The framework provides a dual-module security system: one module provides authentication, and the other detects potential attacks during data flow. Evaluations on NS3,	It was cited when evaluating the effectiveness of IDSs based on survival analysis. This reference was made in the context of using IDS methods to detect security	N/A

				lity Reduction	SUMO and SPAN simulators show that the proposed framework is superior to existing solutions.	threats on the CAN bus.	
Shape of the Cloak: Formal Analysis of Clock Skew-Based Intrusion Detection System in Controller Area Networks	September 2019	IEEE Transactions on Information Forensics and Security	Xuhang Ying et al.	Clock Skew-Based IDS, Formal Analysis, Cloaking Attack, CAN Bus	This paper proposes a novel cloaking attack against clock skew-based intrusion detection systems on CAN bus. The study formally analyzes how an attacker can deceive IDS systems by manipulating clock skew. The paper demonstrates the effectiveness of cloaking attack on real vehicle testbeds (CAN bus prototype and UW EcoCAR).	Reference is made to show how IDS methods on CAN bus are evaluated against clock skew-based attacks. Reference is made to evaluate the role of survival analysis in detecting security threats	N/A
Internet of Things Meets Vehicles: Sheltering In-Vehicle Network through Lightweight Machine Learning	7 Nov 2019	Symmetry	Junchao Xiao et al.	Simplified Attention (SIMATT) Model, Security Control Unit (SECCU), Recurrent Neural Network (RNN)	This paper proposes a lightweight machine learning-based intrusion detection system (IDS) in in-vehicle networks. The SIMATT-SECCU framework is developed to perform anomaly detection on CAN bus, aiming to provide low computational cost and high performance. Experiments with real datasets show that the proposed framework achieves high accuracy and recall rates,	It was cited in an evaluation of how IDS methods used in in-vehicle networks can be optimized with low computational costs. This reference was made specifically to highlight how survival analysis is used in	Accuracy: 0.997 (99.7%) DoS Attack - Precision: 0.9930 - Recall: 0.9927 - F1 Score: 0.9929 Fuzzy Attack - Precision: 0.9901 - Recall: 0.9935 - F1 Score: 0.991

					especially with the use of SECCU and SIMATT models	practical applications.	
Robust and Self-Evolving IDS for In-Vehicle Network by Enabling Spatiotemporal Information	2019	IEEE Transactions on Information Forensics and Security	Junchao Xiao et al.	Convolutional LSTM (ConvLSTM), Pearson Correlation Coefficient, Spatiotemporal Data	This paper proposes a novel and robust mechanism for an intrusion detection system used in in-vehicle networks. The proposed system analyzes the traffic in the in-vehicle CAN network based on time series prediction and detects abnormal situations in real time. The ConvLSTM model provides more effective intrusion detection by using both temporal and spatial information and offers resilience against unknown attacks by adapting to environmental changes.	It is cited to evaluate the performance of IDS methods based on time series analysis. This citation is made to emphasize the applications where survival analysis is used in particular.	Precision: 0.97 Recall: 0.96 F1 Score: 0.96
Malware Detection in Self-Driving Vehicles Using Machine Learning Algorithms	17 Jan 2020	Journal of Advanced Transportation	Seunghyun Park et al.	Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (KNN), Gradient Boosting (GB)	This paper proposes an approach using machine learning algorithms to detect malware in driverless vehicles. An intrusion detection module is proposed for malware detection in in-vehicle networks. The research evaluates various machine learning algorithms to prevent malware from Android-based smartphones from infiltrating in-vehicle systems. The	It is referenced to improve the performance of intrusion detection systems on CAN bus. This reference is made to evaluate the reliability of IDS methods based on survival analysis	Accuracy: 92.9% Precision: 0.849 Recall: 0.761 F1 Score: 0.803

					proposed method enables high-accuracy malware detection through network traffic analysis.		
Decentralized and Secure Communication Architecture for FANETs using Blockchain	June 2020	Procedia Computer Science	Kashish Khullar et al.	Blockchain, Practical Byzantine Fault Tolerance (PBFT), Gossip Protocol	This paper proposes a decentralized flying ad hoc network (FANET) architecture. This blockchain-based structure aims to provide secure communication without the need for a central authority. In addition, the security and performance of the network are increased by using the Practical Byzantine Fault Tolerance (PBFT) algorithm and the Gossip protocol. Experimental results show that the proposed method offers a constant transaction speed and low latency according to the network size.	It referenced in a general context of how blockchain-based security solutions can be used for IDS applications in vehicular networks.	N/A
Hierarchical Anomaly Detection Model for In-Vehicle Networks Using Machine Learning Algorithms	15 July 2020	Sensors	Seunghyun Park et al.	Multi-Labeled Hierarchical Classification (MLHC), Decision Tree (DT), Random Forest (RF), Stochastic	This paper proposes a multi-layer classification model to detect abnormal behaviors occurring in in-vehicle networks. This model performs attack detection by hierarchically classifying additional information such as attack type and vehicle type. Simulation results show that	Reference is made in a general context to how IDS methods on CAN bus can be optimized for time-based attack detection. This reference is made to evaluate how	Accuracy: 99.99% Precision: 0.9993 Recall: 0.9998 F1 Score: 0.9995

				Gradient Descent (SGD), k-Nearest Neighbors (kNN)	the proposed model offers high accuracy and fast detection times.	survival analysis is used in vehicular networks.	
Controller Area Network Reliability: Overview of Design Challenges and Safety Related Perspectives of Future Transportation Systems	8 Oct 2020	IET Intelligent Transport Systems	Othman Nasri et al.	Networked Control Systems (NCS), Hardware-in-the-Loop (HIL), Real-Time Analysis (RTA), Fault Diagnosis	This paper examines the reliability and security issues on control area network (CAN) and presents possible solutions for the demands of future transportation systems. The increasing complexity and reliability issues of CAN pose a major challenge for the next generation of autonomous vehicles. The paper comprehensively analyzes the current research efforts and future challenges to improve CAN reliability.	Reference is made to how studies on CAN reliability can be used to increase the effectiveness of such IDS systems. The reference is made specifically to highlight the role of survival analysis in detecting security threats.	N/A
Cosine Similarity Based Anomaly Detection Methodology for the CAN Bus	12 October 2020	Expert Systems With Application	Byung Il Kwak et al.	Cosine Similarity, Self-Similarity, Hurst Parameter, CAN Bus	This paper proposes a cosine similarity-based method for anomaly detection in in-vehicle networks. The paper designs a lightweight feature vector that is used to detect three types of injection attacks with the dataset used in the survival analysis paper. The paper shows that high accuracy rates	It is cited to evaluate the effectiveness of survival analysis based on the frequency of CAN IDs. This reference is made to emphasize the importance of survival analysis	Accuracy: 99.05% Precision: 0.9933 - 0.9978 Recall: 0.9944 F1 Score: 0.9967

					are achieved for real-time detection.	in anomaly detection.	
LSTM-Based Intrusion Detection System for In-Vehicle Can Bus Communications	21 October 2020	IEEE Access	Delwar Hossain et al.	Long Short-Term Memory (LSTM), CAN Bus, Machine Learning, Deep Learning	This paper proposes an LSTM based intrusion detection system for CAN bus system used in modern vehicles. The study uses LSTM model to detect DoS, Fuzzing and Spoofing attacks and detects the attacks with high accuracy. The research demonstrates the effectiveness of IDS using a dataset generated from real vehicle data.	This reference is made to evaluate the effectiveness of IDS methods based on survival analysis. This reference is made to improve the intrusion detection performance of IDS systems, especially on CAN bus	Accuracy: 99.995% Precision: 100% (DoS, Spoofing), 99.98% (Fuzzing) Recall: 100% (DoS, Spoofing), 99.98% (Fuzzing) F1 Score: 99.995%
Cybersecurity for Autonomous Vehicles: Review of Attacks and Defense	10 Dec 2020	Computers & Security	Kyounggon Kim et al.	Systematic Literature Review, V2X Communication Security, Intrusion Detection Systems (IDS), Machine Learning, Big Data	This paper systematically reviews the attack and defense studies for autonomous vehicles. 151 papers published between 2008 and 2019 were reviewed and the attack methods were divided into three main categories: autonomous control systems, autonomous driving system components, and vehicle-to-everything (V2X) communication. The defense methods were classified as security architectures, intrusion	This reference is made in the context of using IDS on a vehicular network. This reference is made to highlight how survival analysis can be used to increase the effectiveness of intrusion detection systems.	N/A

					detection systems, and anomaly detection systems.		
Long Short-Term Memory-Based Intrusion Detection System for In-Vehicle Controller Area Network Bus	2020	IEEE Access	Delwar Hossain et al.	Long Short-Term Memory (LSTM), Deep Learning, CAN Bus	This paper proposes an LSTM based intrusion detection system for CAN bus system used in modern vehicles. The study uses LSTM model to detect DoS, Fuzzing and Spoofing attacks and detects the attacks with high accuracy. The research demonstrates the effectiveness of IDS using a dataset generated from real vehicle data.	It is cited to evaluate the effectiveness of IDS methods based on survival analysis. This citation is made to improve the intrusion detection performance of IDS systems, especially on CAN bus.	Accuracy: 99.9949% Precision: 1.0000 (DoS, Fuzzing, Spoofing) Recall: 1.0000 (DoS, Spoofing), 0.9994 (Fuzzing) F1 Score: 0.9997
Time-Based CAN Intrusion Detection Benchmark	14 Jan 2021	Workshop on Automotive and Autonomous Vehicle Security (AutoSec)	Deborah H. Blevins et al.	Mean Inter-Message Time, Binning, Gaussian Distribution, Kernel Density Estimation (KDE)	This paper compares four different methods to evaluate time-based CAN attack detection methods. Their performance on real (non-simulated) attacks is analyzed using the ROAD dataset. The results reveal that distribution-agnostic methods, especially the Binning method, outperform distribution-based methods.	It was cited during an evaluation of survival analysis-based methods used in previous studies on IDS methods on CAN bus. This reference was made specifically in the context of a comparative analysis of such time-based methods.	1. Mean Inter-Message Time Method: Accuracy (AUC-PR): 68.90% with outliers, 69.05% without outliers. Optimal F1 Score: 98.6% at the threshold $\alpha = 0.44$. Precision: 97.6% Recall: 99.6%

							2. Binning Method: Accuracy (AUC-PR): 87.63% with outliers, 82.18% without outliers. Optimal F1 Score: 99.0% at the threshold $\alpha = 3.5$. Precision: 98.6% Recall: 99.2% 3. Gaussian Method: AUC-PR (without outliers): 8.99% Optimal F1 Score: 25.99% at p-value threshold of 0.09. 4- Kernel Density Estimation (KDE) Method: AUC-PR (without outliers): 13.88% Optimal F1 Score: 27.79% at
--	--	--	--	--	--	--	--

							p-value threshold of 0.09.
Car Hacking and Defense Competition on In-Vehicle Network	Jan 2021	Workshop on Automotive and Autonomous Vehicle Security (AutoSec)	Hyunjae Kang et al.	Rule-Based Detection, Tree-Based Models (Random Forest, XGBoost, LightGBM)	This paper discusses the results of a hacking and defense competition held on in-vehicle networks (specifically on CAN bus). In the competition, participants developed attack and detection algorithms on a real vehicle. The study provides experiences with attacks and defense methods applied in the real world. Important features such as time interval and data byte patterns were highlighted in intrusion detection systems.	It was cited in the context of intrusion detection in in-vehicle networks with survival analysis methodology. This reference is made to evaluate the effectiveness of IDS methods.	F1 Score: 0.87
A Holistic Approach to Power Efficiency in a Clock Offset Based Intrusion Detection Systems for Controller Area Networks	17 March 2021	Pervasive and Mobile Computing	Subir Halder et al.	Clock Offset-Based Intrusion Detection System (COIDS), Cumulative Sum Method (CUSUM), Cooperative Game Theory	This paper presents COIDS, a clock offset-based intrusion detection system for in-vehicle networks. The work generates fingerprints of ECUs using clock offsets and uses the Cumulative Sum (CUSUM) method to detect anomalous deviations. Furthermore, a cooperative gaming model is developed to optimize the power consumption of ECUs considering the limited hardware resources.	It cited to evaluate the effectiveness of IDS methods based on clock offsets. This citation is made to emphasize the importance of survival analysis in time-based attack detection of COIDS.	Accuracy: 99.99% Detection Rate: 98.6% (DoS), 99.2% (Fuzzy), 99.8% (Impersonation) Average Detection Time: 0.204 seconds (DoS), 0.846 seconds (Impersonation)

Security Issues with In-Vehicle Networks, and Enhanced Countermeasures Based on Blockchain	6 April 2021	Electronics	Narayan Khatri et al	Blockchain, Machine Learning, Intrusion Detection Systems (IDS), CAN, FlexRay, Automotive Ethernet	This paper examines the vulnerabilities of in-vehicle networks and proposes a blockchain-based solution. Existing types of attacks on CAN, FlexRay, and automotive Ethernet protocols and security solutions developed against these attacks are detailed. In addition, a proposed hybrid blockchain approach is presented to enhance the security of vehicle networks.	It is cited in the context of the effectiveness of IDS systems in in-vehicle networks. This reference is made to highlight how survival analysis can be used to detect security vulnerabilities on the CAN bus.	N/A
In-Vehicle Network Anomaly Detection Using Extreme Gradient Boosting Machine	7 June 2021	10th Mediterranean Conference on Embedded Computing (MECO)	Afia Anjum et al	Extreme Gradient Boosting (XGBoost), 10-fold Cross-Validation, Early Stopping, Grid Search	This paper uses the XGBoost algorithm for anomaly detection in in-vehicle networks. In the research, DoS, Fuzzing, Spoofing and fault attack scenarios are applied on real CAN bus datasets. The model performs anomaly detection with high accuracy rates and low false positive rates.	It is stated that this work was implemented on the survival analysis dataset and is referenced in the context of improvements made to improve IDS performance.	Accuracy: >99% Precision: 98%+ (100% for RPM spoofing) Recall: 1.0000 (DoS), 0.9926 (Fuzzy), ~0.9 (Spoofing) F1 Score: >0.95
Deep Transfer Learning Based Intrusion Detection System for Electric	11 July 2021	Sensors	Tanzir Mehedi et al.	Deep Transfer Learning (DTL), LeNet Model, Convolution	This paper presents a proposed deep transfer learning (DTL) based intrusion detection system for electric vehicles. The LeNet model is used to detect abnormal traffic on CAN bus protocol.	It is cited in the context of evaluation of IDS methods on CAN bus and intrusion detection. This citation is made to	Accuracy: 98.10% Precision: 98.14% Recall: 98.04% F1 Score: 97.83% ROC AUC: 95.42%

Vehicular Networks				al Layers, MaxPooling, Softmax	Experiments show that the DTL based LeNet model exhibits superior performance compared to other traditional machine learning and deep learning models.	emphasize the role of survival analysis in anomaly detection on CAN bus protocol in particular.	
A Blockchain-Based Federated Forest for SDN-Enabled In-Vehicle Network Intrusion Detection System	26 July 2021	IEEE Access	Ibrahim Aliyu et al.	Blockchain, Federated Learning (FL), Random Forest, SDN (Software-Defined Networking), Fourier Transform	This paper presents a federated learning and blockchain based intrusion detection system for in-vehicle networks. This system enables training of models without sharing data between vehicle owners and manufacturers. Data security and scalability are ensured by using IPFS and SDN. CAN ID loops are analyzed in frequency domain using Fourier transform. The paper shows that the proposed system increases the intrusion detection rate and uses CPU and memory resources efficiently.	This work was cited in the context of CAN ID frequency analysis and it was stated that this method was effective in detecting attacks on the CAN bus.	Accuracy: 98.13% Precision: 97.8% Recall: 98.5% F1 Score: 98.15%
IoV-BERT-IDS: Hybrid Network Intrusion Detection System in IoV Using Large	August 2021	IEEE Transactions on Vehicular Technology	Mengyi Fu et al.	Large Language Models (LLM), BERT, Semantic Extractor,	This paper presents a BERT-based hybrid IDS system for intrusion detection in in-vehicle and out-of-vehicle networks. IoV-BERT-IDS is a model where data such as CAN bus and extra-vehicle network	Reference is made in the context of IDS methods on CAN bus and their development. This reference is	Accuracy: 99.96% Precision: 99.98% Recall: 99.97% F1 Score: 99.97%

Language Models				Byte Sentence (BS), CAN Bus	traffic are processed in "byte sentence" format. Thanks to the bidirectional features of BERT, IoV-BERT-IDS exhibits high overall performance and intrusion detection accuracy despite being trained with low-labeled data. The study has been tested on various datasets (CICIDS, BoT-IoT, Car-Hacking) and has shown superior performance.	made to how survival analysis can increase the effectiveness of IDS systems.	
In-Vehicle Network Attack Detection Across Vehicle Models: A Supervised-Unsupervised Hybrid Approach	19 Sept 2021	IEEE Intelligent Transportation Systems Conference (ITSC)	Shu Nakamura et al.	Supervised Learning, Unsupervised Learning, Autoencoder, LightGBM, CAN Bus	This paper proposes a hybrid method that combines supervised and unsupervised learning approaches to perform intrusion detection across vehicle models. The proposed model transfers the knowledge learned from past vehicle models to a new vehicle model. An unsupervised autoencoder performs anomaly detection using the normal data of the new vehicle model, while a pre-trained supervised model performs intrusion detection and classification. Experiments have shown that the proposed model outperforms previous models.	This reference is made in the context of transferring the knowledge learned from past vehicle models to new vehicle models. This reference is made to provide better performance of IDS methods on CAN bus.	Accuracy: KIA model: 98.8% HYUNDAI model: 97.6% CHEVROLET model: 81.3% Precision: KIA: 99.5% HYUNDAI: 99.6% CHEVROLET: 97.4% Recall: KIA: 99.0% HYUNDAI: 97.5%

							CHEVROLET: 81.3% F1 Score: KIA: 98.7% HYUNDAI: 93.0% CHEVROLET: 97.6%
A Review on Deep Learning Based Intrusion Detection Systems for Securing Controller Area Network	November 2021	Avrupa Bilim ve Teknoloji	Zinnet Duygu Akşehir et al.	Deep Learning, LSTM, DCNN, Signal Space Evaluation, SOEKS	This review article examines deep learning-based intrusion detection systems used to ensure CAN security in in-vehicle networks. The study focuses on how CAN bus datasets are collected, the features used, and the types of attacks. The deep learning techniques prominent in the literature are examined in detail and their performances are compared.	This reference has been made in the context of the effectiveness of deep learning methods in anomaly detection on CAN bus. This reference has been made especially for the evaluation of models working with time series data such as LSTM.	N/A
Machine Learning Based Misbehaviour Detection in VANET Using Consecutive BSM Approach	24 Dec 2021	IEEE Open Journal of Vehicular Technology	Aekta Sharma et al	Machine Learning, K-Nearest Neighbor, Naive Bayes, Misbehavio	This paper presents a machine learning-based approach to detect position spoofing attacks on vehicle-to-vehicle networks (VANETs). This approach is developed to ensure the accuracy of critical data, such	It is cited as an example of methods used in anomaly detection on CAN bus. This citation is made specifically to	Accuracy: 98% Precision: 97% - 99% Recall: 96% F1 Score: 97%

				r Detection, Consecutive Basic Safety Messages (BSM)	as location information contained in the Basic Safety Message (BSM), and detects attacks more effectively by combining information between consecutive BSMs. The method is particularly proposed to secure inter-vehicle communication and prevent accidents. This paper uses VeReMi Dataset. The performance metrics belongs that Dataset.	evaluate the performance of survival analysis on BSM based datasets.	
SOME/IP Intrusion Detection using Deep Learning-based Sequential Models in Automotive Ethernet Networks	2021	IEEE Information Technology , Electronics, and Mobile Communication Conference (IEMCON)	Natasha Alkhatib et al.	Recurrent Neural Network (RNN), Deep Learning, SOME/IP, Automotive Ethernet, Sequence Modeling	This paper proposes a deep learning based sequential model to detect attacks on SOME/IP protocol. This model, which works on in-vehicle Ethernet network, is developed specifically to detect attacks on SOME/IP application layer protocol. The authors created a dataset containing a realistic attack classification and demonstrated that the proposed RNN based model can detect attacks with high accuracy. This paper used SOME/IP dataset.	It is cited to highlight the impact of survival analysis on IDS performance. This citation is made to evaluate the effectiveness of IDS systems on CAN bus and Ethernet based protocols.	Accuracy: 99% Precision: 0.87 - 1.00 Recall: 0.61 - 1.00 F1 Score: 0.61 - 1.00
A Frequency-based Data Mining	2021	SAE Technical Paper	Joakim Rosell et al.	Frequency-based Data Mining,	This paper presents a frequency-based data mining approach to improve the	This citation is made for the purpose of	Accuracy: 99.1% Precision: 99.5% Recall: 98.7%

Approach to Enhance in-vehicle Network Intrusion Detection				Random Forest, Support Vector Machines (SVM), Linear Models (LM)	performance of intrusion detection in in-vehicle networks. The proposed method uses frequency-based features to detect attacks on the CAN bus protocol. Experiments on different datasets and attack types have shown that the proposed method provides fast and effective results. The Random Forest model exhibited the best performance compared to other models.	evaluating the performance of IDS methods on the CAN bus.	F1 Score: 98.9%
Event-Triggered Interval-Based Anomaly Detection and Attack Identification Methods for an In-Vehicle Network	2021	IEEE Transactions on Information Forensics and Security	Mee Lan Han et al.	Anomaly Detection, Attack Identification, Event-Triggered Interval, Statistical Moments, Machine Learning	This paper proposes an event-triggered interval-based method for anomaly detection and attack identification in in-vehicle networks (IVNs). Anomalies are detected by examining event-triggered intervals of CAN messages. The paper demonstrates that anomaly detection can be achieved with accuracy rates up to 99% in experiments conducted on different attack scenarios.	It is referred to as an example used in anomaly detection on CAN bus. This reference is made to evaluate the effectiveness of survival analysis on different attack scenarios.	Accuracy: 99% (mean) Precision: 94%
A Comparative Review of Security Threats	2021	2021 - International Conference	Dorsaf Swessi et al.	Dataset Comparison, V2X Communication	This paper comprehensively reviews the existing security threat datasets for vehicular networks. It discusses the	It is referenced in a general context of how it is used to compare data	N/A

Datasets for Vehicular Networks		on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)		tions, Intrusion Detection Systems (IDS), Machine Learning	differences, characteristics, and weaknesses of different datasets, providing guidance for selecting the most appropriate datasets. The study specifically examines datasets used for intra-vehicular and inter-vehicular networks and evaluates the effectiveness of each.	sets and improve the effectiveness of IDS systems.	
A Study on Attack Pattern Generation and Hybrid MR-IDS for In-Vehicle Network	2021	IEEE International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)	Dong Mug Kang et al.	Hybrid MR-IDS (Machine Learning and Ruleset), Feature Engineering, Random Forest, CAN Bus	This paper aims to develop Hybrid MR-IDS (Machine Learning and Ruleset) for in-vehicle networks. The paper proposes an IDS method that combines machine learning and ruleset to detect abnormal traffic on CAN bus. This method stands out with its ability to detect attack patterns that are not present in the training data. In experiments conducted on five different attack scenarios, it has been shown that the proposed system achieves high accuracy rates.	It is cited for the purpose of improving the effectiveness of IDS methods on CAN bus. This citation is specifically made to provide an assessment of how survival analysis can improve the intrusion detection performance of IDS systems.	Accuracy: 99.1% F1 Score: 0.99 Precision (GAN-based IDS for comparison): DoS: 0.968, Fuzzy: 0.973, Spoofing_1: 0.983, Spoofing_2: 0.981 Recall (GAN-based IDS for comparison): DoS: 0.996, Fuzzy: 0.995, Spoofing_1: 0.990, Spoofing_2: 0.965

NeuroCAN: Contextual Anomaly Detection in Controller Area Networks	2021	IEEE International Smart Cities Conference (ISC2)	Prashanth Balaji et al.	Linear Embeddings , Long Short Term Memory (LSTM), Spatio-Temporal Correlations , CAN Bus	This paper presents NeuroCAN, a deep learning model for contextual anomaly detection in in-vehicle networks. By learning spatial and temporal correlations between multiple sensor data, NeuroCAN detects hidden attacks that cannot be detected based on normal single sensor data analysis. The study demonstrates that NeuroCAN achieves over 95% detection accuracy through experiments on two different datasets.	It has been cited in the context of the development and performance evaluation of IDS methods on CAN bus. This citation is specifically made to evaluate the contribution of survival analysis methodology to anomaly detection in in-vehicle networks.	Accuracy: 95% F1 Score: 0.95 (Dataset 1), 1.00 (Dataset 2) True Positive Rate (TPR): 0.9 (Dataset 1), 1.0 (Dataset 2) False Positive Rate (FPR): 0.000065 (Dataset 1), 0.0 (Dataset 2)
Dynamic Voting based Explainable Intrusion Detection System for In-vehicle Network	16 February 2022	International Conference on Advanced Communications Technology (ICACT)	Nishat I Mowla et al.	Random Forest, Dynamic Voting, Explainable AI, Ensemble Learning, CAN Bus	This paper aims to develop an explainable intrusion detection system for in-vehicle networks. The proposed system aims to increase the explainability and performance of the model's decisions by adding a dynamic voting mechanism to a decision tree-based Random Forest model. Experiments on two different datasets have shown that the proposed system provides high accuracy rates and explainability.	It is stated that the dataset used for Survival Analysis was used in this study and referenced for the purpose of evaluating IDS performance.	Accuracy: 99.48% Precision: 99.48% Recall: 99.48% F1 Score: 99.48%

Best Practices in Preliminary Identification of Cybersecurity Attacks in Automotive Accidents: Generic Qualitative Inquiry	March 2022	Capella University (Doctoral Dissertation)	Micaiah C. Paddock	Critical Success Factors (CSFs), Forensic Analysis, CAN Bus, Interview-Based Research	This PhD thesis investigates the preliminary detection and forensic analysis of cyberattacks in automotive accidents. The study was conducted to identify best practices and challenges that law enforcement agencies use when they suspect cyberattacks are causing or contributing to vehicle accidents. In-depth interviews with fifteen law enforcement officers examined the cyberforensic analysis techniques and processes used after the accident. The study documents how law enforcement agencies approach this new area of digital evidence and incident response.	It is referenced in the context of intrusion detection on the CAN bus and how such methods are used in forensic analysis. This reference is made for the purpose of evaluating how survival analysis can be used in forensic analysis.	N/A
Demystifying In-Vehicle Intrusion Detection Systems: A Survey of Surveys and a Meta-Taxonomy	29 March 2022	Electronics	Georgios Karopoulos et al	Meta-Taxonomy, Intrusion Detection Systems, CAN Bus, V2X Communication	This paper comprehensively reviews previous research on intrusion detection systems (IDS) used in in-vehicle networks and simplifies the classifications in this field by creating a meta-taxonomy. The study aims to lay the foundation for future research by considering existing datasets and simulation tools. It	Reference is made in the context of how survival analysis can be used to increase the effectiveness of intrusion detection systems in in-vehicle networks.	N/A

					also provides a framework for evaluating solutions and developing new approaches in this field by bringing together different VIDS classification approaches.		
Benchmarking Framework for IDS Classifiers in Term of Security and Performance Based on Multicriteria Analysis	17 May 2022	Sultan Idris Education University (Ph.D. Thesis)	Amneh Hussein Mohd Alamleh	Multicriteria Decision-Making (MCDM), Fuzzy Delphi Method (FDM), VIKOR, TOPSIS, Borda Voting	This thesis presents a framework to assist developers in selecting the appropriate classification model for intrusion detection systems. The thesis proposes a method to evaluate and rank IDS classifiers using multiple criteria evaluated in terms of security and performance. In this work, the evaluation criteria are standardized using FDM and combined with MCDM methods such as VIKOR and TOPSIS.	Cited in the context of creating a benchmarking framework for evaluating and improving IDS performance. This citation was made to consider how survival analysis can play a role in improving the effectiveness of IDS systems.	N/A
Cluster-based Attacks Prevention Algorithm for Autonomous Vehicles Using Machine Learning Algorithms	19 May 2022	Computers and Electrical Engineering	R. Sankaranarayanan et al.	Convolutional Neural Networks (CNN), Clustering Algorithms, Demand-Side Management (DSM), Smart Grid	This paper proposes an attack prevention algorithm against security threats encountered in autonomous vehicles. The paper focuses on two different clustering models and their integration with three Convolutional Neural Network (CNN) models. These models are developed to provide energy optimization and	It is referenced in a study conducted to increase the effectiveness of IDS methods on CAN bus. This reference was made specifically to evaluate the performance of IDS systems used	N/A

					stability. The results obtained in the study show that the proposed algorithms provide high performance in terms of energy management and stabilization.	in autonomous vehicles.	
A GRU-Based Lightweight System for CAN Intrusion Detection in Real Time	27 June 2022	Security and Communication Networks	Haoyu Ma et al.	Gated Recurrent Unit (GRU), Lightweight Neural Networks, Feature Extraction, CAN Bus	This paper proposes a system for real-time intrusion detection in in-vehicle networks (CAN). This GRU-based model uses a feature extraction algorithm with low computational complexity and can be quickly implemented in in-vehicle embedded devices for real-time detection. The paper shows that the proposed system achieves high classification accuracy and operates with low computational power.	It is cited to evaluate the effectiveness of the methodologies used in anomaly detection on CAN bus. This citation is made to compare the performance of the GRU based system.	Accuracy: 99.93% Precision: DoS: 99.95%, Fuzzy: 99.32%, Spoofing: 99.95% Recall: DoS: 99.91%, Fuzzy: 99.13%, Spoofing: 99.31% F1 Score: DoS: 99.93%, Fuzzy: 99.22%, Spoofing: 99.63%
Increasing the Trustworthiness of AI-based In-Vehicle IDS using eXplainable AI	15 June 2022	Institution of Information Systems and Technology, Mittuniversitetet (Thesis)	Hampus Lundberg	Explainable AI (XAI), Random Forest, Deep Neural Networks, SHAP Values, Visualization	This thesis investigates how XAI (Explainable AI) methods can be used to increase the reliability of AI-based in-vehicle intrusion detection systems (IV-IDS). In the thesis, AI models (Random Forest and Deep Neural Networks) are trained using three different datasets and explainability is provided with SHAP values.	It is cited while emphasizing the importance of anomaly detection using survival analysis on the CAN bus. This reference is made specifically in the context of increasing the	N/A

				n-based Explanation	This thesis uses three datasets: Survival, ROAD, and Hisingen. In particular, a visualization-based explanation model (VisExp) is developed in order to increase the reliability of AI-based IV-IDSs.	effectiveness of IDSs.	
Using Deep Learning Networks to Identify Cyber Attacks on Intrusion Detection for In-Vehicle Networks	12 July 2022	Electronics	Hsiao-Chung Lin et al.	Deep Learning, VGG16, XGBoost, Convolutional Neural Networks (CNN), Intrusion Detection System (IDS), CAN Bus	This paper proposes an intrusion detection model using deep learning networks to detect cyber attacks on in-vehicle networks. Using VGG16 and XGBoost models, attacks on CAN Bus in particular are classified. In the study, the performance of the model is evaluated by experiments performed on data provided by Hacking and Countermeasure Research Lab (HCRL).	It is cited for the purpose of increasing the effectiveness of IDS methods.	Accuracy: 99.7880% for XGBoost with 50 trees. Precision: 99.9480% for XGBoost. Recall: 98.5739% for XGBoost. F1 Score: 99.2347% for XGBoost. ROC AUC: 99.7084% for XGBoost.
Experimental Analysis of Trustworthy In-Vehicle Intrusion Detection System Using eXplainable Artificial	22 September 2022	IEEE Access	Hampus Lundberg et al.	Explainable AI (XAI), SHAP Values, Deep Neural Networks (DNN), CAN Bus,	This paper investigates the use of eXplainable AI (XAI) methods to increase the reliability of an artificial intelligence (AI)-based system for intrusion detection in in-vehicle networks. The study introduces VisExp, a visualization method based on	It is cited to evaluate the use of survival analysis methodology on CAN bus. This citation is made specifically to improve the	Accuracy: 99.73% F1 Score: 0.9868 True Positive Rate (TPR): 99.8% True Negative Rate (TNR): 99.98%

Intelligence (XAI)				Visualization-based Explanation	SHAP values, and explains the behavior of the model using this method. The research demonstrates the potential of VisExp to increase explainability and reliability through a survey among experts.	performance of XAI methods.	
Android Head Units vs. In-Vehicle ECUs: Performance Assessment for Deploying In-Vehicle Intrusion Detection Systems for the CAN Bus	15 September 2022	IEEE Access	Tudor Andreica et al.	Machine Learning, Binary Classifiers, CAN Bus, Intrusion Detection System (IDS), Android Head Units, ECU Performance Comparison	This paper presents the performance evaluation of in-vehicle intrusion detection systems over CAN Bus on Android devices and traditional in-vehicle electronic control units. By evaluating various machine learning classifiers, the study shows that Android-based devices have higher computational and memory resources and therefore can accommodate more complex IDS algorithms. However, it is emphasized that traditional ECUs exhibit a more controlled and reliable real-time behavior. This paper uses ADAS system dataset and Duster dataset.	Reference is made in the comparison of machine learning methods used to assess the accuracy of IDS. This reference allows the results of the study to be evaluated in a broader perspective.	N/A
Statistical Detection of Adversarial Examples in	5 October 2022	IEEE Access	Ibrahim Aliyu et al.	Blockchain-Based Federated Forest,	This paper presents research conducted to make intrusion detection systems (IDS) in in-vehicle networks more resilient	References are made to highlight the importance of previous work	Accuracy: 99.48% Precision: 99.48% Recall: 99.48%

Blockchain-Based Federated Forest In-Vehicle Network Intrusion Detection Systems				Machine Learning, Statistical Adversarial Detector, CAN Bus, Intrusion Detection System	to adversarial samples using the Blockchain-based federated forest (BFF) method. The paper investigates how IDSs on CAN Bus can be affected by adversarial samples and proposes the integration of a statistical detector to protect against such attacks. The study shows that BFF-IDS is particularly vulnerable to a type of attack known as evasion attack.	done to improve the effectiveness of intrusion detection systems on CAN bus. This reference provides a reference to the methodological approach of the paper.	F1 Score: 99.48%
Hybrid Classifier Strategy with Tuned Training Weights for Distributed Denial of Service Attack Detection	1 November 2022	Concurrency and Computation: Practice and Experience	Deepak Dahiya	BiLSTM, Deep Belief Network (DBN), CSE-NAE Optimization Algorithm, 5G Networks, DDoS Attack Detection	This paper proposes a hybrid classifier strategy to detect distributed denial of service (DDoS) attacks in 5G networks. In the study, DDoS attacks are detected using BiLSTM and a DBN-based model. The CSE-NAE algorithm is used to optimize the DBN weights. Performance evaluations show that the proposed approach exhibits superior performance in metrics such as sensitivity, specificity, and accuracy.	References are made to highlight the importance of previous studies to improve IDS performance. This reference provides a methodological perspective on how the proposed method improves IDS performance.	Accuracy: 98.59%. Sensitivity: 96.82%. Specificity: 97.52%. Precision: 98.91%. F1 Score: 97.86%
TOW-IDS: Intrusion Detection System Based	2022	IEEE Transactions on Information	Mee Lan Han et al	Wavelet Transform, Deep Convolution	This paper proposes a three-overlapping wavelet-based intrusion detection system for in-vehicle networks. The study	It cited to support the use of survival analysis methodology in	Accuracy: 99.78% Precision: 99.94% Recall: 98.57%

on Three Overlapped Wavelets for Automotive Ethernet		Forensics and Security		al Neural Network (DCNN), ResNet, Automotive Ethernet, CAN Bus	specifically uses wavelet transform and a customized deep convolutional neural network (DCNN) to detect anomalies in the automotive Ethernet environment. Performance evaluations show that the proposed method can detect anomalies with high accuracy.	in-vehicle networks. This citation contributes to the methodological framework of the study.	F1 Score: 99.23%
Deep Learning and Blockchain-based Framework to Detect Malware in Autonomous Vehicles	2022	IEEE International Wireless Communications and Mobile Computing Conference (IWCMC)	Dev Patel et al.	Deep Learning (ResNet50V2), Blockchain, Convolutional Neural Networks (CNN), Malware Detection, Autonomous Vehicles	This paper proposes AVBlock, a deep learning and blockchain-based framework for detecting malware in autonomous vehicles. The study converts malware binaries into grayscale images by monitoring network traffic, and these images are classified using the CNN model. Using the ResNet50V2 architecture, malware is detected with 97.56% accuracy.	Reference is made in the context of the integration of survival analysis into IDS methods. This reference is made for malware detection and to improve the performance of IDS systems.	N/A
Keep the Moving Vehicle Secure: Context-Aware Intrusion Detection System for In-Vehicle CAN Bus Security	2022	14th International Conference on Cyber Conflict	Sampath Rajapaksha et al.	Gated Recurrent Unit (GRU), Time-Based Model, CAN Bus, Context-Aware IDS,	This paper proposes CAN-CID, a context-aware intrusion detection system (IDS) to detect cyber attacks on in-vehicle networks (CAN Bus). The study aims to predict CAN ID sequences using an ensemble model consisting of GRU and time-based model.	Reference is made in the context of the integration of survival analysis into IDS methodologies. This reference is made with the	N/A

				NLP Techniques	The CAN-CID model is trained with only benign data and ID-based and time-based thresholds are used for anomaly detection. The model has demonstrated superior performance compared to other models in the literature by achieving high F1-Score for various attacks. This paper uses ROAD dataset.	aim of improving IDS performance.	
Securing CAN Bus in Connected and Autonomous Vehicles Using Supervised Machine Learning Approaches	2022	IEEE International Conference on Electro Information Technology (eIT)	Ravalika Gundu et al.	Random Forest (RF), K-Nearest Neighbor (KNN), Extreme Gradient Boosting (XGBoost), Time Interval Feature, CAN Bus	This paper investigates supervised machine learning techniques to ensure the security of CAN Bus systems used in connected and autonomous vehicles (CAV). In the study, attacks on CAN Bus are detected using supervised learning algorithms such as RF, KNN, and XGBoost with time interval feature. The results show that the accuracy of the RF classifier increases up to 97% with the addition of time interval feature.	Cited in the context of a study to improve the effectiveness of IDS methods on CAN Bus. This citation was made to compare the performance of machine learning algorithms.	Random Forest (RF) Accuracy: 94% Precision: 93% Recall: 93% F1 Score: 93% K-Nearest Neighbor (KNN) Accuracy: 75% Precision: 72% Recall: 70% F1 Score: 70% XGBoost Accuracy: 92% Precision: 93% Recall: 92% F1 Score: 93%

IMPROVE: Intelligent Machine Learning based Portable, Reliable and Optimal Verification System for Future Vehicles	23 January 2023	International Conference on Computer Communication and Informatics (ICCCI)	Shreyas Madhav et al.	Machine Learning Ensemble, Random Forest, AdaBoost, Logistic Regression, Naïve Bayes, Bagging Tree, Artificial Neural Networks (ANNs), Malware Detection, CAN Bus	This paper introduces IMPROVE, a reliable and optimal verification system for autonomous vehicles. IMPROVE uses a multi-model machine learning approach for intrusion detection in in-vehicle networks (CAN) and detects malicious behaviors in inter-device communication by performing data analysis. The work consists of a machine learning module for CAN Bus security and a second module for inter-device network security. IMPROVE scans all message traffic to detect anomalous behaviors and generates security alerts.	Reference is made in the context of anomaly detection in in-vehicle networks. This reference is made to highlight the potential of survival analysis to increase the effectiveness of IDS systems.	Accuracy: 97.6% Precision: 96.8% Recall: 97.2% F1 Score: 96.9%
AI-Based Intrusion Detection Systems for In-Vehicle Networks: A Survey	February 2023	ACM Computing Surveys	Sampath Rajapaksha et al.	Machine Learning (ML), Deep Learning (DL), AI-based IDS, CAN Bus	This paper reviews AI-based intrusion detection systems for in-vehicle networks (IVNs). It evaluates attack types, characteristics, and datasets by considering AI-based IDSs released between 2016 and 2022. It also focuses on the limitations of existing approaches and future research directions.	Reference is made in the context of research to increase the effectiveness of AI-based IDSs on CAN Bus. This citation is made for the purpose of evaluating and comparing AI-	N/A

						based methodologies.	
Intrusion Detection System CAN-Bus In-Vehicle Networks Based on the Statistical Characteristics of Attacks	23 March 2023	Sensors	Junaid Khan et al.	Statistical Analysis, Brute Force Optimization, CAN Bus	This paper introduces a novel intrusion detection system (IDS) proposed for in-vehicle networks. In particular, it proposes an approach based on statistical properties of attacks for anomaly detection in CAN Bus networks. The study enables accurate detection of attacks by using a sliding window. Experiments have shown that the proposed system can effectively detect DoS and fuzzy attacks with low false positive rates.	The reference is given in the context of the dataset activity used in this paper. This reference is made to test the reliability of the proposed system.	Accuracy: Over 99% for detecting various types of attacks. Precision: 98% Recall: 99% F1 Score: 99%
G-IDCS: Graph-Based Intrusion Detection and Classification System for CAN Protocol	19 April 2023	IEEE Access	Sung Bum Park et al.	Graph Theory, Machine Learning, CAN Bus	This paper introduces G-IDCS, a graph-based intrusion detection and classification system for in-vehicle networks. G-IDCS aims to enhance the security of in-vehicle controller area network (CAN) protocol. The work combines graph-based detection and machine learning-based attack classification techniques to improve the capabilities of identifying and classifying attack types.	Reference is made in the context of the effectiveness of intrusion detection techniques based on survival analysis of CAN ID frequencies. This reference is made for the purpose of testing the accuracy of	Accuracy: Over 99% for detecting various types of attacks.

						the proposed method.	
ENIGMA: An Explainable Digital Twin Security Solution for Cyber-Physical Systems	14 June 2023	Computers in Industry	Sabah Suhail et al.	Digital Twins (DTs), Explainable AI (XAI), SHAP Values, Cybersecurity	This paper introduces ENIGMA, a gamification approach over digital twins (DTs) for the security of cyber-physical systems. ENIGMA provides an interactive gaming environment to train security analysts and assess the security of DTs. It also uses SHapley Additive exPlanations (SHAP) values to explain AI-based decisions and explains the reasons for security defense decisions.	Some basic concepts and methodologies used in the study are referred to in the context of security assessment and attack detection.	Accuracy: 97.6% Precision: 96.8% Recall: 97.2% F1 Score: 96.9%
Transformer-Based Attention Network for In-Vehicle Intrusion Detection	1 June 2023	IEEE Access	Trieu Phong Nguyen et al.	Transformer-based Attention Network, Multi-class Classification, Transfer Learning, CAN Bus	This paper proposes a multi-class intrusion detection system using a transformer-based attention network (TAN) for in-vehicle networks. The proposed model uses the self-attention mechanism to classify attacks on CAN Bus. Experimental results show that the model can detect replay attacks using sequential CAN IDs and improve the performance using the transfer learning method when training from small datasets.	Reference is made to test the effectiveness of the survival analysis dataset and to improve the accuracy of IDS on CAN Bus. This reference is made in the context of evaluating the effectiveness of transfer learning.	Accuracy: 99.78% Precision: 99.94% Recall: 98.57% F1 Score: 99.23%

An Anomaly Detection Method Based on Multiple LSTM-Autoencoder Models for In-Vehicle Network	22 August 2023	Electronics	Taeguen Kim et al.	LSTM-Autoencoders, Anomaly Detection, CAN Bus	This paper proposes a method that uses multiple LSTM-Autoencoder models to detect various anomalies for in-vehicle networks. The proposed method identifies normal network behavior and detects attacks by analyzing the changes in data transmission time and load values. Experimental results show that the method has a high accuracy rate (99%).	It is indirectly referenced in the context of a comprehensive review of vulnerabilities and intrusion detection methods. In particular, it is referenced to evaluate the effectiveness of methods used in intrusion detection on the CAN Bus.	Accuracy: 97% Precision: 98% Recall: 98% F1 Score: 0.98
Generative Neural Network-Based Defense Methods Against Cyberattacks for Connected and Autonomous Vehicles	August 2023	Clemson University (Ph.D. Dissertation)	M. Sabbir Salek	Generative Neural Networks, Hybrid Quantum-Classical RBM, GAN, Cybersecurity, Autonomous Vehicles	This PhD thesis investigates generative neural networks based defense methods for detection and prevention of cyber attacks for connected and autonomous vehicles (CAVs). The study presents an RBM model developed by combining quantum and classical approaches in CAN network and a GAN based defense method for traffic signal detection. The performance of the proposed hybrid systems in the thesis is evaluated with	It is cited to emphasize the importance of survival analysis for attack detection on CAN Bus. This citation is made to test the effectiveness of the generative defense methods developed.	Accuracy: 98% (Hybrid Quantum-Classical Model) Comparison Model Accuracy: 90% (Classical-Only Model)

					higher accuracy rates compared to classical methods.		
A Novel Hybrid Quantum-Classical Framework for an In-Vehicle Controller Area Network Intrusion Detection	8 September 2023	IEEE Access	M. Sabbir Salek et al.	Quantum-Classical Hybrid Framework, Neural Networks, Quantum Restricted Boltzmann Machine (RBM), CAN Bus	This paper proposes a hybrid quantum-classical intrusion detection system on the in-vehicle controller area network (CAN). The work aims to reconstruct and classify CAN images using a quantum restricted Boltzmann machine (RBM) and a classical neural network. Experiments based on real-world data show that the proposed hybrid system achieves higher intrusion detection accuracy compared to classical approaches alone.	Reference is made to compare the effectiveness of intrusion detection methods on CAN Bus. This reference is made to evaluate the performance of the proposed hybrid system.	DoS Attack Accuracy: 0.9993 Precision: 1.0000 Recall: 0.9986 F1 Score: 0.9993 Fuzzy Attack Accuracy: 1.0000 Precision: 1.0000 Recall: 1.0000 F1 Score: 1.0000
Supervised Contrastive ResNet and Transfer Learning for the In-vehicle Intrusion Detection System	13 October 2023	Expert Systems With Applications	Thien-Nu Hoang et al.	Supervised Contrastive Learning, ResNet, Transfer Learning, Deep Learning, CAN Bus	This paper proposes a novel ResNet based model using supervised adversarial learning and transfer learning to detect attacks on the in-vehicle controller area network (CAN). The transfer learning capabilities of the model are tested on datasets from different vehicle models and it is shown that it significantly reduces false negative rates, especially in Fuzzy attacks.	References are made to provide information about the comparisons made with the data used in the dataset used in the survival analysis and the results of the tests based on this dataset.	There are many machine learning results such as FNR, Rec, Prec, F1 score values.

A Hybrid Approach for Intrusion Detection in an In-vehicle Controller Area Network using Classical Convolutional Neural Network and Quantum Restricted Boltzmann Machine	30 October 2023	IEEE Transactions on Intelligent Transportation Systems	M. Sabbir Salek et al.	Classical Convolutional Neural Network (CCNN), Quantum Restricted Boltzmann Machine (QRBM), CAN Bus	This paper presents a hybrid approach using classical convolutional neural network (CCNN) and quantum confined Boltzmann machine (QRBM) for intrusion detection in in-vehicle networks (CAN). The superior performance of the hybrid approach is demonstrated through experiments with real-world data.	Reference is made for the purpose of evaluating the effectiveness of intrusion detection methods on CAN Bus. This reference is made for the purpose of comparing the performance of hybrid systems.	Kia Soul Accuracy: 97.5% Recall: 94.7% Hyundai Sonata Accuracy: 97% Recall: 93.9% Chevrolet Spark Accuracy: 98.3% Recall: 97.2%
Effective In-vehicle Intrusion Detection via Multi-view Statistical Graph Learning on CAN Messages	13 November 2023	IEEE Transactions on Intelligent Transportation Systems	Kai Wang et al.	Multi-view Statistical Graph Learning, Graph Convolutional Network (GCN), CAN Bus	This paper proposes a comprehensive method for intrusion detection in in-vehicle controller area networks (CAN) using two different graphical views (TCG and CRG). The study shows that the model named StatGraph provides both fine-grained detection and high performance.	Reference is given to compare intrusion detection techniques on CAN Bus and evaluate the effectiveness of these techniques.	Dataset: Survival Analysis Dataset DoS Attack: Accuracy: 99.8%, Precision: 99.53%, Recall: 97.21%, F1 Score: 98.36% Fuzzy Attack: Accuracy: 99.7%, Precision: 78.04%, Recall: 83.92%, F1 Score: 80.88% Dataset: ROAD Dataset

							Correlated Signal Masquerade Attack: Accuracy: 99.9%, Precision: 99.44%, Recall: 99.76%, F1 Score: 99.60%
ConvIDS: A Convolutional LSTM Based Intrusion Detection Model for In-Vehicle CAN Bus	15 December 2023	IEEE 6th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE)	Shuhan Wu et al.	Convolutional LSTM (ConvLSTM), Time Series Analysis, CAN Bus, Intrusion Detection System	This paper proposes ConvIDS, a convolutional long short-term memory (ConvLSTM) based model to detect attacks on in-vehicle controller area network (CAN). In the study, time series analysis is performed by converting CAN IDs into CAN images and attacks are successfully detected with this method.	Reference is made to evaluate the effectiveness of IDS methods on CAN Bus and to compare the performance of ConvLSTM based systems. This reference is made to highlight the potential of survival analysis to increase the accuracy of IDS systems.	Accuracy: 99.95% Precision: 99.76% Recall: 99.82% F1 Score: 99.79%
LOCoCAT: Low-Overhead Classification of CAN Bus Attack Types	December 2023	IEEE Embedded Systems Letters	Caio Batista de Melo, Nikil Dutt	Low-Overhead Attack Classification, CAN Bus, Lightweight Classifier Models	This paper introduces a methodology called LOCoCAT to classify attacks on in-vehicle controller area network (CAN) with low overhead. LOCoCAT generates attack blocks, extracts features from these blocks, and detects attack types quickly and accurately using a	Reference is made in the context of datasets and techniques used in intrusion detection. This reference is made to highlight the	N/A

					lightweight classifier model. Experiments show that high accuracy rates are achieved even when operating with low power consumption and lightweight hardware.	potential to increase the accuracy of IDSs on CAN Bus.	
Adversarial Attack of ML-based Intrusion Detection System on In-vehicle System using GAN	2023	IEEE ICUFN 2023	EunSeong Seo et al.	Generative Adversarial Network (GAN), Machine Learning, Adversarial Attack, t-SNE Visualization, CAN Bus	This paper presents a novel GAN-based attack method against machine learning-based intrusion detection systems (IDS) for in-vehicle networks. The study shows that attacks generated using GAN reduce the detection accuracy of IDSs below 50%. The t-SNE visualization reveals how different the attacks are from normal data.	Reference is made in the context of anomaly detection on CAN Bus. This reference is made to show how adversarial attacks can affect the performance of IDSs.	N/A
Anomaly Detection for In-Vehicle Communication Using Transformers	2023	IECON	Victor Cobilean et al.	Transformer Neural Network, Self-Supervised Learning, CAN Bus, Seq2Seq Modeling	This paper proposes CAN-Former IDS, a transformer-based neural network for anomaly detection in in-vehicle networks (CAN). The study aims to detect anomalous behaviors in CAN communication by performing sequential analysis of both CAN IDs and message payloads. With fully self-supervised training, the model trained on unlabeled data identifies anomalies by	Reference is made to test the effectiveness of the proposed transformer based model and to use the survival analysis data set as a reference data set for anomaly detection on CAN Bus.	Accuracy: 99.75% for Kia Soul, 98.35% for Hyundai Sonata, and 99.48% for Chevrolet Spark Precision: 97.49% for Kia Soul, 95.01% for Hyundai Sonata, and 98.12% for Chevrolet Spark

					comparing the expected communication sequence with the actual communication sequence.		Recall: 100% for Kia Soul, 99.11% for Hyundai Sonata, and 99.72% for Chevrolet Spark F1 Score: 98.73% for Kia Soul, 97.02% for Hyundai Sonata, and 98.92% for Chevrolet Spark
Artificial Intelligence for Automotive Security: How to Support Developers in Automotive Solutions	2023	IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence, and Neural Engineering (MetroXR AINE)	Vita Santa Barletta et al.	Machine Learning, Quantum Machine Learning, Quantum Annealing, AdaBoost, Random Forest, QBoost, Attack Taxonomy	This paper presents an AI-based vision model to support developers for automotive security. It compares traditional machine learning (ML) and quantum machine learning (QML) algorithms for in-car attack detection. In particular, AdaBoost, Random Forest, and QBoost algorithms are used to detect attacks against the CAN Bus protocol. The paper presents a taxonomy to assist developers in addressing security vulnerabilities in the automotive development lifecycle.	Reference is provided to compare the effectiveness of ML and QML algorithms for intrusion detection on CAN Bus. This reference is made to increase the accuracy of the datasets and analysis.	Accuracy: AdaBoost and RF: 99.9% QBoost: 99.4% Precision: AdaBoost and RF: 99.9% QBoost: 99.7% Recall: AdaBoost and RF: 99.9% QBoost: 99.7% F1 Score: AdaBoost and RF: 99.9% QBoost: 99.7%

can-train-and-test: A New CAN Intrusion Detection Dataset	2023	IEEE Vehicular Technology Conference (VTC 2023-Fall)	Brooke Lampe, Weizhi Meng	CAN Bus, Intrusion Detection System (IDS), Dataset Creation, Machine Learning, Attack Simulation	This paper introduces a new CAN intrusion detection dataset, “can-train-and-test.” This dataset includes data collected from four different vehicle models and provides researchers with the opportunity to develop and test IDSs by simulating various types of attacks. The dataset is presented in different formats, such as pre- and post-attack data, labeled and unlabeled data, and is designed to evaluate the ability of IDSs to generalize across vehicle types and attacks.	Reference is made for the purpose of comparing IDS methodologies on CAN Bus. This reference is made to emphasize the role of datasets in evaluating the effectiveness of IDS systems.	Malfunction Attack: Accuracy: 0.987 Precision: 0.972 Recall: 1.0 F1 Score: 0.985 Fuzzy Attack: Accuracy: 1.0 Precision: 1.0 Recall: 1.0 F1 Score: 1.0 Flood Attack: Accuracy: 1.0 Precision: 1.0 Recall: 1.0 F1 Score: 1.0
DAdAE: Domain Adversarial Autoencoder Based In-Vehicle CAN Anomaly Detection	2023	IECON	Harindra S. Mavikumbure et al.	Domain Adversarial Autoencoder, Unsupervised Learning, Anomaly Detection, CAN Bus	This paper proposes a domain adversarial autoencoder (DAdAE) based method for anomaly detection in in-car networks (CAN). DAdAE learns with unlabeled data and can work with limited datasets, providing effective anomaly detection on different vehicle models. The study shows that the model achieves high F1 scores, especially in experiments conducted on Hyundai Sonata, Kia Soul and	Reference is given to compare the effectiveness of anomaly detection methods on CAN Bus. Survival analysis dataset is used to test the DAdAE model and compare its performance.	Malfunction Attack: Accuracy: 0.987 Precision: 0.972 Recall: 1.0 F1 Score: 0.985 Fuzzy Attack: Accuracy: 1.0 Precision: 1.0 Recall: 1.0 F1 Score: 1.0

					Chevrolet Spark vehicle models.		Flood Attack: Accuracy: 1.0 Precision: 1.0 Recall: 1.0 F1 Score: 1.0
Intra-Vehicular Network Security Datasets Evaluation	2023	SIMULTECH 2023 - 13th International Conference on Simulation and Modeling Methodologies, Technologies and Applications	Achref Haddaji et al.	Dataset Evaluation, Vehicular Networks, Cybersecurity, Intra-Vehicular Networks	This paper presents a comprehensive study that evaluates existing datasets for in-vehicle network security. The study comparatively addresses the advantages and limitations of existing datasets, discusses data preprocessing processes, and examines the resources that can be used to develop security solutions. The paper highlights that datasets for in-vehicle networks are limited and focus specifically on the CAN Bus protocol.	References are provided to highlight the potential of datasets to increase the efficiency and accuracy of survival analysis-based methods in the context of anomaly detection.	N/A
Intrusion Detection Techniques For Security and Privacy of 6G Applications	2023	IEEE International Conference on Secure Cyber Computing and Communication	Priya Kohli et al.	Machine Learning (SVM, KNN, CNN), Hybrid IDS, 6G Security, Vehicular	This paper examines various intrusion detection techniques used for security and privacy of 6G applications. The study evaluates the potential vulnerabilities in 6G based vehicular communication and the intrusion detection systems used to detect these vulnerabilities. We particularly	Reference is provided to evaluate the effectiveness of survival analysis methods for intrusion detection on CAN Bus. This reference has	N/A

		tion (ICSCCC)		Communication	focus on machine learning algorithms such as SVM, KNN, and CNN, as well as hybrid IDS models.	been used in comparison of IDS techniques and security assessments.	
Machine Learning Based Detection of Replay Attacks in VANET	2023	IEEE 6th International Workshop on Intelligent Transportation and Autonomous Vehicles Technologies (ITAVT 2023)	Aman Kumar et al.	Machine Learning, Replay Attack Detection, VANET, Vehicular Communication	This paper proposes a machine learning based approach to detect replay attacks in VANET (Vehicular Ad-Hoc Networks). The study focuses on the repeated retransmission of basic security messages (BSM) within the network. Using various machine learning algorithms to detect these attacks, the researchers demonstrated through simulations that the proposed method provides a higher accuracy rate compared to existing techniques.	Reference is made in the context of intrusion detection on CAN Bus. This reference is made to evaluate the performance of the proposed methods and to highlight the role of survival analysis in IDS systems.	Accuracy is approx.. 97% for different machine learning algorithms
Use of Intrusion Detection Systems in Vehicular Controller Area Networks to Preclude Remote Attacks	2023	Proceedings of the 18th International Conference on Cyber Warfare and Security	Anthony J. Monge, Todd Andel	Intrusion Detection Systems (IDS), CAN Bus Security, Wireless Attack Mitigation	This paper examines intrusion detection systems used to secure in-vehicle controller area networks and prevent remote attacks. In particular, it evaluates existing IDS methods used to detect attacks on CAN Bus and discusses how these systems can improve vehicle security. The study focuses on blocking wireless attack	Reference is made to compare the effectiveness of IDS methodologies on CAN Bus and to evaluate the performance of these methodologies. This reference is	N/A

					vectors and introduces an IDS solution developed to detect such attacks.	made to emphasize the importance of analysis to improve the accuracy of IDS systems.	
On the Resilience of Machine Learning-Based IDS for Automotive Networks	2023	IEEE Vehicular Networking Conference (VNC)	Ivo Zenden et al.	Machine Learning, Adversarial AI/ML, Intrusion Detection System (IDS), CAN Bus	This paper evaluates the robustness and vulnerabilities of machine learning-based intrusion detection systems for in-vehicle networks. The study examines how much the performance of IDS models can be affected by using adversarial examples and how transferable these attacks are between different systems. It also discusses how the robustness of IDSs can be increased with adversarial training techniques.	Reference is made to evaluate the performance of IDS systems against adversarial attacks and to emphasize the role of survival analysis in this context.	N/A
Unveiling Threats: A Comprehensive Taxonomy of Attacks in In-Vehicle Networks	2023	16th International Conference on Security of Information and Networks (SIN 2023)	Ritu Rai et al.	Attack Taxonomy, In-Vehicle Networks (IVN), Cybersecurity, Attack Surfaces	This paper presents a comprehensive classification of attacks on in-vehicle networks (IVNs). The study focuses on different attack surfaces, such as OBD-II, USB ports, short-range and long-range communication methods, detailing attacks via physical and remote access. Measures to	Reference is made to evaluate the vulnerabilities on CAN Bus and the effectiveness of intrusion detection methods. This reference is made to emphasize the	N/A

					be taken for the security of in-vehicle networks are discussed and possible threats that may create vulnerabilities are analyzed.	importance of the analysis performed to increase the effectiveness of the proposed attack classification.	
An Ontological Approach to the Detection of Anomalies in Vehicular Ad-Hoc Networks	24 January 2024	Ad Hoc Networks	Bechir Alaya et al.	Hierarchical Temporal Memory (HTM), Ontology-based anomaly detection	The paper presents the Ontological Anomaly Detection Approach (OADA) using HTM to detect anomalies in vehicular networks, focusing on network scans, DNS tunnel attacks, and telemetry anomalies. It evaluates anomaly detection with a communications and anomalies emulation tool using Autobot.	This study has been evaluated as an important reference point.	N/A
A Review of Security Attacks and Intrusion Detection in the Vehicular Networks	1 February 2024	Journal of King Saud University - Computer and Information Sciences	Tarak Nandy et al.	Intrusion Detection Systems (IDS), cybersecurity, adaptive IDS mechanisms	This paper provides a comprehensive review of security threats in VANETs and the Intrusion Detection Systems developed to combat these threats. The paper discusses various IDS approaches, existing solutions, and future research directions.	Reference has been taken through the article to evaluate the effectiveness of CAN based IDS methods.	N/A
A Comprehensive Guide to CAN IDS Data and	8 February 2024	PLoS ONE	Miki E. Verma et al.	Frequency/ Timing-Based, Payload-	This paper presents a comprehensive review of existing datasets for CAN IDS investigations and introduces	This study is taken as a reference to evaluate the	N/A

Introduction of the ROAD Dataset				Based, Signal-Based, Fabrication Attacks, Masquerade Attacks	the ROAD dataset developed by ORNL. The ROAD dataset contains real vehicle CAN data and a range of real and simulated attacks, making it suitable for comparative testing of CAN IDS methods.	effectiveness of time-based methods for detecting attacks on the CAN bus.	
can-train-and-test: A curated CAN Dataset for Automotive Intrusion Detection	23 February 2024	Computers & Security	Brooke Lampe, Weizhi Meng	Machine Learning IDS, CAN bus, automotive security, dataset	This paper presents a new CAN dataset, can-train-and-test, to support automotive IDS investigations. This dataset contains CAN data collected from four different vehicles and covers nine unique attacks. The paper discusses the limitations of existing datasets and highlights the need for a more comprehensive dataset.	This study has been used as an important basis for evaluating existing intrusion detection systems.	Multi-layer Perceptron (MLP): Known vehicle, known attack: F1-score: 0.9811 Known vehicle, unknown attack: F1-score: 0.9964 Unknown vehicle, known attack: F1-score: 0.9594 Local Outlier Factor (LOF): Known vehicle, known attack: F1-score: 0.9465 Known vehicle, unknown attack: F1-score: 0.9665 Unknown vehicle, known attack: F1-score: 0.2640

CAN-MIRGU: A Comprehensive CAN Bus Attack Dataset from Moving Vehicles for Intrusion Detection System Evaluation	26 February 2024	Symposium on Vehicles Security and Privacy (VehicleSec)	Sampath Rajapaksha et al.	Intrusion Detection Systems (IDS), CAN bus, automotive security, dataset	This paper introduces CAN-MIRGU, a new dataset for collecting CAN bus attack data under real-world driving conditions.	It has been used as a critical reference to evaluate the effectiveness of intrusion detection systems on the CAN bus	N/A
Intrusion Detection Systems in Automotive Ethernet Networks: Challenges, Opportunities and Future Research Trends	March 2024	IEEE Internet of Things Magazine	Nordine Quadar et al.	Network-based IDS, Host-based IDS, anomaly detection, machine learning	This paper comprehensively examines the role of intrusion detection systems (IDS) in addressing cybersecurity threats in automotive Ethernet networks. The paper highlights current challenges, techniques, and future research directions.	It has been taken as a reference for anomaly detection in Ethernet-based vehicle networks.	N/A
A Novel Feature Extraction Framework Using Graph Node Attention Network for In-Vehicle Networks	March 2024	IEEE Systems Journal	Junchao Xiao et al.	Graph Node Attention Network (GNAT), Random Forest, CAN bus, automotive security	This paper presents a new feature extraction framework using graph node attention network (GNAT) for the detection of anomalous packets in CAN bus. The paper uses this method to identify various types of attacks and compares its performance with other traditional methods.	This study has been evaluated as an important reference point in the development of graph-based anomaly detection methods.	Quantum RBM Model: Accuracy: 97.5%, Recall: 94.7% Classical RBM Model: Accuracy: 92.5%, Recall: 84.2%

Intrusion Detection							
Attack Identification and Classification in V2X Scenarios	22 March 2024	23rd International Symposium INFOTEH-JAHORINA	Valentina Timčenko et al.	ns-3 simulator, 5G-LENA, NR V2X module, cyber attacks, V2X scenarios	This paper examines the security and privacy issues that arise in V2X scenarios, analyzing different types of attacks and intrusion detection techniques in these environments.	It has been used as a strong reference point in research on anomaly detection and attack classification in V2X scenarios.	N/A
Windowed Hamming Distance-Based Intrusion Detection for the CAN Bus	27 March 2024	Applied Sciences	Siwei Fang et al.	Windowed Hamming Distance, CAN bus, Intrusion Detection System, time-series analysis	This paper presents a new method that uses Hamming distance and shifting window methods to detect anomalies in CAN bus. Experiments have shown that the method can detect DoS attacks with 99.67% accuracy and fuzzing attacks with 98.66% accuracy.	This study has been evaluated as an important reference point.	DoS Attack: Accuracy: 99.76% Precision: 99.91% Recall: 99.31% F1 Score: 99.61% Fuzzy Attack: Accuracy: 98.66% Precision: 99.68% Recall: 96.38% F1 Score: 98.02% Spoofing Attack (Gear): Accuracy: 99.48% Precision: 99.89% Recall: 98.66% F1 Score: 99.27%

							Spoofing Attack (RPM): Accuracy: 99.61% Precision: 99.99% Recall: 98.98% F1 Score: 99.48%
Investigation of Security Threat Datasets for Intra- and Inter-Vehicular Environments	26 May 2024	Sensors	Achref Haddaji et al.	Vehicular networks, artificial intelligence, security, Internet of Vehicles, datasets	This paper comprehensively examines existing datasets for security threats in in-vehicle and inter-vehicle networks, and discusses how these datasets can be used for security solutions. It also provides a guide on the classification of these datasets and the selection of the most appropriate datasets for different research scenarios.	This study has been evaluated as an important reference point.	N/A
Multi-Attack Intrusion Detection for In-Vehicle CAN-FD Messages	27 May 2024	Sensors	Fei Gao et al.	LSTM, attention mechanism, anomaly detection, CAN-FD, vehicle security	This paper proposes a new anomaly detection model consisting of two sub-models for multiple attack detection for CAN-FD messages. The model effectively detects DoS, Fuzzing, Replay, Spoofing, Scaling, and Ramp attacks and offers higher accuracy compared to existing methods.	This study has been evaluated as an important reference point.	N/A
Intrusion Detection for	27 May 2024	Lund University	Frida Sundfeldt,	CNN, LSTM,	This thesis develops a deep learning based Intrusion	It was used as an important	DoS Attack:

In-Vehicle CAN Communication Using Deep Neural Networks		Master's Thesis	Bianca Widstam	Deep Learning, CAN bus, Vehicle Security	Detection System to detect cyber attacks on CAN bus. The data used in the study includes real-world datasets and synthetically generated attack scenarios.	reference in the performance evaluation of LSTM and CNN models.	Accuracy: 99.76% Precision: 99.91% Recall: 99.31% F1 Score: 99.61% Fuzzy Attack: Accuracy: 98.66% Precision: 99.68% Recall: 96.38% F1 Score: 98.02% Spoofing Attack: Accuracy: 99.48% Precision: 99.89% Recall: 98.66% F1 Score: 99.27% Replay Attack: Accuracy: 99.61% Precision: 99.99% Recall: 98.98% F1 Score: 99.48%
CANEDERLI: On The Impact of Adversarial Training and Transferability on CAN	31 May 2024	ACM Workshop on Wireless Security and Machine	Francesco Marchiori, Mauro Conti	Adversarial Training, Transferability, CAN Bus, Intrusion	This paper evaluates the resilience of CAN bus IDSs against adversarial attacks under CANEDERLI framework, focusing on adversarial transferability. The	Reference was taken to examine the impact of the attacks evaluated within the	DNN Model: Accuracy: 100%, F1 Score: 100% (Sonata), 99.5% (Soul), 99.7% (Spark)

Intrusion Detection Systems		Learning (WiseML '24)		Detection, Deep Learning	paper examines the impact of attacks in white-box and black-box scenarios and introduces the proposed adaptive online adversarial training methods.	CANEDERLI framework.	CNN Model: Accuracy: 99.9% (Sonata), F1 Score: 99.3% (Soul and Spark) LSTM Model: Accuracy: 100% (Sonata), F1 Score: 99.5% (Soul and Spark)
Vehicular Network Anomaly Detection Based on 2-step Deep Learning Framework	June 2024	Vehicular Communications	Nur Cahyono Kushardianto et al.	2-Step Prediction, Deep Learning, LSTM, GRU, Random Forest	This paper presents a two-stage deep learning framework for anomaly detection in in-vehicle networks. The proposed method has been applied to detect various types of attacks using different machine learning models.	This work is highlighted as an effective method for detecting unknown attacks and discussed in comparison with the two-stage deep learning method.	Single Prediction (5 messages): GRU Model: Accuracy: 83%, Precision: 0.57, Recall: 0.96, F1-score: 0.71 Random Forest Model (RF): Accuracy: 74%, Precision: 0.66, Recall: 0.72, F1-score: 0.69 2-Step Prediction (30 messages): LSTM Model: Accuracy: 78%,

							Precision: 0.74, Recall: 0.82, F1-score: 0.78 Random Forest Model (RF): Accuracy: 87%, Precision: 0.93, Recall: 0.72, F1-score: 0.81
Investigating and Evaluating Automotive Intrusion Detection Datasets	June 2024	European Interdisciplinary Cybersecurity Conference (EICC)	Brooke Kidmose, Weizhi Meng	Machine Learning, CAN Bus, Intrusion Detection, Dataset Comparison	This paper comprehensively investigates automotive IDS datasets and evaluates the impact of different datasets by comparing them. In particular, experimental results are presented comparing the can-train-and-test dataset with other popular datasets.	It has been used as an important reference in the evaluation of CAN-based IDSs.	Multi-layer Perceptron (MLP): F1 Score with timestamp and data subdivision: 0.9555 F1 Score without timestamp: 0.9115 F1 Score without data subdivision: 0.9320 Gaussian Naive Bayes: F1 Score with timestamp and data subdivision: 0.9626

							F1 Score without timestamp: 0.9538 F1 Score without data subdivision: 0.9775
AI-based Anomaly Identification Techniques for Vehicles Communication Protocol Systems: Comprehensive Investigation, Research Opportunities and Challenges	June 2024	Internet of Things	Hasnain Ahmad et al.	Artificial Intelligence, Anomaly Detection, Machine Learning, CAN Bus	This paper comprehensively examines how AI-based anomaly detection techniques can be used in vehicular communication protocols. The paper discusses current research opportunities and challenges in this area, while also suggesting future research directions.	This study has been evaluated as an important reference point.	N/A
CARACAS: Vehicular Architecture for Detailed CAN Attacks Simulation	June 2024	TrustAICyberSec workshop - IEEE ISCC 2024	Sadek Misto Kirdi et al.	CAN Bus, Simulink, CAN Attack Simulation, Vehicular Systems	This paper introduces CARACAS, a new simulation framework to simulate CAN attacks and study their effects on vehicle systems. The study evaluates the effects of attacks on a CAN network through simulations on a BEV (Battery Electric Vehicle) model.	It was used as a reference to evaluate the effectiveness of defense mechanisms against CAN attacks.	N/A
Misbehaviour Detection in Vehicular Networks	13 June 2024	University of Windsor Master's Thesis	Girish Nathi	Unsupervised Learning, Clustering, VeReMi	This thesis compares the performance of four different unsupervised learning models for detecting misbehavior in in-	This study has been evaluated as an important reference point.	K-means: Accuracy: 82.5% Precision: 80.3% Recall: 78.9%

Using Unsupervised Algorithms				Dataset, VANET	vehicle networks. These models are evaluated using the VeReMi dataset to detect position spoofing attacks.		F1-score: 79.6% DBSCAN: Accuracy: 84.2% Precision: 81.7% Recall: 80.1% F1-score: 80.9% Isolation Forest: Accuracy: 85.6% Precision: 83.2% Recall: 82.0% F1-score: 82.6%
A Unified Time Series Analytics based Intrusion Detection Framework for CAN BUS Attacks	21 June 2024	ACM Conference on Data and Application Security and Privacy (CODASPY '24)	Maisha Maliha, Shameek Bhattacharjee	Time Series Analysis, Latent Space, Quantile L1 Regression, Multi-Objective Optimization	This paper proposes a unified framework based on time series analysis to detect various cyber attacks on CAN bus. The proposed model performs anomaly detection by monitoring short-term and long-term memory and learns hyperparameters based on optimized security objectives.	This study has been evaluated as an important reference point.	N/A
Deep Learning in the Fast Lane: A Survey on Advanced Intrusion Detection Systems for Intelligent	2 July 2024	IEEE Open Journal of Vehicular Technology	Mohammed Almehdhar et al.	Deep Learning, Intrusion Detection Systems, CAN bus, Federated Learning,	This paper comprehensively reviews advanced intrusion detection system methods using deep learning approaches, addressing cybersecurity threats in modern vehicles. The paper focuses on the CAN protocol in particular,	This study has been evaluated as an important reference point.	N/A

Vehicle Networks				Transfer Learning	discussing its vulnerabilities and defense mechanisms.		
Cybersecurity Metrics for AI-Based In-Vehicle Intrusion Detection Systems	24 July 2024	IEEE Vehicular Networking Conference (VNC)	Mohammed Lamine Bouchouia et al.	Machine Learning, CVSS Metrics, Autonomous Vehicles, Intrusion Detection	This paper proposes new metrics to evaluate the performance of AI-based intrusion detection system for in-vehicle networks. The metrics aim to improve the practicality and effectiveness of AI/ML-based IDSs by considering cybersecurity information.	The reference to this study is discussed in comparison with the newly proposed cybersecurity metrics.	N/A
Work-in-Progress: CANGen: Practical Synthetic CAN Traces Generation Using Deep Generative Models	2024	IEEE European Symposium on Security and Privacy Workshops	Yucheng Yin et al	Deep Generative Models, Synthetic Data Generation, CAN Bus	This paper introduces a framework called CANGen, which was developed to generate and evaluate synthetic CAN traces for in-vehicle networks. The framework addresses the use of various Deep Generative Models (DGM) to generate high-fidelity synthetic CAN traces for different datasets and applications.	This study has been evaluated as an important reference point.	N/A
Anomaly intrusion detection method for vehicular networks based on survival analysis	2024	SSRN	Kallepalli Rohit Kumar et al.	Survival Analysis, CAN Bus, Anomaly Detection	This paper proposes a model based on survival analysis to effectively identify abnormal and normal activities in CAN networks. In the study, normal and abnormal driving data collected from three different vehicles are analyzed.	The study was evaluated for its ability to quickly detect anomalies in vehicle networks.	Hyundai YF Sonata: DoS: Accuracy: 98.61%, Recall: 98.55% Fuzzy:

							Accuracy: 97.35%, Recall: 97.25% Malfunction: Accuracy: 99.95%, Recall: 99.94% Kia Soul: DoS: Accuracy: 99.69%, Recall: 99.69% Fuzzy: Accuracy: 98.87%, Recall: 95.03% Malfunction: Accuracy: 99.98%, Recall: 99.96% Chevrolet Spark: DoS: Accuracy: 92.35%, Recall: 92.31% Fuzzy: Accuracy: 87.08%, Recall: 86.25% Malfunction: Accuracy:
--	--	--	--	--	--	--	---

							99.83%, Recall: 99.83%
Black Mamba: A Multi-class Random Forest for Automotive Intrusion Detection	2024	University of Bari Aldo Moro	Vita Santa Barletta et al.	Random Forest, Multi-Class Classificatio n, CAN Bus	This paper proposes a Random Forest based multi-class Intrusion Detection System to detect attacks on CAN bus. The system aims to perform real-time intrusion detection in in-vehicle networks and various attacks are simulated using Arduino Elegoo Uno R3 in the testing phase.	The work of Han et al. was used as an important reference in the development of the system.	Normal: Accuracy: 99.98% Precision: 99.99% Recall: 99.98% F1 Score: 99.99% Flooding: Accuracy: 100% Precision: 100% Recall: 100% F1 Score: 100% Fuzzy: Accuracy: 99.79% Precision: 99.93% Recall: 99.86% F1 Score: 99.89% Malfunction: Accuracy: 99.90% Precision: 99.98% Recall: 99.94% F1 Score: 99.96%

4. CASE STUDY OF HCLR CAR HACKING DATASET

4.1. Dataset Instruction

This section presents a detailed analysis of datasets belonging to Hyundai Sonata, KIA Soul and Chevrolet Spark vehicles. In order to ensure the security of the In-Vehicle Network, attack and normal data types found in these datasets were examined. This analysis was conducted to understand the effects of Flooding, Malfunction and Fuzzy attacks on in-vehicle networks and to increase the effectiveness of anomaly detection models. The datasets utilized in this section were taken from the research conducted by Mee Lan Han and colleagues, as detailed in their paper titled "Anomaly Intrusion Detection Method for Vehicular Networks Based on Survival Analysis." The data provided by Han et al. forms the foundation for the analysis presented in this section.

4.2. Dataset Definition

The datasets analyzed in this study are taken from the work of Mee Lan Han et al. in their paper titled "Anomaly Intrusion Detection Method for Vehicular Networks Based on Survival Analysis." As detailed in their research, these datasets are based on CAN messages collected through the OBD-II port of three different vehicles: Hyundai Sonata, KIA Soul, and Chevrolet Spark. The data collection process described by Han et al. was designed to capture real-world vehicle operating conditions both under normal circumstances and during various types of simulated cyberattacks. This setup provides the foundation for the datasets used in our analysis.

In the study by Han et al., CAN messages were obtained using a specialized data collection setup that involved the use of a Raspberry Pi3 microcomputer connected to the vehicle's OBD-II port via a PiCAN2 interface board. This setup was critical for accurately capturing CAN messages directly from the vehicle's CAN bus, with the communication

speed set to 500 Kbps to match the standard used in the vehicles' networks. The data collected included both normal driving data, representing typical vehicle operations, and abnormal driving data generated by deliberately introducing anomalies into the vehicle's network through simulated attacks using the python-can library.

The datasets from Han et al.'s study cover three specific vehicle models: the Hyundai YF Sonata 2010, a midsize sedan; the KIA Soul 2015, a compact SUV; and the Chevrolet Spark 2015, a mini-compact car. These vehicles were chosen to provide a representative sample across different vehicle types and manufacturers, allowing for a comprehensive evaluation of intrusion detection algorithms in diverse automotive contexts.

Each CAN message in the datasets is characterized by several key fields, as described by Han et al. These include the Timestamp, which records the exact time each CAN message was captured, and the CAN ID, which identifies the source or function of each message within the vehicle's network. The DLC (Data Length Code) specifies the length of the data field in each message, while the Payload contains the actual data transmitted within the CAN message. Finally, the Class label distinguishes between normal ("R") and attack ("T") messages, enabling the differentiation between benign and malicious network behavior. More detailed analysis can be done as follows:

Timestamp: The Timestamp field records the exact time each CAN message was logged. This field is crucial for analyzing the timing and sequence of messages, which can reveal patterns or irregularities in the communication flow. In the context of attacks, timestamps can help identify abnormal message rates, delays, or bursts of activity that may indicate an ongoing attack.

CAN ID: The CAN ID is a unique identifier assigned to each CAN message. It indicates the priority of the message and the specific component or system within the vehicle that is sending the data. In a typical vehicle network, different CAN IDs are associated with different functions such as engine control, transmission, or braking systems. During an attack, certain CAN IDs may be spoofed, or unexpected CAN IDs may appear in the network traffic, signaling a potential security breach.

DLC (Data Length Code): The Data Length Code (DLC) specifies the length of the data field in the CAN message. This code is typically a small integer (ranging from 0 to 8) that indicates the number of bytes of data being transmitted. Consistent DLC values are expected under normal conditions, corresponding to the specific data requirements of each CAN ID. However, during an attack, the DLC values might deviate from their usual patterns, either becoming unusually large or small, which can help in detecting an intrusion.

Payload: The Payload field contains the actual data being transmitted in the CAN message. This data is usually represented as a sequence of hexadecimal bytes. The payload carries the information necessary for the vehicle's components to perform their functions, such as sensor readings, control signals, or status updates. Analyzing the payload is critical in detecting anomalies because attacks often involve injecting malicious or spoofed data into the network. For example, in a Flooding attack, the payload might consist of repetitive or nonsensical data aimed at overwhelming the network.

Class: The Class label is used to distinguish between normal and attack messages within the dataset. In these datasets, the label "R" indicates a normal (attack-free) message, while "T" signifies a message associated with an attack. This field is vital for training machine learning models to differentiate between benign and malicious behavior within the vehicle network. The distribution of "R" and "T" labels in a dataset directly affects the model's ability to accurately detect and respond to network anomalies.

These datasets, as referenced from the work of Mee Lan Han et al., provide a robust basis for developing and testing anomaly detection methods in vehicular networks. By utilizing real-world data from different vehicle models and including a variety of attack scenarios, these datasets are invaluable for advancing the field of automotive cybersecurity and improving the resilience of vehicle networks against cyber threats.

4.3. Dataset Analysis by Vehicles

Determining whether datasets are balanced or imbalanced is critical for training anomaly detection models. A balanced dataset is a dataset where there is approximately an equal distribution between normal and attack data. In contrast, an imbalanced dataset occurs when there is a significant imbalance between normal and attack data. Imbalanced datasets can often cause machine learning models to be biased in detecting attacks, making it harder to detect attacks.

The analysis of the datasets for Hyundai Sonata, KIA Soul, and Chevrolet Spark vehicles reveals that these datasets are largely imbalanced. This imbalance is evident in the significantly higher number of normal data compared to attack data. A detailed examination of each dataset for the three vehicles clearly shows the disproportion between the number of normal and attack data points.

For the Hyundai Sonata datasets, the Flooding attack dataset contains 78,906 normal data points and 17,093 attack data points. This indicates that attack data is relatively sparse compared to the total dataset, with normal data being much more prevalent. Similarly, the Malfunction dataset has 78,797 normal data points and only 8,202 attack data points. The Fuzzy attack dataset includes 78,904 normal data points and 9,095 attack data points. Each of these datasets displays a significant imbalance, where the attack data is considerably less frequent than the normal data.

The datasets for KIA Soul also exhibit a similar imbalance. The Flooding attack dataset contains 103,927 normal data points and 16,072 attack data points. The Malfunction dataset includes 108,229 normal data points and 4,770 attack data points, while the Fuzzy dataset contains 122,386 normal data points and 21,613 attack data points. These datasets demonstrate a clear imbalance, with attack data being significantly outnumbered by normal data.

The datasets for Chevrolet Spark further illustrate this imbalance. The Flooding dataset has 70,000 normal data points and 14,999 attack data points, while the Malfunction dataset contains 47,004 normal data points and only 3,995 attack data points. The Fuzzy dataset includes 37,956 normal data points and 3,043 attack data points. These datasets

show that normal data overwhelmingly outnumbers attack data, indicating a significant imbalance.

This analysis reveals that all three vehicle models exhibit a considerable imbalance between attack and normal data in their respective datasets. This imbalance is characterized by a much higher number of normal data points compared to attack data. The large disparity between the number of normal and attack data points makes these datasets predominantly imbalanced. This is an important aspect to consider during the analysis of these datasets.

The presence of such imbalance in these datasets offers a critical observation for vehicle network security studies. The imbalance between attack and normal data in the datasets for Hyundai Sonata, KIA Soul, and Chevrolet Spark directly affects the distribution of data and the detectability of attacks. This observation is crucial for understanding the impact of attacks on in-vehicle networks and for evaluating the effectiveness of detection methods developed to counter these attacks.

Table 4.1 – Dataset Distribution and Balance Status by Vehicles

Vehicle Model	Dataset Type	Normal Data Count	Attack Data Count	Balance Status
Hyundai Sonata	Attack-Free	78,904	0	Balanced
Hyundai Sonata	Flooding	78,906	17,093	Imbalanced
Hyundai Sonata	Malfunction	78,797	8,202	Imbalanced
Hyundai Sonata	Fuzzy	78,904	9,095	Imbalanced
KIA Soul	Attack-Free	122,386	0	Balanced
KIA Soul	Flooding	103,927	16,072	Imbalanced
KIA Soul	Malfunction	108,229	4,770	Imbalanced
KIA Soul	Fuzzy	122,386	21,613	Imbalanced
Chevrolet Spark	Attack-Free	37,956	0	Balanced
Chevrolet Spark	Flooding	70,000	14,999	Imbalanced

Chevrolet Spark	Malfunction	47,004	3,995	Imbalanced
Chevrolet Spark	Fuzzy	37,956	3,043	Imbalanced

This table summarizes the balance status of the datasets for each vehicle model. The Attack-Free datasets, which contain only normal data, are balanced due to the absence of attack data. However, all the datasets containing attack data are imbalanced, with attack data being significantly less frequent than normal data. This is a critical aspect to consider in the overall structure and analysis of the datasets.

Flooding attacks target CAN networks by sending messages from a specific CAN ID at excessively high frequencies in short intervals, disrupting network communication and potentially impairing ECU functionality. This type of attack can congest the network, blocking or delaying the transmission of legitimate messages. To detect flooding attacks in this dataset, an analysis was conducted by evaluating the transmission frequency of each CAN ID. The method involved data preprocessing, calculating time differences between messages, determining the average frequency, and flagging CAN IDs with an abnormally high message frequency based on a defined threshold.

In the data preparation phase, key columns in the dataset were renamed for clarity, and timestamps and CAN ID values were converted to suitable formats for analysis. Timestamps were converted to numeric format, while CAN IDs were transformed to string format to ensure consistency. Following this, each CAN ID group was sorted by timestamp, and the time differences (or intervals) between consecutive messages were calculated. These time differences provided a frequency profile for each CAN ID, serving as a primary indicator of flooding attacks. Any initial NaN values in the time difference column (caused by the first message in each CAN ID group) were removed to maintain clean data for the analysis.

To establish a baseline for flooding detection, the average message transmission interval was calculated for each CAN ID. This average frequency served as a reference point, representing each ID's typical messaging behavior. A threshold was then set at 10% of

the overall average frequency across all CAN IDs. This low threshold was chosen to highlight CAN IDs with transmission frequencies significantly above normal, thereby identifying those with potential flooding activity. CAN IDs with average frequencies below this threshold were flagged as potential flooding attacks and marked as True in the “flooding_detected” column in the results table.

The results obtained from this method offer a practical approach for detecting flooding attacks in IDS applications for CAN networks. By evaluating each CAN ID’s message frequency relative to an established threshold, this analysis effectively highlights potential attack patterns. Integrating this frequency-based analysis into IDS algorithms can contribute to more robust security mechanisms within automotive networks. Future work may involve refining the threshold and incorporating additional analytical methods to improve the detection accuracy of such attack patterns, ultimately enhancing CAN-based IDS solutions for secure automotive environments.

Table 4.2 – Hyundai Dataset’s First 10 row for Flooding Detection

CAN_ID	avg_frequency	flooding_detected
0000	0.001190	True
0002	0.010000	False
00A0	0.099915	False
00A1	0.099916	False
0130	0.009999	False
0131	0.009999	False
0140	0.009999	False
0153	0.010000	False
018F	0.009992	False
01F1	0.019999	False

In this part, frequency analysis was applied to detect potential fuzzy attacks in Hyundai vehicles’ CAN dataset. Fuzzy attacks are among the common threats to automotive

networks, disrupting ECU functionality by altering the message flow in the system. These attacks involve sending messages more frequently or irregularly than expected. In this analysis, the message transmission frequency for each CAN ID was calculated, and those with an average frequency below a statistically determined threshold were flagged as potential fuzzy attacks.

To detect fuzzy attacks, a frequency analysis based on the transmission frequency of CAN IDs was applied. The average time difference between messages was calculated for each CAN ID. Under normal circumstances, each CAN ID sends messages at a specific frequency, forming a “traffic profile.” According to this profile, if the messages of some CAN IDs arrive much more frequently or irregularly than expected, this situation can be evaluated as a potential fuzzy attack.

The time differences were calculated for each CAN ID to determine the frequency. This step allowed us to observe the frequency with which each CAN ID sent messages, forming a statistical profile for each ID's expected behavior. These time differences provided the basis for detecting fuzzy attacks by highlighting variations in transmission intervals.

The time differences were then grouped to calculate the average frequency for each CAN ID. These average frequencies served as a statistical indicator of each ID's message transmission rate and provided a reference for normal operation. This reference value served as the basis for setting a threshold to identify fuzzy attacks in the system.

The threshold value used in the analysis was set as 50% of the average frequency of all CAN IDs. In determining this value, the average frequency across all IDs was calculated, and a threshold was set at half of this average (i.e., messages arriving more frequently than half the general average were flagged as anomalies).

For example, if the average time difference across all CAN IDs is calculated as 1 second, the threshold is set to 0.5 seconds. If a CAN ID's average message transmission time is below 0.5 seconds, this CAN ID is marked as anomalous, as it indicates a transmission frequency higher than expected. With this method, CAN IDs that sent messages at abnormal frequencies were detected. Rows with CAN IDs identified for fuzzy attacks

were marked as True in the "fuzzing_detected" column. In such analysis, the threshold can be adjusted according to the normal behavior of the system, which may vary slightly for each dataset.

In the analysis results, certain CAN IDs with average transmission frequencies below the set threshold were identified as potential fuzzy attacks. This situation indicates that these CAN IDs sent messages more frequently than usual, suggesting a security threat to the system. The frequency analysis provides valuable data for IDS algorithms to identify fuzzy attacks and supports the development of security systems.

Frequency-based fuzzy detection has proven to be an effective method for identifying abnormal message transmission frequencies in CAN networks. This analysis offers a rapid detection mechanism, especially for frequency-based attacks. However, detection based solely on frequency analysis may not cover more complex attack types. Therefore, it is recommended to use additional analysis methods such as time-based correlation and statistical anomaly detection to provide more comprehensive security in detecting fuzzy attacks. Additionally, adjusting the threshold according to each dataset's normal behavior is crucial for improving detection accuracy.

Table 4.3 – Hyundai Dataset's First 10 Row for Fuzzing Detection

CAN_ID	avg_frequency	fuzzing_detected
0000	2.242646	False
0001	4.531060	False
0002	0.009984	True
0003	2.302166	False
0004	1.833942	False
0005	2.228271	False
0006	3.287526	False
0007	2.614014	False
0008	1.792779	False

0009

1.758290

False

The malfunction detection analysis for this dataset aimed to identify irregularities or unexpected patterns within CAN messages, which may indicate potential faults or anomalies within the automotive network. In CAN networks, malfunctioning conditions often manifest as deviations in message frequency or irregularities in data content. By calculating each CAN ID's average transmission frequency and identifying deviations from these patterns, we aimed to systematically detect these malfunctioning signals.

The analysis began with data preparation and preprocessing to ensure a consistent structure across the dataset. After loading the data, columns were renamed for clarity, and key fields, such as timestamps and CAN IDs, were converted to numeric and string formats respectively. This ensured uniformity and allowed for accurate frequency calculations. To maintain data integrity, rows with missing timestamp values were removed, which is essential to avoid skewing interval calculations and preserve accuracy.

Following data preprocessing, we calculated the time intervals between consecutive messages for each CAN ID. Sorting the data by CAN ID and timestamp allowed us to compute these intervals, which represent the time differences, or gaps, between messages sent by each CAN ID. This step was crucial as it provided insight into the typical frequency and regularity of each CAN ID's transmission under normal conditions. Monitoring these intervals is important because sudden or significant changes in timing could indicate a malfunction.

After obtaining the time intervals, the next step was to calculate the average frequency for each CAN ID. By grouping data by CAN ID and averaging the time intervals, we established a baseline frequency measure for each ID. This average frequency served as a reference for normal operational behavior, allowing us to identify any CAN ID that exhibited irregular transmission patterns. Detecting malfunctions relies on these reference values, as they offer a comparative metric to flag deviations in message timing.

To identify potential malfunctions, we established a statistical threshold. This threshold was set at 15% deviation from the mean frequency, a range chosen to account for typical,

expected fluctuations in message timing while still capturing significant outliers. In practical terms, any CAN ID with an average message interval that varied by more than 15% from the overall mean was flagged as a potential malfunction. This approach provided a balance, ensuring that natural variations would not trigger false positives while still capturing potential faults.

Finally, CAN IDs with message frequencies that fell outside this acceptable range were marked as potential malfunctions in the results. These flagged IDs represent signals with transmission patterns significantly different from the established baseline, suggesting abnormal behavior that could indicate a malfunction or fault within the network. By following these systematic steps, the analysis provided a clear method for detecting malfunctioning behavior within the CAN network based on deviations from expected message timing patterns. This approach is a valuable tool for improving network reliability and early fault detection in automotive systems.

Table 4.4 – Hyundai Dataset’s First 10 Row for Malfunction Detection

CAN_ID	avg_frequency	malfunction_detected
0002	0.010019	True
00A0	0.100164	False
00A1	0.100163	False
0130	0.010031	True
0131	0.010031	True
0140	0.010031	True
0153	0.010019	True
018F	0.010014	True
01F1	0.020078	True
0260	0.010014	True

The increasing reliance on connected and automated vehicles (CAVs) has introduced new cybersecurity challenges. With vehicles now incorporating advanced communication protocols and sophisticated electronic control systems, they are becoming prime targets for cyber-attacks. These attacks can disrupt normal vehicle operations, compromise passenger safety, and even jeopardize broader transportation networks. To address these challenges, developing robust intrusion detection systems (IDS) tailored for vehicular environments has become essential.

This part evaluates various machine learning and deep learning algorithms for detecting cyber-attacks in datasets collected from three different vehicle: KIA, Hyundai, and Chevrolet. The datasets encompass a variety of attack scenarios, including Flooding, Fuzzy, and Malfunction, alongside normal traffic data. Features such as CAN ID, DLC, and Payload, representing critical vehicle communication attributes preprocessed, and analyzed to assess algorithm performance.

The performance of each algorithm was measured using key metrics such as accuracy, precision, recall, and F1-score, providing a comprehensive understanding of their suitability for vehicular cybersecurity applications. The study also highlights the limitations of specific algorithms, such as Naive Bayes and LinearSVC, in handling the complexities of vehicular datasets. Furthermore, tree-based methods like XGBoost and CatBoost consistently emerged as top performers across the datasets, demonstrating their robustness and adaptability.

Through this comparative analysis, the report aims to provide valuable insights into the most effective machine learning and deep learning approaches for real-world deployment in connected vehicles. These findings contribute to the development of more reliable intrusion detection systems, enhancing the cybersecurity posture of next-generation vehicles.

4.4. Machine Information

The application was conducted on a Windows 10 Home Single Language 64-bit operating system computer. The computer has an Intel Core i7-6700HQ processor and provides high performance with 4 core and 8 virtual cores operating at 2.60 GHz. This system, which has a memory capacity of 16 GB (16384 MB), provides sufficient memory space for data processing and model training processes. Internal Intel HD Graphics 530 and external NVIDIA GeForce GTX 960M were used as the graphics processing unit. NVIDIA GeForce GTX 960M, with its 12 GB memory capacity, increases the computational speed and shortens the processing times in the training of deep learning models.

In the application, interactive analysis and visualization were performed using Jupyter Notebook. Jupyter Notebook made data preprocessing, model training, and evaluation processes more efficient. Popular libraries such as TensorFlow, Scikit-learn, Keras, NumPy, and Pandas were used to implement machine learning and deep learning algorithms. These libraries allowed the analysis and comparison of different algorithms' performances in detail.

The powerful hardware and software components of my work environment allowed me to effectively evaluate the performances of various machine learning and deep learning algorithms used in anomaly detection. This environment increased the reliability and validity of the results obtained with high accuracy, precision, recall and F1 score. In particular, the speed and computational capacity provided by the GPU provided a great advantage in the training processes of deep learning models. In this way, the opportunity was found to determine the most appropriate detection methods by comprehensively examining the effectiveness of different algorithms in the study of in-vehicle network attack detection.

4.5. Application of Machine Learning Algorithms

This section delves into the performance analysis of various machine learning and deep learning algorithms used to detect cyber-attacks on Kia vehicle. The study focuses on attacks such as Flooding, Fuzzy, and Malfunction, using a labeled dataset specifically tailored for vehicular communication analysis. The dataset includes critical features such

as CAN ID, DLC, and Payload, which represent the essential communication parameters in the vehicle's Controller Area Network (CAN). To ensure the dataset's quality and improve the models' predictive capabilities, preprocessing steps such as feature engineering, payload decomposition, and handling missing values were implemented.

The evaluation relied on metrics like accuracy, precision, recall, and F1-score to provide a comprehensive understanding of the models' effectiveness. Each algorithm's performance is discussed in detail, highlighting their strengths and limitations within the context of vehicular anomaly detection.

4.5.1 Hyundai Dataset Application Results

The Hyundai attacks dataset was analyzed using various machine learning and deep learning algorithms to evaluate their performance under different training-to-test ratios: 60-40, 70-30, and 80-20. Key metrics, including accuracy, precision, recall, and macro F1-score, were used to assess the models. Additionally, the training and testing times for each algorithm were recorded to examine their runtime costs. This comprehensive analysis aimed to understand how training data proportions influence algorithm performance and runtime efficiency.

In the 60-40 split, where the training data was limited, ensemble methods such as CatBoost and Random Forest demonstrated superior performance. CatBoost achieved the highest accuracy (0.83) and F1-score (0.75), followed by Random Forest with an accuracy of 0.81 and an F1-score of 0.76. Deep learning models such as CNN and LSTM, on the other hand, struggled to perform well in recall, scoring 0.51 and 0.60, respectively. This indicates the challenges these models face in identifying true positives when trained on limited data.

Table 4.5 – Hyundai Dataset Algorithm Performance Results (60-40)

Algorithm	Accuracy	Precision	Recall	F1-Score (macro)
Decision Tree	0.82	0.79	0.74	0.78

Random Forest	0.81	0.85	0.72	0.76
XGBoost	0.80	0.83	0.72	0.75
K-Nearest Neighbors	0.76	0.74	0.74	0.76
Multilayer Perceptron	0.78	0.73	0.67	0.70
CatBoost	0.83	0.84	0.74	0.75
Gradient Boosting	0.77	0.81	0.64	0.71
CNN	0.75	0.83	0.51	0.62
RNN	0.73	0.69	0.74	0.74
LSTM	0.71	0.72	0.60	0.67

When the training set proportion increased to 70%, most algorithms showed noticeable improvements in their performance metrics. CatBoost again emerged as the top-performing model, with its accuracy increasing to 0.85 and its F1-score to 0.78. Ensemble methods like XGBoost and Random Forest also benefited significantly, achieving F1-scores of 0.77 and 0.78, respectively. Deep learning models, including CNN and LSTM, exhibited marginal improvements. CNN's recall improved slightly to 0.55, but its overall performance remained lower than that of the ensemble methods. These results reaffirm that deep learning models require larger and more complex datasets for optimal learning.

Table 4.6 – Hyundai Dataset Algorithm Performance Results (70-30)

Algorithm	Accuracy	Precision	Recall	F1-Score (macro)
Decision Tree	0.84	0.81	0.77	0.79
Random Forest	0.83	0.87	0.75	0.78
XGBoost	0.82	0.85	0.74	0.77

K-Nearest Neighbors	0.79	0.77	0.77	0.77
Multilayer Perceptron	0.80	0.76	0.70	0.73
CatBoost	0.85	0.86	0.76	0.78
Gradient Boosting	0.80	0.83	0.69	0.75
CNN	0.78	0.85	0.55	0.67
RNN	0.77	0.73	0.76	0.75
LSTM	0.75	0.74	0.63	0.69

Under the 80-20 split, where the training set was at its largest, further performance improvements were observed across most models. CatBoost achieved the highest metrics, with an accuracy of 0.88 and an F1-score of 0.83, showcasing its ability to generalize effectively with increased training data. Random Forest and XGBoost followed closely, achieving accuracies of 0.87 and 0.85 and F1-scores of 0.82 and 0.80, respectively. KNN, despite its relatively simple nature, demonstrated balanced performance with an accuracy and F1-score of 0.83 and 0.81, respectively. However, the deep learning models CNN and LSTM continued to show limited gains, with CNN achieving a precision of 0.88 but still lagging in recall and F1-score (0.60 and 0.72, respectively). This underscores the difficulty of optimizing deep learning architectures on datasets of this scale.

Table 4.7 – Hyundai Dataset Algorithm Performance Results (80-20)

Algorithm	Accuracy	Precision	Recall	F1-Score (macro)
Decision Tree	0.85	0.83	0.78	0.81
Random Forest	0.87	0.89	0.79	0.82
XGBoost	0.85	0.87	0.77	0.80
K-Nearest Neighbors	0.83	0.81	0.81	0.81

Multilayer Perceptron	0.82	0.79	0.73	0.76
CatBoost	0.88	0.89	0.80	0.83
Gradient Boosting	0.85	0.86	0.73	0.79
CNN	0.82	0.88	0.60	0.72
RNN	0.81	0.77	0.79	0.78
LSTM	0.79	0.78	0.66	0.72

Runtime analysis revealed that training times increased proportionally with the size of the training set, while testing times decreased due to smaller test sets. In the 60-40 split, Decision Tree exhibited the shortest training time (0.64 seconds), while KNN incurred the longest testing time (0.021 seconds). As the training data increased to 80%, Decision Tree's training time rose to 0.80 seconds, while CatBoost maintained efficient runtime performance, reducing its testing time to 0.008 seconds. Deep learning models such as CNN and LSTM exhibited the largest increases in training time, reaching 0.95 and 0.90 seconds, respectively. However, their testing times decreased slightly, benefiting from the smaller test set.

Table 4.8 – Runtime Costs for Hyundai Dataset Algorithm (60-40)

Algorithm	Training Time (s)	Testing Time (s)
Decision Tree	0.64	0.02
Random Forest	0.63	0.017
XGBoost	0.70	0.016
K-Nearest Neighbors	0.76	0.021
Multilayer Perceptron	0.75	0.020
CatBoost	0.68	0.015
Gradient Boosting	0.64	0.019
CNN	0.73	0.017

RNN	0.64	0.020
LSTM	0.68	0.032

Table 4.9 – Runtime Costs for Hyundai Dataset Algorithm (70-30)

Algorithm	Training Time (s)	Testing Time (s)
Decision Tree	0.72	0.015
Random Forest	0.70	0.013
XGBoost	0.78	0.013
K-Nearest Neighbors	0.85	0.017
Multilayer Perceptron	0.83	0.016
CatBoost	0.75	0.012
Gradient Boosting	0.72	0.015
CNN	0.80	0.013
RNN	0.72	0.016
LSTM	0.75	0.027

Table 4.10 – Runtime Costs for Hyundai Dataset Algorithm (80-20)

Algorithm	Training Time (s)	Testing Time (s)
Decision Tree	0.80	0.012
Random Forest	0.78	0.010
XGBoost	0.85	0.010
K-Nearest Neighbors	1.00	0.013
Multilayer Perceptron	0.90	0.011
CatBoost	0.83	0.008
Gradient Boosting	0.82	0.010
CNN	0.95	0.011

RNN	0.85	0.011
LSTM	0.90	0.020

These results demonstrate that increasing the proportion of training data generally enhances algorithm performance. Ensemble methods, particularly CatBoost and Random Forest, consistently outperformed other models on the Hyundai dataset, delivering high accuracy and F1-scores while maintaining reasonable runtime efficiency. Deep learning models showed limited improvements compared to ensemble methods, indicating their reliance on larger and more complex datasets for optimal performance. The runtime analysis further highlighted the trade-offs between training and testing times, emphasizing the importance of selecting models that balance predictive power with computational efficiency. This study underscores the critical role of training-to-test ratios in optimizing algorithm performance and provides a robust framework for selecting the most suitable algorithm for similar datasets.

4.5.2 Kia Dataset Application Results

The KIA dataset was analyzed using various machine learning and deep learning algorithms, and their performance was evaluated under three different training-to-test ratios: 60-40, 70-30, and 80-20. The analysis was based on key metrics, including accuracy, precision, recall, and macro F1-score, as well as runtime costs, such as training and testing times. This study provides insights into how the models respond to varying training data proportions and highlights the trade-offs between performance and computational efficiency.

For the 60-40 split, ensemble methods such as Random Forest and CatBoost demonstrated strong performance. Random Forest achieved an accuracy of 0.83 and an F1-score of 0.78, while CatBoost recorded an accuracy of 0.82 and an F1-score of 0.79. These results highlight the robustness of ensemble methods when trained on limited data. Among deep learning models, CNN showed reasonable precision (0.86) but struggled in recall (0.53) and F1-score (0.65), indicating challenges in generalizing to the test data.

LSTM, another deep learning model, performed moderately with an accuracy of 0.74 and an F1-score of 0.67.

Table 4.11 – Kia Dataset Algorithm Performance Results (60-40)

Algorithm	Accuracy	Precision	Recall	F1-Score (macro)
Decision Tree	0.80	0.78	0.74	0.78
Random Forest	0.83	0.88	0.74	0.78
XGBoost	0.79	0.84	0.72	0.75
K-Nearest Neighbors	0.77	0.76	0.76	0.78
Multilayer Perceptron	0.80	0.75	0.69	0.71
CatBoost	0.82	0.84	0.76	0.79
Gradient Boosting	0.78	0.83	0.65	0.73
CNN	0.78	0.86	0.53	0.65
RNN	0.76	0.71	0.73	0.74
LSTM	0.74	0.73	0.61	0.67

Increasing the training proportion to 70% resulted in performance improvements across most algorithms. CatBoost emerged as the top-performing model, achieving an accuracy of 0.84 and an F1-score of 0.81, reflecting its adaptability to larger training sets. Random Forest and XGBoost also benefitted from the additional data, with F1-scores increasing to 0.80 and 0.78, respectively. CNN improved its precision to 0.88, but its recall and F1-score remained comparatively low at 0.58 and 0.69, respectively. Decision Tree, while less complex than ensemble methods, demonstrated notable gains with an F1-score of 0.79, showing its capacity to adapt to an increased training set size.

Table 4.12 – Kia Dataset Algorithm Performance Results (70-30)

Algorithm	Accuracy	Precision	Recall	F1-Score (macro)
Decision Tree	0.82	0.80	0.76	0.79
Random Forest	0.85	0.89	0.76	0.80
XGBoost	0.81	0.86	0.74	0.78
K-Nearest Neighbors	0.79	0.78	0.78	0.78
Multilayer Perceptron	0.82	0.78	0.72	0.74
CatBoost	0.84	0.86	0.78	0.81
Gradient Boosting	0.81	0.85	0.70	0.76
CNN	0.80	0.88	0.58	0.69
RNN	0.79	0.75	0.76	0.76
LSTM	0.77	0.75	0.64	0.69

Under the 80-20 split, where the training set was at its largest, the performance of all algorithms improved further. Random Forest and CatBoost achieved the highest metrics, with accuracies of 0.87 and 0.86 and F1-scores of 0.83 each. XGBoost followed closely with an accuracy of 0.85 and an F1-score of 0.81. Among deep learning models, CNN exhibited the greatest improvement, reaching a precision of 0.90 and an F1-score of 0.73. However, it still lagged behind ensemble methods in overall performance. LSTM and RNN also showed improvements, with F1-scores of 0.72 and 0.78, respectively, but their performance remained lower compared to Random Forest and CatBoost.

Table 4.13 – Kia Dataset Algorithm Performance Results (80-20)

Algorithm	Accuracy	Precision	Recall	F1-Score (macro)
Decision Tree	0.84	0.83	0.78	0.81

Random Forest	0.87	0.89	0.79	0.83
XGBoost	0.85	0.87	0.77	0.81
K-Nearest Neighbors	0.82	0.80	0.80	0.80
Multilayer Perceptron	0.84	0.81	0.75	0.77
CatBoost	0.86	0.88	0.80	0.83
Gradient Boosting	0.84	0.86	0.72	0.78
CNN	0.82	0.90	0.62	0.73
RNN	0.81	0.77	0.79	0.78
LSTM	0.80	0.79	0.67	0.72

The runtime analysis revealed significant variations in training and testing times across algorithms and data splits. For the 60-40 split, Decision Tree required the shortest training time (0.72 seconds), while Random Forest exhibited the shortest testing time (0.02 seconds). Deep learning models like CNN and LSTM demonstrated higher training times (0.76 and 0.70 seconds, respectively) due to their computational complexity. In the 70-30 split, training times increased proportionally with the larger training set, with KNN and CNN showing significant increases to 0.84 and 0.85 seconds, respectively. Testing times decreased slightly due to the smaller test set, with CatBoost achieving the shortest time of 0.012 seconds. Under the 80-20 split, training times rose further, with CNN reaching 1.05 seconds, highlighting the computational cost of deep learning models. Testing times continued to decrease, with ensemble methods like CatBoost and Random Forest completing predictions in 0.010 and 0.012 seconds, respectively.

Table 4.14 – Runtime Costs for Kia Dataset Algorithm (60-40)

Algorithm	Training Time (s)	Testing Time (s)
Decision Tree	0.72	0.03
Random Forest	0.65	0.02

XGBoost	0.70	0.018
K-Nearest Neighbors	0.74	0.024
Multilayer Perceptron	0.73	0.019
CatBoost	0.70	0.014
Gradient Boosting	0.66	0.021
CNN	0.76	0.019
RNN	0.63	0.019
LSTM	0.70	0.035

Table 4.15 – Runtime Costs for Kia Dataset Algorithm (70-30)

Algorithm	Training Time (s)	Testing Time (s)
Decision Tree	0.80	0.025
Random Forest	0.75	0.015
XGBoost	0.78	0.015
K-Nearest Neighbors	0.84	0.018
Multilayer Perceptron	0.81	0.016
CatBoost	0.78	0.012
Gradient Boosting	0.75	0.016
CNN	0.85	0.015
RNN	0.72	0.015
LSTM	0.80	0.030

Table 4.16 – Runtime Costs for Kia Dataset Algorithm (80-20)

Algorithm	Training Time (s)	Testing Time (s)
Decision Tree	0.90	0.020
Random Forest	0.85	0.012
XGBoost	0.90	0.012

K-Nearest Neighbors	1.00	0.015
Multilayer Perceptron	0.95	0.013
CatBoost	0.90	0.010
Gradient Boosting	0.88	0.013
CNN	1.05	0.012
RNN	0.90	0.013
LSTM	0.95	0.025

These results underscore the robustness and efficiency of ensemble methods such as CatBoost and Random Forest, which consistently outperformed other models across all splits, achieving high accuracy and F1-scores while maintaining efficient runtime performance. Deep learning models, while showing improvements with larger training sets, struggled to match the ensemble methods in terms of overall metrics. The runtime analysis further highlights the computational demands of deep learning algorithms, particularly in training, making ensemble methods a preferable choice for scenarios where both performance and efficiency are critical considerations.

4.5.3 Chevrolet Dataset Application Results

The Chevrolet dataset was analyzed using various machine learning and deep learning algorithms, with performance evaluated across three different training-to-test ratios: 60-40, 70-30, and 80-20. The evaluation metrics included accuracy, precision, recall, and macro F1-score, along with runtime costs, namely training and testing times. The aim was to understand how these algorithms respond to varying data proportions and to identify their strengths and limitations for this dataset.

Under the 60-40 split, the performance of the algorithms reflected the limited availability of training data. Ensemble methods such as Random Forest, XGBoost, and CatBoost achieved the best results, each recording an accuracy of 0.80 and an F1-score of 0.75. Decision Tree followed closely with an accuracy of 0.79 and an F1-score of 0.75. Deep

learning models, including CNN, RNN, and LSTM, showed comparatively weaker performance, particularly in recall and F1-score metrics. CNN struggled the most, with a recall of 0.45 and an F1-score of 0.59, highlighting its challenges in effectively generalizing with less data.

Table 4.17 – Chevrolet Dataset Algorithm Performance Results (60-40)

Algorithm	Accuracy	Precision	Recall	F1-Score (macro)
Decision Tree	0.79	0.80	0.70	0.75
Random Forest	0.80	0.81	0.69	0.75
XGBoost	0.80	0.81	0.70	0.75
K-Nearest Neighbors	0.77	0.74	0.74	0.74
Multilayer Perceptron	0.73	0.69	0.68	0.69
CatBoost	0.80	0.81	0.70	0.75
Gradient Boosting	0.77	0.80	0.62	0.70
CNN	0.73	0.85	0.45	0.59
RNN	0.73	0.68	0.72	0.70
LSTM	0.73	0.77	0.55	0.64

The 70-30 split yielded improved results across most algorithms due to the larger training dataset. Ensemble methods once again outperformed the rest, with CatBoost, Random Forest, and XGBoost achieving identical accuracy (0.82) and F1-scores (0.78). Decision Tree also benefited from the increased training data, achieving an accuracy of 0.81 and an F1-score of 0.77. Deep learning models showed slight improvements; CNN, for instance, increased its precision to 0.88 and F1-score to 0.63. However, the overall gap between ensemble methods and deep learning models remained significant, reaffirming the former's superiority in handling this dataset.

Table 4.18 – Chevrolet Dataset Algorithm Performance Results (70-30)

Algorithm	Accuracy	Precision	Recall	F1-Score (macro)
Decision Tree	0.81	0.82	0.73	0.77
Random Forest	0.82	0.83	0.72	0.78
XGBoost	0.82	0.83	0.72	0.78
K-Nearest Neighbors	0.79	0.76	0.76	0.76
Multilayer Perceptron	0.75	0.72	0.71	0.72
CatBoost	0.82	0.83	0.72	0.78
Gradient Boosting	0.79	0.82	0.65	0.73
CNN	0.76	0.88	0.50	0.63
RNN	0.75	0.70	0.74	0.72
LSTM	0.75	0.79	0.58	0.67

When the training-to-test ratio was increased to 80-20, the algorithms demonstrated further performance enhancements. Random Forest, XGBoost, and CatBoost achieved the highest accuracy (0.84) and F1-scores (0.79), showcasing their ability to effectively leverage the additional training data. Decision Tree also performed well, with an accuracy of 0.83 and an F1-score of 0.79. Among deep learning models, CNN showed the greatest improvement, reaching a precision of 0.90 and an F1-score of 0.68. This indicates that while deep learning models benefit from more data, they still fall behind ensemble methods in terms of overall performance metrics.

Table 4.19 – Chevrolet Dataset Algorithm Performance Results (80-20)

Algorithm	Accuracy	Precision	Recall	F1-Score (macro)
------------------	-----------------	------------------	---------------	-------------------------

Decision Tree	0.83	0.84	0.75	0.79
Random Forest	0.84	0.85	0.74	0.79
XGBoost	0.84	0.85	0.74	0.79
K-Nearest Neighbors	0.81	0.78	0.78	0.78
Multilayer Perceptron	0.78	0.75	0.74	0.74
CatBoost	0.84	0.85	0.74	0.79
Gradient Boosting	0.82	0.84	0.68	0.75
CNN	0.78	0.90	0.55	0.68
RNN	0.78	0.73	0.76	0.75
LSTM	0.78	0.81	0.61	0.70

Runtime analysis revealed a proportional increase in training times with larger training sets and a decrease in testing times with smaller test sets. Under the 60-40 split, Decision Tree had the shortest training time (0.88 seconds), while Random Forest had the longest (1.24 seconds). Testing times were relatively uniform, with most algorithms completing within 0.03 to 0.04 seconds. In the 70-30 split, training times increased slightly across all algorithms, with Random Forest and K-Nearest Neighbors requiring the longest times at 1.32 and 1.10 seconds, respectively. Testing times decreased marginally, reflecting the smaller test set size. In the 80-20 split, training times increased further; CNN, RNN, and LSTM reached approximately 1 second, reflecting their higher computational demands. Testing times dropped further, with CatBoost achieving the shortest duration of 0.024 seconds.

Table 4.20 – Runtime Costs for Kia Dataset Algorithm (60-40)

Algorithm	Training Time (s)	Testing Time (s)
Decision Tree	0.88	0.04

Random Forest	1.24	0.04
XGBoost	0.96	0.037
K-Nearest Neighbors	0.98	0.037
Multilayer Perceptron	0.90	0.032
CatBoost	0.85	0.030
Gradient Boosting	0.86	0.032
CNN	0.85	0.030
RNN	0.87	0.034
LSTM	0.90	0.040

Table 4.21 – Runtime Costs for Kia Dataset Algorithm (70-30)

Algorithm	Training Time (s)	Testing Time (s)
Decision Tree	0.91	0.03
Random Forest	1.32	0.03
XGBoost	1.05	0.028
K-Nearest Neighbors	1.10	0.030
Multilayer Perceptron	0.95	0.029
CatBoost	0.88	0.028
Gradient Boosting	0.92	0.029
CNN	0.88	0.028
RNN	0.91	0.031
LSTM	0.95	0.035

Table 4.22 – Runtime Costs for Kia Dataset Algorithm (80-20)

Algorithm	Training Time (s)	Testing Time (s)
Decision Tree	0.95	0.02
Random Forest	1.40	0.02

XGBoost	1.12	0.025
K-Nearest Neighbors	1.20	0.026
Multilayer Perceptron	1.00	0.025
CatBoost	0.95	0.024
Gradient Boosting	1.00	0.025
CNN	0.92	0.024
RNN	0.96	0.028
LSTM	1.00	0.033

These results highlight the importance of training-to-test ratios in optimizing algorithm performance. Ensemble methods such as Random Forest, XGBoost, and CatBoost consistently outperformed other models in all configurations, combining strong predictive capabilities with reasonable runtime efficiency. Deep learning models, while benefiting from larger training sets, demonstrated limited gains compared to ensemble methods, suggesting their reliance on larger and more complex datasets for optimal performance. Runtime efficiency analysis also underscores the computational cost of deep learning models, particularly in training, making ensemble methods more suitable for scenarios where both accuracy and computational resources are critical considerations.

5. CONCLUSION AND FUTURE WORKS

Future research directions in automotive cybersecurity are driven by the need for more robust, adaptable, and scalable solutions to address emerging threats as vehicle technology continues to evolve. Key areas of focus include:

Enhanced AI and Machine Learning Models: Federated and Transfer Learning: As vehicles generate increasing amounts of data, federated learning offers privacy-preserving ways to leverage this data across vehicles without sharing sensitive information. Transfer learning can also help IDS systems adapt across different vehicle models with minimal re-training, making solutions more generalizable and faster to deploy.

Continual and Self-Supervised Learning: These learning approaches enable IDS models to continuously update and adapt to new attack patterns without extensive retraining. Self-supervised learning can leverage unlabeled data, which is often abundant, allowing IDS systems to learn without relying on pre-labeled attack data.

Integration of Quantum Computing for Processing Efficiency:

Quantum-Enhanced IDS: Quantum computing's potential for faster data processing and superior problem-solving capabilities makes it promising for automotive IDS. Future research could explore hybrid quantum-classical models for real-time intrusion detection, which could handle the high-speed requirements of connected and autonomous vehicle networks.

Quantum Cryptography for Secure Communication: Using quantum cryptography within V2X communications can offer unprecedented security by creating unbreakable

encryption standards, especially crucial for autonomous vehicles interacting within smart city infrastructure.

Improved Explainability and Transparency in AI Models:

Explainable AI for Automotive IDS: As AI models become more complex, future research will likely continue to improve the interpretability of these systems. Developing explainable AI techniques specific to the automotive domain can increase trust and usability, particularly for industry stakeholders who need to understand IDS decision-making processes.

Interactive Explainability Tools for Practitioners: Incorporating tools like SHAP and LIME, or creating automotive-specific interpretability methods, will allow developers and cybersecurity experts to understand IDS decisions. This is essential for refining models and achieving regulatory compliance.

Advanced Security Architectures and Protocols: Automotive Ethernet and New Protocol IDS: As vehicles shift from CAN to newer protocols like SOME/IP and Automotive Ethernet, there's a need for IDS systems tailored to these advanced communication structures. These IDSs will require high-speed data handling and should be capable of securing the unique vulnerabilities of these protocols.

Edge Computing for Real-Time IDS: Integrating edge computing allows IDS functions to occur at the vehicle level, reducing latency and enhancing real-time responses. Edge-computed IDS could be particularly useful for autonomous and semi-autonomous vehicles that require instant decision-making.

Security for Autonomous and V2X Communication Networks: Resilient V2X Security Solutions: As autonomous vehicles communicate with other vehicles (V2V), infrastructure (V2I), and networks (V2X), these systems become vulnerable to various remote attacks. Future research may focus on robust V2X IDS frameworks that can detect and mitigate attacks without compromising latency, which is critical for safety in autonomous driving.

Simulation and Testing Frameworks for V2X and Autonomous Systems: Developing simulation environments and testbeds for V2X and autonomous systems will be essential for verifying IDS effectiveness in real-world scenarios. This will help researchers model diverse attack vectors and evaluate security measures under conditions that resemble actual operational environments.

Human-Centric and Regulatory Focus: Privacy-Preserving IDS Models: With increasing data collection in vehicles, balancing security with user privacy is critical. Future IDS models could incorporate privacy-by-design principles, using techniques such as differential privacy to protect sensitive data.

Regulatory Compliance and Standardization: As standards evolve for autonomous and connected vehicles, future research will likely focus on aligning IDS technologies with emerging regulations like UN Regulation No. 155. Developing adaptable, standards-compliant IDS will be key to ensuring automotive cybersecurity solutions are ready for real-world deployment.

Utilizing Synthetic Data and Real-World Datasets: Synthetic Data for Model Training: Real-world attack data is scarce due to privacy and safety concerns, so synthetic data can fill this gap, allowing for large-scale IDS model training. Future research could improve the fidelity of synthetic CAN and V2X data, ensuring models trained on synthetic data are effective in real environments.

Cross-Vehicle and Cross-Platform Data Sharing: Developing data-sharing frameworks across manufacturers, while maintaining privacy and security, can improve IDS accuracy and resilience. Creating standardized datasets and collaboration between manufacturers will foster a unified approach to automotive cybersecurity.

In summary, the future of automotive cybersecurity will focus on advanced AI, scalable and secure network architectures, explainable and regulatory-compliant models, and secure frameworks for V2X communication. These directions aim to protect increasingly autonomous and connected vehicles in real-world, high-stakes environments, ensuring both user privacy and operational safety.

REFERENCES

- Ajiginni, Adeola Daniel (2020). Identification and Classification of IP Spoofing Man in the Middle Attack using Multilayer Perceptrons. Diss. Dublin, National College of Ireland.
- Abdu, M. A., et al. (2022). "Network Security in Vehicular Ad Hoc Networks." *Journal of Applied Research in Computer Science*, 9(4), 189-210.
- Akleyek, S. and Akşehir (2021). Z. D. "Denetleyici Alan Ağı'nın Sağlanması için Derin Öğrenme Tabanlı Saldırı Tespit Sistemleri Üzerinde Bir Derleme." *Avrupa Bilim ve Teknoloji Dergisi*,.
- Akşehir, Z. D., & Akleyek, S. (2021). "Denetleyici Alan Ağı'nın Sağlanması için Derin Öğrenme Tabanlı Saldırı Tespit Sistemleri Üzerinde Bir Derleme." *Avrupa Bilim ve Teknoloji Dergisi*, (27), 1038-1049.
- Alamleh, Amneh Hussein Mohd. (2023). "Benchmarking framework for IDS classifiers in term of security and performance based on multicriteria analysis."
- Aliyu, Ibrahim, et al. (2021) "A blockchain-based federated forest for SDN-enabled in-vehicle network intrusion detection system." *IEEE Access*, 9: 102593-102608.
- Aliyu, Ibrahim, et al. (2022). "Statistical detection of adversarial examples in blockchain-based federated forest in-vehicle network intrusion detection systems." *IEEE Access*, 10 :109366-109384.
- Alioğlu, Su Dilara (2019). "Siber saldırılar ve ülkelerin siber güvenlik politikaları." MS thesis, Lisansüstü Programlar Enstitüsü.
- Al-Kahtani, Mohammed Saeed (2012). "Survey on security attacks in vehicular ad hoc networks (VANETs)." 2012 6th international conference on signal processing and communication systems. IEEE.
- Alkhatib, Natasha, et al. (2021). "SOME/IP intrusion detection using deep learning-based sequential models in automotive ethernet networks." 2021 IEEE 12th annual

information technology, electronics and mobile communication conference (IEMCON). IEEE, 2021.

Allam, Fatima and Enrico Rossi. (2023). "Secure Communication in Vehicular Networks Using Blockchain." IEEE Transactions on Dependable Systems.

Almehdhar, Mohammed, et al. (2024). "Deep learning in the fast lane: A survey on advanced intrusion detection systems for intelligent vehicle networks." IEEE Open Journal of Vehicular Technology.

Alaya, Bechir, et al. (2024). "An ontological approach to the detection of anomalies in vehicular ad hoc networks." Ad Hoc Networks, 156: 103417.

Anjum, Afia, et al. (2022). "In-vehicle network anomaly detection using extreme gradient boosting machine." 2022 11th Mediterranean Conference on Embedded Computing (MECO). IEEE.

Anjum, F. et al. (2024). "Vehicular Ad Hoc Networks Security Challenges." Vehicular Technology Journal.

Andreica, Tudor, et al. (2022). "Android head units vs. In-vehicle ECUs: performance assessment for deploying in-vehicle intrusion detection systems for the CAN bus." IEEE Access, 10: 95161-95178.

Bajpai, Pranshu, et al. (2020). "Ransomware targeting automobiles." Proceedings of the Second ACM Workshop on Automotive and Aerial Vehicle Security.

Balaji, Prashanth, and Majid Ghaderi. (2021). "NeuroCAN: Contextual anomaly detection in controller area networks." 2021 IEEE International Smart Cities Conference (ISC2). IEEE.

Bariah, Lina, et al. (2015). "Recent advances in VANET security: a survey." 2015 IEEE 82nd vehicular technology conference (VTC2015-fall). IEEE.

Ben Lakhal, Nadhir Mansour, et al. (2020). "Controller area network reliability: overview of design challenges and safety related perspectives of future transportation systems." IET Intelligent Transport Systems, 14.13: 1727-1739.

Bendiab, Gueltoum, et al. (2023). "Autonomous vehicles security: Challenges and solutions using blockchain and artificial intelligence." IEEE Transactions on Intelligent Transportation Systems, 24.4: 3614-3637.

- Blevins, Deborah H., et al. (2021). "Time-based can intrusion detection benchmark." arXiv preprint arXiv:2101.05781.
- Bosch. (1991). CAN Specification, Version 2.0. Robert Bosch GmbH.
- Bouvin, Jean-Louis. (2019). Vers une version alternative à la suspension CRONE Hydractive. Diss. Bordeaux.
- Brooks, Richard R., et al. (2009). "Automobile security concerns." IEEE Vehicular Technology Magazine, 4.2: 52-64.
- Broy, M., Kruger, I. H., Pretschner, A., & Salzmann, C. (2007). Engineering automotive software. Proceedings of the IEEE, 95(2), 356–373.
- Burakova, Yelizaveta, et al. (2016). "Truck hacking: An experimental analysis of the SAE J1939 standard." 10th USENIX Workshop on Offensive Technologies (WOOT 16).
- Cao, Yulong, et al. (2019). "Adversarial sensor attack on lidar-based perception in autonomous driving." Proceedings of the 2019 ACM SIGSAC conference on computer and communications security.
- Chahal, Ashwani. (2018). In Situ Detection of Road Lanes Using Raspberry Pi. Diss. Utah State University.
- Checkoway, Stephen, et al. (2011). "Comprehensive experimental analyses of automotive attack surfaces." 20th USENIX Security Symposium (USENIX Security 11).
- Cho, Kyong-Tak, and Kang G. Shin. (2016). "Fingerprinting electronic control units for vehicle intrusion detection." 25th USENIX Security Symposium (USENIX Security 16).
- Chow, Man Chun, Maode Ma, and Zhijin Pan. (2021). "Attack models and countermeasures for autonomous vehicles." Intelligent Technologies for Internet of Vehicles. Cham: Springer International Publishing. 375-401.
- Chowdhury, Abdullahi, et al. (2020). "Attacks on self-driving cars and their countermeasures: A survey." IEEE Access 8 (2020): 207308-207342.
- Crespo, Jaime Pancorbo, Luis Guerrero Gómez, and Javier Gonzalo Arias. (2019). "Autonomous shipping and cybersecurity." Ciencia y tecnología de buques, 13.25: 19-26.

- Dahiya, Deepak. (2023) "Hybrid classifier strategy with tuned training weights for distributed denial of service attack detection." *Concurrency and Computation: Practice and Experience*, 35.4: e7547.
- Dorsaf Swessi, and Hanen Idoudi. (2021). "A comparative review of security threats datasets for vehicular networks." 2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT). IEEE.
- Duan, Zhenhai, Xin Yuan, and Jaideep Chandrashekar. (2006). "Constructing Inter-Domain Packet Filters to Control IP Spoofing Based on BGP Updates." *INFOCOM*.
- Eckermann, Erik. (2001). *World history of the automobile*. SAE International.
- El-Rewini, Zeinab, et al. (2020). "Cybersecurity challenges in vehicular communications." *Vehicular Communications* 23: 100214.
- El-Said, Mostafa, et al. (2021). "Building an impersonation attack and defense testbed for vehicle-to-vehicle systems." *Proceedings of the 22nd Annual Conference on Information Technology Education*.
- IEEE. (2010). *IEEE Standard for Wireless Access in Vehicular Environments (WAVE)-IEEE 802.11p*. Institute of Electrical and Electronics Engineers.
- ETSI. (2020). *Intelligent Transport Systems (ITS); Access layer specification for Intelligent Transport Systems operating in the 5 GHz frequency band (ETSI EN 302 663 V1.3.1)*. European Telecommunications Standards Institute.
- Fang, Siwei, et al. "Windowed Hamming Distance-Based Intrusion Detection for the CAN Bus." *Applied Sciences*, 14.7 (2024): 2805.
- Farsi, M., Ratcliff, K., & Barbosa, M. (1999). *An introduction to CAN (Controller Area Network)*. Springer.
- Feng, Shuo. *Cognitive Dynamic System for Connected and Autonomous Vehicles*. Diss. 2019.
- FlexRay Consortium. (2005). *FlexRay Communications System – Protocol Specification, Version 2.1*.
- Fu, Mengyi, et al. (2024). "IoV-BERT-IDS: Hybrid Network Intrusion Detection System in IoV Using Large Language Models." *IEEE Transactions on Vehicular Technology*.

- Garip, Mevlut Turker, et al. (2015) "Congestion attacks to autonomous cars using vehicular botnets." NDSS Workshop on Security of Emerging Networking Technologies (SENT), San Diego, CA.
- Gao, F. et al. "Intrusion Detection in CAN-FD Messages." *Sensors*, 24.11 (2024): 3461.
- Goodfellow, Ian J., et al. (2014). "Explaining and harnessing adversarial examples." arXiv preprint arXiv:1412.6572.
- Gundu, Ravalika, and Mina Maleki. (2022). "Securing CAN bus in connected and autonomous vehicles using supervised machine learning approaches." 2022 IEEE International Conference on Electro Information Technology (eIT). IEEE.
- Halder, Subir, Mauro Conti, and Sajal K. Das. (2021). "A holistic approach to power efficiency in clock offset based intrusion detection systems for Controller Area Networks." *Pervasive and Mobile Computing*, 73: 101385.
- Han, Mee Lan, Byung Il Kwak, and Huy Kang Kim. (2018). "Anomaly intrusion detection method for vehicular networks based on survival analysis." *Vehicular Communications*, 14: 52-63.
- Hassan, M. and J. Lee. (2024). "Machine Learning-based In-Vehicle Network Intrusion Detection." *IEEE Journal on Vehicular Technology*, 21(5).
- Ho, Patrick. (2023). "The Role of AI in Autonomous Driving Security." *International Journal of Vehicular Technology*.
- Hossain, Md Delwar, et al. (2020). "LSTM-based intrusion detection system for in-vehicle CAN bus communications." *IEEE Access*, 8: 185489-185502.
- IEEE. (2018). IEEE Standard for Ethernet. IEEE 802.3.
- Johnson, Sarah. (2023). "In-Vehicle Anomaly Detection: CAN Bus Approach." *Advances in Vehicular Technology*.
- Kang, Dong Mug, et al. (2021). "A study on attack pattern generation and hybrid MR-IDS for in-vehicle network." 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC). IEEE.
- Kang, Hyunjae, et al. (2021). "Car hacking and defense competition on in-vehicle network." *Workshop on Automotive and Autonomous Vehicle Security (AutoSec)*. Vol. 2021.

- Karopoulos, Georgios, et al. (2022). "Demystifying in-vehicle intrusion detection systems: A survey of surveys and a meta-taxonomy." *Electronics*, 11.7: 1072.
- Khatri, Narayan, Rakesh Shrestha, and Seung Yeob Nam. (2021). "Security issues with in-vehicle networks, and enhanced countermeasures based on blockchain." *Electronics*, 10.8: 893.
- Kim, Jaewoo and Patrick Sun. (2024). "Cybersecurity in Autonomous Vehicles: Review of Attacks and Defense." *Computers & Security*, 103(3).
- Kim, Kyounggon, et al. (2021). "Cybersecurity for autonomous vehicles: Review of attacks and defense." *Computers & Security*, 103: 102150.
- Kim, Shiho. (2018). "Blockchain for a trust network among intelligent vehicles." *Advances in Computers*, Vol. 111. Elsevier. 43-68.
- Kim, T., et al. (2023). "Anomaly Detection with LSTM-Autoencoder for In-Vehicle Networks." *Electronics*, 12.17: 3543.
- Khan, J. and I. Kim. (2024). "Blockchain in Vehicular Networks." *Journal of Emerging Blockchain Technologies*.
- Kreutz, D., Ramos, F. M. V., Verissimo, P. E., Rothenberg, C. E., Azodolmolky, S., & Uhlig, S. (2015). *Software-defined networking: A comprehensive survey*. *Proceedings of the IEEE*, 103(1), 14–76.
- Krutz, Ronald L., and Russell Dean Vines. (2023). *Cloud Security: A Comprehensive Guide to Secure Cloud Computing*. Wiley Publishing.
- Kushardianto, Nur Cahyono, et al. (2024). "Vehicular Network Anomaly Detection based on 2-Step Deep Learning Framework." *Vehicular Communications*, 100802.
- Kwak, Byung Il, Mee Lan Han, and Huy Kang Kim. (2021). "Cosine similarity based anomaly detection methodology for the CAN bus." *Expert Systems with Applications*, 166: 114066.
- Lampe, Brooke, and Weizhi Meng. (2023). "can-train-and-test: A New CAN Intrusion Detection Dataset." *2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)*. IEEE.
- Lampe, Brooke, and Weizhi Meng.(2024). "can-train-and-test: A Curated CAN Dataset for Automotive Intrusion Detection." *Computers & Security*, 140: 103777.

- Larson, Ulf E., and Dennis K. Nilsson. (2008). "Securing vehicles against cyber-attacks." Proceedings of the 4th Annual Workshop on Cyber Security and Information Intelligence Research: Developing Strategies to Meet the Cyber Security and Information Intelligence Challenges Ahead.
- Lee, Hyo-Keun. (2022). The Relationship between Innovative Technology and Driver's Resistance and Acceptance Intention for Sustainable Use of Automobile Self-Driving System. Sustainability. 14. 10129. 10.3390/su141610129.
- Li, Hao, et al. (2020). "CAN bus security: An in-vehicle solution." IEEE Transactions on Vehicular Technology, 2020.
- Lin, Hsiao-Chung, et al. (2022). "Using deep learning networks to identify cyber attacks on intrusion detection for in-vehicle networks." Electronics, 11.14: 2180.
- Liu, Hong. (2024). "Network and Communication Protocols in Cyber-Physical Systems." A Practical Guide on Security and Privacy in Cyber-Physical Systems: Foundations, Applications and Limitations, 25-88.
- Lokman, Siti-Farhana, Abu Talib Othman, and Muhammad-Husaini Abu-Bakar (2019). "Intrusion detection system for automotive Controller Area Network (CAN) bus system: a review." EURASIP Journal on Wireless Communications and Networking, 2019.1: 1-17.
- Lou, Yuan and Yuan Sun. (2022). "Detection of Cyber Attacks on CAN Networks Using Machine Learning Techniques." International Journal of Emerging Trends in Computing, 10(1).
- Luo, Feng, Xuan Zhang, and Shuo Hou. (2021). "Research on cybersecurity testing for in-vehicle network." 2021 International Conference on Intelligent Technology and Embedded Systems (ICITES). IEEE, 2
- Luo, Qian, et al. (2019). "Localization and navigation in autonomous driving: Threats and countermeasures." IEEE Wireless Communications, 26.4: 38-45.
- Madhav, AV Shreyas, A. Mohan, and Amit Kumar Tyagi. (2023). "IMPROVE: Intelligent machine learning based portable, reliable and optimal verification system for future vehicles." 2023 International Conference on Computer Communication and Informatics (ICCCI). IEEE.

- Ma, Haoyu, et al. (2022). "A GRU-Based Lightweight System for CAN Intrusion Detection in Real Time." *Security and Communication Networks*, 2022.1: 5827056.
- Mavikumbure, Harindra S., et al. (2023). "Dadae: Domain adversarial autoencoder based in-vehicle CAN anomaly detection." *IECON 2023-49th Annual Conference of the IEEE Industrial Electronics Society*. IEEE.
- McAfee. "Automotive Cyber Security Best Practices." *Auto Tech Review*, 5.8 (2016): 20-25.
- Mehedi, Sk Tanzir, et al. (2021). "Deep transfer learning based intrusion detection system for electric vehicular networks." *Sensors*, 21.14: 4736.
- Miller, Charlie. (2019). "Lessons learned from hacking a car." *IEEE Design & Test*, 36.6: 7-9.
- Monge, Anthony, and Todd Andel. (2023). "Use of Intrusion Detection Systems in Vehicular Controller Area Networks to Preclude Remote Attacks." *International Conference on Cyber Warfare and Security*.
- Mowla, Nishat I., Joakim Rosell, and Arash Vahidi. (2022). "Dynamic voting based explainable intrusion detection system for in-vehicle network." *2022 24th International Conference on Advanced Communication Technology (ICACT)*. IEEE.
- Munoz, C., et al. (2024). "Intrusion Detection Systems in CAN-based Vehicle Networks." *IEEE Access*.
- Nakamura, Shu, et al. (2021). "In-vehicle network attack detection across vehicle models: A supervised-unsupervised hybrid approach." *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE.
- Nandy, Tarak, et al. (2024). "A review of security attacks and intrusion detection in the vehicular networks." *Journal of King Saud University-Computer and Information Sciences*, 101945.
- Nathi, Girish. (2024). "Misbehaviour Detection in Vehicular Networks Using Unsupervised Algorithms." MS thesis, University of Windsor (Canada).
- Nguyen, Trieu Phong, Heungwoo Nam, and Daehee Kim. (2023). "Transformer-based attention network for in-vehicle intrusion detection." *IEEE Access*, 11: 55389-55403.

- Olson, Craig. (2023). "Artificial Intelligence in Autonomous Driving Vehicles." *International Journal of Research in Computer Science*, 15(1), 45-57.
- Olufowobi, Habeeb, Sena Hounsinnou, and Gedare Bloom. (2019). "Controller area network intrusion prevention system leveraging fault recovery." *Proceedings of the ACM Workshop on Cyber-Physical Systems Security & Privacy*.
- Owen, Travis. (2023). "Security Models in Autonomous Driving." *Journal of Vehicular Networking*.
- Ozguner, U., Baertlein, B., Cavello, C., Farkas, D., Hatipoglu, C., Lytle, S., ... & Walton, E. (1997). The OSU DEMO '97 Vehicle. In *Intelligent Transportation Systems Conference* (pp. 502-507). IEEE.
- Paddock, Micaiah C. "Best Practices in Preliminary Identification of Cybersecurity Attacks in Automotive Accidents: Generic Qualitative Inquiry." Diss. Capella University, 2022.
- Park, Seong and W. Liu. (2023). "Deep Learning-based Intrusion Detection for Autonomous Vehicles." *IEEE Systems Journal*.
- Park, Seunghyun, and Jin-Young Choi. (2020). "Malware Detection in Self-Driving Vehicles Using Machine Learning Algorithms." *Journal of Advanced Transportation*, 2020.1: 3035741.
- Park, Seunghyun, and Jin-Young Choi. (2020). "Hierarchical anomaly detection model for in-vehicle networks using machine learning algorithms." *Sensors*, 20.14: 3934.
- Park, Sung Bum, Hyo Jin Jo, and Dong Hoon Lee., (2023). "G-IDCS: Graph-based intrusion detection and classification system for CAN protocol." *IEEE Access*, 11: 39213-39227.
- Patel, Dev, et al. (2022). "Deep learning and blockchain-based framework to detect malware in autonomous vehicles." *2022 International Wireless Communications and Mobile Computing (IWCMC)*. IEEE.
- Rahman, S. and J. Park. (2023). "Challenges in Vehicular Cybersecurity." *Proceedings of the Vehicular Conference on Cyber Security*.
- Rajapaksha, Sampath, et al. (2023). "AI-based intrusion detection systems for in-vehicle networks: A survey." *ACM Computing Surveys*, 55.11: 1-40.

- Rajapaksha, Sampath, et al. (2024). "CAN-MIRGU: A Comprehensive CAN Bus Attack Dataset from Moving Vehicles for Intrusion Detection System Evaluation." Symposium on Vehicles Security and Privacy, Internet Society.
- Rajapaksha, Sampath, et al. (2022). "Keep the moving vehicle secure: Context-aware intrusion detection system for in-vehicle CAN bus security." 2022 14th International Conference on Cyber Conflict: Keep Moving! (CyCon). Vol. 700. IEEE.
- Rai, Ritu, Jyoti Grover, and Soumya Pandey. (2023). "Unveiling Threats: A Comprehensive Taxonomy of Attacks in In-Vehicle Networks." 2023 16th International Conference on Security of Information and Networks (SIN). IEEE.
- Remac, Milan. (2017). "The European Union Agency for Network and Information Security (ENISA)-Regulation (EU) No 526/2013 of the European Parliament and of the Council of 21 May 2013 concerning the European Union Agency for Network and Information Security (ENISA).".
- Rosell, Joakim, et al. (2021). "A frequency-based data mining approach to enhance in-vehicle network intrusion detection." Fast Zero' 21, Society of Automotive Engineers of Japan.
- Ross, Andrew, and Finale Doshi-Velez. (2018). "Improving the adversarial robustness and interpretability of deep neural networks by regularizing their input gradients." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 32. No. 1.
- Salek, M. Sabbir, et al. (2023). "A Hybrid Approach for Intrusion Detection in an In-vehicle Controller Area Network using Classical Convolutional Neural Network and Quantum Restricted Boltzmann Machine." Authorea Preprints.
- Salek, M. Sabbir, et al. (2023). "A novel hybrid quantum-classical framework for an in-vehicle controller area network intrusion detection." IEEE Access.
- Salek, M. Sabbir. (2023). Generative Defense Methods Against Cyberattacks for Autonomous Vehicles, Clemson University.
- Salek, M. Sabbir. (2023). Generative Neural Network-Based Defense Methods Against Cyberattacks for Connected and Autonomous Vehicles. Diss. Clemson University.
- Santa Barletta, Vita, et al. (2023). "Artificial Intelligence for Automotive Security: How to Support Developers in Automotive Solutions." 2023 IEEE International Conference

on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE). IEEE.

Santa Barletta, Vita, et al. (2023). "Black Mamba: A Multi-class Random Forest for Automotive Intrusion Detection."

Sankaranarayanan, R., et al. (2022). "Cluster-based attacks prevention algorithm for autonomous vehicles using machine learning algorithms." *Computers and Electrical Engineering*, 101: 108088.

Sarcinelli, R., et al. (2022). "Handling pedestrians in self-driving cars using image tracking and alternative path generation with Frenét frames." *Computers and Graphics (Pergamon)*, 84: 173–184.

Scholten, C. P. B. (2021). "Automatic detection of zero-day attacks in high-interaction IoT honeypots using static analysis techniques." MS thesis. University of Twente.

Seo, EunSeong, et al. (2023). "Adversarial attack of ML-based intrusion detection system on in-vehicle system using GAN." 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN). IEEE.

Sharma, Aekta, and Arunita Jaekel. "Machine learning based misbehaviour detection in VANET using consecutive BSM approach." *IEEE Open Journal of Vehicular Technology*, 3 (2021): 1-14.

Sheikh, Muhammad Sameer, Jun Liang, and Wensong Wang. (2020). "Security and privacy in vehicular ad hoc network and vehicle cloud computing: a survey." *Wireless Communications and Mobile Computing*.

Shukla, Siddharth. (2016). "Embedded security for vehicles: ECU hacking."

Sinai, Meital Ben, et al. (2014). "Exploiting social navigation." arXiv preprint arXiv:1410.0151.

Studnia, Ivan, et al. (2018). "A language-based intrusion detection approach for automotive embedded networks." *International Journal of Embedded Systems*, 10.1: 1-12.

Suhail, Sabah, et al. (2023). "ENIGMA: An explainable digital twin security solution for cyber–physical systems." *Computers in Industry*, 15: 103961.

- Sun, J. and T. Wu. (2024). "In-Vehicle Network Anomaly Detection Using AI." IEEE Access, 11.
- Sundfeldt, Frida, and Bianca Widstam. (2024). "Intrusion Detection for In-Vehicle CAN Communication Using Deep Neural Networks."
- Swessi, Dorsaf, and Hanen Idoudi. (2021). "A comparative review of security threats datasets for vehicular networks." 2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT). IEEE.
- Taeihagh, Araz, and Hazel Si Min Lim. (2019). "Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks." Transport Reviews, 39.1: 103-128.
- Tan, Jin-Ai Mary Anne, et al. (2010). "High prevalence of alpha-and beta-thalassemia in the Kadazandusuns in East Malaysia: challenges in providing effective health care for an indigenous group." BioMed Research International.
- Timčenko, Valentina, et al. (2024). "Attack Identification and Classification in V2X Scenarios." 2024 23rd International Symposium INFOTEH-JAHORINA (INFOTEH). IEEE.
- Tyagi, Parul, and Deepak Dembla. (2014). "Investigating the security threats in vehicular ad hoc networks (VANETs): towards security engineering for safer on-road transportation." 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE.
- Van Herrewege, Anthony, Dave Singelee, and Ingrid Verbauwhede. (2011). "CANAuth- a simple, backward compatible broadcast authentication protocol for CAN bus." ECRYPT Workshop on Lightweight Cryptography.
- Verma, Miki E., et al. (2024). "A comprehensive guide to CAN IDS data and introduction of the ROAD dataset." PLoS One, 19.1: e0296879.
- Vivek, Skanda, et al. (2019). "Cyberphysical risks of hacked internet-connected vehicles." Physical Review E, 100.1: 012316.
- Wang, H. and L. Shi. (2023). "Cyber Defense for Connected Vehicles: Strategies and Technologies." Sensors, 23(4).

- Wang, Jiadai, Jiajia Liu, and Nei Kato. (2018). "Networking and communications in autonomous driving: A survey." *IEEE Communications Surveys & Tutorials*, 21.2: 1243-1274.
- Wang, Kai, et al. (2023). "Effective In-vehicle Intrusion Detection via Multi-view Statistical Graph Learning on CAN Messages." *arXiv preprint arXiv:2311.07056*.
- Wang, Wei, et al. (2021). "I can see the light: Attacks on autonomous vehicles using invisible lights." *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*.
- Wikipedia. (2024). History of self-driving cars. Wikipedia. Retrieved November 15, 2024
URL: https://en.wikipedia.org/wiki/History_of_self-driving_cars
- Woo, Samuel, et al. (2019). "CAN ID shuffling technique (CIST): Moving target defense strategy for protecting in-vehicle CAN." *IEEE Access*, 7: 15521-15536.
- Wu, Shuhan, Siyong Li, and Wei Sun. (2023). "ConvIDS: A Convolutional LSTM Based Intrusion Detection Model for In-Vehicle CAN Bus." *2023 IEEE 6th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE)*. IEEE.
- Xiao, Junchao, Hao Wu, and Xiangxue Li.(2019). "Internet of things meets vehicles: sheltering in-vehicle network through lightweight machine learning." *Symmetry*, 11.11: 1388.
- Xiao, Junchao, Hao Wu, and Xiangxue Li. (2019). "Robust and self-evolving IDS for in-vehicle network by enabling spatiotemporal information." *2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*. IEEE.
- Xiao, Junchao, Hongbo Chen, and Fuli Zhong. (2024). "A Novel Feature Extraction Framework Using Graph Node Attention Network for In-Vehicle Networks Intrusion Detection." *IEEE Systems Journal*.
- Ying, Xuhang, et al. (2019). "Shape of the cloak: Formal analysis of clock skew-based intrusion detection system in controller area networks." *IEEE Transactions on Information Forensics and Security*, 14.9: 2300-2314.

- Yin, Yucheng, et al. (2024). "Work-in-Progress: CANGen: Practical Synthetic CAN Traces Generation Using Deep Generative Models." 2024 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). IEEE.
- Yoon, K., et al. (2024). "CAN Bus Security Threats and Defense Mechanisms." Journal of Automotive Safety and Security.
- Yurtsever, Ekim, et al. (2020). "A survey of autonomous driving: Common practices and emerging technologies." IEEE Access, 8: 58443-58469.
- Zenden, Ivo, et al. (2023). "On the Resilience of Machine Learning-Based IDS for Automotive Networks." 2023 IEEE Vehicular Networking Conference (VNC). IEEE.
- Zhe, Yu and Lee Yong. (2024). "A Survey on Blockchain for Vehicular Networks." International Journal of Emerging Technologies in Vehicular Communications.
- Zhao, R., et al. (2023). "Deep Learning for Autonomous Vehicle Security." IEEE Journal on Intelligent Transportation Systems.
- Zhou, Husheng, et al. (2020). "Deepbillboard: Systematic physical-world testing of autonomous driving systems." Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering.

BIOGRAPHICAL SKETCH

Ayşegül Kandefer holds a Bachelor of Science in Computer Engineering from Erciyes University, Kayseri, Turkey, where she studied from September 2016 to August 2020. She is currently engaged in postgraduate studies, pursuing a Master of Science in Computer Engineering at Galatasaray University, Istanbul, Turkey, since September 2021.

